Transcriptome profiling and longitudinal cohort studies of myositis subsets

Iago Pinal Fernández

A thesis presented for the degree of

Doctor of Philosophy in Bioinformatics

Faculty of Health Sciences

Open University of Catalonia

Spain

2020

To Noa

"Do or do not, there is no try!"

Yoda

Transcriptome profiling and longitudinal cohort studies of myositis subsets

Iago Pinal Fernández

Abstract

Inflammatory myopathies are a heterogeneous family of rare autoimmune diseases affecting multiple organs and systems, including the muscle, skin, lung, and/or the joints. Accurately defining its pathogenesis and classifying them correctly are key for understanding and managing these diseases.

This doctoral thesis has two main objectives. First, to develop a research framework to explore specific autoantibody-defined myositis subsets and determine if the autoantibodies are superior to the clinical classification systems to predict the phenotype of patients with myositis. Second, to define which are the most important pathogenic pathways and the specific expression profiles in the muscle tissue of patients with different types of myositis.

To achieve the first objective we developed a clinical research framework in the Johns Hopkins myositis cohort. This framework included designing a database to store and validate the data that was entered, and developing tools to parse information from clinical reports, merge tables for analysis, and automate table and graph creation. With this framework we explored specific autoantibody-defined myositis subsets and quantitatively compared the ability of autoantibodies to the 2017 EULAR/ACR classification standard to predict the phenotype of patients with myositis. To complete the second objective we performed RNA sequencing on 119 muscle biopsies of patients with different types of myositis and 20 controls. We studied the differential expression, performed pathway analysis and developed exploratory machine learning pipelines to define the specific expression profiles and pathogenic pathways in each disease subgroup.

With these studies we determined that the autoantibodies outperform current clinical criteria to predict the phenotype of myositis patients and discovered unique expression profiles in the muscle tissue of patients with different types of myositis.

Acknowledgements

I would like to express my profound gratitude to all the people and institutions that made possible this doctoral thesis.

First to Dr. Mammen, the director of this thesis and my mentor, for all his support, patience, continuous teaching and great advice. None of all the great things that have happened to me during the last years of my life would have been possible without him.

To Dr. Carrión Ribas, the tutor and co-director this Ph.D. for all his great guidance and encouragement in moving forward this project. Her infinite care and understanding made this process as easy as it could possibly be.

To Dr. Selva-O'Callaghan, great mentor and friend, for introducing me to the fascinating world of myositis, for his wise, and kind guidance along the initial steps of my academic career.

To Maria, for her loving comprehension during not just one, but two doctoral theses. Without her by my side I would not have dared to initiate this project or been able to finish it.

Finally, to Noa. She probably will not remember this period of her life, but without her good character and the boost of joy that she gives me every morning, doing this Ph.D. would have been infinitely more complicated.

Contents

Abbreviations

ADM	amyopathic dermatomyositis
AIC	akaike information criteria
AS	antisynthetase syndrome
AZA	azathioprine
BIC	bayesian information criteria
CK	creatine kinase
CT	computed tomography
DM	dermatomyositis
EMG	electromyography
ENMC	european neuromuscular center
FPKM	fragments per kilobase million
GSEA	gene-set enrichment analysis
HLA	human leukocyte antigen
HRCT	high-resolution computed tomography

IBM . . . . . inclusion body myositis

IFN . . . . . interferon

ILD . . . . . interstitial lung disease

IMNM . . . . . immune-mediated necrotizing myositis

LOWESS . . . . . locally weighted scatterplot smoothing

MAAs . . . . . myositis-associated autoantibodies

MAC . . . . . membrane attack complex

MHC . . . . . major histocompatibility complex

MRI . . . . . magnetic resonance imaging

MSAs . . . . . myositis-specific autoantibodies

MTX . . . . . methotrexate

PET . . . . . positron emission tomography

PM . . . . . polymyositis

TPM . . . . . transcripts per kilobase million

1. Introduction

1.1 A brief history of myositis

The inflammatory myopathies are a group of rare systemic autoimmune diseases characterized by variable involvement of the muscle, skin, lungs and/or joints.[1]

In the 19th century, Wagner described the first case of myositis.[2] Over the next years, these diseases were collectively named as polymyositis (PM), myositis universalis acuta or pseudotrichinosis.[3–9] At the end of the 19th century, Unverricht used the term dermatomyositis (DM) to define those patients showing both muscle and skin involvement.[10]

During the 20th century, several key clinical observations were made in patients with myositis. Thus, it was identified that neoplasms were common in patients with DM, that some patients had subcutaneous calcinosis and others characteristic erythematous lesions on the knuckles that were called Gottron's papules and are pathognomonic of DM.[6–9] Also, in 1940 it was found that DM was also present in children and that in juvenile cases it had an important vascular component which could lead to fatal gastrointestinal complications.[11] Finally, in 1956 it was discovered that myositis was also associated with interstitial lung disease.[12]

Years later, in 1967, it was observed that some patients with PM showed inclusions in the muscle that looked like those caused by myxovirus infections.[13] This type of myositis was named inclusion body myositis (IBM) and the first case series, from 1978, clearly defined that these patients were refractory to immunosuppressant med-

ications, showed important distal weakness, and experienced marked anterior thigh compartment muscle atrophy.[14]

In 1975 Krain[15] described some patients developing the characteristic skin features of DM without muscle involvement at the onset of the disease. Four years later, Pearson coined this phenomenon amyopathic dermatomyositis.[16] Given that Krain's patients eventually developed muscle involvement after an amyopathic onset, it was believed that this type of patients would invariably develop muscle weakness during the first years of their evolution. However, Sontheimer and cols. proved that a subset of dermatomyositis patients never developed clinically relevant muscle involvement.[17] Later on, the term amyopathic dermatomyositis was expanded to clinically amyopathic dermatomyositis to acknowledge the fact that some of these patients did not have clinically relevant muscle weakness but had minor muscle involvement detectable by elevation of muscle enzymes, electromyography (EMG), muscle biopsy or muscle magnetic resonance imaging (MRI).[18]

Between 1976 and 1985, a series of studies from Reichlin, Targoff, Nishikai, Hochberg, Arnett, and others established that many patients with myositis were positive for specific autoantibodies.[19–22] Years later, in 1991, Drs. Lori Love and Fred Miller suggested that these autoantibodies defined more homogeneous subsets of patients and could have classificatory value.[23]

Finally, the last major clinical group to be added to the field of myositis was the immune-mediated necrotizing myositis. The original description in 1991[24] included three cases of necrotizing myositis without significant inflammatory cell infiltration and microangiopathy with thickened "pipestem" capillaries, microvascular deposits of complement, and capillary depletion. Over the years, the vascular component of the syndrome has been deemphasized and this entity has been re-defined as presenting exclusive muscle involvement, often severe, with necrosis but without significant inflammatory infiltrates.

1.2 Classification

Given the heterogeneity of clinical manifestations and the multiplicity of serological groups in myositis, the classification of these patients has not been an easy

task and it is still currently a work in progress.

Originally, Bohan and Peter suggested a single set of criteria to diagnose all patients with myositis based on the presence of muscle weakness, elevation of muscle enzymes, an irritable pattern in the EMG, an inflammatory muscle biopsy and, in those with dermatomyositis, the presence of the characteristic skin rash.[25] These criteria were highly influential and have been the gold standard to develop new criteria ever since. The main criticism of this criteria was that they did not include IBM as a distinct category and, thus, all the IBM patients would be classified as PM.

Since 1975 there have been multiple systems proposed to classify patients with myositis. From those, perhaps the most influential were (Table 1.1):

1. Tanimoto's criteria in 1995[26] expanding the classificatory clinical features of Bohan and Peter's criteria.

2. Griggs' criteria for IBM in 1995[27] which required a combination of clinical, epidemiological and pathological features to establish the diagnosis of IBM. Unfortunately, those compliant with the epidemiological and clinical criteria but with incomplete biopsy findings were labeled as "possible" IBM. Studies tended to exclude IBM patients falling in Griggs' "possible" criteria even if they were very similar to "probable" IBM categorizations using subsequent classification proposals.

3. Targoff's ingenious criteria in 1997[28] including the myositis-specific autoantibodies (MSAs) in a manner that would allow the criteria to be automatically updated as new autoantibodies were discovered.

4. Badrising's criteria in 2000[29] modifying a previous set of criteria from the European neuromuscular center (ENMC) in 1997.[30] They require muscle weakness and a mononuclear inflammatory infiltrates with invasion of non-necrotic muscle fibers combined with either a set of clinical features or a combination of clinical and pathologic features. Includes two categories: probable and definite.

5. Sontheimer's criteria in 2002[18] proposing to consider amyopathic DM as a new

major clinical group.

6. Dalakas and Hohlfeld criteria in 2003[31] which were heavily reliant on muscle biopsy findings to classify myositis subtypes.

7. 119th ENMC criteria in 2004,[32] also heavily reliant on the muscle biopsy findings.

8. First Medical Research Council workshop on IBM in 2010[33] proposed a set of criteria where pathologically defined IBM was equal to Griggs' definite categorization. Alternatively, both clinically defined and possible IBM categories required the presence of either rimmed vacuoles, increased major histocompatibility complex (MHC)-I, or invasion of non-necrotic fibers by mononuclear cells plus a set of clinical features.

9. Pestronk's criteria in 2011,[34] proposing a novel classification exclusively based on histologic findings.

10. 188th ENMC IBM criteria in 2011[35] requiring a set of clinical and epidemiological features accompanied by the characteristic pathological features.

Importantly, in 2017, the joint EULAR/ACR criteria for myositis were published.[36] These criteria used a weighted score based on a set of epidemiologic, clinical, and laboratory variables to classify patients as myositis. Those patients classified as having myositis could be further subclassified in four different categories: PM/immune-mediated necrotizing myositis (IMNM), IBM, amyopathic dermatomyositis (ADM), DM, juvenile dermatomyositis, and juvenile myositis other than JDM.[36]

Notwithstanding the support of the two main rheumatology associations in the world, the 2017 EULAR/ACR myositis criteria has multiple methodological problems that limit its applicability in clinical research:

1. Although it was already accepted in the field that MSAs help to define phenotypically distinct sets of myositis patients, the 2017 EULAR/ACR myositis classification criteria included only anti-Jo1 autoantibodies (of note, the presence of anti-Jo1 antibodies was the variable with the highest value to classify a patient as myositis).

2. These criteria did not classify patients better than Targoff's criteria and based on the composition of their artificial cohort of patients it is not clear that they would outperform other proposals in real clinical settings (e.g. for a cohort with of patients with a 90% prevalence of myositis Bohan and Peter would classify correctly more patients than the 2017 EULAR/ACR classification).

3. The criteria were developed using a data-driven approach but had to be modified based on the opinion of experts because they did not fit the current state of the art.

4. A convenience sampling was used to develop the criteria, which makes the interpretation of the probability scores complicated and invalidates the predictive values reported in the manuscript.

5. The gold-standard to build the classification was the opinion of experts and, thus, the data-driven approach was in reality equivalent to the opinion of experts that they used as a starting point.

6. The complexity of the criteria made it hard to use and even more complicated to memorize.

7. Some of the variables lacked proper definition. For example, the lack of response to treatment in the classificatory tree.

8. Finally, the cutoff was selected by experts and there was no complete external validation of the criteria (only a validation of the sensitivity).

After the 2017 EULAR/ACR criteria were published, the ENMC sponsored a series of workshops to develop criteria for each one of the main groups of patients with myositis. Thus, in 2017 the ENMC published their proposal to classify IMNM[37] and in 2019 another to classify DM.[38] Importantly, these two sets of criteria heavily weighted on the importance of MSAs, and for the first time, allowed patients positive for MSAs to be diagnosed as IMNM or DM requiring only that they showed either muscle or skin involvement, respectively. Unfortunately, by being developed independently, both the IMNM and the DM criteria were not mutually exclusive and it would be feasible for a patient to fulfill both of them simultaneously, which complicates classificatory tasks.

Also, in 2018, a French group proposed a new set of criteria based on performing unsupervised multiple correspondence analysis and hierarchical clustering to aggregate patients in subgroups.[39] However, most of the study is focused on describing the characteristics of the artificial clusters that their methodology defined. Besides methodologic concerns on the technique that they use to determine the number of clusters,[40] the classification criteria that they propose is unrealistically simplistic, using only the presence of DM rash, presence of antisynthetase autoantibodies and finger flexor weakness to classify patients. Moreover, these criteria are questionable from a practical standpoint since, for example, they would classify a patient with anti-Jo1 autoantibodies and DM rash as the cluster corresponding to DM and not as an antisynthetase syndrome (AS).

Persistent unsolved controversies among experts regarding myositis classification include:

1. If the different criteria should be used exclusively for research studies or if they should aim to be also useful for clinical diagnosis.
2. If it should be only one set of criteria to fit all types of myositis or one for each type of myositis.
3. If only one set of criteria is used, if the criteria should only try to distinguish myositis vs. non-myositis patients or they should also define the myositis subgroup.
4. If criteria are developed individually for each type of myositis how can we ensure that they are mutually exclusive.

1.3 Major myositis subgroups, pathogenesis, and autoantibodies

In this section I will review the features of the main clinical subgroups currently recognized within myositis: IBM, IMNM, DM, overlap myositis (including the AS), and PM (Table 1.2). Most MSAs are generally associated with one of these broad clinical subgroups (except for IBM that does not have any known MSAs). Thus,

Table 1.1: Most influential classification criteria in myositis.

Authors	Year	Characteristics
Bohan and Peter[25]	1975	Most influential classification in myositis, based on the presence of a combination of clinical and laboratory findings.
Tanimoto[26]	1995	Expanded the classificatory clinical features of Bohan and Peter criteria.
Griggs[27]	1995	Most influential criteria for inclusion body myositis (IBM). Based on a combination of clinical, epidemiological and pathological features.
Targoff[28]	1997	Early criteria including the myositis-specific autoantibodies (MSAs).
Badrising[29]	2000	Modified version of a previous set of criteria from the ENMC in 1997.
Sontheimer[18]	2002	Criteria proposing to include amyopathic dermatomyositis (DM) as a new major clinical group.
Dalakas and Hohfeld[31]	2003	Heavily reliant on the muscle biopsy findings to do the diagnosis of most myositis subtypes.
Hoogendijk[32]	2004	Also heavily reliant on the muscle biopsy findings.
Hilton-Jones[33]	2010	IBM criteria. Pathologically defined IBM equal to Griggs', both clinically defined and possible IBM require the characteristic pathological features.
Pestronk[34]	2011	Classification exclusively based on muscle biopsy findings.
Rose[35]	2013	IBM criteria requiring a set of clinical and epidemiological features accompanied by the characteristic pathological features.
Lundberg[36]	2017	Current EULAR/ACR consensus criteria. Considerable methodological issues.
Allenbach[37]	2017	Immune-mediated necrotizing myositis criteria emphasizing the importance of MSAs.
Mariampillai[39]	2018	Concerns about the methodology. Unrealistically simplistic and questionably practical.
Mammen[38]	2020	DM criteria emphasizing the importance of MSAs.

within each section, I will review the most relevant autoantibodies that have been described so far. Moreover, in the last part, I will briefly discuss a group of autoantibodies that are not specific for a particular clinical phenotype but may act as disease modifiers.

1.3.1 Sporadic inclusion body myositis

As with other myositis subtypes, IBM patients show muscle weakness and are usually found to have elevated creatine kinase (CK) levels and myopathic EMG features. However, IBM patients are usually over 50 years-old while the other myositis tend to affect younger patients, including children.[1] Also, other myositis are more frequent in women, but in IBM men are affected twice as frequently as women.[1].

Regarding the pattern of muscle weakness, patients with IBM usually have distal weakness, including the finger flexors, wrist flexors, and ankle dorsiflexors, which is rarely prominent for other types of myositis.[1,41] Also, symmetric weakness is the rule in patients with other types of myositis, but many IBM patients have an asymmetric pattern of weakness.[1] Moreover, weakness can occur over weeks or months in other myositis, but the course of the disease in IBM is usually slow with weakness occurring over the course of years.[1] Finally, compared with other myositis, IBM patients have the most characteristic muscle MRI pattern, with severe involvement of the anterior thigh compartment.[42,43]

Importantly, there is no clear evidence that immunosuppression benefits patients with IBM whereas other myositis usually do respond to it.[1] Moreover, IBM patients also present progressive dysphagia,[1] that can lead to bronchoaspiration and can be studied using videofluoroscopy.

Also, unlike other types of myositis, IBM is not associated with any MSAs. Notwithstanding this, autoantibodies recognizing NT5C1a are present in 30-60% of IBM patients, but they are also found in 5-10% PM and 15-20% of DM patients and patients with lupus and Sjögren's syndrome.[44–47] Anti-NT5C1a autoantibodies have been associated with increased severity and mortality in these patients.[47,48] Additionally, a recent report has suggested that anti-NT5C1a autoantibodies may directly cause muscle damage.[49]

As for the muscle biopsies of patients with IBM, they characteristically include co-existing inflammation, abnormal protein aggregation, and mitochondrial dysfunction.[50] The inflammatory infiltrate is comprised of CD8+ T cells that surround and invade non-necrotic fibers (a.k.a. primary inflammation). Importantly, it has been recently found that these terminally differentiated T-cells express the surface marker KLRG1 and the carbohydrate epitope CD57 and that in most (22/38 [58%]) patients these cells meet criteria for T-cell large granular lymphocytic leukemia.[51,52] Although this association is yet to be confirmed, it would explain the refractoriness and advanced age of these patients.[51–53]

Rimmed vacuoles, best visualized by Gomori trichrome staining, are a hallmark of IBM muscle biopsies. Although some patients with hereditary myopathies also have rimmed vacuoles, their presence can help in distinguishing IBM from other myositis.[50] How IBM rimmed vacuoles are formed remains unknown. However, nuclear membrane proteins are found within rimmed vacuoles, suggesting they could be the remnants of degenerated myonuclei.[54,55] A more recent study revealed that proteins accumulating in rimmed vacuoles are related to protein folding and autophagy, suggesting that impaired autophagic function may be implicated in their formation.[56] Cytoplasmic accumulations also contain Congo red staining material ("amyloid"), p62 and TDP-43.[50] Although it is a widely spread notion that the cytoplasmic inclusions contain β-amyloid, studies specifically measuring the expression of this protein in the muscle of patients with IBM failed to prove this fact.[55,57]

An increased number of cytochrome oxidase negative muscle fibers and the presence of "ragged red fibers" suggest that mitochondrial damage plays a significant role in IBM.[50] Accordingly, a recent study showed that mitochondrial DNA is depleted and that mitochondrial fusion proteins are dysregulated in IBM muscle.[58] Furthermore, an increased frequency of mitochondrial DNA deletions has been reported in IBM muscle.[59]

Important in IBM, but relevant to all types of myositis, performing a muscle MRI to select the location of the muscle biopsy increases the diagnostic accuracy of the pathology.[60]

1.3.2 Immune-mediated necrotizing myositis

IMNM can be described as a distinct type of myositis characterized by proximal muscle weakness, exceptionally high muscle enzyme levels, myopathic EMG findings, and muscle biopsies showing necrosis and/or regeneration with minimal lymphocytic infiltrates and no perifascicular atrophy. Typical IMNM muscle biopsies also include MHC type I upregulation, M2-macrophage infiltration, and membrane attack complex (MAC) deposition on non-necrotic fibers.[61,62] Extramuscular manifestations are rare and generally mild when they occur.[63–65]

Around two-thirds of the patients with IMNM have autoantibodies recognizing either the SRP or the HMGCR proteins. However, 20% anti-SRP-positive and anti-HMGCR-positive patients have lymphocytic infiltrates in their muscle biopsies but are otherwise indistinguishable from their counterparts with necrotizing biopsies.[64,66,67]

Anti-SRP and anti-HMGCR myopathy share many features, including similar muscle biopsy findings, high CK levels, and minimal extramuscular manifestations.[61] Furthermore, in both, younger patients seem to have a more aggressive and refractory muscle disease.[63,65] However, differences between these two IMNM subtypes have been documented. First, anti-HMGCR myopathy is associated with statin exposure,[68] while anti-SRP myopathy is not associated with statins.[63,64] Second, anti-SRP-positive patients have more severe weakness and a higher number of necrotic muscle fibers than anti-HMGCR-positive patients.[61,63,67] Third, the presence of interstitial lung disease, although uncommon in both groups, is more frequent in those with anti-SRP autoantibodies (13-22%) than in anti-HMGCR (<5%).[61,63–65] Fourth, a single report suggested that anti-HMGCR myopathy and autoantibody-negative IMNM may have an increased risk of malignancy.[69] However, autoantibody-negative IMNM is an ill-defined entity and other cohorts of anti-HMGCR patients did not confirm an association of this autoantibody with cancer.[61,65] Fifth, several studies have confirmed DRB1*11:01 as an immunogenetic risk factor for developing anti-HMGCR myopathy (present in 70% of those with anti-HMGCR autoantibodies but only in 15% of the general population) and one report suggested that class II human leukocyte antigen (HLA) allele DRB1*08:03 is associated with anti-SRP

myopathy.[70–72] Finally, anti-HMGCR myopathy has rarely been associated with cardiac involvement.[61] In contrast, early cross-sectional studies in anti-SRP suggested a high prevalence of cardiac manifestations in these patients,[73,74] although this has not been confirmed in recent cohort studies.[61,63,64] In patients with suspicion of cardiac involvement, an electrocardiogram and an echocardiogram should be performed. Gadolinium-enhanced MRI can assess for active myocardial inflammation and in selected cases, an endomyocardial biopsy can confirm the diagnosis.[75]

The mechanisms underlying myofiber necrosis in IMNM remain to be elucidated. However, some clues have emerged. For example, given the MAC deposits on the surface of non-necrotic fibers, it has been proposed that anti-SRP and anti-HMGCR autoantibodies could be directly pathogenic.[67] In this regard, a recent study suggested that these autoantibodies may induce muscle atrophy, increase levels of reactive oxygen species and cytokines (e.g., tumor necrosis factor and IL-6), and impair myoblast fusion (by decreasing the production of IL-4 and IL-13) of cultured muscle cells.[76] However, these IMNM-associated autoantibodies did not induce necrosis and further studies may be needed to show that they are pathogenic in vivo.[76]

1.3.3 Dermatomyositis

DM patients classically show proximal muscle weakness and characteristic cutaneous manifestations that develop over weeks to months. However, some patients with DM rash have little or no muscle involvement as demonstrated by lack of weakness, muscle enzyme elevation, MRI, EMG, or muscle biopsy findings. Clinically amyopathic DM is often considered as a different subtype of myositis[18] but, for simplicity, we will include them in this section.

The pathognomonic skin rash of DM includes a violaceous periorbital, often edematous, rash (i.e., heliotrope rash) as well as erythematous lesions on the extensor surfaces of the joints (i.e., Gottron's papules). Usually, muscle enzymes are elevated and the EMG reveals a myopathic pattern (myopathic motor units with fibrillations and spontaneous sharp waves). As in other types of myositis, the MRI in DM may reveal intramuscular T2 hyperintensities caused by muscle inflammation and/or necrosis.[43] In addition, DM patients often have T2 hyperintensities around individual

muscles as a result of fascial involvement, a feature seen less frequently in other myositis.[43]

Perifascicular atrophy is a highly specific feature of muscle biopsies from DM patients (specificity >90%), but it lacks sensitivity (25-50%).[77,78] Limited data support that perifascicular MX1 (human myxovirus resistance protein 1) and RIG-1 (retinoic acid-inducible gene I) expression have higher diagnostic sensitivity (71% and 50%) than perifascicular atrophy in DM.[77,79] Additionally, DM biopsies often have cellular infiltrates consisting predominantly of CD4+ T cells, plasmacytoid dendritic cells, B cells, and macrophages.[50] These cells often surround medium-sized blood vessels (perivascular inflammation) and invade the perimysium.[50] However, up to 16% of DM biopsies lack infiltrates and have prominent necrosis that is pathologically indistinguishable from IMNM.[78] Capillary loss can occur and it can be detected deposition of MAC and presence of microtubular inclusions on intramuscular capillaries.[50] Furthermore, as in other myositis, there is usually upregulation of class I MHC on the sarcolemma of muscle fibers. In DM patients, class I MHC upregulation, and other pathological findings (e.g., myofiber de/regeneration and necrosis) may be especially prominent in perifascicular regions.[50]

Approximately 70% of DM patients have one MSA.[80] Each DM-specific autoantibody is associated with a unique clinical phenotype. Thus, autoantibodies recognizing Mi2 have been associated with "classic" DM features including proximal muscle weakness and severe skin manifestations.[81] DM patients with autoantibodies recognizing nuclear matrix protein (NXP)-2 are more likely than other DM patients to present with both proximal and distal muscle weakness, subcutaneous edema, and/or dysphagia.[82] Furthermore, anti-NXP2-positive patients are more prone than other DM patients to develop calcinosis, which are painful deposits of calcium in the soft tissues, often refractory to immunosuppressant treatment.[82] DM patients with anti-transcription intermediary factor (TIF)-1γ and, to a lesser degree, those with anti-NXP2 autoantibodies are at increased risk of malignancy within three years of their diagnosis; as such, these patients may require comprehensive cancer screening.[82–84] The traditional approach to cancer screening is to perform a complete physical examination, general laboratory tests, tumor markers, thoracoabdominal computed tomography (CT), and a gynecologic exam, including ultrasonography

and mammography plus any other age and gender-appropriate screening tests. Alternatively, a single positron emission tomography (PET)-CT has been shown to have an equivalent sensitivity for detecting malignancy as the traditional approach.[85]

Patients with DM and autoantibodies recognizing small ubiquitin-like modifier activating enzyme (SAE) or melanoma differentiation-associated gene 5 (MDA5) tend to have more significant skin than muscle involvement.[86–89] Along with the typical skin manifestations of DM, anti-MDA5-positive patients are prone to develop ulcers, often on the flexor surface of the digits and palms.[89] Most anti-MDA5 patients are hypo or amyopathic.[87–89] Furthermore, unlike patients with other DM autoantibodies, anti-MDA5-positive patients frequently develop a rapidly progressive and sometimes lethal form of interstitial lung disease (ILD).[87,88] All myositis patients with suspicion of ILD should initially be evaluated using pulmonary function tests (including CO diffusion and ins/expiratory pressures) and a chest high-resolution computed tomography (HRCT). ILD monitoring should rely on periodical pulmonary function tests and subsequent HRCT should be restricted to evaluating those with evolving pulmonary issues.[1]

Some combination of genetic risk factors and environmental exposures are presumably required to trigger DM. Indeed, several immunogenetic risk factors, including certain class II HLA alleles, have been implicated in DM pathogenesis.[90] Interestingly, ultraviolet light exposure is also a known risk factor for developing DM.[91] However, the majority of people with known genetic risk factors, even those with high ultraviolet light exposure, never develop DM. An increased number of mutations and loss of heterozygosity in TIF1 genes from tumors in anti-TIF1γ-positive DM patients have recently been reported.[92] This observation suggests the possibility that mutations in TIF1 genes may generate neoantigens that could trigger autoimmunity by means of molecular mimicry.

Whatever the cause, once a patient has developed DM, it's unclear what mechanisms maintain muscle damage and weakness. Notwithstanding this, there is strong evidence that the interferon (IFN) pathway is relevant to DM pathogenesis.[93] Specifically, a marked overexpression of IFN-inducible genes has been demonstrated in the muscle,[93] peripheral blood,[94,95] and skin[96] of DM patients. Moreover, the expression

levels of IFN-inducible genes correlate with indicators of DM disease activity.[94,95] The presence of plasmacytoid dendritic cells, potent sources of interferon, along with the increased expression of type-I-interferon-inducible proteins in the perifascicular area, suggest that interferon could somehow mediate perifascicular atrophy.[77,93]

1.3.4 Overlap myositis

Autoimmune myopathy may also occur in patients presenting features of other autoimmune diseases, such as lupus, rheumatoid arthritis, Sjögren's syndrome or systemic sclerosis.[97–99] Many of these patients also have autoantibodies that are associated with characteristic phenotypes.[97–99]

The most representative form of overlap myositis is the antisynthetase syndrome (AS), with autoantibodies targeting the aminoacyl tRNA synthetases, enzymes that conjugate an amino acid to its cognate tRNA.[100,101] Those recognizing histidyl-tRNA synthetase (anti-Jo1), threonyl-tRNA synthetase (anti-PL7), and alanyl-tRNA synthetase (anti-PL12) are the most common.[100,101] Patients with any one of these autoantibodies can be defined as having AS and typically present with one or more of the following features: myositis, ILD, arthritis, Raynaud's syndrome, fever, and hyperkeratotic radial fingers lesions known as "mechanic's hands".[1] AS patients may also have skin rashes similar to DM.[1] Of note, not all AS patients have muscle weakness. Indeed, whereas 90% anti-Jo1 patients have myositis, up to 50% anti-PL12 patients present with ILD but no muscle involvement.[100] Furthermore, anti-Jo1-positive patients have more severe weakness while anti-PL7 and anti-PL12 have more severe lung involvement.[100,101]

When present, myopathic AS features are very similar to DM, including proximal muscle weakness, elevated muscle enzymes, and myopathic EMG.[1] AS patients often have intramuscular T2 MRI hyperintensities, but a specific MRI pattern has not been described.[102] Muscle biopsies may reveal perifascicular atrophy similar to DM. However, compared to DM, AS may have an increased number of perifascicular necrotic fibers.[103,104] Furthermore, it has been reported that AS biopsies show nuclear actin aggregation, an electron microscopy feature that is not seen in other myositis.[105] To date, very little is known about what triggers and maintains autoim-

munity in the AS.

Anti-PM/Scl autoantibodies are associated with myositis in patients with systemic sclerosis.[97] Similarly, anti-U1RNP-positive mixed-connective tissue disease patients and anti-Ku may have myositis along with additional systemic sclerosis (e.g. sclerodactyly) and lupus (e.g. glomerulonephritis or serositis) features.[98,99]

1.3.5 Polymyositis

PM is defined by the presence of muscle weakness, elevated CK levels, myopathic EMG features and an inflammatory muscle biopsy with none of the characteristic accompanying features of the other abovementioned groups. Many patients previously classified as having PM could now be considered to have AS without a rash, IMNM, or IBM based on characteristic clinical manifestations, serological features, and muscle biopsy findings.[106–108] Even if some true PM patients may still exist,[109] PM remains a diagnosis of exclusion and PM patients should be closely monitored for new clinical features suggesting alternative diagnoses.

1.3.6 Myositis-associated autoantibodies

Myositis autoantibodies have been traditionally classified as MSAs and myositis-associated autoantibodies (MAAs) depending on their association to "pure" forms of myositis (e.g. anti-Mi2, anti-NXP2, or anti-SRP) or to myositis accompanied by features of other autoimmune diseases (e.g. anti-PM/Scl, anti-Ku, or anti-U1RNP) respectively.[110] However, the term MAAs has also been used for a group of autoantibodies that appear concomitantly with others in patients with very different clinical phenotypes, often acting as disease modifiers. Among this type of MAAs the most common is anti-Ro52,[100,111–113] but also anti-FHL1,[114] anti-cortactin,[115] anti-PUF60,[116,117] or the abovementioned anti-NT5c1a.[44–48]

Anti-Ro52 often co-occurs with anti-Jo1 autoantibodies. Patients with both antibodies have more frequent and severe ILD, poorer response to immunosuppressive drugs and decreased survival.[100,111–113] Moreover, high anti-Ro52 titers are associated with more severe ILD, myositis and joint involvement in adult anti-Jo1 patients.[100,111–113] Alternatively, anti-FHL1 is present in up to 25% of patients with

Table 1.2: Clinical features and grouping of the most main myositis-specific autoantibodies.

Group	Muscle	Lung	Skin
Inclusion body myositis	+++	∅	∅
Immune-mediated necrotizing myositis			
Anti-SRP	+++	+	∅
Anti-HMGCR	+++	∅	∅
Dermatomyositis			
Anti-Mi2	++	∅	++
Anti-NXP2	++	∅	++
Anti-TIF1	+	∅	++
Anti-SAE	+	∅	++
Anti-MDA5	+	+++	+++
Overlap myositis			
Antisynthetase syndrome			
Anti-Jo1	++	++	+
Anti-PL7	++	+++	+
Anti-PL12	+	+++	+
Anti-Pm/Scl	+	+	+
Anti-Ku	+	+	+
Anti-U1RNP	+	+	+

myositis and is associated with the presence of muscle atrophy, dysphagia, pronounced muscle fiber damage and vasculitis.[114] As for the anti-cortactin antibodies, they were found in 12% of the patients with myositis, were more common in PM or IMNM, and were not associated with any specific clinical feature.[115] Anti-PUF60 antibodies were found in 15% of patients with myositis, were associated with anti-TIF1g autoantibodies, and were associated with a higher prevalence of skin ulcerations.[116,117] Finally, as it was mentioned earlier, anti-NT5c1a is found predominantly in patients with IBM and has been associated with increased severity and mortality in these patients.[44–48]

The current terminology of MSAs and MAAs is rather confusing since the MAAs category groups together two different populations of autoantibodies, those that are specific of a certain phenotype and those that are not. Also, establishing which phenotypes constitute "pure" forms of myositis is often complicated and of questionable practical importance. An alternative way to classify autoantibodies conceptually and solve the above-mentioned issues would be to consider them as disease-specific or disease-independent. Disease-specific antibodies would include all the MSAs and those MAAs that are usually mutually exclusive (the presence of one is associated with the absence of others) and linked to a specific phenotype (e.g. anti-PM/Scl, anti-Ku, or anti-U1RNP). Alternatively, disease-independent antibodies would be those that may act as disease modifiers but are not linked to any particular combination of clinical features (e.g. anti-Ro52, anti-FHL1, anti-cortactin, anti-PUF60, or anti-NT5c1a).

2. Motivation and Objectives

2.1 Motivation

The inflammatory myopathies are a heterogeneous family of rare diseases affecting multiple organs and systems, including the muscle, skin, lung, and/or the joints. Accurately defining its pathogenesis and classifying them are key to understand and manage these diseases.

Studying myositis pathogenesis based on an incorrect classification of its different individual diseases may lead to incorrect conclusions. For this reason, we believed that it was necessary to approach defining the myositis classification and its pathogenesis in parallel.

We hypothesize that disease-specific autoantibodies in myositis are tightly associated with the cause of the disease. If this is the case, different autoantibody groups will show distinct activation of pathogenic pathways, clinical manifestations, prognosis, and response to treatment.

Moreover, other autoantibodies not associated to a specific clinical syndrome may act as disease modifiers and increase the risk of developing certain clinical manifestations.

Finally, even if no disease-specific autoantibodies have been identified in IBM, its clinical and epidemiological features are characteristic enough that we can study its pathogenesis separately from other types of myositis.

2.2 Objectives

- To develop a research framework to study longitudinal cohorts of specific myositis subgroups. With this framework we will:
  - Determine if the autoantibodies are superior to current clinical classification systems to predict the phenotype of patients with myositis.
  - Study the characteristic clinical features, prognosis, and response to therapy of patients with different myositis autoantibodies.
- Define which are the most important pathogenic pathways and the specific expression profiles in the muscle tissue of patients with different types of myositis.

3. Methods

In this section, I will review the methodology that was used in the two sections of this doctoral thesis. First, I will explain the components and structure of the research framework that was developed to organize and analyze efficiently the data of the Johns Hopkins Myositis Center longitudinal cohort study. Second, I will detail the most relevant statistical techniques that were used to perform the series of epidemiologic analysis of this clinical cohort. Finally, I will summarize the design of the Myositis muscle RNA sequencing studies, focusing on the bioinformatic analysis of the resulting sequencing data that was obtained. The detailed information about the methods of each of the individual studies that compose this doctoral thesis is included in Appendix A.

3.1 Longitudinal cohort study framework

3.1.1 The Johns Hopkins Myositis Center longitudinal cohort study

In 2007 the Johns Hopkins Hospital founded a monographic center to treat and do research in patients with different types of myositis. It is a multidisciplinary unit combining the expertise of Rheumatologists, Neurologists, Pulmonologists, and Physical Therapy specialists. Over the years, more than two thousand patients suspected of having an inflammatory myositis have been evaluated and followed over time, becoming the largest longitudinal myositis cohort in the world.

3.1.2 Database design

In 2013, a research database was created to organize the information that was available for the patients of the myositis cohort. The first challenge was to collect all the information, that was originally spread over different files with data of variable quality, included inconsistencies, duplicated observations, was not normalized, and relied heavily on text descriptions.

To organize this heterogeneous collection of data, first, we had to choose what would be the main blocks of information that were relevant from a myositis research standpoint and how to organize them in a way that will be useful not just for a single project but for many studies to come.

It was decided that the minimum set of data that we needed included a unique identifier for each patient and epidemiological data, like the gender, date of birth, or the race. Moreover, it was necessary to include, among others: the general phenotype of the patients; the family history; the medications; the clinical features as they were longitudinally recorded; the results of the laboratory tests, MRI, muscle biopsy, pulmonary function tests, and chest CT. Finally, we would need to include the information of each one of the autoantibodies, both the consensus interpretation and all the individual serologic results.

Also, we had to select the best technical solution to store the data according to the expected level of usage, limited time to maintain it, and availability of the software in the terminals of the different users. It was decided that using Microsoft Access fulfilled the required criteria for its initial intended usage (Figure 3.1). It could be used concurrently by up to 20 users, both the front-end and the back-end could be easily maintained by a single researcher, and it had institutional support in all the user terminals. The database was designed so that only a limited number of people would have access to the back-end, but all the researchers in the protocol could view, query, and modify the data in the front-end. Over the years this framework stopped being compliant with the institutional Hopkins policies and the back-end was migrated to a Microsoft SQL Server, with ongoing plans for the front-end to be moved to a web framework.

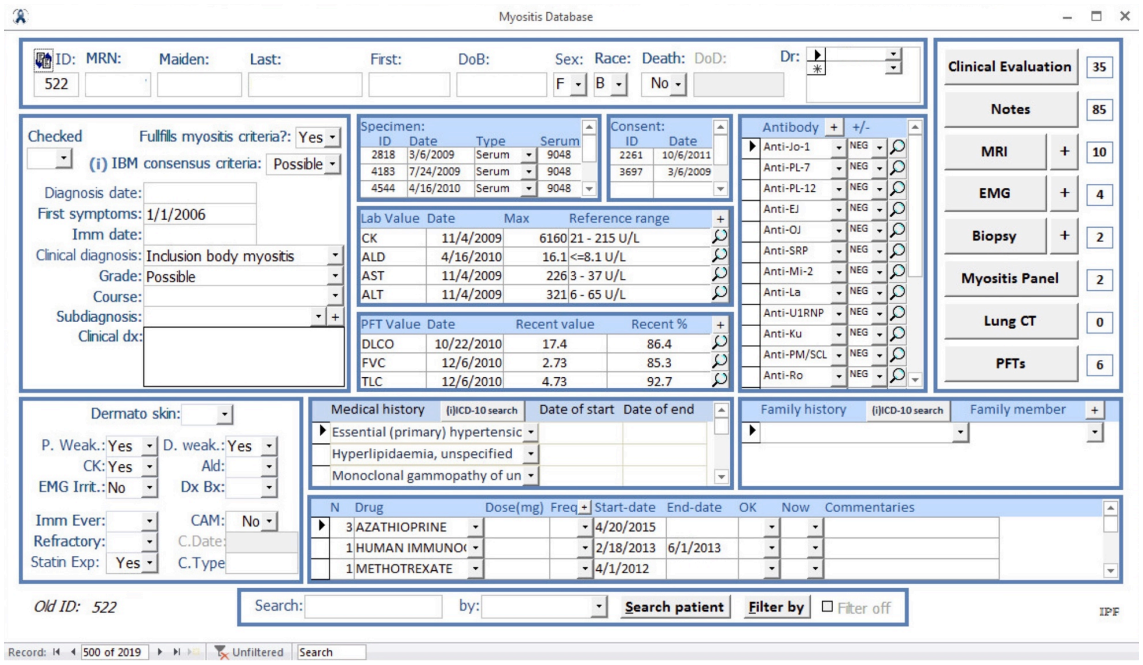


Figure 3.1: Appearance of the front-end of the myositis database.

A central database with un-duplicated information was shared by all the concurrent research projects, but each researcher was able to work in different subsets of patient and specific sets of variables by defining filters on each project's front-end. By having a central data repository, different projects would review and enter new data for themselves but, by doing so, they would be helping to curate the data for other concurrent and future projects, increasing the overall productivity of the group.

3.1.3 Data mining

Once the database was built, it was obvious that importing certain types of data could be automated. Originally, it was not feasible to access the back-end of the Hopkins electronic patient record system directly, but users had access to the front-end of the application to collect the data of those patients consented for the study. Also, there was no application programming interface to automate the data mining process. Thus, the only option, if we wanted to automate importing the thousands of registries, would be to do it through mouse and keyboard automation. A first crude data import was successful at automating the extraction of all the relevant clinical and laboratory data (e.g. creatine kinase, aldolase, AST, ALT,

and autoantibody reports). This approach was further perfected through a script with modules to include each one of the sequences of key-strokes and mouse-clicks necessary to import each section of the data for every patient. Finally, contingency systems were put in place to reset and tag abnormal data imports when the system did not get the correct set of information.

A portion of the data required expert interpretation from researchers trained to evaluate medical jargon and familiarized with the type of disease. The front-end of the database was progressively modified to simplify these data-entry tasks and include quality control variables that could be used to detect errors during the process.

3.1.4 Data wrangling

Once the information was available in the database, a key part of any data analysis effort was to clean the raw data to adapt it to the specific analyses. It was key to automate these steps in order to be able to update the results if new data was added, be able to design the analysis at the same time that the data was being collected, and reuse the data cleaning steps to optimize productivity. Most analyses of clinical data were performed using Stata and thus, Stata scripts used the raw Access tables as the starting point of the process, which facilitated updating the data by replacing the raw tables. It was first necessary to integrate all the different tables in a single comprehensive dataset while allowing for enough flexibility so the process of integrating the tables was modular. A set of functions was built to merge each auxiliary table to the main one including the epidemiological features of the patients. Also, the variables were labeled and calculated fields were generated. Finally, a limited set of "working datasets" were built for the different families of analysis in order to avoid delays each time it was necessary to modify a piece of the code if the data to be used had not changed.

3.1.5 Data analysis

Univariate analysis

Univariate analysis is often the first step in an exploratory analysis. Hypothesis contrast using Fisher's exact test, Chi-squared, Student's T-tests, and Wilcoxon's rank-sum test is the bread and butter of any epidemiological study. However, performing these tests is often extremely labor-intensive due to the massive amount of comparisons that may be needed when multiple variables are being explored simultaneously. Also, generating publication-quality tables is often a source of productivity loss in research. For this reason, a Stata program called *table_ compare* was built to simplify performing the hypothesis testing, deciding automatically when to perform Fisher's or Chi-square tests and manually to indicate if continuous variables should be considered normally-distributed or nonparametric tests should be applied. Moreover, this tool allows performing multivariate analysis, handles paired data, and can export the results ready for publication. This tool allowed us to speed up performing and updating the different analyses.

Multilevel regression models

A complication of observational longitudinal cohort studies, when performed in a clinical practice setting, is that each patient has a different number of observations (e.g. visits or laboratory determinations). In order to avoid bias, it is necessary to use techniques to take into account this uneven length of follow-up. A simplistic solution would be averaging the periods of observation and adjusting for the time of follow-up, but this would result in an unnecessary loss in statistical power. A valid alternative in this situation is to use a family of regression methods called multilevel regression models or random effects mixed models. These models include the fixed effects of conventional regression analysis but then allow for random components grouped in categories (e.g. each individual patient will have a different evolution of the CK levels over time). Multilevel regression models allow for random intercepts (e.g. each patient having higher or lower overall CK) and random slopes (e.g. each patient can have a faster or slower decrease in the CK levels). Strictly speaking, the full model with random intercepts and random slopes is just necessary if it is

significantly different from the model with random intercepts, and the model with random intercepts only if it is significantly different from the standard regression model. However, doing these model evaluations adds a layer of complexity that may be eliminated by using the most complex model with random slopes and random intercepts that will control for all possible variation caused by the heterogeneity of the patient population and the differential follow-up.

Factor analysis of mixed data

To model the phenotype of the patients we used factor analysis of mixed data. The phenotype is a latent variable, i.e., a variable that cannot be directly observed but has to be inferred through a mathematical model.

Factor analysis of mixed data is a principal component method dedicated to exploring data containing both continuous and categorical variables. The continuous variables are scaled to unit variance and the categorical variables are transformed into a disjunctive data table and then scaled using the specific scaling of multiple correspondence analysis. This ensures to balance the influence of both continuous and categorical variables in the analysis. It means that both variables are on an equal foot to determine the dimensions of variability. We used the package FactoMinerR v.2.1 to perform the factor analysis of mixed data and factoextra v. 1.0.6 to obtain the scree plots and the variable weight plots.

As input for the factor analysis of mixed data we selected a set of clinical, epidemiological, and laboratory parameters that were: 1) well documented in the literature to be associated with the phenotype of patients with myositis, 2) systematically collected in our cohort, and 3) well defined. Thus, we included:

1. Epidemiologic variables: Gender, race, age at onset.
2. Clinical variables: presence or absence during the course of the disease of muscle weakness, interstitial lung disease, arthritis, heliotrope or Gottron's rash, calcinosis, Raynaud's phenomenon, mechanic's hands, dysphagia, and fevers.
3. Laboratory values: Maximum CK, presence of anti-Ro52 autoantibodies.

Anti-Ro52 autoantibodies were included because they are associated with the severity of the disease and with specific clinical features in patients with myositis.[100,111–113]

The detailed distribution of the muscle weakness, biopsy features, MRI patterns, or EMG findings were excluded from this analysis because they were not available for all patients, and restricting the sample size could bias the study.

Model comparison

Both the akaike information criteria (AIC) and the bayesian information criteria (BIC) are estimators of out-of-sample prediction error and thereby the relative quality of statistical models for a given set of data.[118,119] Given a collection of models for the data, AIC and BIC estimate the quality of each model, relative to each of the other models. Thus, they provide a means for model selection.

The formula for AIC is:

$$AIC = 2 - 2\ln(\hat{L})$$

Whereas BIC is formally defined as:

$$BIC = k\ln(n) - 2\ln(\hat{L})$$

Where:

$\hat{L}$=the maximized value of the likelihood function of the model

n=the number of observations

k=the number of parameters estimated by the model

AIC and BIC are founded on information theory. When a statistical model is used to represent the process that generated the data, the representation will rarely be exact; so some information will be lost by using the model to represent the process. AIC and BIC estimate the relative amount of information lost by a given model: the less information a model loses, the higher the quality of that model.

In estimating the amount of information lost by a model, AIC and BIC deal with the trade-off between the goodness of fit of the model and the simplicity of the model. In other words, AIC and BIC deal with both the risk of overfitting and the risk of underfitting.

Given a set of candidate models for the data, the preferred model is the one with the minimum AIC or BIC value. Thus, AIC and BIC reward goodness of fit (as assessed by the likelihood function), but it also includes a penalty that is an increasing function of the number of estimated parameters. The penalty discourages overfitting, which is desired because increasing the number of parameters in the model almost always improves the goodness of the fit.

The main difference between AIC and BIC is that the penalty for the number of parameters is larger for BIC. Thus, if 'k 'is the number of parameters and 'n' is the number of observations, the penalty for AIC is $2k$, whereas for BIC is $\ln(n)k$.

Various authors have suggested guidelines to interpret the magnitude of the differences in AIC and BIC between two models (ΔAIC and ΔBIC respectively). Thus, Burnham and Anderson[120] suggested that models having $\Delta\text{AIC} \leq 2$ have substantial support (evidence) to believe that are equivalent, those in which $4 \leq \Delta\text{AIC} \leq 7$ have considerably less support, and models having $\Delta\text{AIC} > 10$ have essentially no support. Alternatively, Raftery[121] suggested that the evidence that two models are not equivalent would be weak with a ΔBIC between 0-2, positive with ΔBIC between 2-6, strong with ΔBIC between 6-10, and very strong with ΔBIC over 10. Based on this we selected a threshold of ΔAIC and ΔBIC of 10 to consider one model superior to others.

Graphical analysis of longitudinal data

A key part of any scientific study is to represent the results visually in a way that is easy and quick to understand even for non-experts. To do so in an efficient manner, functions were build to make the most common graphical analysis, including individual and aggregated patient evolution plots and Kaplan-Meier curves. A useful graphical method to represent the evolution of nonlinear parameters longitudinally is the locally weighted scatterplot smoothing (LOWESS). LOWESS fits simple models to localized segments of the data to build up a function that describes the deterministic part of the variation in the data, point by point. The advantage of LOWESS over other methods (e.g. quadratic regression) is that it does not need to specify a function to fit the model to the data, making it simple and flexible for complex graphical representations. Alternatively, it requires a dense cloud of

observations to be stable and it can be easily biased by outliers if the local density of data is low. Also, LOWESS does not return a simple mathematical function and, thus, is complicated to use for predictive purposes. Finally, LOWESS is relatively computationally intensive, but with the range of observations that we used (in the thousands), this was a negligible issue.

Survival analysis

Many clinical questions in our dataset were related to the rate of development of the clinical manifestations from the onset of the disease. Also, death is, in all severe systemic diseases, an important prognostic predictor to analyze. Both the timing to develop a clinical manifestation and the time to death were appropriate questions to analyze using COX regression and Kaplan-Meier curves.

Standardized cancer and mortality rates

A recurrent question in cohort studies of patients with myositis is if certain groups of patients have higher cancer or mortality rates. Internal comparisons can be helpful by defining if some groups have higher or lower mortality rates than others or than the rest of the individuals in the cohort. However, the most relevant comparison is often with the general population through indirect standardization. A limitation of these types of comparisons is that they require well-annotated epidemiological data. In the US, both survival and cancer data can be acquired through the Centers for Disease and Control Prevention. Specifically, mortality data can be obtained from the United States Cancer Statistics (UCSC) registry (`www.cdc.gov/cancer/uscs/`) while the survival data can be acquired from the Compressed Mortality File (`wonder.cdc.gov/wonder/help/cmf.html`).

3.2 Myositis muscle RNA sequencing studies

Obtaining insight into the pathogenesis of the disease is the first basic step to understand the variability in the clinical manifestations, tailor specific treatments for the disease, and estimate how many disease categories there are. In order to do so, Next Generation Sequencing techniques are extremely powerful since they allow

us to get great volumes of data on the transcriptional changes of the affected tissues. For this section of the study, we used RNA sequencing from myositis muscle biopsies to understand the mechanisms of muscle damage in these autoimmune diseases. Moreover, for certain studies, we also used RNA sequencing data from differentiated human skeletal muscle cells and injured mouse muscle.

3.2.1 Cultured human skeletal muscle cells

Human skeletal muscle myoblasts (Lonza) were cultured according to the manufacturer's protocol. When 80% confluent, the cultures were induced to differentiate into myotubes by replacing the growth medium with differentiation medium (DMEM, 2% horse serum, and L-glutamine). Two plates of cells were collected for RNA extraction at 7 separate time points: immediately before differentiation and then daily for 6 days.

3.2.2 Mouse Muscle Injury

Muscle injury and regeneration were induced in mice using cardiotoxin. Briefly, 6 week-old C57BL/6 mice were unilaterally injured by intramuscular injection of 0.1 mL of 10 uM CTX into the tibialis anterior muscle. Injured tibialis anterior muscles were harvested at days 3 (n=2), 5 (n=2), 7 (n=2), 10 (n=4), 14 (n=4), and 28 (n=3) post-injury. Contralateral (uninjured) tibialis anterior muscles were also collected (n=9). Muscle tissue was snap-frozen and stored at -80 degrees Celsius.

3.2.3 Human muscle biopsy processing

Open muscle biopsies were placed in an aluminum foil envelope. 2-methylbutane (isopentane) was pre-chilled using liquid nitrogen and the aluminum foil envelopes were submerged in the isopentane for 15 seconds. After this, the samples were placed in cryovials at -80^{o} for long-term storage. Samples collected at other institutions were shipped in dry ice to the NIH Muscle Disease Unit.

3.2.4 RNA extraction

The first step to perform RNA sequencing is to get the RNA from the samples of interest. To do so we had first to homogenize the muscle tissue using ceramic beads and then separate the RNA from the rest of the components of the cell using the TRIzol protocol. Briefly, muscle biopsies were homogenized in TRIzol using 1.4 mm ceramic bead low-binding tubes, and the RNA was extracted using the regular TRIzol protocol. Concentration and quality of the resulting RNA were assessed using standard NanoDrop and TapeStation protocols, respectively.

3.2.5 RNA library preparation

Libraries were prepared using the NeoPrepTM system according to the TruSeqM Stranded mRNA Library Prep protocol (Illumina) and sequenced using the Illumina HiSeq 2500 or 3000. This methodology had the advantage of being mostly automated, decreasing possible human errors in this step.

3.2.6 RNA sequencing analysis

Demultiplexing

Since modern DNA sequencers have a huge sequencing capacity, several samples can be run in each lane of a flow cell. Thus, the first step after we have sequenced the RNAseq library will be to separate all the different experiments in individual folders based on their characteristic index. This was performed using bcl2fastq v.2.17.1 parallelizing it to speed up the process. The output of this step is a file in .fastq format consisting of repeating blocks of 4 lines containing the annotation, sequence, comments, and quality of each one of the reads of the experiment.

Sequence cleanup

After the sequences of each sample are demultiplexed it is convenient to clean up the data to eliminate sequences with low quality and remove adapter contamination that may bias the results. This was performed using trimmomatic v.0.36.

Alignment

Once the sequences were cleaned each read had to be aligned to the reference genome in order to be able to identify what genes are expressed and how much. A technical complication for this step in RNAseq libraries is that mature RNA has been already spliced and thus, sequences containing sections of two or more exons will not be aligned to any region of the reference DNA since it includes the intronic regions. To overcome this limitation, specific RNA aligners have been developed that work by splitting reads that cannot be aligned confidently and trying to align each fragment of the split separately. This will be repeated recursively until the whole read could be aligned or it was concluded that there was no good match in the reference genome. Limited studies have compared the multiple available tools to perform RNA alignment, concluding that STAR offers slightly better accuracy compared to other options.[122] Thus, to perform the alignment we used STAR v.2.5 using the hg19 (GRCh37) or the mm10 (GRCm38) reference genomes.

Gene expression quantitation

In RNA sequencing libraries the number of reads of each gene is proportional to the level of expression of that gene in the sample of origin. However, the number of reads will be influenced by the length of the gene and the total number of reads of the library. Thus, to quantify the levels of expression of each gene the resulting number of reads has to be normalized to the total number of reads of the library and the length of each gene. Two basic statistics are used to measure the level of expression of RNAseq experiments depending on how the normalization is performed, fragments per kilobase million (FPKM), and transcripts per kilobase million (TPM). FPKMs are calculated by dividing the read counts of each gene by the total number of read counts of the library and the result is divided by the length of the gene in kilobases. Alternatively, the TPM is calculated by dividing first by the length of the gene and later on by the total number of reads of the library. The advantage of TPM over FPKM would be that the total number of TPMs per library will be constant and this will allow us to compare the levels of expression of different genes across different samples. Alternatively, FPKM is more popular and is valid for comparing the same gene across different samples, which is one of the most common comparisons to

make.

RNAseq libraries can be single-end or pair-end depending on the number of RNA reads that are obtained from each strand of RNA (one in single-end and two at each side in pair-end). In single-end experiments, FPKM can be also named RPKM or reads per kilobase million. However, for paired-end experiments the term read is not correct because each pair of reads identifies a fragment, and for this reason, is preferable to use the term FPKM that will be correct both for single and pair-end experiments.

There are multiple libraries that allow quantitation of the level of expression of each gene and most of them result in similar results provided that the same gene coordinates are used. In our case, we used Stringtie v.1.3.3, which is a tool contained in the "new Tuxedo package".[123]

Quality control

A key part of any Next Generation Sequencing pipeline is to confirm that the quality of the reads is adequate for the purposes of the analysis and there are no sequencing artifacts. This can be performed right after the sequencing, the demultiplexing, the alignment, or after the gene quantification steps. Given that examining the output of the quality control for hundreds of samples can be tedious and time-consuming, in our experience it was more efficient to perform the quality control at the end of the pipeline to capture any failures that the process could have produced. In case of problems with the sample, further steps of quality control after each one of the steps can help to identify the error. For doing the quality control we used fastqc v.0.11.2.

Differential expression

The main results of an RNAseq analysis are the expression levels of each one of the 20,000 genes in a biological sample. However, the relevant questions usually require comparing the differences between conditions. The process of comparing the levels of expression among biologically relevant groups is called differential expression. This is, arguably, one of the most critical steps in the RNAseq pipeline given the multiple models that can be used to do the comparisons and the variety

of tools available for the task. Some of these tools use the Poisson distribution, others the negative binomial or the beta-binomial while there are some that use Bayesian non-parametric models. Thus, both edgeR and DESeq2 use a variation of the Fisher exact test adopted for the negative binomial distribution returning exact p-values computed from the derived probabilities. CuffDiff uses the test statistics $T = \frac{E[\log(y)]}{Var[\log(y)]}$, where y is the ratio of the normalized counts between two conditions, and this ratio approximately follows a normal distribution. Hence a t-test is used to calculate the p-value for the differential expression. Finally, limma uses a moderated t-statistic to compute p-values in which both the standard error and the degrees of freedom are modified. All of these tools used the Benjamini and Hochberg approach for multiple hypothesis correction.[124]

The results of these tools vary considerably and there is yet no gold standard on which is the best to do these types of analyses. However, there are studies comparing them showing that DESeq may have a slight advantage in terms of detection accuracy.[125] Moreover, our own benchmark tests showed that CuffDiff had a very poor correspondence with previous gene expression data in our field, while both EdgeR, DESeq, and limma had similar and good results. In conclusion, we decided to use the DESeq2 v.1.20 algorithm for differential expression.[126]

DESeq2 performs an internal normalization where the geometric mean is calculated for each gene across all samples. The counts for a gene in each sample are then divided by this mean. The median of these ratios in a sample is the size factor for that sample. This procedure corrects for library size and RNA composition bias, which can arise, for example, when only a small number of genes are very highly expressed in one experiment condition but not in the other.

Additionally, DESeq2 automatically detects count outliers using the Cooks' distance and removes these genes from the analysis. DESeq2 v.1.20 also performs independent filtering which maximizes the number of genes which will have a Benjamini and Hochberg-adjusted p–value[124] less than a critical value set by default to 0.1. Removing these genes with low counts improves the detection power by making the multiple testing adjustment of the p-values less severe. To speed up the computations we prefiltered genes with a total count across conditions below 10. Since

these genes would have been excluded from the analysis afterward anyways, this did not influence the calculations at all.

DESeq2 uses shrinkage estimation for dispersions and fold changes. A dispersion value is estimated for each gene through a model fit procedure. Using these estimations, the package fits a negative binomial generalized linear model for each gene and uses the Wald test for significance testing. The Wald test p-values from the subset of genes that pass the independent filtering step are adjusted for multiple testing using the procedure of Benjamini and Hochberg.[124]

To ensure the stability of the central tendency and dispersion values of each biological group between different sections of the study, the normalization process included the totality of the samples even if that specific comparison did not include some of those samples.

We assigned equal weights to each autoantibody subgroups within DM and IMNM to avoid giving more importance to differentially expressed transcriptomic features of autoantibody subgroups with a higher number of biopsies at this stage of the analysis.

Pathway analysis

Once we have identified the list of genes explaining the differences between two or more conditions it is often necessary to find what are the pathways that are activated or suppressed. Doing so has the main difficulty of being based on manual annotations of the pathways. Thus, if the annotation is poor, incomplete, or contains wrong data our pathway analysis will be biased. There are several tools to perform pathway analysis, most of them basing their results on the Fisher's exact test, the hypergeometric distribution, or a chi-square analysis of the 2x2 contingency tables comparing the number of genes observed in our dataset with the genes present in each pathway. DAVID and MetaCore have similar approaches to the problem by doing a simple frequentist assessment of genes present in the case dataset vs. genes observed in each pathway. Alternatively, IPA Pathway analysis also performs this basic analysis but it also provides information about the activation or suppression of the pathway. Another option is to use gene-set enrichment analysis (GSEA). This

approach has the advantages of not needing an arbitrary threshold to define a list of "significant" genes and that it uses the actual significance of the associations of the genes with the trait being analyzed. However, it is limited to a single pathway comparison at a time and, thus, it is less suitable for pathway discovery purposes.

RNAseq-based classification

RNAseq is an extremely powerful technique that quantifies simultaneously the expression of more than twenty thousand genes. This has the potential to be used predictively to classify complex diseases based on its pathogenesis. Multiple algorithms can be used for this task, from simplistic approaches using logistic regression and simple classification trees to bagging or boosting strategies (like the random forest or AdaBoost respectively). Moreover, neural networks and other algorithms like the support vector machines can also be useful for this task. However, given that there are no good studies assessing when to use each algorithm based on the characteristics of the data, we took an agnostic approach to this problem by testing all these mathematical tools in the same training datasets and testing their accuracy in those samples that were not used for training purposes.

In short, the sample was split into a training set containing 2/3 of the observations and a test set containing the remaining 1/3. The training set was used to build the classificatory models and the testing set to evaluate the accuracy of the model. The classifiers that we tested were the linear support vector machine (SVM), the radial basis function SVM, random forest, nearest neighbor, Gaussian process, decision tree, multi-layer perceptron neural network, adaptative boosting (AdaBoost), gaussian naïve Bayes and quadratic discriminant analysis. Models were built in 2/3 random resamples of the data and tested in the remaining 1/3. The accuracy of classifying correctly each one of the myositis subsets was determined based on the mean and 95% CI of one thousand resampling cycles.

We decided to use all the genes that were significantly different (with a cutoff q-value <0.05) in each group compared to the rest using all the samples. An alternative would have been to include the differentially expressed genes contained in the training set of each cycle. However, this approach was excessively computationally expensive. Nonetheless, to demonstrate the equivalency of these approaches, we

modified our pipeline to train 100 cross-validation cycles using only the differentially expressed genes resulting from each training set and the performance of the models was equivalent using both methods.

Ranking genes in classificatory models

Oftentimes it is interesting to rank the most important features for the classification model. This can be done, independently of the algorithm, using recursive feature elimination.[127] This technique iteratively constructs new models removing the features with low weights and thus, estimating the importance of each feature for the prediction task.

4. Discussion

In this section, I will do an integrative discussion of the different manuscripts that compose this doctoral thesis. Here, my objective was not to substitute but to introduce and complement the discussion of each one of the manuscripts that are included in Appendix A. First, I will review the projects that were conducted using the research framework to study the Johns Hopkins longitudinal cohort of myositis patients. Then, I will comment on the different studies that were completed using the RNAseq profiling of myositis muscle biopsies.

As it was mentioned earlier, given the uncertainty regarding the etiology of the different types of myositis, we believed that it was necessary to approach studying the classification and its pathogenesis in parallel. This way we would be certain to be studying homogeneous groups of patients.

It may be argued that the clinical section of this doctoral dissertation does not fall within the realm of bioinformatics. However, bioinformatics was originally defined as the study of informatic processes in biotic systems.[128] It is a broad multi-disciplinary field that combines biology, computer science, information engineering, mathematics, and statistics to analyze and interpret biological data. Accordingly, the informatic methodology that we have used to retrieve, organize, analyze, and synthesize efficiently the longitudinal clinical information of our myositis patients fits the original definition of what is bioinformatics.

Particularly, this doctoral dissertation would be a good example of research in translational bioinformatics, a young discipline defined as "the development of stor-

age, analytic, and interpretive methods to optimize the transformation of increasingly voluminous biomedical data, and genomic data, into proactive, predictive, preventive, and participatory health".[129] Particularly, we are using a combination of clinical and basic bioinformatic research to get a deeper understanding on various aspects of a disease, including diagnosis, classification, therapeutics, prognosis and pathophysiology of different types of myositis.

4.1 Longitudinal cohort studies of myositis subsets

Arguably, the main methodological advantage of cohort studies over other types of observational studies is that they allow exploring multiple exposures in different subsets of patients. Alternatively, they usually require a costly infrastructure and a considerable investment of time to obtain consistently high-quality data over a prolonged period of time. In all the studies that are included in this section, we show the full power of combining a relatively well-managed cohort of patients with an efficient data management and data analysis research framework.

By planning complex projects (e.g. comparing the performance of MSAs to the 2017 EULAR/ACR criteria) as a collection of smaller and more manageable studies (e.g. anti-Mi2 study) we were able to gradually explore the best data-gathering strategies, analytical tools, and inherent limitations of our cohort, progressively adapting and incrementally increasing the quality and efficiency of our studies over time.

Thanks to this stepwise approach to large epidemiologic projects we were able to analyze the clinical features, survival, association with cancer, prognosis of particular MSAs (e.g. anti-Mi2), and both disease-specific (e.g. anti-U1RNP) and non-disease-specific (anti-Ro52) MAAs. Moreover, we were able to systematically compare the effectiveness of certain treatments within prevalent myositis subsets (e.g. methotrexate (MTX) vs. azathioprine (AZA) in the AS). Finally, we could aggregate and expand the data of various studies to answer relevant questions to our whole area of research (e.g. comparing the performance of MSAs to the 2017 EULAR/ACR criteria).

First, regarding the analysis of MSAs it is important to emphasize that these

autoantibodies are generally mutually exclusive. This implies that each one of them can be compared with the other with very little concern for patients being included simultaneously in more than one category. In the study analyzing the clinical phenotype of anti-Mi2 autoantibodies (Appendix section A.4) we used this characteristic to our advantage comparing a group of 58 anti-Mi2 patients with other relevant autoantibody-defined groups including: (1) non-Mi2 DM comprising the anti-NXP2, anti-TIF1γ, and anti-MDA5; (2) AS including anti-Jo1, anti-PL7, and anti-PL12, and (3) IMNM including anti-SRP and anti-HMGCR antibodies.

With this design, we could define that patients with anti-Mi2 autoantibodies have more weakness and higher CK levels than patients with anti-Mi2 negative DM. Also, we could establish that the muscle involvement was more severe in patients with anti-Mi2 autoantibodies than in the AS and similar in the upper extremities to IMNM. In contrast with the frequent and severe muscle disease, we found that the extramuscular manifestations were less common in anti-Mi2 patients. Moreover, we could not find a significant association with cancer compared to the general population. Importantly, our data suggested that the levels of anti-Mi2 autoantibodies were associated with the muscle enzyme levels and the strength and that it decreased with treatment, occasionally normalizing. This last finding may be of practical importance since it suggests that anti-Mi2 autoantibodies may be useful not just as a useful biomarker of disease activity but to decide whether it is safe to completely withdraw immunosuppressive medication from patients when the anti-Mi2 levels decrease.

Besides, multiple serologic techniques can be used to determine anti-Mi2 autoantibodies. A key part of our work on defining the characteristics of specific autoantibody-defined subsets of patients is to validate the serologic techniques that we are using so we can know if they are reliable and, if they are not, how to modify the manufacturer's recommendations to be used confidently. In this case, we validated the EUROLine Myositis Profile 4 line blot by immunoprecipitation and ELISA (Appendix section A.5) finding that only those subjects that are positive for both anti-Mi2α and anti-Mi2β can be reliably considered anti-Mi2 without further validation. Alternatively, those who are positive just for anti-Mi2β are usually false positives and those who are only positive for anti-Mi2α have only about 50% chance

of being true anti-Mi2 positives.

As it was mentioned in the introduction, the current terminology of MAAs groups together two different populations of autoantibodies, those that are not specific for any given phenotype and those that, despite being associated with a homogeneous phenotype, show features that are found in non-myositis autoimmune diseases. Relevant to the bioinformatic analysis of these groups of patients, those that are specific for a homogeneous phenotype and mutually exclusive can be analyzed similarly to MSAs groups of patients, but those that are not specific for any given phenotype have to be analyzed by comparing MAAs-positive with MAAs-negative patients both globally and for each of the different subgroups. In this doctoral thesis, we studied anti-U1RNP patients as an example of a MAAs that is associated with a particular phenotype (Appendix section A.1). By analyzing a cohort of 20 anti-U1RNP patients we found that these individuals typically present with proximal weakness and necrotizing muscle biopsies, showing arthritis, dermatitis, and ILD as the most common extramuscular clinical features. Also, pericarditis and glomerulonephritis were uniquely found in patients with coexisting anti-U1RNP autoantibodies and anti-Ro52 autoantibodies.

We also used our research framework to study anti-Ro52 autoantibodies in a cohort of 371 juvenile DM (Appendix section A.3). These autoantibodies are MAAs that are not specific of any particular phenotype but have been suggested to act as disease modifiers. In our study we found that anti-Ro52 autoantibodies are present in 14% of patients with juvenile myositis and are strongly associated with anti-MDA5 and AS autoantibodies. Also, we determined that in patients with juvenile DM, those with anti-Ro52 autoantibodies were more likely to have ILD, had more severe disease and poorer prognosis.

Finally, by aggregating and expanding the data of various of our earlier studies we were able to summarize the phenotype of 524 MSAs-positive patients using factor analysis of mixed data and demonstrate the utility of MSAs to subclassify myositis patients (Appendix section A.6). Moreover, AIC and BIC of the different regression models showed that MSAs outperform the 2017 EULAR/ACR criteria to predict the factor-analysis-derived phenotype in adult MSAs-positive myositis patients. Based

Table 4.1: Proposal of myositis classification based on myositis-specific autoantibodies.

Myositis-specific autoantibody +	Muscle weakness or Creatine kinase elevation or Interstitial lung disease or Arthritis or Gottron's sign or papules or Heliotrope

on our results we proposed a classification (Table 4.1) using the MSAs to inform patient selection for assembling myositis cohorts of the most phenotypically and clinically homogeneous groups.

Importantly for this study is the methodology that we used to handle the phenotype of the patients statistically. In other projects, the authors defined questionable homogeneous clusters of patients linking them to clinical classifications based on the opinion of experts.[39] Alternatively, in our study we used the clinical characteristics of the patients to answer a relevant clinical question: do the autoantibodies outperform clinical classifications systems in myositis? The rationale behind this question is that if the target of the immune response is defined by the autoantigen, in those diseases with a humoral response, the autoantibodies will completely define the disease. Therefore, there will be no possible combination of clinical features that will be able to outperform the autoantibodies to predict the phenotype of the patients.

4.2 Transcriptome profiling of myositis muscle biopsies

The sheer amount of information that we obtained by performing RNAseq in the muscle biopsies of patients with different types of inflammatory myopathy has the potential to answer multiple questions relevant to the pathogenesis of myositis simultaneously. In this doctoral thesis, we show the results of addressing three spe-

cific issues. First, we wanted to explore the expression of the different autoantigens in patients with different autoantibodies; second, we did an in-depth study of the interferon pathway (both type I and type II) in the different clinical and autoantibody myositis groups; finally, we developed a methodology to identify unique gene expression profiles in each clinical and autoantibody myositis group.

Regarding the first of the studies(Appendix section A.7), after the first MSAs were discovered in myositis, it was proposed that regenerating muscle cells in biopsy tissue from human myositis muscle express high levels of several myositis autoantigens, including Mi2, TIF1γ, Jo1, HMGCR, and SRP.[67,130–132] Given this observation, it has been proposed that an increased expression of myositis autoantigens may initiate and/or maintain autoimmunity against these particular proteins. However, it has not been determined if autoantigens other than Mi2, TIF1γ, Jo1, HMGCR and SRP are expressed at high levels in regenerating muscle, if autoantigen expression patterns differ between myositis subgroups, or if there is a relationship between the expression level of an autoantigen and the presence of its corresponding autoantibody. Thus, we explored these questions in our dataset finding that myositis autoantigens are in fact highly expressed during muscle regeneration, however, we could not find any significant association between the increased expression of a given autoantigen and its corresponding autoantibody. This is, patients with anti-HMGCR antibodies did not show higher levels of HMGCR than patients with anti-SRP antibodies, and so on.

As per the second study using these data,(Appendix section A.8) prior studies had established the preferential activation of the IFN1 pathway in DM muscle.[93] However, activation of the IFN1 pathway has not been compared between patients with DM with different MSAs. Furthermore, the IFN1 pathway activation was found to be relatively low in IBM but has not been systematically explored in AS or IMNM.[93,133,134] Similarly, although IFN2 pathway activation has been implicated in IBM muscle,[135,136] activation of IFN2 pathways in muscle biopsies from patients with DM, IMNM, AS, and IBM has not been systematically analyzed. Our study confirms that DM muscle biopsies are characterized by high levels of both IFN1- and IFN2-inducible genes. In contrast, biopsies from patients with AS and IBM reveal gene expression patterns consistent with prominent IFN2 activation. Finally,

RNA sequencing analysis reveals that IMNM biopsies show relatively low activation of the IFN pathway. Moreover, all the different MSAs DM groups that we explored showed a similar level of activation of both IFN pathways.

Finally, microarray analysis led to the discovery that type I and type II IFN-inducible genes are upregulated in muscle biopsies from patients with DM[93] and IBM,[135,136] respectively. However, disease-specific gene expression profiles have not been fully described in patients with IMNM, AS or any of the MSAs-defined subtypes of DM. Furthermore, little attention has been given to genes that are differentially expressed between patients with different types and subtypes of myositis.[93,134,137,138] In this last study included in the doctoral thesis,(Appendix section A.9) we trained machine learning algorithms to classify muscle biopsies using transcriptomic data from normal, IBM and MSAs-positive muscle biopsies. We then used recursive feature elimination to identify novel disease-specific gene expression patterns that may be pathologically relevant in DM, AS, IMNM, IBM and MSAs-defined subtypes of myositis. With this approach we could determine that DM, AS, IMNM, and IBM are best distinguished based on their gene expression pattern by using linear support vector machines. Furthermore, by applying recursive feature elimination to these classification models, we not only confirmed known pathological pathways in myositis, such as the role of type I IFN in DM, but we also identified novel genes that are uniquely upregulated in other types and MSAs-defined subtypes of myositis.

This study contained results of key biological and clinical importance but also relevant information for future bioinformatic studies in the field. Most importantly, it defined that the linear support vector machines outperform other models that theoretically should behave well in scenarios with high dimensionality and low sample size, like the random forests. Also, we were able to validate the recursive feature elimination to sort the different genes in order of importance. Finally, we compared different strategies to restrict the number of genes determining that doing it before the cross-validation is equivalent and less computationally expensive than doing it afterward.

4.3 Limitations and future directions

Our studies suggest that autoantibodies outperform clinical classification systems to predict the phenotype of myositis patients and found a variety of specific pathogenic pathways, and clinical associations. Notwithstanding this, it will be necessary to validate our results in different cohorts and using different techniques to confirm them and understand better its importance. Also, some of these studies have been the largest ever conducted in our field, including often samples from different epidemiologic backgrounds. However, the sample size in some of our autoantibody groups was limited and we may have lost relevant signals for this reason. Future efforts will aim to increase the sample size of these patient groups. Finally, our clinical studies used retrospectively collected clinical data, and our transcriptional research focused on studying muscle tissue. Prospectively, it will be key to expand the clinical information that we collect and to study other biological myositis samples such as skin, lung, or blood cells.

At this moment we are conducting several studies to validate particular transcriptional signals that we detected in our studies at the protein level. Moreover, international collaborative efforts are being conducted to validate the importance of autoantibodies in myositis classification through the International Myositis Assessment and Clinical Studies Group and the ENMC.

Also, we are conducting single-cell and single-nuclei RNAseq studies in myositis muscle biopsies to understand which cells in the muscle are expressing each one of the specific markers that we detected. Moreover, we are also analyzing the distribution of the expression of the different markers is myositis muscle biopsies using spatial transcriptomics. Besides, we are studying the transcriptome of blood cells at a single-cell resolution to understand the distinct inflammatory response of each one of the myositis subgroups.

Finally, any of the classical types of myositis and myositis-overlap syndromes that we have studied have a clear etiology. However, there are types of inflammatory myositis, like checkpoint-inhibitor-induced or graft vs. host myositis that have a known cause and may help us to understand the rest. We are establishing

collaborations with experts all around the world to study these types of man-made myositis.

5. Conclusions

- We developed a research framework that was able to generate an efficient research pipeline to study longitudinal cohorts of specific myositis subgroups. With this framework, we could prove that:
  - Patients with anti-U1RNP myositis typically present with proximal weakness and necrotizing muscle biopsies. Arthritis, dermatitis, and ILD are the most common extramuscular clinical features. Pericarditis and glomerulonephritis are uniquely found in patients with anti-U1-RNP- positive myositis.
  - Azathioprine and methotrexate have similar efficacy and adverse events in patients with AS. Pneumonitis is a rare but important event in patients receiving methotrexate.
  - Anti-Ro52 autoantibodies are present in 14% of patients with juvenile myositis and are strongly associated with anti-MDA5 and antisynthetase autoantibodies. Anti-Ro52 autoantibodies in juvenile myositis are strongly associated with ILD. Furthermore, patients with anti-Ro52 autoantibodies have more severe disease and poorer prognosis.
  - Patients with anti-Mi2-positive DM have more severe muscle disease than patients with anti-Mi2-negative DM or patients with AS. Anti-Mi2 autoantibody levels correlate with disease severity and may normalize in patients who enter remission.
  - Only those subjects that are positive for both anti-Mi2α and anti-Mi2β by

line blot can be reliably considered anti-Mi2 without further validation.

  - MSAs outperform the 2017 EULAR/ACR criteria to predict the phenotype of patients with myositis.

  - MSAs can be used to build phenotypically homogeneous groups in myositis research.

- We performed a systematic analysis of the transcriptome of muscle biopsies from patients with different types of myositis showing that:

  - Most myositis autoantigens are highly expressed during muscle regeneration but are not associated with autoantibody specificity.

  - IFN1 is most upregulated in DM, with intermediate activation of the pathway in AS and lower levels of activation in IBM and IMNM. Alternatively, IFN2 is robustly activated in DM, AS, and IBM but not in IMNM.

  - Unique gene expression profiles in muscle biopsies from patients with MSAs-defined subtypes of myositis and IBM suggest that different pathological mechanisms underly muscle damage in each of these diseases.

References

1. Selva-O'Callaghan A, Pinal-Fernandez I, Trallero-Araguas E, Milisenda JC, Grau-Junyent JM, and Mammen AL. Classification and management of adult inflammatory myopathies. Lancet Neurol 2018;17:816–28.
2. Wagner E. Fall einer seltenen Muskelkrankheit. Arch Heilkd 1863;4:282–3.
3. Pottain PC. Morve chronique de forme anormal. Bull Et Mem Hop Paris 1875;12:314–8.
4. Hepp P. Über einen Fall von acuter parenhymatoser Myositis, welche Geschwulste bildete und Fluctuation vortauschte. Berlin Klin Wochenschr 1887;24:389.
5. Unverricht H. Polymyositis acuta progressive. Z Klin Med 1887;12.
6. Stertz G. Polymyositis. Berl Klin Wochenshr 1916;53:489.
7. Kankeleit H. Über primaire nichteitrige Polymyositis. Dtsch Arch Klin Med 1916;120:335–9.
8. Patjes G and Clejat C. Sclerose atrophic de la peau et myosite. Ann Derm Et Syphil Paris 1926:550–68.
9. Hautveränderungen HG. bei Dermatomyositis. VIII Congres Intern. Dermatol. et Syphilogr.,Copenhagen, 5–8 Agout, 1930. Compt Rend Seancy 1930:826–30.
10. Unverricht H. Dermatomyositis acuta. Dtsch Med Wochenschr 1891;17:41–4.
11. Hecht MS. Dermatomyositis in children. J Pediatr 1940;17:791.
12. Mills ES and Mathews WH. Interstitial pneumonitis in dermatomyositis. JAMA 1956;160:1467–70.

13. Chou SM. Myxovirus-like structures in a case of human chronic polymyositis. Science 1967;158:1453–5.
14. Carpenter S, Karpati G, Heller I, and Eisen A. Inclusion body myositis: a distinct variety of idiopathic inflammatory myopathy. Neurology 1978;28:8–17.
15. Krain L. Dermatomyositis in six patients without initial muscle involvement. Arch Dermatol 1975;111:241–5.
16. Pearson CM. Polymyositis and dermatomyositis. In: *Arthritis (and Allied Conditions)*. Ed. by DJ M. 8th ed. Philadelphia: Lea and Febiger Publishers, 1979. Chap. 52:742–61.
17. Euwer RL and Sontheimer RD. Amyopathic dermatomyositis (dermatomyositis sine myositis). Presentation of six new cases and review of the literature. J Am Acad Dermatol 1991;24:959–66.
18. Sontheimer RD. Would a new name hasten the acceptance of amyopathic dermatomyositis (dermatomyositis sine myositis) as a distinctive subset within the idiopathic inflammatory dermatomyopathies spectrum of clinical illness? J Am Acad Dermatol 2002;46:626–36.
19. Nishikai M and Reichlin M. Heterogeneity of precipitating antibodies in polymyositis and dermatomyositis. Characterization of the Jo-1 antibody system. Arthritis Rheum 1980;23:881–8.
20. Targoff IN and Reichlin M. The association between Mi-2 antibodies and dermatomyositis. Arthritis Rheum 1985;28:796–803.
21. Reichlin M and Mattioli M. Description of a serological reaction characteristic of polymyositis. Clin Immunol Immunopathol 1976;5:12–20.
22. Hochberg MC, Feldman D, Stevens MB, Arnett FC, and Reichlin M. Antibody to Jo-1 in polymyositis/dermatomyositis: association with interstitial pulmonary disease. J Rheumatol 1984;11:663–5.
23. Love LA, Leff RL, Fraser DD, et al. A new approach to the classification of idiopathic inflammatory myopathy: myositis-specific autoantibodies define useful homogeneous patient groups. Medicine (Baltimore) 1991;70:360–74.

24. Emslie-Smith AM and Engel AG. Necrotizing myopathy with pipestem capillaries, microvascular deposition of the complement membrane attack complex (MAC), and minimal cellular infiltration. Neurology 1991;41:936–9.
25. Bohan A and Peter JB. Polymyositis and dermatomyositis (second of two parts). N Engl J Med 1975;292:403–7.
26. Tanimoto K, Nakano K, Kano S, et al. Classification criteria for polymyositis and dermatomyositis. J Rheumatol 1995;22:668–74.
27. Griggs RC, Askanas V, DiMauro S, et al. Inclusion body myositis and myopathies. Ann Neurol 1995;38:705–13.
28. Targoff IN, Miller FW, Medsger TA, and Oddis CV. Classification criteria for the idiopathic inflammatory myopathies. Curr Opin Rheumatol 1997;9:527–35.
29. Badrising UA, Maat-Schieman M, Duinen SG van, et al. Epidemiology of inclusion body myositis in the Netherlands: a nationwide study. Neurology 2000;55:1385–7.
30. Emery AEH. Diagnostic criteria for neuromuscular disorders. Royal society of medicine press London, 1997.
31. Dalakas MC and Hohlfeld R. Polymyositis and dermatomyositis. Lancet 2003;362:971–82.
32. Hoogendijk JE, Amato AA, Lecky BR, et al. 119th ENMC international workshop: trial design in adult idiopathic inflammatory myopathies, with the exception of inclusion body myositis, 10-12 October 2003, Naarden, The Netherlands. Neuromuscul Disord 2004;14:337–45.
33. Hilton-Jones D, Miller A, Parton M, Holton J, Sewry C, and Hanna MG. Inclusion body myositis: MRC Centre for Neuromuscular Diseases, IBM workshop, London, 13 June 2008. Neuromuscul Disord 2010;20:142–7.
34. Pestronk A. Acquired immune and inflammatory myopathies: pathologic classification. Curr Opin Rheumatol 2011;23:595–604.
35. Rose MR and Group ENMCIBMW. 188th ENMC International Workshop: Inclusion Body Myositis, 2-4 December 2011, Naarden, The Netherlands. Neuromuscul Disord 2013;23:1044–55.

36. Lundberg IE, Tjärnlund A, Bottai M, et al. 2017 European League Against Rheumatism/American College of Rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups. Ann Rheum Dis 2017;76:1955–64.
37. Allenbach Y, Mammen AL, Stenzel W, Benveniste O, and Immune-Mediated Necrotizing Myopathies Working G. 224th ENMC International Workshop:: Clinico-sero-pathological classification of immune-mediated necrotizing myopathies Zandvoort, The Netherlands, 14-16 October 2016. Neuromuscul Disord 2017.
38. Mammen AL, Allenbach Y, Stenzel W, Benveniste O, and 239th Workshop Study Group ENMC. 239th ENMC International Workshop: Classification of dermatomyositis, Amsterdam, the Netherlands, 14-16 December 2018. Neuromuscul Disord 2020;30:70–92.
39. Mariampillai K, Granger B, Amelin D, et al. Development of a New Classification System for Idiopathic Inflammatory Myopathies Based on Clinical Manifestations and Myositis-Specific Autoantibodies. JAMA neurology 2018;75:1528–37.
40. Pinal-Fernandez I and Mammen AL. On using machine learning algorithms to define clinically meaningful patient subgroups. Ann Rheum Dis 2019.
41. Lloyd TE, Mammen AL, Amato AA, Weiss MD, Needham M, and Greenberg SA. Evaluation and construction of diagnostic criteria for inclusion body myositis. Neurology 2014;83:426–33.
42. Tasca G, Monforte M, De Fino C, Kley RA, Ricci E, and Mirabella M. Magnetic resonance imaging pattern recognition in sporadic inclusion-body myositis. Muscle Nerve 2015;52:956–62.
43. Pinal-Fernandez I, Casal-Dominguez M, Carrino JA, et al. Thigh muscle MRI in immune-mediated necrotising myopathy: extensive oedema, early muscle damage and role of anti-SRP autoantibodies as a marker of severity. Ann Rheum Dis 2017;76:681–7.
44. Lloyd TE, Christopher-Stine L, Pinal-Fernandez I, et al. Cytosolic 5'-Nucleotidase 1A As a Target of Circulating Autoantibodies in Autoimmune Diseases. Arthritis Care Res (hoboken) 2016;68:66–71.

45. Herbert MK, Stammen-Vogelzangs J, Verbeek MM, et al. Disease specificity of autoantibodies to cytosolic 5'-nucleotidase 1A in sporadic inclusion body myositis versus known autoimmune diseases. Ann Rheum Dis 2016;75:696–701.
46. Muro Y, Nakanishi H, Katsuno M, Kono M, and Akiyama M. Prevalence of anti-NT5C1A antibodies in Japanese patients with autoimmune rheumatic diseases in comparison with other patient cohorts. Clin Chim Acta 2017;472:1–4.
47. Lilleker JB, Rietveld A, Pye SR, et al. Cytosolic 5'-nucleotidase 1A autoantibody profile and clinical characteristics in inclusion body myositis. Ann Rheum Dis 2017;76:862–8.
48. Goyal NA, Cash TM, Alam U, et al. Seropositivity for NT5c1A antibody in sporadic inclusion body myositis predicts more severe motor, bulbar and respiratory involvement. J Neurol Neurosurg Psychiatry 2016;87:373–8.
49. Tawara N, Yamashita S, Zhang X, et al. Pathomechanisms of anti-cytosolic 5'-nucleotidase 1A autoantibodies in sporadic inclusion body myositis. Ann Neurol 2017;81:512–25.
50. Dalakas MC. Muscle biopsy findings in inflammatory myopathies. Rheum Dis Clin North Am 2002;28:779–98, vi.
51. Greenberg SA, Pinkus JL, Amato AA, Kristensen T, and Dorfman DM. Association of inclusion body myositis with T cell large granular lymphocytic leukaemia. Brain 2016;139:1348–60.
52. Greenberg SA, Pinkus JL, Kong SW, Baecher-Allan C, Amato AA, and Dorfman DM. Highly differentiated cytotoxic T cells in inclusion body myositis. Brain 2019;142:2590–604.
53. Greenberg SA. Inclusion body myositis: clinical features and pathogenesis. Nature reviews. Rheumatology 2019;15:257–72.
54. Greenberg SA, Pinkus JL, and Amato AA. Nuclear membrane proteins are present within rimmed vacuoles in inclusion-body myositis. Muscle Nerve 2006;34:406–16.

55. Nalbantoglu J, Karpati G, and Carpenter S. Conspicuous accumulation of a single-stranded DNA binding protein in skeletal muscle fibers in inclusion body myositis. Am J Pathol 1994;144:874–82.
56. Guttsches AK, Brady S, Krause K, et al. Proteomics of rimmed vacuoles define new risk allele in inclusion body myositis. Ann Neurol 2017;81:227–39.
57. Greenberg SA. Theories of the pathogenesis of inclusion body myositis. Curr Rheumatol Rep 2010;12:221–8.
58. Catalan-Garcia M, Garrabou G, Moren C, et al. Mitochondrial DNA disturbances and deregulated expression of oxidative phosphorylation and mitochondrial fusion proteins in sporadic inclusion body myositis. Clin Sci (Lond) 2016;130:1741–51.
59. Rygiel KA, Tuppen HA, Grady JP, et al. Complex mitochondrial DNA rearrangements in individual cells from patients with sporadic inclusion body myositis. Nucleic Acids Res 2016;44:5313–29.
60. Van De Vlekkert J, Maas M, Hoogendijk JE, De Visser M, and Van Schaik IN. Combining MRI and muscle biopsy improves diagnostic accuracy in subacute-onset idiopathic inflammatory myopathy. Muscle Nerve 2015;51:253–8.
61. Watanabe Y, Uruha A, Suzuki S, et al. Clinical features and prognosis in anti-SRP and anti-HMGCR necrotising myopathy. J Neurol Neurosurg Psychiatry 2016;87:1038–44.
62. Chung T, Christopher-Stine L, Paik JJ, Corse A, and Mammen AL. The composition of cellular infiltrates in anti-HMG-CoA reductase-associated myopathy. Muscle Nerve 2015;52:189–95.
63. Pinal-Fernandez I, Parks C, Werner JL, et al. Longitudinal Course of Disease in a Large Cohort of Myositis Patients With Autoantibodies Recognizing the Signal Recognition Particle. Arthritis Care Res (hoboken) 2017;69:263–70.
64. Suzuki S, Nishikawa A, Kuwana M, et al. Inflammatory myopathy with anti-signal recognition particle antibodies: case series of 100 patients. Orphanet J Rare Dis 2015;10:61.
65. Tiniakou E, Pinal-Fernandez I, Lloyd TE, et al. More severe disease and slower recovery in younger patients with anti-3-hydroxy-3-methylglutaryl-coenzyme

A reductase-associated autoimmune myopathy. Rheumatology (Oxford) 2017;56:787–94.

66. Mammen AL, Chung T, Christopher-Stine L, et al. Autoantibodies against 3-hydroxy-3-methylglutaryl-coenzyme A reductase in patients with statin-associated autoimmune myopathy. Arthritis Rheum 2011;63:713–21.
67. Allenbach Y, Arouche-Delaperche L, Preusse C, et al. Necrosis in anti-SRP(+) and anti-HMGCR(+)myopathies: Role of autoantibodies and complement. Neurology 2018;90:e507–e517.
68. Christopher-Stine L, Casciola-Rosen LA, Hong G, Chung T, Corse AM, and Mammen AL. A novel autoantibody recognizing 200-kd and 100-kd proteins is associated with an immune-mediated necrotizing myopathy. Arthritis Rheum 2010;62:2757–66.
69. Allenbach Y, Keraen J, Bouvier AM, et al. High risk of cancer in autoimmune necrotizing myopathies: usefulness of myositis specific antibody. Brain 2016;139:2131–5.
70. Ohnuki Y, Suzuki S, Shiina T, et al. HLA-DRB1 alleles in immune-mediated necrotizing myopathy. Neurology 2016;87:1954–5.
71. Mammen AL, Gaudet D, Brisson D, et al. Increased frequency of DRB1*11:01 in anti-hydroxymethylglutaryl-coenzyme A reductase-associated autoimmune myopathy. Arthritis Care Res (hoboken) 2012;64:1233–7.
72. Limaye V, Bundell C, Hollingsworth P, et al. Clinical and genetic associations of autoantibodies to 3-hydroxy-3-methyl-glutaryl-coenzyme a reductase in patients with immune-mediated myositis and necrotizing myopathy. Muscle Nerve 2015;52:196–203.
73. Targoff IN, Johnson AE, and Miller FW. Antibody to signal recognition particle in polymyositis. Arthritis Rheum 1990;33:1361–70.
74. Kao AH, Lacomis D, Lucas M, Fertig N, and Oddis CV. Anti-signal recognition particle autoantibody in patients with and patients without idiopathic inflammatory myopathy. Arthritis Rheum 2004;50:209–15.
75. Chen F, Peng Y, and Chen M. Diagnostic Approach to Cardiac Involvement in Idiopathic Inflammatory Myopathies. Int Heart J 2018;59:256–62.

76. Arouche-Delaperche L, Allenbach Y, Amelin D, et al. Pathogenic role of anti-signal recognition protein and anti-3-Hydroxy-3-methylglutaryl-CoA reductase antibodies in necrotizing myopathies: Myofiber atrophy and impairment of muscle regeneration in necrotizing autoimmune myopathies. Ann Neurol 2017;81:538–48.

77. Suarez-Calvet X, Gallardo E, Pinal-Fernandez I, et al. RIG-I expression in perifascicular myofibers is a reliable biomarker of dermatomyositis. Arthritis Res Ther 2017;19:174.

78. Pinal-Fernandez I, Casciola-Rosen LA, Christopher-Stine L, Corse AM, and Mammen AL. The Prevalence of Individual Histopathologic Features Varies according to Autoantibody Status in Muscle Biopsies from Patients with Dermatomyositis. J Rheumatol 2015;42:1448–54.

79. Uruha A, Nishikawa A, Tsuburaya RS, et al. Sarcoplasmic MxA expression: A valuable marker of dermatomyositis. Neurology 2017;88:493–500.

80. Betteridge Z and McHugh N. Myositis-specific autoantibodies: an important tool to support diagnosis of myositis. J Intern Med 2016;280:8–23.

81. Ghirardello A, Zampieri S, Iaccarino L, et al. Anti-Mi-2 antibodies. Autoimmunity 2005;38:79–83.

82. Albayda J, Pinal-Fernandez I, Huang W, et al. Antinuclear Matrix Protein 2 Autoantibodies and Edema, Muscle Disease, and Malignancy Risk in Dermatomyositis Patients. Arthritis Care Res (hoboken) 2017;69:1771–6.

83. Fiorentino DF, Chung LS, Christopher-Stine L, et al. Most patients with cancer-associated dermatomyositis have antibodies to nuclear matrix protein NXP-2 or transcription intermediary factor 1gamma. Arthritis Rheum 2013;65:2954–62.

84. Trallero-Araguas E, Rodrigo-Pendas JA, Selva-O'Callaghan A, et al. Usefulness of anti-p155 autoantibody for diagnosing cancer-associated dermatomyositis: a systematic review and meta-analysis. Arthritis Rheum 2012;64:523–32.

85. Selva-O'Callaghan A, Grau JM, Gámez-Cenzano C, et al. Conventional cancer screening versus PET/CT in dermatomyositis/polymyositis. Am J Med 2010;123:558–62.

86. Ge Y, Lu X, Shu X, Peng Q, and Wang G. Clinical characteristics of anti-SAE antibodies in Chinese patients with dermatomyositis in comparison with different patient cohorts. Sci Rep 2017;7:188.
87. Labrador-Horrillo M, Martinez MA, Selva-O'Callaghan A, et al. Anti-MDA5 antibodies in a large Mediterranean population of adults with dermatomyositis. J Immunol Res 2014;2014:290797.
88. Sato S, Hoshino K, Satoh T, et al. RNA helicase encoded by melanoma differentiation-associated gene 5 is a major autoantigen in patients with clinically amyopathic dermatomyositis: Association with rapidly progressive interstitial lung disease. Arthritis Rheum 2009;60:2193–200.
89. Narang NS, Casciola-Rosen L, Li S, Chung L, and Fiorentino DF. Cutaneous ulceration in dermatomyositis: association with anti-melanoma differentiation-associated gene 5 antibodies and interstitial lung disease. Arthritis Care Res (hoboken) 2015;67:667–72.
90. Miller FW, Chen W, O'Hanlon TP, et al. Genome-wide association study identifies HLA 8.1 ancestral haplotype alleles as major genetic risk factors for myositis phenotypes. Genes Immun 2015;16:470–80.
91. Mamyrova G, Rider LG, Ehrlich A, et al. Environmental factors associated with disease flare in juvenile and adult dermatomyositis. Rheumatology (Oxford) 2017;56:1342–7.
92. Pinal-Fernandez I, Ferrer-Fabregas B, Trallero-Araguas E, et al. Tumour TIF1 mutations and loss of heterozygosity related to cancer-associated myositis. Rheumatology (Oxford) 2018;57:388–96.
93. Greenberg SA, Pinkus JL, Pinkus GS, et al. Interferon-alpha/beta-mediated innate immune mechanisms in dermatomyositis. Ann Neurol 2005;57:664–78.
94. Walsh RJ, Kong SW, Yao Y, et al. Type I interferon-inducible gene expression in blood is present and reflects disease activity in dermatomyositis and polymyositis. Arthritis Rheum 2007;56:3784–92.
95. Baechler EC, Bauer JW, Slattery CA, et al. An interferon signature in the peripheral blood of dermatomyositis patients is associated with disease activity. Mol Med 2007;13:59–68.

96. Wong D, Kea B, Pesich R, et al. Interferon and biologic signatures in dermatomyositis skin: specificity and heterogeneity across diseases. PLoS One 2012;7:e29161.

97. Guillen-Del Castillo A, Pilar Simeon-Aznar C, Fonollosa-Pla V, et al. Good outcome of interstitial lung disease in patients with scleroderma associated to anti-PM/Scl antibody. Semin Arthritis Rheum 2014;44:331–7.

98. Rigolet A, Musset L, Dubourg O, et al. Inflammatory myopathies with anti-Ku antibodies: a prognosis dependent on associated lung disease. Medicine (Baltimore) 2012;91:95–102.

99. Escola-Verge L, Pinal-Fernandez I, Fernandez-Codina A, et al. Mixed Connective Tissue Disease and Epitope Spreading: An Historical Cohort Study. J Clin Rheumatol 2017;23:155–9.

100. Pinal-Fernandez I, Casal-Dominguez M, Huapaya JA, et al. A longitudinal cohort study of the anti-synthetase syndrome: increased severity of interstitial lung disease in black patients and patients with anti-PL7 and anti-PL12 autoantibodies. Rheumatology (Oxford) 2017;56:999–1007.

101. Trallero-Araguas E, Grau-Junyent JM, Labirua-Iturburu A, et al. Clinical manifestations and long-term outcome of anti-Jo1 antisynthetase patients in a large cohort of Spanish patients from the GEAS-IIM group. Semin Arthritis Rheum 2016;46:225–31.

102. Andersson H, Kirkhus E, Garen T, Walle-Hansen R, Merckoll E, and Molberg O. Comparative analyses of muscle MRI and muscular function in anti-synthetase syndrome patients and matched controls: a cross-sectional study. Arthritis Res Ther 2017;19:17.

103. Mescam-Mancini L, Allenbach Y, Hervier B, et al. Anti-Jo-1 antibody-positive patients show a characteristic necrotizing perifascicular myositis. Brain 2015;138:2485–92.

104. Noguchi E, Uruha A, Suzuki S, et al. Skeletal Muscle Involvement in Anti-synthetase Syndrome. JAMA Neurol 2017;74:992–9.

105. Stenzel W, Preusse C, Allenbach Y, et al. Nuclear actin aggregation is a hallmark of anti-synthetase syndrome-induced dysimmune myopathy. Neurology 2015;84:1346–54.

106. Meulen MF van der, Bronner IM, Hoogendijk JE, et al. Polymyositis: an overdiagnosed entity. Neurology 2003;61:316–21.

107. Chahin N and Engel AG. Correlation of muscle biopsy, clinical course, and outcome in PM and sporadic IBM. Neurology 2008;70:418–24.

108. Vilela VS, Prieto-Gonzalez S, Milisenda JC, Selva OCA, and Grau JM. Polymyositis, a very uncommon isolated disease: clinical and histological re-evaluation after long-term follow-up. Rheumatol Int 2015;35:915–20.

109. Amato AA and Griggs RC. Unicorns, dragons, polymyositis, and other mythological beasts. Neurology 2003;61:288–9.

110. McHugh NJ and Tansley SL. Autoantibodies in myositis. Nat Rev Rheumatol 2018;14:290–302.

111. Vancsa A, Csipo I, Nemeth J, Devenyi K, Gergely L, and Danko K. Characteristics of interstitial lung disease in SS-A positive/Jo-1 positive inflammatory myopathy patients. Rheumatol Int 2009;29:989–94.

112. Bauhammer J, Blank N, Max R, et al. Rituximab in the Treatment of Jo1 Antibody-associated Antisynthetase Syndrome: Anti-Ro52 Positivity as a Marker for Severity and Treatment Response. J Rheumatol 2016;43:1566–74.

113. Marie I, Hatron PY, Dominique S, et al. Short-term and long-term outcome of anti-Jo1-positive patients with anti-Ro52 antibody. Semin Arthritis Rheum 2012;41:890–9.

114. Albrecht I, Wick C, Hallgren Å, et al. Development of autoantibodies against muscle-specific FHL1 in severe inflammatory myopathies. J Clin Invest 2015;125:4612–24.

115. Labrador-Horrillo M, Martínez MA, Selva-O'Callaghan A, et al. Identification of a novel myositis-associated antibody directed against cortactin. Autoimmun Rev 2014;13:1008–12.

116. Fiorentino DF, Presby M, Baer AN, et al. PUF60: a prominent new target of the autoimmune response in dermatomyositis and Sjögren's syndrome. Ann Rheum Dis 2016;75:1145–51.

117. Zhang YM, Yang HB, Shi JL, et al. The prevalence and clinical significance of anti-PUF60 antibodies in patients with idiopathic inflammatory myopathy. Clin Rheumatol 2018;37:1573–80.

118. Akaike H. A new look at the statistical model identification. IEEE Trans Autom Control 1974;19:716–23.

119. Schwarz G et al. Estimating the dimension of a model. Ann Statist 1978;6:461–4.

120. Burnham KP and Anderson DR. Multimodel inference: understanding AIC and BIC in model selection. Sociological methods & research 2004;33:261–304.

121. Raftery AE. Bayesian model selection in social research. Sociological methodology 1995:111–63.

122. Baruzzo G, Hayer KE, Kim EJ, Di Camillo B, FitzGerald GA, and Grant GR. Simulation-based comprehensive benchmarking of RNA-seq aligners. Nat Methods 2017;14:135–9.

123. Pertea M, Kim D, Pertea GM, Leek JT, and Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc 2016;11:1650–67.

124. Benjamini Y and Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat 1995;57:289–300.

125. Seyednasrollah F, Laiho A, and Elo LL. Comparison of software packages for detecting differential expression in RNA-seq studies. Briefings Bioinf 2015;16:59–70.

126. Love MI, Huber W, and Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.

127. Guyon I, Weston J, Barnhill S, and Vapnik V. Gene selection for cancer classification using support vector machines. Machine Learning 2002;46:389–422.

128. Hogeweg P and Hesper B. Interactive instruction on population interactions. Comput Biol Med 1978;8:319–27.

129. Tenenbaum JD. Translational Bioinformatics: Past, Present, and Future. Genomics Proteomics Bioinformatics 2016;14:31–41.

130. Mohassel P, Rosen P, Casciola-Rosen L, Pak K, and Mammen AL. Expression of the dermatomyositis autoantigen transcription intermediary factor 1gamma in regenerating muscle. Arthritis Rheumatol 2015;67:266–72.

131. Mammen AL, Casciola-Rosen LA, Hall JC, Christopher-Stine L, Corse AM, and Rosen A. Expression of the dermatomyositis autoantigen Mi-2 in regenerating muscle. Arthritis Rheum 2009;60:3784–93.

132. Casciola-Rosen L, Nagaraju K, Plotz P, et al. Enhanced autoantigen expression in regenerating muscle cells in idiopathic inflammatory myopathy. J Exp Med 2005;201:591–601.

133. Salajegheh M, Kong SW, Pinkus JL, et al. Interferon-stimulated gene 15 (ISG15) conjugates proteins in dermatomyositis muscle with perifascicular atrophy. Ann Neurol 2010;67:53–63.

134. Greenberg SA, Sanoudou D, Haslett JN, et al. Molecular profiles of inflammatory myopathies. Neurology 2002;59:1170–82.

135. Ivanidze J, Hoffmann R, Lochmüller H, Engel AG, Hohlfeld R, and Dornmair K. Inclusion body myositis: laser microdissection reveals differential upregulation of IFN-γ signaling cascade in attacked versus nonattacked myofibers. Am J Pathol 2011;179:1347–59.

136. Allenbach Y, Chaara W, Rosenzwajg M, et al. Th1 response and systemic treg deficiency in inclusion body myositis. PLoS One 2014;9:e88788.

137. Hamann PD, Roux BT, Heward JA, et al. Transcriptional profiling identifies differential expression of long non-coding RNAs in Jo-1 associated and inclusion body myositis. Sci Rep 2017;7:8024.

138. Raju R and Dalakas MC. Gene expression profile in the muscles of patients with inflammatory myopathies: effect of therapy with IVIg and biological validation of clinically relevant genes. Brain 2005;128:1887–96.

A. Published works

This section includes the full version of all the manuscripts that compose this doctoral thesis. Of note, in all these studies Dr. Pinal-Fernandez conducted the entirety of the bioinformatic analysis, contributing also to coordinating the study, gathering the data, performing the laboratory procedures, and drafting the manuscripts.

A summary of the study was included at the beginning of each section. Moreover, all the corresponding appendixes and supplementary figures are attached at the end of each manuscript.

A.1 Muscular and extramuscular features of myositis patients with anti-U1-RNP autoantibodies. Neurology 2018 (PMID: 30824556).

In this longitudinal cohort study, we define the clinical phenotype of patients with myositis with anti-U1RNP autoantibodies. We analyzed the prevalence and severity of clinical features at disease onset and during follow-up in patients with anti-U1RNP myositis comparing it to DM, IMNM, and the AS.

Twenty anti-U1RNP patients, 178 patients with DM, 135 patients with IMNM, and 132 patients with AS were included. Anti-U1RNP patients were younger (37 years) and more likely to be black (60%) than patients with AS, DM, or IMNM. Muscle weakness was a presenting feature in 15% of anti-U1RNP patients; 80% eventually developed weakness. Four of 7 anti-U1RNP patients had necrotizing muscle biopsies. Arthritis occurred in 60% of anti-U1RNP patients; this was increased compared to DM (18%) or IMNM (6%) (all $p < 0.01$). DM-specific skin features developed in 60% of anti-U1RNP patients. ILD occurred in 45% of anti-U1RNP patients; fewer patients with DM (13%) and IMNM (6%) and more patients with AS (80%) developed ILD (all $p < 0.01$). Glomerulonephritis and pericarditis occurred in 25% and 40% of anti-U1RNP patients, respectively, but rarely in the other groups; these features occurred only in those with coexisting anti-Ro52 autoantibodies. No anti-U1RNP patient had cancer-associated myositis or died during the study period.

In conclusion, patients with anti-U1RNP myositis typically present with proximal weakness and necrotizing muscle biopsies. Arthritis, dermatitis, and ILD are the most common extramuscular clinical features. Pericarditis and glomerulonephritis are uniquely found in patients with anti-U1RNP myositis.

Muscular and extramuscular features of myositis patients with anti-U1-RNP autoantibodies.

Maria Casal-Dominguez, M.D., Ph.D.[1,2*], Iago Pinal-Fernandez, M.D., Ph.D.[1,2,3*], Andrea M. Corse, M.D.[2], Julie Paik, M.D., M.H..[4], Jemima Albayda, M.D.[4], Livia Casciola-Rosen, Ph.D.[4], Cheilonda Johnson, M.D.[4], Sonye K. Danoff, M.D., Ph.D.[4], Lisa Christopher-Stine, M.D., M.P.H.[2,4], Eleni Tiniakou, M.D.[4#], Andrew L. Mammen, M.D., Ph.D.[1,2,4#]

[1]Muscle Disease Unit, Laboratory of Muscle Stem Cells and Gene Regulation, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland.

[2]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland.

[3]Faculty of Health Sciences, Universitat Oberta de Catalunya.

[4]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland.

[*,#] These authors contributed equally to this study.

Funding sources:

This work was financially supported by the Intramural Research Program of the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health. IPF was supported by a fellowship from The Myositis Association. ET was supported by the Scientist Development Award from the Rheumatology Research Foundation and the Jerome-Greene Foundation. The

Johns Hopkins Rheumatic Disease Research Core Center, which processed and banked the sera and performed some of the autoantibody assays, is supported by NIH grants P30-AR-053503 and P30- AR-070254.

Address correspondence to Andrew L. Mammen, Muscle Disease Unit, Laboratory of Muscle Stem Cells and Gene Expression, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, 50 South Drive, Room 1146, Building 50, MSC 8024, Bethesda, MD 20892. E-mail:andrew.mammen@nih.gov Phone: 240-234-0054. Fax: 301-594-0305.

Word count: 3531 words

Disclosures: Maria Casal-Dominguez, Iago Pinal-Fernandez, Andrea M. Corse, Julie Paik, Jemima Albayda, Livia Casciola-Rosen, Cheilonda Johnson, Sonye K. Danoff, Lisa Christopher-Stine, Eleni Tiniakou, and Andrew L. Mammen report no disclosures.

ABSTRACT

Objective: To define the clinical phenotype of myositis patients with anti-U1-ribonucleoprotein (RNP) autoantibodies.

Methods: In this longitudinal cohort study, the prevalence and severity of clinical features at disease onset and during follow-up in anti-U1-RNP-positive myositis patients were compared to those with dermatomyositis (DM), immune-mediated necrotizing myopathy (IMNM), and the antisynthetase syndrome (AS).

Results: Twenty anti-U1-RNP-positive, 178 DM, 135 IMNM, and 132 AS patients were included. Anti-U1-RNP-positive patients were younger (~37 years) and more likely to be black (60%) than AS, DM, or IMNM patients. Muscle weakness was a presenting feature in 15% of anti-U1-RNP-positive patients; 80% eventually developed weakness. Four of seven anti-U1-RNP-positive patients had necrotizing muscle biopsies. Arthritis occurred in 60% of anti-U1-RNP-positive patients; this was increased compared to DM (18%) or IMNM (6%) (all $p<0.01$). DM-specific skin features developed in 60% of anti-U1-RNP-positive patients. Interstitial lung disease (ILD) occurred in 45% of anti-U1-RNP-positive patients; fewer patients with DM (13%) and IMNM (6%) and more AS patients (80%) developed ILD (all $p<0.01$). Glomerulonephritis and pericarditis occurred in 25% and 40% of anti-U1-RNP-positive patients, respectively, but rarely in the other groups; these features occurred only in those with co-existing anti-Ro52 autoantibodies. No anti-U1-RNP patient had cancer-associated myositis or died during the study period.

Conclusions: Anti-U1-RNP myositis patients typically present with neck sparing proximal weakness and necrotizing muscle biopsies. Arthritis, dermatitis, and ILD are the most common extramuscular clinical features. Pericarditis and glomerulonephritis are uniquely found in anti-U1-RNP-positive myositis patients.

INTRODUCTION

The autoimmune myopathies are a heterogeneous family of diseases that affect both skeletal muscle and other organ systems. The most common forms of autoimmune myopathy include dermatomyositis (DM), immune-mediated necrotizing myopathy (IMNM), and the antisynthetase syndrome (AS). In addition, myositis may overlap with other autoimmune diseases (1).

Myositis autoantibodies are associated with unique clinical phenotypes in patients with various forms of myositis, including overlap myositis. For example, autoantibodies recognizing U1-ribonucleoprotein (RNP) have been reported to occur in myositis patients, including those who also have systemic lupus erythematosus, systemic sclerosis, or mixed connective tissue disease (MCTD). To date, the prevalence and severity of the muscular and extramuscular clinical features at disease onset and during follow-up have not been well-described in myositis patients with anti-U1-RNP autoantibodies. Furthermore, no studies have directly compared the clinical features of anti-U1-RNP-positive myositis patients to those with DM, IMNM, and AS (2).

In the present study, we conducted a longitudinal cohort study of myositis patients with anti-U1-RNP autoantibodies. The demographic, clinical, and laboratory features of these myositis patients was compared to those with DM, IMNM, and AS.

MATERIALS AND METHODS

Patients and autoantibody testing

All patients enrolled in the Johns Hopkins Myositis Center longitudinal cohort study between 2002 and 2017 were included if they were positive for anti-U1-RNP or myositis-specific autoantibodies as described below.

Patient sera were screened for anti-U1-RNP antibodies by enzyme-linked immunosorbent assay (ELISA) at the Johns Hopkins laboratory and/or by S35-immunoprecipitation (IP), immunodiffusion and/or RNA-IP at the Oklahoma Medical Research Foundation; positive samples were confirmed by RNA-IP at the NIH Muscle Disease Unit lab. Patient sera were screened for anti-Ro52 by EUROLINE myositis profile blot. Serum samples from anti-U1-RNP-positive patients were subsequently tested for (a) anti-dsDNA and anti-Sm autoantibodies by ELISA, (b) anti-centromere, anti-topoisomerase, anti-RNA polymerase III, and anti-U3-RNP autoantibodies using the EUROLINE systemic sclerosis profile, (c) and myositis-specific autoantibodies (as described for the comparison groups below).

The comparison groups included patients who were positive for myositis-specific autoantibodies by at least two different immunologic techniques from among the following: ELISA, *in vitro* transcription and translation (IVTT)-IP, line blotting (EUROLINE myositis profile), IP from S35-labeled HeLa cell extracts, immunodiffusion, and/or RNA-immunoprecipitation as previously described (3). The AS group included all patients with an antisynthetase autoantibody. The DM group included all patients positive for anti-Mi2, anti-NXP2, anti-TIF1γ or anti-MDA5 autoantibodies. The IMNM group included all patients positive for anti-SRP or anti-HMGCR autoantibodies.

Muscle strength was assessed by the examining physician using the Medical Research Council scale; serial strength measurements for each patient were made by the same physician. The MRC scale was transformed to Kendall's 0-10 scale (4) for analysis. The average of right-and left-side measurements for arm abduction and hip flexion strength was used for calculations (possible range 0–10).

At their initial visit to the Johns Hopkins Myositis Center, the presence or absence of clinical signs and symptoms at disease onset was established retrospectively based on a review of prior patient records and patient recollection. Interstitial lung disease (ILD) at the onset of the disease (often prior to the first visit at the Myositis Center) was assessed by retrospective chart review. At the first visit and on subsequent visits to the Johns Hopkins Myositis Center, the presence or absence of DM-specific rashes (i.e., heliotrope or Gotron's rashes), SSc-specific skin involvement (i.e., sclerodactyly), esophageal symptoms (i.e., reflux and dysphagia), and AS associated clinical symptoms (i.e. arthralgia) and signs, either observed by the clinician (i.e. mechanics hands and arthritis) or reported by the patient (i.e. fever and Raynaud's phenomenon) were assessed prospectively. During follow-up, ILD was defined through a multidisciplinary approach as recommended by the American Thoracic Society (ATS) (5). All patients with suspicion of pulmonary hypertension (PH) (compatible clinical and echocardiographical features) underwent a right heart catheterization. Those with a mean pulmonary arterial pressure (PAPm) ≥ 25 mmHg at rest were considered as having PH (6). Pulmonary function testing (PFT) included spirometry, lung volumes measured by helium dilution, and diffusing capacity by single breath carbon

monoxide (DLCO) based on ATS criteria (7). Muscle enzyme levels and PFTs were included for analysis if obtained within a period of 6 weeks before or after strength testing (except for peak, minimum and mean values, where all available data was included). Glomerulonephritis was assessed by kidney biopsy and pericarditis by echocardiography.

The cumulative features recorded at all visits were used to classify the anti-U1-RNP-positive patients. Each patient was classified for myositis type based on the Bohan and Peter criteria (8) and for MCTD using the Sharp (2), Kasukawa (9), Alarcon-Segovia (10) and Khan (11) criteria. The patients were also classified using the 2013 ACR/EULAR classification criteria for systemic sclerosis (12) and the 1997 ACR classification criteria for lupus (13).

All available muscle biopsies were interpreted at the Johns Hopkins Neuromuscular Pathology Laboratory by pathologists blinded to autoantibody status. The pathologists consistently reported on the presence or absence of perifascicular atrophy, perivascular inflammation, primary inflammation (i.e., the invasion of non-necrotic fibers by mononuclear cells), and necrotizing myopathy (i.e., prominent myofiber necrosis in the absence of perifascicular atrophy or primary inflammation).

Standard protocol approvals and patient consents. This study was approved by the Johns Hopkins Institutional Review Board and written informed consent was obtained from each participant.

Statistical analysis

Dichotomous variables were expressed as percentage (count) and continuous variables as mean (SD). Bivariate comparisons of continuous variables were made using Student´s t-test while bivariate comparisons of dichotomous variables were made either using chi-squared test or Fisher´s exact test, as appropriate. CK, a highly positively skewed variable, was expressed as median, first, and third quartile for descriptive purposes and transformed through a base-10 logarithm for the statistical analysis. Each one of the study groups was compared to the sample of anti-U1-RNP patients.

To account for the different number of visits per patient, the evolution of the pulmonary function tests, CK levels and muscle strength were studied using multilevel linear regression models with random slopes and random intercepts. The mean of hip flexor and arm abductor strength (range 0-10) was used as the strength outcome for regression analysis. Logistic regression was used to analyze dichotomous variables across groups adjusting by possible confounders.

The influence of non-modifiable risk factors (sex, race, length of illness and age at the onset of the first symptoms), the corticosteroid dose and the administration of intravenous immunoglobulins (IVIG), rituximab, methotrexate, azathioprine and mycophenolate were used as adjusting covariates. Other treatments administered to less than 10% of the cohort were not included in the analysis.

All statistical analyses were performed using Stata/MP 14.1. A 2-sided p-value of 0.05 or less was considered significant with no correction for multiple comparisons.

Data Availability

No unpublished data related to this study are publicly available.

RESULTS

General features

Among 437 patients enrolled in the Johns Hopkins Myositis Center Longitudinal Cohort Study who underwent testing for anti-U1-RNP autoantibodies, 20 (4.6%) were positive. Of note, patients with inclusion body myositis, genetic muscle disease, toxic myopathies, and other non-myositis diagnoses did not routinely undergo testing for anti-U1-RNP autoantibodies. The comparator groups included 178 patients with DM, 135 patients with IMNM, and 132 patients with AS.

The mean follow-up time for the anti-U1-RNP-positive patient group was 6.4 years, which was longer than that for the DM or IMNM patient groups, who were followed for a mean time of 4.2 and 4.0 years, respectively (Table 1). Anti-U1-RNP-positive patients had a mean of 22 visits per patient, which was more than double the mean number of visits for each of the other three groups (Table 1). The median time between the onset of the disease and the first visit at Hopkins was 1 year (Q1-Q3: 0.6-6.4). 18 out of the 20 anti-U1-RNP patients were treated with immunosuppressand drugs before arriving at the Myositis Center.

Anti-U1-RNP-positive patients were younger (37.3 years old) at disease onset compared to patients with DM (47.1 years old; $p<0.01$), IMNM (51.5 years old; $p<0.001$), or AS (45.0 years old; $p<0.05$). As in the other groups, there was a marked female predominance among anti-U1-RNP participants (80%). Of note, 60% of anti-U1-RNP-positive patients were black, which was greater than the one of black patients in the other groups; only 12% of DM, 30% of AS and 24% of IMNM were black (all $p<0.05$ compared to the anti-U1-RNP-positive group). All groups were exposed to similar treatment modalities, although mycophenolate was

more commonly used in anti-U1-RNP-positive patients (45%) compared to IMNM patients (20%).

No anti-U1-RNP-positive patient died during the study period or developed any malignancy within 3 years of the onset of the first disease symptoms.

Muscle involvement

At the onset of disease, muscle weakness was less prevalent among anti-U1-RNP-positive patients (15%) compared to those with DM (47%), AS (55%), or IMNM (83%) (p values all <0.01) (Table 2). Weakness emerged in 80% of anti-U1-RNP-positive patients during the course of the disease, which was less frequent only in comparison to the IMNM patients (80% vs. 96%, p=0.02) (Table 3).

At their first visit to the Johns Hopkins Myositis Center, the severity of weakness in anti-U1-RNP-positive patients was similar to the DM and AS patients, with proximal weakness predominantly in hip flexors and arm abductors (Table 4). At this point, 9 out of the 20 patients had measurable weakness in arm abductors or hip flexors but just 5 of the 20 mantained full strength over the follow-up period. Interestingly, U1-RNP-positive patients were the only group with no detectable neck weakness. Patients with anti-U1-RNP autoantibodies had a higher median CK (229 vs. 117 IU/L, p=0.02) and aldolase levels (26.9 vs. 9.4 IU/L, p<0.001) compared to DM participants (Table 5). In contrast, anti-U1-RNP-positive patients had lower median CK levels compared to IMNM patients (CK 229 vs. 1401 IU/L, p<0.001) and were stronger, particularly in the hip flexors (mean Kendall score of 8.8 vs. 6.7, p=0.001) (Table 5).

Multilevel regression analysis showed that higher CK levels were associated with lower strength (β=-0.5, p=0.03) in anti-U1-RNP-positive patients. This analysis also confirmed that, independent of the length of illness, age at onset, race, sex or immunosuppressant treatment, IMNM was the only group weaker (β=-1.3, p<0.001) and with higher CK levels (β=0.7, p<0.001) than anti-U1-RNP patients.

Of the 7 anti-U1-RNP patients with biopsies available for review at Johns Hopkins, 4 (57%) had a predominantly necrotizing pattern. The other three patients had biopsies revealing combinations of myofiber regeneration (in one biopsy), perimysial inflammation (in two biopsies), and perivascular inflammation (in two biopsies) (Figure 1). The prevalence of a necrotizing muscle biopsy was not different in patients with AS (6/27, 22%, p=0.2) or IMNM (56/70, 80%, p=0.2), but was more common than in DM (5/45, 11%, p=0.01). Perifascicular atrophy was not observed in any anti-U1-RNP patient nor in IMNM, but was common in muscle biopsies from DM (25/45, 56%, p=0.01) and AS (14/27, 52%, p=0.03) patients. Similarly, lymphocytic invasion of non-necrotic muscle fibers was not found in any of the anti-U1-RNP-positive patient muscle biopsies. By contrast, muscle biopsies from 9% of DM, 30% of AS and 16% of IMNM showed this feature (Table 6). Predominant perifascicular necrosis was not detected in any of the U1-RNP patients.

Lung involvement

Although ILD was present in just one (5%) anti-U1-RNP patient at the onset of disease (Table 2), 45% developed ILD during follow-up (Table 3). In contrast,

more AS (80%, p=0.002) and fewer DM (13%, p=0.002) or IMNM patients (6%, p<0.001) developed ILD during the course of disease. The mean FVC and DLCO was lower (71.6% and 64.2%) in anti-U1-RNP-positive patients compared to those with DM (89.2% and 96.4%, p<0.005), or IMNM (87.7% and 103.3%, p<0.02) (Table 5).

While none of the 20 anti-U1-RNP-positive patients were noted to have pulmonary hypertension (PH) at disease onset (Table 2), 5 (25%) developed this during the course of their illness (Table 3). The prevalence of PH was higher in anti-U1-RNP-positive patients than in DM (3%, p=0.001) or IMNM (1%, p<0.001) patients, but was similar to that observed in AS patients (20%, p=0.8). Three of 5 patients with PH had concomitant ILD. Of note, immunosuppressive therapy improved, but did not completely reverse, PH in all three anti-U1-RNP-positive patients with available longitudinal echocardiographic information.

Logistic regression confirmed that, independent of the age of onset, length of illness, sex or race, ILD occurred less frequently in anti-U1-RNP-positive patients than in AS patients (OR 0.2, p<0.001) and more frequently in anti-U1-RNP-positive patients than in DM (OR 4, p=0.01) or IMNM (OR 13, p<0.001) participants. Moreover, logistic regression showed that PH was more common in anti-U1-RNP-positive patients compared to those with DM (OR 10.2, p=0.001) or IMNM (OR 23, p=0.01). There was no difference in the prevalence of PH between those with anti-U1-RNP autoantibodies and those with AS (OR 1.1, p=0.8). Multilevel regression models also confirmed that, independent of the abovementioned confounding variables and treatments received, DLCO in patients with anti-U1-RNP autoantibodies was similar to patients with the AS (DLCO

β=13%, p=0.1), but was more severe when compared to those with IMNM (DLCO β=33%, p=0.006) or DM (DLCO β=34%, p<0.001).

Skin involvement

Heliotrope rashes and/or Gottron's sign, the characteristic cutaneous features of DM, were present in only 15% of patients with anti-U1-RNP autoantibodies at the onset of disease (Table 2). However, the cumulative presence of these features rose to 60% during the follow-up period (Table 3). Sclerodactyly, a typical skin feature of systemic sclerosis, was not present in any anti-U1-RNP-positive patient at the onset of the disease (Table 2), but ultimately affected 25% of these patients (Table 3). Raynaud phenomenon and mechanic's hands, characteristic cutaneous manifestations of AS patients, were present during the course of disease in 80% and 50% of U1-RNP-positive patients, respectively; both of these features were less common in DM patients (22% and 28%, respectively; both p<0.04) (Table 3). Of note, Raynaud´s phenomenon occurred more often in anti-U1-RNP-positive patients than in those with AS (39%, p<0.001) (Table 3). As expected, compared to anti-U1-RNP-positive patients, those with IMNM had a markedly lower prevalence of skin involvement.

Other extramuscular involvement

Glomerulonephritis was not present in any of the 20 anti-U1-RNP-positive myositis patients at the onset of disease and pericarditis was initially present in just 1 (5%) patient (Table 2). However, glomerulonephritis occurred in 5 (25%) patients

during the course of disease (Table 3) and renal biopsies revealed membranous glomerulonephritis in 4 (80%) of these cases. Similarly, pericarditis eventually complicated the clinical course in 8 (40%) of those with anti-U1-RNP autoantibodies (Table 3). In contrast, aside from a single AS patient, neither glomerulonephritis nor pericarditis were diagnosed in patients with other forms of myositis.

The prevalence of gastroesophageal reflux and/or dysphagia was uncommon at the onset of the disease (5-10%) (Table 2) but eventually affected ~50% of the anti-U1-RNP-positive patients (Table 3). Dysphagia was more prevalent among those with anti-U1-RNP autoantibodies when compared to the AS patients (18%, $p<0.01$) (Table 3).

At the onset of disease, arthritis and arthralgia were present in 15% and 30% of anti-U1-RNP-positive patients, respectively (Table 2); the prevalence of joint involvement increased during the course of the disease, eventually affecting 60-65% of patients (Table 3). Arthritis was more common in anti-U1-RNP-positive patients compared to those with DM or IMNM (18% and 6%, both $p<0.001$) (Table 3).

Co-existing anti-Ro52 autoantibodies

Since anti-Ro52 autoantibodies are commonly found in myositis patients and may be associated with more severe disease (14), all patients were tested for these autoantibodies. Anti-Ro52 autoantibodies were more frequent in anti-U1-RNP-positive patients (75%) than in patients with DM (22%) or IMNM (17%) (both $p < 0.001$). Co-existing anti-Ro52 autoantibodies were found in 80% of AS

patients, which was similar when compared to the prevalence of these autoantibodies in anti-U1-RNP-positive patients (Table 1).

Muscle weakness was not significantly different among anti-U1-RNP-positive patients with and without anti-Ro52 autoantibodies. Similarly, the severity of ILD was not significantly different among anti-U1-RNP-positive patients with and without anti-Ro52 autoantibodies. Despite that, the prevalence of ILD was non-significantly increased in anti-Ro52 positive patients (53% vs. 20%, p=0.3). Of note, 5 of 15 (33%) patients who had both anti-U1-RNP and anti-Ro52 autoantibodies were found to have PH, whereas no anti-U1-RNP-positive patient without anti-Ro52 autoantibodies was diagnosed with PH (p=0.3).

Interestingly, glomerulonephritis and pericarditis occurred also only in patients with both anti-U1-RNP and anti-Ro52 autoantibodies. Glomerulonephritis occurred in 5 of 15 (42%) patients with both anti-U1-RNP and anti-Ro52 autoantibodies (p=0.3). Pericarditis occurred in 8 of 12 (67%) patients with these two autoantibodies; this was increased compared to anti-U1-RNP-positive patients without anti-Ro52 autoantibodies (p=0.05).

Co-existing autoantibodies associated with scleroderma, lupus, and myositis

As some patients with anti-U1-RNP autoantibodies have clinical features of scleroderma or lupus, we tested for autoantibodies classically associated with these diseases. No anti-U1-RNP-positive patient had anti-topoisomerase, anti-centromere, or anti-polymerase III autoantibodies (i.e., scleroderma-associated autoantibodies). Co-existing autoantibodies recognizing Sm, dsDNA, and U3-RNP

(i.e., lupus-associated autoantibodies) were found in 5 (25%), 3 (15%), and 1 (5%), of the anti-U1-RNP-positive patients, respectively.

Anti-U1-RNP-positive patients with co-existing anti-dsDNA autoantibodies had more pericarditis (100% vs. 29%, p=0.05), glomerulonephritis (100% vs. 12%, p=0.009) and subcutaneous edema (100% vs. 24%, p=0.03) than those without anti-dsDNA autoantibodies. We did not identify other significant clinical differences between anti-U1-RNP-positive patients with and without anti-Sm or anti-dsDNA autoantibodies.

Three anti-U1-RNP-positive patients had co-existing anti-Jo1 autoantibodies; these patients were excluded from the AS group for the purposes of this study. Otherwise, no patient with anti-U1-RNP autoantibodies had a myositis-specific autoantibody.

Clinical classification of anti-U1-RNP-positive myositis patients

All 20 anti-U1-RNP-positive patients fulfilled Bohan and Peter's criteria for either DM (60%) or polymyositis (40%). Given their diverse clinical manifestations, most of these patients also fulfilled diagnostic criteria for one or more other systemic autoimmune diseases. For example, 9 (45%) met the 2013 ACR/EULAR classification criteria for scleroderma and 11 (55%) met the 1997 ACR classification criteria for lupus. Ninety percent of anti-U1-RNP-positive patients met at least one set of criteria for MCTD; 18 (90%) met Kasukawa's, 16 (80%) met Khan's, 14 (70%) met Alarcon´s, and 5 (25%) met Sharp's criteria.

DISCUSSION

In this study, we have defined the distinctive clinical phenotype of myositis patients with anti-U1-RNP autoantibodies. These patients are younger and more likely to be black then those with DM, IMNM, or AS. This is consistent with a prior report that a high proportion of anti-U1-RNP-positive patients are black (15). Like the other myositis groups, those with anti-U1-RNP autoantibodies have proximal pattern of muscle weakness. However, the neck muscles are spared only in those with anti-U1-RNP autoantibodies. Anti-U1-RNP-positive patients are also notable for their prominent extramuscular manifestations. These include Raynaud phenomenon (80%), arthralgia/arthritis (60%), DM skin features (60%), necrotizing muscle biopsies (57%), mechanic's hands (50%), and dysphagia (50%). Other common clinical manifestations in these patients include ILD (45%), pericarditis (40%), subcutaneous edema (35%), fever (35%), glomerulonephritis (25%), pulmonary hypertension (25%), sclerodactyly (25%), and calcinosis (25%).

The unique clinical phenotype of anti-U1-RNP-positive patients can be further appreciated by comparing them to each of the three other myositis groups separately. Compared to DM patients, those with anti-U1-RNP autoantibodies are more likely to have sclerodactyly, Raynaud phenomenon, mechanic's hands, ILD, arthritis, pericarditis, and glomerulonephritis; as expected, they are less likely to have heliotrope or Gottron sign. Compared with AS patients, anti-U1-RNP-positive patients are more likely to have Raynaud phenomenon, dysphagia, pericarditis, and glomerulonephritis; they are less likely to have ILD and sicca syndrome. Finally, anti-U1-RNP-positive patients are more likely to have all of the studied extramuscular manifestations of disease compared to those with IMNM. In

contrast, significantly more IMNM patients have weakness and weakness is more severe in IMNM patients.

Given that IMNM seems to have the least in common with anti-U1-RNP-positive patients, it may be surprising to find that muscle biopsies from both groups can be strikingly similar with prominent myofiber necrosis and scant lymphocytic infiltration. Of note, others have also reported myofiber necrosis and regeneration in muscle biopsies from anti-U1-RNP-positive patients (16, 17). Since the prognosis of patients with anti-U1-RNP autoantibodies is very different than those with anti-HMGCR myopathy or anti-SRP myopathy, testing for each of these autoantibodies is indicated in patients presenting with a necrotizing muscle biopsy.

Pericarditis with or without glomerulonephritis occurred in 40% of myositis patients with anti-U1-RNP autoantibodies; these complications were exceedingly rare in the other myositis groups. Of note, all 8 anti-U1-RNP-positive myositis patients with pericarditis/glomerulonephritis had co-existing anti-Ro52 autoantibodies. Similarly, PH was detected only in RNP-positive patients who were also positive for anti-Ro52 autoantibodies. Taken together, 60% of anti-U1-RNP-positive patients with co-existing anti-Ro52 antibodies developed pulmonary hypertension, pericarditis and/or glomerulonephritis, while no patient without anti-Ro52 developed any of these manifestations ($p=0.04$). Although it requires confirmation in other cohorts, based on these observations, clinicians could consider testing anti-U1-RNP-positive myositis patients for anti-Ro52 to identify those patients most at risk for developing these serious extramuscular manifestations of disease.

This study has several limitations. First, most of the conclusions are based

on signs and symptoms that were recorded prospectively from the beginning of the study in 2002. Consequently, we could not include activity and damage tools that were not available when the study started. Second, data used in this study is based on patients presenting to a multidisciplinary Myositis Center and may be biased towards including patients with active muscle and lung disease. Third, due to the relatively small sample size of our anti-U1-RNP-positive population, our study may have been underpowered to detect differences in some key features like the association between ILD and anti-Ro52 autoantibodies. Future studies including larger numbers of anti-U1-RNP-positive patients will be of value.

These limitations notwithstanding, we have shown that patients with anti-U1-RNP autoantibodies appear to have a unique syndrome different from patients with DM, AS, or IMNM. This syndrome is characterized by proximal muscle weakness, necrotizing muscle biopsies, and frequent extramuscular manifestations. Glomerulonephritis, pericarditis, and pulmonary hypertension are relatively common in anti-U1-RNP-positive patients with co-existing anti-Ro52 autoantibodies but rare in the other myositis groups. We propose that testing for anti-Ro52 autoantibodies may be useful to determine which anti-U1-RNP-positive patients are at most risk for these complications.

APPENDIX 1: AUTHORS

Name	Location	Role	Contribution
Maria Casal-Dominguez, MD, PhD	NIH, Bethesda	Author	Designed and conceptualized study; analyzed the data; drafted the manuscript for intellectual content
Iago Pinal-Fernandez, MD, PhD	NIH, Bethesda	Author	Designed and conceptualized study; analyzed the data; drafted the manuscript for intellectual content
Andrea Corse, MD	Johns Hopkins, Baltimore	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Julie Paik, MD, MHS	Johns Hopkins, Baltimore	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Jemima Albayda, MD	Johns Hopkins, Baltimore	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Livia Casciola-Rosen, PhD	Johns Hopkins, Baltimore	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Cheilonda Johnson, MD	Johns Hopkins, Baltimore	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Sonye K. Danoff, MD, PhD	Johns Hopkins, Baltimore	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Lisa Christopher-Stine, MD, MPH	Johns Hopkins, Baltimore	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Eleni Tiniakou, MD	Johns Hopkins, Baltimore	Author	Designed and conceptualized study; interpreted the data; major role in the acquisition of data; revised the manuscript for intellectual content
Andrew L. Mammen, MD, PhD	NIH, Bethesda	Author	Designed and conceptualized study; interpreted the data; major role in the acquisition of data; drafted the manuscript for intellectual content

REFERENCES

1. Dalakas MC. Inflammatory muscle diseases. N Engl J Med. 2015;372(18):1734-47.

2. Sharp GC, Irvin WS, Tan EM, Gould RG, Holman HR. Mixed connective tissue disease--an apparently distinct rheumatic disease syndrome associated with a specific antibody to an extractable nuclear antigen (ENA). Am J Med. 1972;52(2):148-59.

3. De Lorenzo R, Pinal-Fernandez I, Huang W, Albayda J, Tiniakou E, Johnson C, et al. Muscular and extramuscular clinical features of patients with anti-PM/Scl autoantibodies. Neurology. 2018;90(23):e2068-e76.

4. Rider LG, Werth VP, Huber AM, Alexanderson H, Rao AP, Ruperto N, et al. Measures of adult and juvenile dermatomyositis, polymyositis, and inclusion body myositis: Physician and Patient/Parent Global Activity, Manual Muscle Testing (MMT), Health Assessment Questionnaire (HAQ)/Childhood Health Assessment Questionnaire (C-HAQ), Childhood Myositis Assessment Scale (CMAS), Myositis Disease Activity Assessment Tool (MDAAT), Disease Activity Score (DAS), Short Form 36 (SF-36), Child Health Questionnaire (CHQ), physician global damage, Myositis Damage Index (MDI), Quantitative Muscle Testing (QMT), Myositis Functional Index-2 (FI-2), Myositis Activities Profile (MAP), Inclusion Body Myositis Functional Rating Scale (IBMFRS), Cutaneous Dermatomyositis Disease Area and Severity Index (CDASI), Cutaneous Assessment Tool (CAT), Dermatomyositis Skin Severity Index (DSSI), Skindex, and Dermatology Life Quality Index (DLQI). Arthritis Care Res (Hoboken). 2011;63 Suppl 11:S118-57.

5. Travis WD, Costabel U, Hansell DM, King TE, Jr., Lynch DA, Nicholson AG, et al. An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. Am J Respir Crit Care Med. 2013;188(6):733-48.

6. Galie N, Humbert M, Vachiery JL, Gibbs S, Lang I, Torbicki A, et al. 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: The Joint Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS): Endorsed by: Association for European Paediatric and Congenital Cardiology (AEPC), International Society for Heart and Lung Transplantation (ISHLT). Eur Heart J. 2016;37(1):67-119.

7. Macintyre N, Crapo RO, Viegi G, Johnson DC, van der Grinten CP, Brusasco V, et al. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. Eur Respir J. 2005;26(4):720-35.

8. Bohan A, Peter JB. Polymyositis and dermatomyositis (second of two parts). N Engl J Med. 1975;292(8):403-7.

9. Kasukawa R, Tojo T, Miyawaki S. Preliminary diagnostic criteria for classification of mixed connective tissue disease. In: Kasukawa R, Sharp G, editors. Mixed Connective Tissue Disease and Antinuclear Antibodies. Amsterdam: Elsevier; 1987. p. 41.

10. Alarcon Segovia D, Villareal M. Classification and diagnostic criteria for mixed connective tissue disease. In: Kasukawa R, Sharp G, editors. Mixed Connective Tissue Disease and Anti-nuclear Antibodies. Amsterdam: Elsevier; 1987. p. 33.

11. Kahn M. Syndrom de Sharp. In: Kahn M, Peltier A, Meyer O, Piette J, editors. Les maladies systemiques, 3rd. Amsterdam: Elsevier; 1991. p. 545.

12. van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative. Arthritis Rheum. 2013;65(11):2737-47.

13. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum. 1997;40(9):1725.

14. Marie I, Hatron PY, Dominique S, Cherin P, Mouthon L, Menard JF, et al. Short-term and long-term outcome of anti-Jo1-positive patients with anti-Ro52 antibody. Semin Arthritis Rheum. 2012;41(6):890-9.

15. Coppo P, Clauvel JP, Bengoufa D, Oksenhendler E, Lacroix C, Lassoued K. Inflammatory myositis associated with anti-U1-small nuclear ribonucleoprotein antibodies: a subset of myositis associated with a favourable outcome. Rheumatology (Oxford). 2002;41(9):1040-6.

16. Oxenhandler R, Hart M, Corman L, Sharp G, Adelstein E. Pathology of skeletal muscle in mixed connective tissue disease. Arthritis Rheum. 1977;20(4):985-8.

17. Vianna MA, Borges CT, Borba EF, Caleiro MT, Bonfa E, Marie SK. Myositis in mixed connective tissue disease: a unique syndrome characterized by immunohistopathologic elements of both polymyositis and dermatomyositis. Arq Neuropsiquiatr. 2004;62(4):923-34.

Figure 1. Muscle biopsy from an anti-U1-RNP-positive patient (H&E). The arrow shows a necrotic cell undergoing early-stage myophagocytosis. Regenerating myofibers, characterized by basophilic cytoplasm and enlarged nuclei, are indicated by the arrow heads.

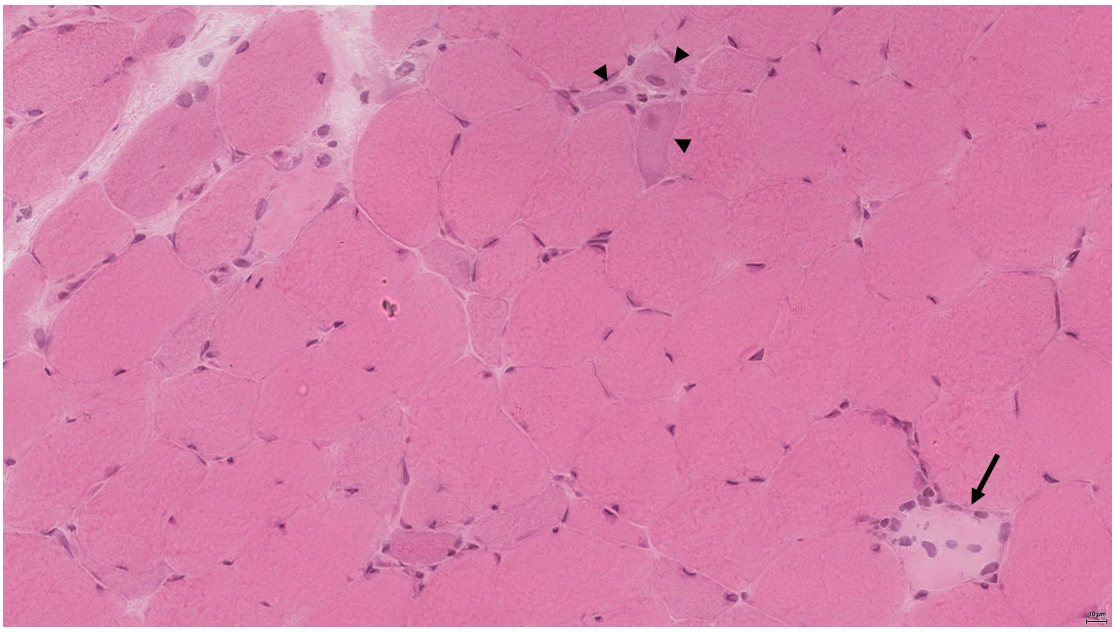


Table 1: General features of anti-U1-RNP-positive patients and control groups.

	Anti-U1-RNP (n=20)	DM (n=178)	AS (n=132)	IMNM (n=135)
Female sex	80% (16)	75% (134)	73% (97)	63% (85)
Race				
White	40% (8)	77% (137)***	61% (81)	67% (91)*
Black	60% (12)	12% (22)***	30% (39)**	24% (32)***
Other races	0% (0)	11% (19)	9% (12)	9% (12)
Age of onset	37.3 (17.9)	47.1 (15.6)**	45.0 (13.3)*	51.5 (14.9)***
Time of follow-up (years)	6.4 (3.9)	4.2 (3.5)**	4.7 (3.9)	4.0 (3.9)*
Number of visits	22.0 (18.8)	9.9 (7.4)***	9.6 (7.2)***	9.1 (9.3)***
Cancer associated myositis	0% (0)	8% (15)	3% (4)	5% (7)
Anti-Ro52	75% (15)	22% (39)***	80% (106)	17% (23)***
Treatments				
Corticosteroids	90% (18)	83% (147)	96% (127)	75% (101)
Azathioprine	40% (8)	26% (47)	58% (76)	27% (36)
Methotrexate	45% (9)	51% (90)	47% (62)	50% (67)
Mycophenolate	45% (9)	35% (63)	38% (50)	20% (27)*
IVIG	40% (8)	48% (86)	37% (49)	37% (50)
Rituximab	35% (7)	16% (28)	20% (27)	24% (32)

* p<0.05, ** p<0.01, *** p<0.001

Dichotomous variables were expressed as percentage (count) and continuous variables as mean (SD). Bivariate comparisons of continuous variables were made using Student´s t-test while bivariate comparisons of dichotomous variables were made either using chi-squared test or Fisher´s exact test, as appropriate. Each one of the clinical groups was compared to the sample of anti-U1-RNP patients.

Table 2: Clinical features of anti-U1-RNP-positive patients and control groups at the onset of the disease.

	Anti-U1-RNP (n=20)	DM (n=178)	AS (n=132)	IMNM (n=135)
Muscle involvement				
Muscle weakness	15% (3)	47% (83)**	55% (73)***	83% (112)***
Myalgia	20% (4)	20% (36)	24% (32)	18% (24)
Skin involvement				
DM-specific skin involvement	15% (3)	72% (129)***	16% (21)	1% (2)*
SSc-specific skin involvement	0% (0)	0% (0)	0% (0)	0% (0)
Raynaud phenomenon	20% (4)	4% (8)*	19% (25)	4% (5)*
Telangectasias	0% (0)	0% (0)	1% (1)	0% (0)
Ulcers	10% (2)	3% (5)	0% (0)*	0% (0)*
Carpal tunnel	5% (1)	1% (1)	5% (6)	1% (1)
Livedo reticularis	5% (1)	0% (0)	0% (0)	0% (0)
Mechanics hands	0% (0)	4% (7)	11% (15)	1% (1)
Calcinosis	0% (0)	2% (4)	3% (4)	0% (0)
Subcutaneous edema	10% (2)	6% (10)	5% (6)	0% (0)*
Puffy hands	5% (1)	2% (3)	1% (1)	0% (0)
Lung involvement				
Interstitial lung disease	5% (1)	6% (10)	52% (68)***	1% (1)
Pulmonary hypertension	0% (0)	0% (0)	1% (1)	0% (0)
Esophageal involvement				
Gastroesophageal reflux	5% (1)	0% (0)	10% (13)	0% (0)
Dysphagia	10% (2)	10% (17)	8% (11)	7% (10)
Joint involvement				
Arthritis	15% (3)	6% (11)	20% (27)	1% (1)**
Arthralgia	30% (6)	16% (28)	44% (58)	7% (9)**
Systemic involvement				
Fever	20% (4)	6% (11)	11% (15)	1% (1)***
Sicca syndrome	0% (0)	1% (2)	2% (3)	0% (0)
Pericarditis	5% (1)	0% (0)	0% (0)	0% (0)
Glomerulonephritis	0% (0)	0% (0)	0% (0)	0% (0)

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Chi-squared or Fisher´s exact tests were used to compare each one of the clinical groups with the anti-U1-RNP patients.

Table 3: Cumulative clinical features of anti-U1-RNP-positive patients and control groups.

	Anti-U1-RNP (n=20)	DM (n=178)	AS (n=132)	IMNM (n=135)
Muscle involvement				
Muscle weakness	80% (16)	85% (152)	90% (119)	96% (130)*
Myalgia	60% (12)	56% (100)	65% (86)	52% (70)
Skin involvement				
DM-specific skin involvement	60% (12)	96% (171)***	60% (79)	4% (5)***
SSc-specific skin involvement	25% (5)	2% (3)***	13% (17)	0% (0)***
Raynaud phenomenon	80% (16)	22% (40)***	39% (52)***	15% (20)***
Telangectasias	20% (4)	21% (37)	20% (26)	8% (11)
Ulcers	15% (3)	14% (25)	7% (9)	0% (0)**
Carpal tunnel	15% (3)	8% (15)	20% (27)	10% (13)
Livedo reticularis	20% (4)	12% (22)	10% (13)	4% (5)*
Mechanics hands	50% (10)	28% (49)*	58% (77)	5% (7)***
Calcinosis	25% (5)	21% (38)	9% (12)	1% (1)***
Subcutaneous edema	35% (7)	18% (32)	27% (35)	4% (6)***
Puffy hands	20% (4)	8% (15)	10% (13)	0% (0)***
Lung involvement				
Interstitial lung disease	45% (9)	13% (24)**	80% (106)**	6% (8)***
Pulmonary hypertension	25% (5)	3% (5)**	20% (27)	1% (2)***
Esophageal involvement				
Gastroesophageal reflux	45% (9)	29% (51)	29% (38)	25% (34)
Dysphagia	50% (10)	53% (95)	18% (24)**	39% (53)
Joint involvement				
Arthritis	60% (12)	18% (32)***	55% (72)	6% (8)***
Arthralgia	65% (13)	51% (90)	62% (82)	36% (49)*
Systemic involvement				
Fever	35% (7)	18% (32)	24% (32)	7% (10)**
Sicca syndrome	15% (3)	31% (55)	48% (63)**	19% (26)
Pericarditis	40% (8)	0% (0)***	1% (1)***	0% (0)***
Glomerulonephritis	25% (5)	0% (0)***	1% (1)***	0% (0)***

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Chi-squared or Fisher´s exact tests were used to compare each one of the clinical groups with the anti-U1-RNP patients.

Table 4. Pattern of weakness at the first visit of anti-U1-RNP-positive patients and control groups.

	Anti-U1-RNP (n=20)	DM (n=178)	AS (n=132)	IMNM (n=135)
Neck flexors	10.0 (0.0)	9.0 (1.9)*	9.7 (1.1)	8.4 (2.5)*
Neck extensors	10.0 (0.0)	9.8 (0.6)	9.9 (0.6)	9.7 (1.1)
Arm abductors	8.8 (2.0)	8.6 (2.2)	9.2 (1.4)	8.0 (2.0)
Elbow flexors	9.6 (1.0)	9.2 (1.3)	9.6 (0.7)	8.8 (1.5)*
Elbow extensors	9.5 (1.2)	9.0 (1.5)	9.5 (1.0)	8.7 (1.5)*
Wrist flexors	9.8 (0.7)	9.8 (0.7)	9.9 (0.4)	9.8 (0.7)
Wrist extensors	9.8 (0.7)	9.7 (0.9)	9.9 (0.4)	9.8 (0.6)
Finger flexors	9.8 (0.7)	9.8 (0.6)	9.8 (0.6)	9.7 (0.8)
Finger extensors	9.7 (0.8)	9.7 (1.0)	9.9 (0.3)	9.9 (0.4)
Hip flexors	8.4 (2.3)	8.3 (2.3)	8.7 (2.0)	5.8 (3.1)***
Hip extensors	9.7 (0.8)	9.6 (1.4)	9.7 (0.9)	8.9 (2.1)
Knee flexors	9.9 (0.4)	9.8 (0.5)	9.9 (0.4)	9.0 (1.6)*
Knee extensors	9.3 (1.9)	9.7 (0.9)	9.8 (0.5)	9.2 (1.3)
Ankle flexors	10.0 (0.0)	9.8 (1.0)	9.9 (0.3)	9.8 (0.6)
Ankle extensors	9.9 (0.2)	9.9 (0.4)	10.0 (0.0)	9.8 (0.6)

* p<0.05, ** p<0.01, *** p<0.001

Strength values were expressed as means (SD) and bivariate comparisons were made using Student´s t-test.
This table includes strength data from all patients, both with and without weakness, at their initial visit to the Johns Hopkins Myositis Center.

Table 5. Muscle strength, muscle enzyme levels, and pulmonary function testing in anti-U1-RNP-positive patients and control groups.

	Anti-U1-RNP (n=20)	DM (n=178)	AS (n=132)	IMNM (n=135)
Mean hip flexor strength	8.8 (1.8)	8.8 (1.8)	9.0 (1.5)	6.7 (2.7)**
Hip flexors strength at last visit	9.7 (0.8)	9.1 (1.9)	9.1 (1.6)	6.9 (3.5)**
Mean arm abductor strength	9.1 (1.2)	9.1 (1.6)	9.4 (1.1)	8.5 (1.9)
Arm abductors strength at last visit	9.5 (1.4)	9.3 (1.9)	9.4 (1.3)	8.8 (2.3)
Median CK	229 (121-632)	117 (68-290)*	282 (114-963)	1401 (502-2969)***
Maximum CK	708 (399-4108)	719 (139-3508)	1352 (396-5850)	4706 (2000-8990)**
Mean aldolase	26.9 (35.9)	9.4 (7.5)***	24.4 (43.8)	29.3 (29.7)
Maximum aldolase	53.3 (86.7)	13.4 (16.3)***	54.4 (184.0)	49.9 (60.3)
Mean FVC	71.6 (27.7)	89.2 (21.0)**	72.5 (19.5)	87.8 (20.0)*
Minimum FVC(%)	67.1 (28.0)	86.2 (23.3)**	65.4 (22.6)	86.6 (20.6)**
Mean DLCO	64.2 (34.7)	96.4 (22.6)***	69.5 (23.2)	103.3 (24.2)***
Minimum DLCO(%)	60.6 (35.7)	92.2 (25.2)***	59.6 (25.4)	101.3 (26.3)***

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Strength and FVC values were expressed as means (SD) and CK as medians (Q1-Q3). Bivariate comparisons were made using Student´s t-test for the strength and Wilcoxon rank-sum test for CK. Follow-up strength was defined as the mean strength of all the visits, excluding the first one. Each one of the clinical groups was compared to the sample of anti-U1-RNP patients.

Table 6: Muscle biopsy features of anti-U1RNP and control groups.

	Anti-U1RNP (n=7)	DM (n=45)	ASyS (n=27)	IMNM (n=70)
Necrotizing myopathy	57% (4)	11% (5)*	22% (6)	80% (56)
Perifascicular atrophy	0% (0)	56% (25)*	52% (14)*	0% (0)
Perivascular inflammation	43% (3)	62% (28)	63% (17)	27% (19)
Primary inflammation	0% (0)	9% (4)	30% (8)	16% (11)

* p<0.05, ** p<0.01, *** p<0.001

Chi-squared or Fisher´s exact tests were used to compare each one of the clinical groups with the anti-U1RNP patients.

A.2 Efficacy and adverse effects of methotrexate compared with azathioprine in the Antisynthetase Syndrome. Clin Exp Rheumatol 2019 (PMID: 31074729).

In this study, we analyzed the efficacy in terms of muscle strength, and corticosteroid tapering as well as the prevalence of adverse effects in patients with the AS treated with AZA compared to those treated with MTX.

We compared the clinical outcomes in AS patients treated with AZA versus MTX including change in corticosteroid dose, strength, and CK as well as the prevalence of adverse effects.

Among 169 patients with AS, 102 were treated at some point exclusively with either AZA or MTX (± corticosteroids). There were no significant differences in the rate of muscle strength recovery, CK decrease or corticosteroid tapering between those AS patients treated with MTX versus AZA. The prevalence of adverse events in patients treated with AZA and MTX was similar (29% vs. 25%, $p>0.05$); elevated liver enzymes (17% AZA vs. 12% MTX) and gastrointestinal involvement (10% AZA vs. 8% MTX) were the most common adverse events. While no patients treated with AZA developed lung complications, two of the patients treated with MTX experienced reversible pneumonitis with MTX cessation.

In conlussion, AZA and MTX showed similar efficacy and adverse events in patients with AS. Pneumonitis is a rare but important event in patients receiving MTX.

Efficacy and adverse effects of methotrexate compared with azathioprine in the Antisynthetase Syndrome.

Casal-Dominguez M[1,2*], Pinal-Fernandez I[1,2,3*], Huapaya J[2], Albayda J[2], Paik J J[2], Johnson C[2], Silhan L[2], Mammen AL[1,2], Danoff S K[2#], Christopher-Stine L[2#].

*[#] These authors contributed equally to this project**.**

Iago Pinal-Fernandez, M.D, Ph.D., Maria Casal-Dominguez, M.D. and Andrew L. Mammen, M.D., Ph.D.: National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD. Iago Pinal-Fernandez, M.D, Ph.D., Maria Casal-Dominguez, Julio Huapaya M.D., Jemima Albayda M.D., Julie J Paik M.D., Cheilonda Johnson M.D., Leann Silhan M.D., Andrew L. Mammen M.D Ph.D., Sonye K Danoff, M.D., Ph.D. and Lisa Christopher-Stine, M.D., M.P.H.: Johns Hopkins University School of Medicine, Baltimore, MD. Iago Pinal-Fernandez: Faculty of Health Sciences, Universitat Oberta de Catalunya, Barcelona, Spain.

Address correspondence to:

Lisa Christopher-Stine, MD, MPH

5200 Eastern Avenue;MFL Center Tower #4500

Baltimore, MD 21224

Sonye K. Danoff, MD, PhD

1800 E. Monument Street, 5th Floor

Baltimore, MD 21287

ABSTRACT:

Objectives: To study the efficacy in terms of muscle strength, and corticosteroid tapering as well as the prevalence of adverse effects in patients with the antisynthetase syndrome (ASyS) treated with azathioprine (AZA) compared to those treated with methotrexate (MTX).

Methods: We compared the clinical outcomes in ASyS patients treated with AZA versus MTX including change in corticosteroid dose, strength, and creatine kinase (CK) as well as the prevalence of adverse effects.

Results: Among 169 patients with ASyS, 102 were treated at some point exclusively with either AZA or MTX (+/- corticosteroids). There were no significant differences in the rate of muscle strength recovery, CK decrease or corticosteroid tapering between those ASyS patients treated with MTX versus AZA. The prevalence of adverse events in patients treated with AZA and MTX was similar (29% vs. 25%, $p>0.05$); elevated liver enzymes (17% AZA vs. 12% MTX) and gastrointestinal involvement (10% AZA vs. 8% MTX) were the most common adverse events. While no patients treated with AZA developed lung complications, two of the patients treated with MTX experienced reversible pneumonitis with MTX cessation.

Conclusions: AZA and MTX showed similar efficacy and adverse events in patients with ASyS. Pneumonitis is a rare but important event in patients receiving MTX.

Keywords: Antisynthetase syndrome, Methotrexate, Azathioprine, Myositis.

INTRODUCTION

The antisynthetase syndrome (ASyS), first described as an entity in 1990,(1) is characterized by the presence of an antisynthetase antibody which targets cytoplasmic enzymes that catalyze the formation of the aminoacyl-tRNA complex. Clinically, this syndrome is characterized by myositis, interstitial lung disease (ILD) or both. Other features including Raynaud's phenomenon, arthritis, fever and mechanic's hands are also common clinical features of the ASyS syndrome.(2, 3)

Corticosteroids are considered first line treatment in the ASyS, but most of the time, other immunosuppressive agents are needed. Methotrexate (MTX) or Azathioprine (AZA) are common first line therapies in ASyS patients,(4) but mycophenolate,(5, 6) the calcineurin inhibitors, cyclosporine and tacrolimus,(7-9) as well as cyclosphospamide(10) and rituximab,(11, 12) have also been used for the treatment of these patients with good results.

The high risk of patients with the ASyS developing ILD makes MTX use controversial due its potential to induce hypersensitivity pneumonitis, which may be mistaken for ILD related to the underlying ASyS. Unlike ASyS associated ILD, MTX pneumonitis is typically reversible with MTX cessation. It has been suggested that MTX may be more beneficial than AZA in some groups of patients who are refractory to prednisone.(13) Although MTX and azathioprine are two of the most widely used immunosuppressant drugs for the ASyS, the efficacy to treat the manifestations of the disease, comparative efficacy as steroid-sparing drugs and secondary effects are, to a large extent, unknown.

Our main objective was to study the differences in muscle strength and changes in the dose of corticosteroids, as well as the profile of adverse effects between ASyS patients treated with AZA versus those treated with MTX.

MATERIALS AND METHODS

In this longitudinal cohort study, we included all Johns Hopkins Myositis Center patients who were positive for one of the ASyS antibodies (anti-Jo1, anti-PL7, anti-PL12, anti-OJ or anti-EJ) and presented with at least two of the following clinical manifestations: myositis, ILD, polyarthritis or mechanic's hands. All the treatments administered at each clinical evaluation were recorded, and those patients treated with AZA or MTX without concomitant use of other steroid sparing agents were included for analysis. The sera from all patients was screened for anti-Jo1, anti-PL7, anti-PL12, anti-EJ, and anti-OJ by ELISA, line blotting (Euroline Myositis Profile 4; Euroimmun), by immunoprecipitation at the Oklahoma Medical Research Foundation and/or using Quest Diagnostics myositis panels.

This study was approved by the Johns Hopkins Institutional Review Board, and written informed consent was obtained from each participant.

The change in strength, CK and dose of corticosteroids during the period that patients were exposed exclusively to AZA or MTX (+/- corticosteroids) were analyzed using multilevel regression models adjusted for age at onset, sex, race, dose of corticosteroids, type of antisynthetase antibody and time from the onset to the clinical evaluation.

At each visit, arm abduction and hip flexion strength, were evaluated by the examining physician using the Medical Research Council (MRC) scale. This scale was transformed to Kendall´s 0-10 scale for analysis purposes as previously described.(14) Several investigators examined the patients, but serial strength measurements for each patient were made by the same physician.

Adverse effects were recorded as reported by the attending physician. Accordingly, laboratory abnormalities, like elevation of the liver enzymes, leukopenia or pancytopenia were based on the normality cutoff of the corresponding facility where the tests were performed. Also, MTX associated pneumonitis was defined by the occurrence of cough or dyspnea in a time course consistent with exposure to MTX which resolved with stopping this medication. All the episodes of possible MTX pneumonitis were reviewed by three of the authors (SD, MCD and IPF). The probability of the adverse effect was quantitated using Naranjo's method.(15)

Dichotomous variables were expressed as percentages and absolute frequencies, and continuous features were reported as means and standard deviations (SD). Pairwise comparisons for dichotomous variables between groups were made using chi-square test or Fisher's exact test, as appropriate. Student´s t-test was used to compare continuous variables among groups. CK, a highly positively skewed variable, was compared using Wilcoxon rank-sum test for the univariate analysis and transformed through a base-10 logarithm for regression analysis. Locally weighted regression was applied to analyse graphically the evolution of the strength and dose of corticostestoids over time.

All statistical analyses were performed using Stata/MP 14.1. A 2-sided p value of 0.05 or less was considered significant with no correction for multiple comparisons.

RESULTS

Of 169 patients with the ASyS (73% women), 124 of them were positive for anti-Jo1, 23 for anti-PL12, 16 for anti-PL7 and 3 for anti-EJ and anti-OJ respectively. Of these patients, 63 (37%) were treated with AZA exclusively, 26 (15%) were treated with MTX exclusively and 26 (15%) were treated with both AZA and MTX at some point of their evolution (total of 115 patients). The average length of exposure to these medications was 24 months for AZA and 29 months for MTX. In general, AZA was administered to patients with less muscle involvement (lower CK and higher strength) and more severe lung involvement (lower FVC) while MTX was given to patients with milder lung involvement. MTX and AZA in combination were used in patients with more severe muscle disease (lower strength and higher CK). Patients treated with MTX were mostly white and presented anti-Jo1 autoantibodies more commonly than the other treatment groups (Supplementary Table 1).

Twenty-nine percent of all the patients who were treated with AZA showed adverse effects to this drug, compared with 25% of the patients that were treated with MTX ($p>0.05$). The most common adverse effects with both drugs were elevated liver function tests (17% AZA vs. 12% MTX), gastrointestinal symptoms such as nausea and diarrhea (10% AZA vs. 8% MTX) and cytopenias (6% AZA vs. 4% MTX), but none of these were significantly different between both drugs. Of note, most reported adverse effects were mild. While no patient with AZA experienced pulmonary adverse effects related to the use of the immunosuppressant treatment, four patients treated with MTX (8%) reported pulmonary events ($p=0.02$) but only two presented clear evidence of MTX pneumonitis ($p=0.1$). These two patients did not have pre-existing lung involvement. One had cough and lung CT involvement that reverted rapidly after MTX discontinuation and the other one was challenged twice with MTX developing cough

and shortness of breath that reverted quickly after stopping the drug both times. (Table 1). Both patients had a Naranjo's score(15) of 5 which corresponds to a probable adverse effect. Complementarily, the two other patients that reported pulmonary events were patients with pre-existing ILD reporting worsening of their respiratory symptoms (one cough and one dyspnea) during MTX treatment. However, the time course was considered inconsistent with MTX pneumonitis and there were no objective tests available to show worsening of the ILD.

Of the 115 patients treated with AZA or MTX, 102 received either of the drugs combined with no other immunosuppressant drug than corticosteroids (59 AZA, 20 MTX and 23 AZA and MTX at different time points). These patients accounted for 450 visits under treatment with AZA or MTX +/- corticosteroids (mean of 4.4 visits per patient) that were used to compare the rate of change of strength, and corticosteroid tapering. There were no significant differences in the rate of muscle strength recovery ($p=0.9$), CK decrease ($p=0.6$) or corticosteroid tapering ($p=0.9$) during treatment with AZA or MTX (Figure 1).

DISCUSSION

This study demonstrated that MTX and AZA are comparable in terms of rate of muscle strength recovery, rate of corticosteroid tapering, and rate of CK decrease with similar rates of adverse events. We did identify two episodes of MTX pneumonitis which reversed with discontinuation of therapy.

MTX has been reported to cause pneumonitis in 4-8% of the patients exposed to this drug(16) and this may dissuade clinicians from prescribing MTX in patients with ASyS autoantibodies or preexisting ILD.(17) Our study confirms previous data regarding the low prevalence of MTX pneumonitis (4%).

Some authors have suggested an increased efficacy of MTX over AZA in selected groups of patients.(13) Our study found that MTX was comparable to AZA in terms of efficacy in patients with the ASyS.

The data that we report is based on a cohort study followed longitudinally in the context of routine clinical care and not a clinical trial. The assignment of therapy to the individual patient was based on physician preference, thus, it is possible that some of the analyses were subject to unaccounted bias. Patients underwent PFTs and CT imaging as part of clinical care, therefore, we cannot comment on the appearance of some features such as ILD except as detected based on clinical symptoms and findings Likewise, adverse events were both patient-reported and surveyed by the treating clinicians but not necessarily in a routine manner for all patients. Moreover, the small number of patients in each group precludes a cautious interpretation of our results.

In conclusion, in our real-world clinical experience, we found that compared with AZA, MTX had a similar prevalence of adverse effects and efficacy. MTX pneumonitis occurred in 4% of patients started on this medication, but was entirely

reversible with stopping therapy, thus, attention to this potential adverse event is important with rapid discontinuation of therapy if symptoms occur.

CONFLICT OF INTERESTS: No potential conflict of interest was reported by the authors.

FUNDING: This research was supported in part by the Intramural Research Program of the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health. SKD, LCS and the the Myositis Research Database are supported by the The Huayi and Siuling Zhang Discovery Fund. IPF´s research is supported by a Fellowship from the Myositis Association.

REFERENCES

1. MARGUERIE C, BUNN CC, BEYNON HL, et al. Polymyositis, pulmonary fibrosis and autoantibodies to aminoacyl-tRNA synthetase enzymes. Q J Med 1990; 77: 1019-38.

2. TRALLERO-ARAGUAS E, GRAU-JUNYENT JM, LABIRUA-ITURBURU A, et al. Clinical manifestations and long-term outcome of anti-Jo1 antisynthetase patients in a large cohort of Spanish patients from the GEAS-IIM group. Semin Arthritis Rheum 2016; 46: 225-31.

3. MARASCO E, CIOFFI E, COMETI L, et al. One year in review 2018: idiopathic inflammatory myopathies. Clin Exp Rheumatol 2018; 36: 937-47.

4. ODDIS CV. Update on the pharmacological treatment of adult myositis. J Intern Med 2016; 280: 63-74.

5. SAKETKOO LA, ESPINOZA LR. Experience of mycophenolate mofetil in 10 patients with autoimmune-related interstitial lung disease demonstrates promising effects. Am J Med Sci 2009; 337: 329-35.

6. SWIGRIS JJ, OLSON AL, FISCHER A, et al. Mycophenolate mofetil is safe, well tolerated, and preserves lung function in patients with connective tissue disease-related interstitial lung disease. Chest 2006; 130: 30-6.

7. KOTANI T, TAKEUCHI T, MAKINO S, et al. Combination with corticosteroids and cyclosporin-A improves pulmonary function test results and chest HRCT findings in dermatomyositis patients with acute/subacute interstitial pneumonia. Clin Rheumatol 2011; 30: 1021-8.

8. ODDIS CV, SCIURBA FC, ELMAGD KA, STARZL TE. Tacrolimus in refractory polymyositis with interstitial lung disease. Lancet 1999; 353: 1762-3.

9. WILKES MR, SEREIKA SM, FERTIG N, LUCAS MR, ODDIS CV. Treatment of antisynthetase-associated interstitial lung disease with tacrolimus. Arthritis Rheum 2005; 52: 2439-46.

10. YAMASAKI Y, YAMADA H, YAMASAKI M, et al. Intravenous cyclophosphamide therapy for progressive interstitial pneumonia in patients with polymyositis/dermatomyositis. Rheumatology (Oxford) 2007; 46: 124-30.

11. ANDERSSON H, SEM M, LUND MB, et al. Long-term experience with rituximab in anti-synthetase syndrome-related interstitial lung disease. Rheumatology (Oxford) 2015; 54: 1420-8.

12. KEIR GJ, MAHER TM, MING D, et al. Rituximab in severe, treatment-refractory interstitial lung disease. Respirology 2014; 19: 353-9.

13. JOFFE MM, LOVE LA, LEFF RL, et al. Drug therapy of the idiopathic inflammatory myopathies: predictors of response to prednisone, azathioprine, and methotrexate and a comparison of their efficacy. Am J Med 1993; 94: 379-87.

14. PINAL-FERNANDEZ I, CASAL-DOMINGUEZ M, HUAPAYA JA, et al. A longitudinal cohort study of the anti-synthetase syndrome: increased severity of interstitial lung disease in black patients and patients with anti-PL7 and anti-PL12 autoantibodies. Rheumatology (Oxford) 2017.

15. NARANJO CA, BUSTO U, SELLERS EM, et al. A method for estimating the probability of adverse drug reactions. Clin Pharmacol Ther 1981; 30: 239-45.

16. SATHI N, CHIKURA B, KAUSHIK VV, WISWELL R, DAWSON JK. How common is methotrexate pneumonitis? A large prospective study investigates. Clin Rheumatol 2012; 31: 79-83.

17. AMATO AA, GRIGGS RC. Treatment of idiopathic inflammatory myopathies. Curr Opin Neurol 2003; 16: 569-75.

Supplementary Table 1. Clinical features of the antisynthetase patients taking either azathioprine (AZA), methotrexate (MTX) or both at any given point in their evolution.

	AZA (n=63)	MTX (n=26)	AZA and MTX# (n=26)	Total (n=115)
Type of autoantibody				
Anti-Jo1	70% (44)*	100% (26)**	77% (20)	78% (90)
Anti-PL7	10% (6)	0% (0)	8% (2)	7% (8)
Anti-PL12	16% (10)	0% (0)*	15% (4)	12% (14)
Anti-EJ	3% (2)	0% (0)	0% (0)	2% (2)
Anti-OJ	2% (1)	0% (0)	0% (0)	1% (1)
Age of onset	47.8 (13.3)	45.5 (11.9)	42.5 (11.9)	46.1 (12.8)
Time of follow-up (years)	4.1 (3.0)	4.9 (3.8)	4.9 (4.0)	4.4 (3.4)
Number of visits	10.3 (6.9)	8.7 (6.6)	10.5 (7.7)	10.0 (7.0)
Female sex	75% (47)	81% (21)	85% (22)	78% (90)
Race				
White	54% (34)	85% (22)**	46% (12)	59% (68)
Black	40% (25)	12% (3)*	35% (9)	32% (37)
Other races	6% (4)	4% (1)	19% (5)*	9% (10)
Mortality	5% (3)	0% (0)	12% (3)	5% (6)
Proximal weakness	86% (54)	85% (22)	92% (23)	87% (99)
Strength				
Mean strength	9.3 (0.8)*	9.0 (1.4)	8.8 (1.5)	9.1 (1.1)
Minimum strength	8.8 (1.2)*	8.2 (2.1)	7.9 (1.8)*	8.4 (1.6)
Creatine kinase (CK)				
Median CK	185 (99-630)*	283 (148-658)	428 (157-1164)*	260 (122-826)
Maximum CK	896 (254-3360)*	1412 (647-4144)	2508 (473-7904)	1331 (301-5000)
Interstitial lung disease	92% (58)**	62% (16)**	77% (20)	82% (94)
FVC				
Mean FVC	64.4 (20.0)**	86.7 (18.9)***	67.3 (15.9)	70.1 (20.9)
Minimum FVC	55.4 (21.4)***	80.7 (20.3)***	60.6 (18.9)	62.4 (22.9)
Treatments				
Corticosteroids	97% (61)	92% (24)	100% (26)	97% (111)
Mycophenolate	32% (20)	23% (6)	42% (11)	32% (37)
IVIG	32% (20)	31% (8)	50% (13)	36% (41)
Rituximab	25% (16)	31% (8)	27% (7)	27% (31)

** p<0.05, ** p<0.01, *** p<0.001*
 # Methotrexate and azathioprine were not necessarily administered concomitantly.

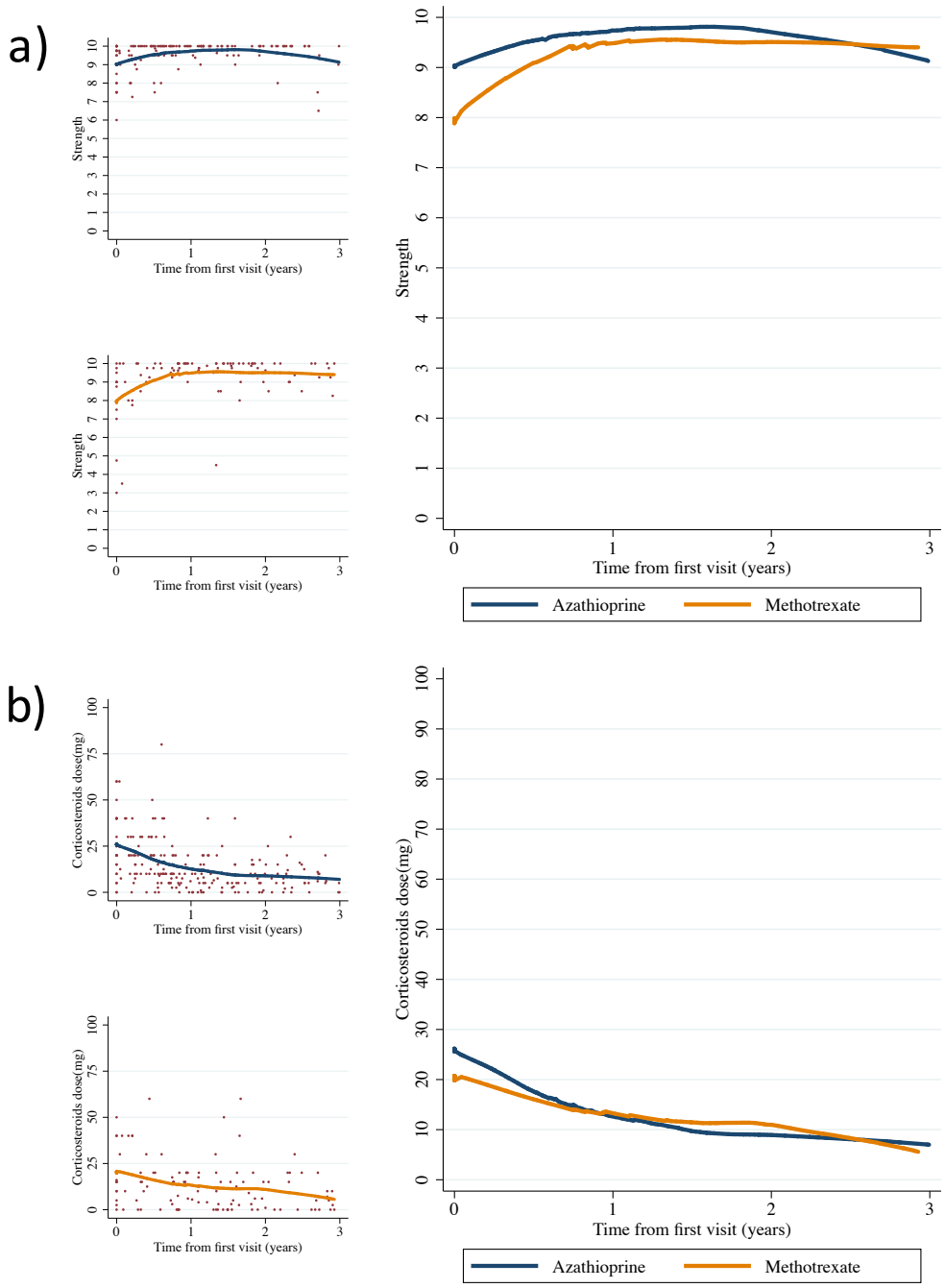
Each one of the treatment groups was compared to the rest of patients. Dichotomous variables were shown as %(n) and compared using the chi-squared or Fisher´s exact tests while continuous variables were reported as mean (SD) and compared using Student's test. The CK, a positively skewed variable, was presented as median (Q1-Q3) and compared using Wilcoxon rank-sum test.

Table 1. Adverse effects of methotrexate and azathioprine in patients with the antisynthetase syndrome.

	Azathioprine (n=89)	Methotrexate (n=52)	p-value
Elevated liver function tests	17% (15)	12% (6)	0.4
Gastrointestinal	10% (9)	8% (4)	0.8
Nausea	7% (6)	8% (4)	1.0
Diarrhea	2% (2)	0% (0)	0.5
Abdominal pain	1% (1)	0% (0)	1.0
Leukopenia	4% (4)	2% (1)	0.7
Pancytopenia	1% (1)	2% (1)	1.0
Other adverse effects	6% (5)	13% (7)	0.1
Methotrexate pneumonitis	0% (0)	4% (2)	0.1
Rate of adverse effects	29% (26)	25% (13)	0.5

Dichotomous variables were compared using Chi-squared or Fisher´s test as appropriate.

Figure 1. Strength recovery (a) and corticosteroid tapering (b) in patients with the antisynthetase syndrome treated with azathioprine and methotrexate.



A.3 Anti-Ro52 autoantibodies are associated with interstitial lung disease and more severe disease in patients with juvenile myositis. Ann Rheum Dis 2019 (PMID: 31018961).

Anti-Ro52 autoantibodies are associated with more severe ILD in adult myositis patients with AS autoantibodies. However, few studies have examined anti-Ro52 autoantibodies in juvenile myositis. The purpose of this study was to define the prevalence and clinical features associated with anti-Ro52 autoantibodies in a large cohort of patients with juvenile myositis.

For this, we screened sera from 302 patients with juvenile DM, 25 patients with juvenile PM, and 44 patients with juvenile connective tissue disease–myositis overlap for anti-Ro52 autoantibodies by ELISA. Clinical characteristics were compared between myositis patients with and without anti-Ro52 autoantibodies.

Anti-Ro52 autoantibodies were found in 14% patients with juvenile DM, 12% with juvenile PM, and 18% with juvenile connective tissue disease–myositis overlap. anti-Ro52 autoantibodies were more frequent in patients with AS (64%, $p<0.001$) and anti-MDA5 (31%, $p<0.05$) autoantibodies. After controlling for the presence of myositis-specific autoantibodies, anti-Ro52 autoantibodies were associated with the presence of ILD (36% vs 4%, $p<0.001$). The disease course was more frequently chronic, remission was less common, and an increased number of medications was received in anti-Ro52 positive patients.

In conclusion, anti-Ro52 autoantibodies are present in 14% of patients with juvenile myositis and are strongly associated with anti-MDA5 and AS autoantibodies. In all patients with juvenile myositis, those with anti-Ro52 autoantibodies were more likely to have ILD. Furthermore, patients with anti-Ro52 autoantibodies have more severe disease and poorer prognosis.

Anti-Ro52 autoantibodies are associated with interstitial lung disease and more severe disease in patients with juvenile myositis

Sara Sabbagh[#], Iago Pinal-Fernandez[#], Takayuki Kishi, Ira N. Targoff, Frederick W. Miller, Lisa G. Rider[#], Andrew L. Mammen[#] Childhood Myositis Heterogeneity Collaborative Study Group*

[#] These authors contributed equally to this project.

Sara Sabbagh, D.O., Iago Pinal-Fernandez, M.D., Ph.D., and Andrew L. Mammen, M.D., Ph.D.: Muscle Disease Unit, Laboratory of Muscle Stem Cells and Gene Regulation, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health (NIH), Bethesda, MD. Takayuki Kishi M.D., Ph.D., Frederick W. Miller M.D., Ph.D. and Lisa G. Rider M.D.: Environmental Autoimmunity Group, National Institute of Environmental Health Sciences, NIH, Bethesda, MD. Ira N. Targoff, M.D.: VA Medical Center and Oklahoma Medical Research Foundation, Oklahoma City, OK. Iago Pinal-Fernandez, M.D., Ph.D. and Andrew L. Mammen, M.D., Ph.D.: Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD. Andrew L. Mammen, M.D., Ph.D.: Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD. Iago Pinal-Fernandez: Faculty of Health Sciences, Universitat Oberta de Catalunya, Barcelona, Spain.

Running Title: Ro52 autoantibodies in juvenile myositis.

Key words: myositis, juvenile idiopathic inflammatory myopathies, anti-Ro52 autoantibodies, myositis associated autoantibodies, interstitial lung disease.

Address correspondence to: Andrew L. Mammen, M.D., Ph.D., Muscle Disease Unit, Laboratory of Muscle Stem Cells and Gene Expression, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, 50 South Drive, Room 1141, Building 50, MSC 8024, Bethesda, MD 20892. E-mail: andrew.mammen@nih.gov. Phone: 301-451-1199. Fax: 301-594-0305.

Manuscript word count: 3162/3000 words

Abstract word count: 229/250 words

*Members of the Childhood Myositis Heterogeneity Collaborative Study Group who contributed to this project:

Bita Arabshahi, Lilliana Barillas-Arias, Mara Becker, April Bingham, Ruy Carrasco, Victoria Cartwright, Rodolfo Curiel, Marietta M. DeGuzman, Barbara Anne Eberhard, Barbara S. Edelheit, Terri Finkel, Stephen W. George, Ellen A. Goldmuntz, William Hannan, Michael Henrickson, Adam M. Huber, Anna Jansen, James Jarvis, Lawrence Jung, Ildy M. Katona, Steven J. Klein, W Patrick Knibbe, Bianca A. Lang, Carol B. Lindsley, Gulnara Mamyrova, Linda Myers, Stephen R. Mitchell, Kabita Nanda, Terrance P. O'Hanlon, Murray H. Passo, Maria D. Perez, Donald A. Person, Linda I. Ray, Rafael F. Rivas-Chacon, Tova Ronis, Deborah Rothman, Adam Schiffenbauer, Bracha Shaham,

David Sherry, Abigail Smukler, Matthew L. Stoll, Sangeeta H. Sule, Scott A. Vogelgesang, Rita Volochayev, Jennifer C. Wargula, Pamela Weiss.

Funding: This research was supported by the Intramural Research Programs of the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS; ZIA AR041203) and the National Institute of Environmental Health Sciences (NIEHS) of the National Institutes of Health (NIH).

KEY MESSAGES

What is already known about this subject?

The influence of anti-Ro52 autoantibodies in the clinical features and prognosis of juvenile myositis patients was mostly unkown.

What does this study add?

-Juvenile myositis patients with anti-Ro52 autoantibodies are more likely to develop interstitial lung disease.

-Anti-Ro52 autoantibodies are more common in juvenile myositis patients with anti-MDA5 and antisynthetase autoantibodies.

-Juvenile myositis patients with anti-Ro52 autoantibodies have more severe disease and a poorer prognosis.

How might this impact on clinical practice?

Anti-Ro52 autoantibodies are useful prognostic markers for ILD and severity of the disease in juvenile myositis patients.

ABSTRACT (229 words, 250 limit)

Objectives: Anti-Ro52 autoantibodies are associated with more severe interstitial lung disease (ILD) in adult myositis patients with anti-aminoacyl tRNA synthetase autoantibodies. However, few studies have examined anti-Ro52 autoantibodies in juvenile myositis. The purpose of this study was to define the prevalence and clinical features associated with anti-Ro52 autoantibodies in a large cohort of patients with juvenile myositis.

Methods: We screened sera from 302 patients with juvenile dermatomyositis (JDM), 25 patients with juvenile polymyositis (JPM), and 44 patients with juvenile connective tissue disease-myositis overlap (JCTM) for anti-Ro52 autoantibodies by ELISA. Clinical characteristics were compared between myositis patients with and without anti-Ro52 autoantibodies.

Results: Anti-Ro52 autoantibodies were found in 14% of JDM, 12% of JPM, and 18% of JCTM patients. Anti-Ro52 autoantibodies were more frequent in patients with anti-aminoacyl tRNA synthetase (64%, $p<0.001$) and anti-MDA5 (31%, $p<0.05$) autoantibodies. After controlling for the presence of myositis-specific autoantibodies, anti-Ro52 autoantibodies were associated with the presence of ILD (36% vs 4%, $p<0.001$). Disease course was more frequently chronic, remission was less common, and an increased number of medications was received in anti-Ro52 positive patients.

Conclusions: Anti-Ro52 autoantibodies are present in 14% of juvenile myositis patients and are strongly associated with anti-MDA5 and anti-aminoacyl tRNA synthetase autoantibodies. In all juvenile myositis patients, those with anti-Ro52 autoantibodies were more likely to have ILD. Furthermore, patients with anti-Ro52 autoantibodies have more severe disease and a poorer prognosis.

INTRODUCTION

Idiopathic inflammatory myopathies (IIM) are a heterogeneous group of systemic autoimmune diseases characterized by weakness, chronic inflammation of skeletal muscles, and elevated serum muscle enzyme levels.[1] Many patients also have extramuscular manifestations, including involvement of the skin, lungs, and/or joints. Most IIM patients have a myositis-specific autoantibody (MSA), defined as an autoantibody found only in IIM patients, which are typically mutually exclusive.[2] In contrast, myositis-associated autoantibodies (MAAs) are found in IIM, but may also be present in patients with other autoimmune diseases and may be seen in association with an MSA or other MAAs.

MSAs are associated with specific phenotypes and can be used to define unique subgroups.[2, 3] For instance, anti-melanoma differentiation-associated gene 5 (MDA5) autoantibodies are associated with cutaneous ulceration and palmar papules, minimal muscle involvement, arthritis, interstitial lung disease (ILD) which may be rapidly progressive, and a high fatality rate.[4-7] In contrast, patients with autoantibodies recognizing a tRNA-synthetase, such as histidyl-tRNA synthetase (i.e., Jo1), have anti-synthetase syndrome, a unique multisystem autoimmune disease characterized by a combination of myositis, ILD, arthritis, Raynaud's phenomenon, fever, and/or mechanic's hands.[8] Of note, while many phenotypic features are similar between juvenile and adult IIM with the same MSAs, there are some important differences. For example, adults with anti-p155/140 (TIF-1) autoantibodies have an increased risk of malignancy, whereas anti-p155/140 (TIF-1) autoantibody positive children do not.[2, 9]

The phenotypes associated with MAAs are less well-described than those associated with MSAs. In adult IIM patients, the most common MAA is anti-Ro52.[10] Interestingly, anti-Ro52 autoantibodies often co-occur with anti-Jo1 autoantibodies[11] and adult patients with both autoantibodies have more severe ILD and more frequently develop lung fibrosis than those with anti-

Jo1 autoantibodies alone.[12, 13] In addition, higher anti-Ro52 autoantibody titers are associated with the development of more severe ILD[14], myositis, and joint impairment in anti-Jo1-positive adult patients.[15] Patients with both anti-Jo1 and anti-Ro52 autoantibodies have a poorer response to various immunosuppressive drugs and a decrease in survival.[13, 15]

To date, only one study has examined the prevalence of anti-Ro52 autoantibodies in children with IIM. A recent analysis of 22 children with myositis revealed that 23% had anti-Ro52 autoantibodies, although specific clinical associations were not examined.[16] As anti-Ro52 autoantibodies have not previously been well described in children with myositis, the purpose of this study was to define the prevalence of and clinical features associated with anti-Ro52 autoantibodies in a large cohort of patients with juvenile myositis.

PATIENTS AND METHODS

Patients and serum samples

Of the 543 patients from the Childhood Myositis Heterogeneity Collaborative Study who were enrolled between 1989 and 2016 with probable or definite myositis by Bohan and Peter criteria[17] those 371 (68%) with a serum sample available for anti-Ro52 autoantibody testing were included in the study. Patients with juvenile myositis (n=371) included 302 (81.4%) with juvenile dermatomyositis (JDM), 25 (6.7%) with juvenile polymyositis (JPM) and 44 (11.9%) with juvenile connective tissue disease–myositis (JCTM) overlap. The JCTM subgroup included patients meeting criteria for myositis and another autoimmune disease, including 13 with juvenile systemic lupus erythematosus, 11 with juvenile systemic sclerosis, 7 patients with juvenile idiopathic arthritis, and 13 with other autoimmune conditions including autoimmune hepatitis, eosinophilic fasciitis, diabetes mellitus, lichen sclerosis, linear morphea, psoriasis, Sjögren's syndrome, and ulcerative colitis. Sera was drawn for each patient at time of study enrollment. Sera from 90 healthy control children enrolled in the same studies were available.

All subjects were enrolled in institutional review board-approved natural history studies as previously described,[18] and all provided informed consent. A standardized physician questionnaire captured demographics, clinical and laboratory features, environmental exposures at illness onset or diagnosis, as well as therapeutic usage and responses.[18] Seven organ system symptom scores at diagnosis, defined as the number of symptoms present divided by the number of symptoms assessed, and an overall clinical symptom score as the average of the seven individual organ symptom scores, were calculated as previously described.[19-21] In 7 of 33 patients, the presence of ILD was diagnosed by high resolution computed tomography (HRCT) and lung biopsy. In 11 of 33 patients, ILD was diagnosed by HRCT alone and in 5/33 patients, ILD was diagnosed by biopsy alone. In those patients

that did not undergo HRCT or lung biopsy, 7 patients were diagnosed with ILD by chest radiographic imaging combined with pulmonary function testing. Three patients did not have imaging records available and the diagnosis of ILD was based on physician documentation in the medical record. Complete clinical response and remission were defined as at least 6 months of inactive disease on or off therapy, respectively.[20] A course of treatment was defined as a single episode from beginning of administration of a given medication to the termination of treatment with that medication, or combination of medications, in each patient. Medical record review, conducted in >75% of patients, verified the clinical, demographic, laboratory and therapeutic data contained in the physician questionnaires. Follow up visits occurred in 55% of patients, with an average time from enrollment date to final evaluation of 4.3 years. Patient characteristics in our cohort are comparable with other registry based JDM cohorts with regard to the female predominance, age at diagnosis, symptom duration from disease onset to diagnosis, and disease manifestations.[22-25]

Myositis autoantibody assays

Anti-Ro52 autoantibody detection

An enhanced performance Ro52 enzyme-linked immunosorbent assay (ELISA) [SS-A 52 ELISA, Quanta Lite, INOVA Diagnostics, San Diego, CA] was performed according to the manufacturer's instructions. The reactivity for each sample was calculated by dividing the optical density (OD) of the sample by the OD of the low positive control and multiplying by the number of units (25 units) assigned to that control.

Myositis specific autoantibody detection

Other myositis autoantibodies were tested by validated methods, including protein and RNA immunoprecipitation (IP) using radiolabeled HeLa or K562 cell extracts and double immunodiffusion.[18] For anti-p155/140 (TIF-1), anti-MJ (Nuclear Matrix Protein 2, NXP2) and anti-MDA5 autoantibodies, serum samples were screened by IP with confirmatory testing by IP immunoblotting.[18] Anti-3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGCR) autoantibodies were screened by ELISA and confirmed by immunoprecipitation using a ^{35}S-methionine-labelled HMGCR protein produced by *in vitro* transcription and translation as previously described.[26]

Analysis

Dichotomous variables were expressed as percentages and absolute frequencies, and continuous features were reported as means and SD. Pairwise comparisons for categorical variables between groups were made using χ^2 test or Fisher's exact test, as appropriate, while continuous variables were compared using Student's t-test. Logistic and linear regression were used to adjust the comparisons for possible confounding variables, including the year of diagnosis, length of follow-up and MSAs. Creatine kinase, a highly positively skewed variable, was expressed as median, first and third quartile for descriptive purposes and transformed through a base-10 logarithm for analysis. All statistical analyses were performed using Stata/MP V.14.1 (StataCorp LLC, College Station, Texas). As this was an exploratory study, a two-sided P value of ≤ 0.05 was considered statistically significant.

RESULTS

Anti-Ro52 autoantibodies were more prevalent in patients with juvenile IIM (JIIM) than in healthy control children (14% vs 1%, p<0.001). Sera from 14% of patients with JDM, 12% with JPM, and 18% with JCTM had anti-Ro52 autoantibodies (Figure 1, Table 1). There were no significant differences in gender, race, age at diagnosis, or delay to diagnosis between juvenile myositis patients with and without anti-Ro52 autoantibodies (Table 2).

Figure 1. Swarm plot of anti-Ro52 autoantibody ELISA results for juvenile healthy controls and JIIM patients divided into JDM, JPM, and JCTM. The dashed line of 20 units indicates the cut-off value for anti-Ro52 autoantibody positivity. Out of 371 JIIM patients, 53 (14%) were positive for anti-Ro52 autoantibodies by ELISA. Of these patients, 42 had JDM, 3 had JPM, and 8 had JCTM. Out of 90 juvenile healthy controls, one patient (1.1%) was positive for anti-Ro52 autoantibodies by ELISA. JCTM, juvenile connective tissue myositis; JDM, juvenile dermatomyositis; JIIM, juvenile idiopathic inflammatory myopathy; JPM, juvenile polymyositis.

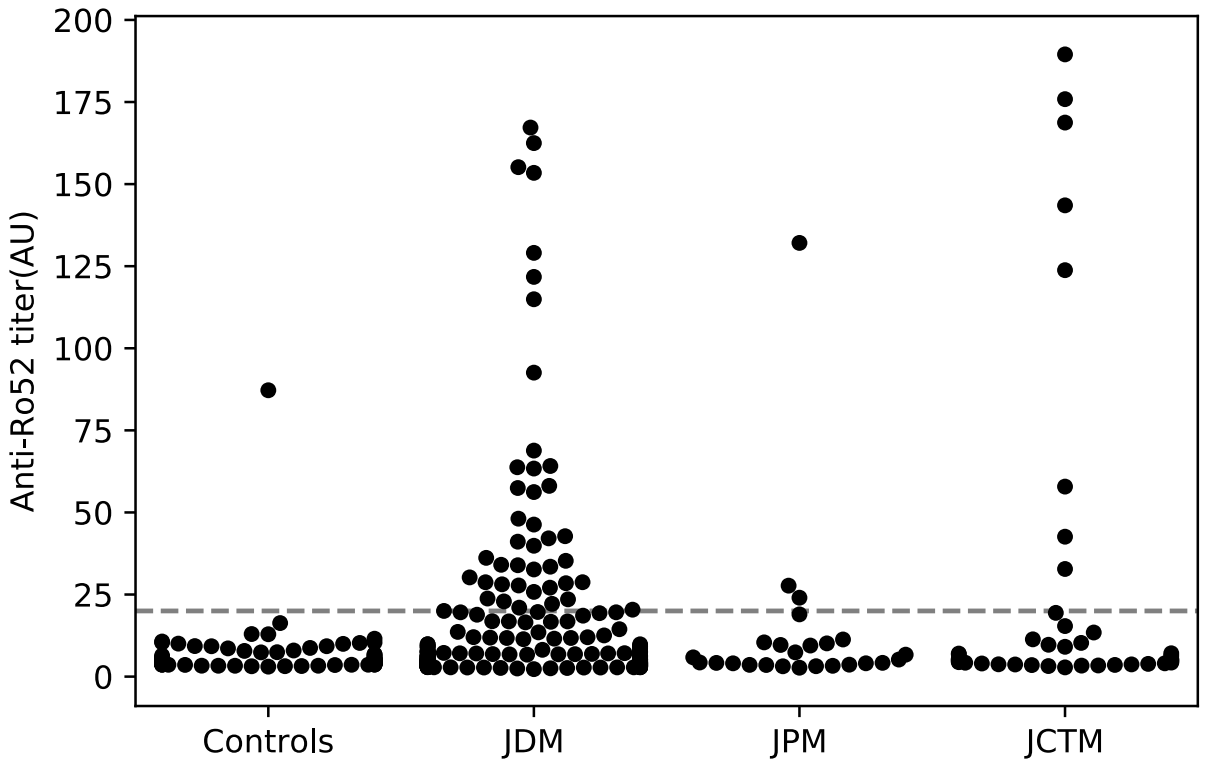


Table 1. Prevalence of anti-Ro52 autoantibodies among patients with juvenile myositis.

Clinical subgroup	Anti-Ro52 autoantibody positive % (n/N)
Juvenile healthy controls (N=90)	1% (n=1)
Juvenile myositis (N=371)	14% (n=53) ***
Juvenile dermatomyositis (N=302)	14% (n=42) ***
Juvenile polymyositis (N=25)	12% (n=3) *
Juvenile connective tissue-disease myositis (N=44):	18% (n=8) ***
Juvenile lupus erythematosus (N=13)	23% (n=3) **
Juvenile systemic sclerosis (N=11)	0% (n=0)
Juvenile idiopathic arthritis (N=7)	29% (n=2) *
Other autoimmune diseases[a] (N=13)	23% (n=3) **
Myositis specific autoantibody subgroup	
Anti-p155/140 (TIF-1) (N=119)	11% (n=13)
Anti-NXP2 (N=77)	14% (n=11)
Anti-MDA5 (N=32)	31% (n=10) *
Anti-aminoacyl tRNA synthetase (N=14)	64% (n=9) ***
Anti-SRP (N=7)	0% (n=0)
Anti-Mi2 (N=13)	15% (n=2)
Anti-HMGCR (N=4)	50% (n=2)
MSA negative (N=96)	5% (n=5) **

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

Chi-squared or Fisher´s exact tests were used to compare the percentage of positive patients compared with the percentage of negative patients within each myositis clinical and autoantibody subgroup.

Abbreviations: TIF-1: transcription intermediary factor 1, NXP2: nuclear matrix protein-2, MDA5: melanoma differentiation associated protein-5, SRP: signal recognition particle, HMGCR: 3-Hydroxy-3-Methylglutaryl-CoA Reductase, MSA: myositis specific autoantibody

[a] autoimmune hepatitis, eosinophilic fasciitis, fasciitis, juvenile diabetes mellitus, lichen sclerosis, linear morphea, psoriasis, Sjögren's syndrome, ulcerative colitis.

Table 2. General features of juvenile myositis patients with and without anti-Ro52 autoantibodies.

	Total (N=371) % (n/N) or Mean (SD)	Anti-Ro52 autoantibody positive (N=53) % (n/N) or Mean (SD)	Anti-Ro52 autoantibody negative (N=318) % (n/N) or Mean (SD)	p-value
Age at diagnosis	9.0 (4.4)	9.5 (4.7)	8.9 (4.3)	0.3
Age at enrollment	12.5 (7.1)	12.6 (7.7)	12.5 (7.0)	1.0
Delay to diagnosis (years)	0.7 (1.2)	0.55 (0.56)	0.75 (1.27)	0.3
Follow-up (years)	5.8 (6.4)	4.3 (6.4)	6.0 (6.4)	0.09
Female	71% (263/371)	74% (39/53)	70% (224/318)	0.6
Race				
White	65% (240/371)	57% (30/53)	66% (210/318)	0.2
Black	16% (59/371)	21% (11/53)	15% (48/318)	0.3
Hispanic	6% (24/371)	6% (3/53)	7% (21/318)	1.0
Other races [a]	13% (48/371)	17% (9/53)	12% (39/318)	0.3
Myositis-specific autoantibodies				
Anti-p155/140 (TIF-1)	33% (119/359)	26% (13/50)[b]	34% (106/309) [c]	0.2
Anti-NXP2	21% (77/366)	21% (11/52) [b]	21% (66/314) [c]	1.0
Anti-MDA5	9% (32/368)	19% (10/53)	7% (22/315) [c]	0.01
Anti-aminoacyl tRNA synthetase	4% (14/360)	18% (9/49) [b]	2% (5/311) [c]	< 0.001
Anti-SRP	2% (7/360)	0% (0/49) [b]	2% (7/311) [c]	0.6
Anti-Mi2	4% (13/354)	4% (2/49) [b]	4% (11/305) [c]	0.7
Anti-HMGCR	1% (4/371)	4% (2/53)	1% (2/318)	0.10
MSA negative	27% (96/362)	9% (5/53)	29% (91/309) [c]	0.002

Dichotomous variables were represented as percentage (count/total) and continuous variables as mean (SD). Chi-squared or Fisher´s exact tests were used to compare dichotomous variables, as appropriate, while continuous variables were compared using Student´s t-test.

Abbreviations: TIF-1: transcription intermediary factor 1, NXP2: nuclear matrix protein-2, MDA5: melanoma differentiation associated protein-5, SRP: signal recognition particle, HMGCR: 3-Hydroxy-3-Methylglutaryl-CoA Reductase, MSA: myositis specific autoantibody.

[a] Asian (Korean, Japanese, Chinese, Indian, Filipino), Pacific Islands, American Indian.
[b] N ≠ 53 due to missing data.
[c] N ≠ 318 due to missing data.

Prevalence of anti-Ro52 autoantibodies among myositis-specific autoantibody subgroups

Of those patients positive for anti-Ro52 autoantibodies, 26% had co-existing anti-p155/140 (TIF-1) autoantibodies, 21% had anti-NXP-2 autoantibodies, 19% had anti-MDA5 autoantibodies, 18% had anti-aminoacyl tRNA synthetase autoantibodies, 4% had anti-Mi2 autoantibodies, 4% had anti-HMGCR autoantibodies, and 9% were MSA negative (Table 2). Anti Ro-52 autoantibodies were significantly increased in the anti-MDA5 and anti-aminoacyl tRNA synthetase autoantibody subgroups than in other MSA subgroups (Table 1). For instance, anti-Ro52 autoantibodies co-existed in 31% of juvenile IIM sera with anti-MDA5 autoantibodies ($p<0.05$) and 64% of those with anti-aminoacyl tRNA synthetase autoantibodies ($p<0.001$) (Table 1). Similarly, anti-MDA5 autoantibodies co-existed in 19% of anti-Ro52 autoantibody positive sera and 7% of anti-Ro52 autoantibody negative sera ($p=0.01$). Anti-aminoacyl tRNA synthetase autoantibodies co-existed in 18% of anti-Ro52 autoantibody positive sera and 2% of anti-Ro52 autoantibody negative sera ($p<0.001$) (Table 2). Less than 15% of those with anti-p155/140 (TIF1), anti-nuclear matrix protein-2 (NXP2), anti-signal recognition particle (SRP), or anti-Mi2 autoantibodies, and only 5% of those without an MSA were anti-Ro52 positive (Table 1).

Pulmonary manifestations among patients with anti-Ro52 autoantibodies

After controlling for the presence of MSAs (including anti-aminoacyl tRNA synthetase and anti-MDA5 autoantibodies) a multivariate analysis showed anti-Ro52 autoantibodies were highly associated with pulmonary involvement. Overall, patients with anti-Ro52 autoantibodies more often had ILD (36% vs 4%, $p<0.001$), dyspnea on exertion (59% vs 25% $p<0.001$), and a higher early pulmonary score (mean 0.18 vs 0.08, $p=0.002$) than those without these autoantibodies (Table 3). Within the anti-MDA5 autoantibody positive subgroup, Ro52 reactivity was even more strongly associated with ILD: 70% of those with co-existing anti-Ro52 autoantibodies had ILD compared to only 9% of those who were anti-Ro52 negative ($p=0.001$) (Table 4). Similarly, among the anti-

aminoacyl tRNA synthetase autoantibody subgroup, 100% of anti-Ro52 autoantibody positive and 40% of anti-Ro52 negative patients had ILD (p=0.03) (Table 4). Other pulmonary manifestations were also associated with Ro52 reactivity within the anti-MDA5 and anti-aminoacyl tRNA synthetase autoantibody subgroups. Specifically, among those patients with anti-MDA5 autoantibodies, patients who also were positive for anti-Ro52 autoantibodies more often had dyspnea on exertion (90% vs 27%, p=0.002) and higher early pulmonary scores (mean 0.29 vs 0.02, p<0.001) than those who were anti-Ro52 autoantibody negative. In patients with anti-aminoacyl tRNA synthetase autoantibodies, anti-Ro52 autoantibody positive patients had increased frequency of dyspnea on exertion (89% vs 40%), although this did not reach statistical significance. Patients with co-existing anti-p155/140 (TIF-1) and anti-Ro52 autoantibodies also had an increased frequency of ILD (15% vs 1%, p=0.03) and dyspnea on exertion (50% vs 16%, p=0.01) compared to anti-p155/140 (TIF-1) autoantibody positive patients who were anti-Ro52 autoantibody negative (Table 4). Of note, in the MSA negative subgroup, none of 5 Ro52-positive patients had ILD (Table 4). The association of anti-Ro52 autoantibodies with ILD was significant within the JDM clinical subgroup: 33% of JDM patients with anti-Ro52 autoantibodies had ILD compared to 1% of anti-Ro52 negative JDM patients (p<0.001) (Table 4).

Table 3. Clinical features of juvenile myositis patients with and without anti-Ro52 autoantibodies.

Signs/symptoms ever present	Total (N=371) % (n/N) or Mean (SD)	Anti-Ro52 autoantibody positive (N=53) % (n/N) or Mean (SD)	Anti-Ro52 autoantibody negative (N=318) % (n/N) or Mean (SD)	univariate p-value	multivariate p-value
Muscle involvement					
Proximal weakness	99% (369/371)	98% (52/53)	100% (317/318)	0.3	0.3
Myalgia	64% (234/363)	62% (32/52) [a]	65% (202/311) [b]	0.6	0.1
Distal weakness	47% (170/363)	46% (24/52) [a]	47% (146/311) [b]	0.9	0.9
Muscle atrophy	37% (136/367)	44% (23/52) [a]	36% (113/315) [b]	0.2	0.3
Falling episodes	45% (164/367)	44% (23/52) [a]	45% (141/315) [b]	0.9	1.0
Lung involvement					
Dyspnea on exertion	30% (109/366)	59% (30/51) [a]	25% (79/315) [b]	< 0.001	< 0.001
Interstitial lung disease	9% (33/369)	36% (19/53)	4% (14/316) [b]	< 0.001	< 0.001
Dysphonia	32% (118/367)	32% (17/53)	32% (101/314) [b]	1.0	0.7
Joint involvement					
Arthralgia	64% (236/369)	70% (37/53)	63% (199/316) [b]	0.3	0.4
Joint contractures	61% (224/370)	63% (33/52) [a]	60% (191/318)	0.6	0.7
Arthritis	51% (189/370)	60% (31/52) [a]	50% (158/318)	0.2	0.7
Skin involvement					
Heliotrope	79% (293/369)	87% (46/53)	78% (247/316) [b]	0.2	0.2
Gottron´s papules	82% (305/370)	77% (41/53)	83% (264/317) [b]	0.3	0.3
Malar rash	70% (259/371)	68% (36/53)	70% (223/318)	0.7	0.6
Photosensitivity	48% (172/362)	49% (25/51) [a]	47% (147/311) [b]	0.8	0.9
V or Shawl sign rash	31% (113/369)	42% (22/53)	29% (91/316) [b]	0.06	0.07
Linear extensor erythema	36% (130/363)	31% (16/52) [a]	37% (114/311) [b]	0.4	0.3
Calcinosis	29% (109/371)	28% (15/53)	30% (94/318)	0.9	0.1
Raynaud´s phenomenon	15% (55/369)	23% (12/53)	14% (43/316) [b]	0.09	0.04
Mechanic´s hands	7% (27/366)	9% (5/53)	7% (22/313) [b]	0.6	0.5
Gastrointestinal involvement					
Dysphagia	41% (151/370)	38% (20/53)	41% (131/317) [b]	0.6	1.0
Regurgitation	21% (77/370)	26% (14/53)	20% (63/317) [b]	0.3	0.5
Systemic involvement					
Weight loss	42% (155/369)	52% (27/52) [a]	40% (128/317) [b]	0.1	0.8
Fever	31% (112/358)	41% (21/51) [a]	30% (91/307) [b]	0.10	0.8
Muscle Enzymes					
Peak creatine kinase, IU/L	781 (252-5142)	1121 (225-3971)	750 (256-5249)	0.7	0.9
Peak aldolase, IU/L	20.0 (34.5)	18.0 (22.5)	20.3 (36.1)	0.6	0.3
Severity at onset	2.2 (1.1)	2.2 (1.7)	2.2 (0.9)	0.9	0.4
Early total symptom score	0.2 (0.1)	0.27 (0.14)	0.23 (0.11)	0.03	0.8
Early muscle score	0.4 (0.2)	0.37 (0.18)	0.38 (0.20)	0.7	0.5
Early joint score	0.5 (0.4)	0.48 (0.38)	0.45 (0.43)	0.6	0.1
Early cutaneous score	0.3 (0.1)	0.26 (0.15)	0.25 (0.14)	0.6	0.4
Early gastrointestinal score	0.1 (0.1)	0.08 (0.11)	0.07 (0.11)	0.6	1.0
Early pulmonary score	0.1 (0.2)	0.18 (0.23)	0.08 (0.14)	< 0.001	0.002
Early cardiac score	0.0 (0.1)	0.05 (0.12)	0.02 (0.07)	0.04	0.05
Early constitutional symptoms	0.4 (0.3)	0.48 (0.34)	0.38 (0.26)	0.02	1.0

Dichotomous variables were represented as percentage (count/total), continuous variables as mean (SD) and the creatine kinase was presented as median (Q1-Q3). For the univariate analysis, dichotomous variables were compared using chi-squared or Fisher´s exact tests, as appropriate while continuous variables were compared using Student´s t-test. Multivariate analysis used linear or logistic regression adjusted for length of follow-up, year of onset and autoantibodies. Creatine kinase was log-transformed prior to statistical analysis.

[a] $N \neq 53$ due to missing data.
[b] $N \neq 318$ due to missing data.

Table 4: Pulmonary features of juvenile myositis patients with and without anti-Ro52 autoantibodies within juvenile myositis clinical and autoantibody subgroups

	Anti-Ro52 autoantibody positive % (n/N) or Mean (SD)	Anti-Ro52 autoantibody negative % (n/N) or Mean (SD)	p-value
JDM subgroup (N=302)			
Interstitial lung disease	33% (14/42)	1% (3/258)[a]	< 0.001
Dyspnea on exertion	62% (26/42)	19% (50/258)[a]	< 0.001
Early pulmonary score	0.20 (0.22)	0.07 (0.13)	< 0.001
JPM subgroup (N=25)			
Interstitial lung disease	33% (1/3)	18% (4/22)	0.5
Dyspnea on exertion	50% (1/2)[a]	67% (14/21)[a]	1.0
Early pulmonary score	0.17 (0.29)	0.19 (0.20)	0.8
JCTM subgroup (N=44)			
Interstitial lung disease	50% (4/8)	19% (7/36)	0.09
Dyspnea on exertion	43% (3/7)[b]	42% (15/36)	1.0
Early pulmonary score	0.12 (0.25)	0.09 (0.16)	0.7
Anti-MDA5 autoantibody subgroup (N=32)			
Interstitial lung disease	70% (7/10)	9% (2/22)	0.001
Dyspnea on exertion	90% (9/10)	27% (6/22)	0.002
Early pulmonary score	0.29 (0.19)	0.02 (0.06)	< 0.001
Anti-aminoacyl tRNA synthetase autoantibody subgroup (N=14)			
Interstitial lung disease	100% (9/9)	40% (2/5)	0.03
Dyspnea on exertion	89% (8/9)	40% (2/5)	0.09
Early pulmonary score	0.31 (0.31)	0.27 (0.30)	0.8
Anti-p155/140 (TIF-1) autoantibody subgroup (N=119)			
Interstitial lung disease	15% (2/13)	1% (1/106)	0.03
Dyspnea on exertion	50% (6/12)[a]	16% (17/105)[a]	0.01
Early pulmonary score	0.16 (0.24)	0.06 (0.12)	0.01
Anti-NXP2 autoantibody subgroup (N=76)			
Interstitial lung disease	9% (1/11)	0% (0/65)[b]	0.1
Dyspnea on exertion	45% (5/11)	27% (18/66)	0.3
Early pulmonary score	0.16 (0.19)	0.10 (0.14)	0.2
MSA negative subgroup (N=96)			
Interstitial lung disease	0% (0/5)	10% (9/90)	1.0
Dyspnea on exertion	25% (1/4)	33% (30/90)	1.0
Early pulmonary score	0.04 (0.09)	0.08 (0.15)	0.5

Dichotomous variables were represented as percentage (count/total), continuous variables as mean (SD). For the univariate analysis, dichotomous variables were compared using chi-squared or Fisher´s exact tests, as appropriate while continuous variables were compared using Student´s t-test.

Abbreviations: JDM: juvenile dermatomyositis, JPM: juvenile polymyositis, JCTM: juvenile connective tissue myositis; MDA5: melanoma differentiation associated protein-5, TIF-1: transcription intermediary factor 1, NXP2: nuclear matrix protein-2, SRP: signal recognition particle.

[a] Data missing for two patients within juvenile myositis clinical or autoantibody subgroup.
[b] Data missing for one patient within juvenile myositis clinical or autoantibody subgroup.

Other clinical manifestations among patients with anti-Ro52 autoantibodies

Independent of MSA status, anti-Ro52 autoantibodies were also associated with Raynaud's phenomenon (23% vs 14%, p=0.04) (Table 3). Furthermore, within the anti-NXP2 subgroup, Ro52 reactivity was associated with more cutaneous involvement: patients with both anti-NXP2 and anti-Ro52 autoantibodies had a higher prevalence of V- or Shawl-sign rashes (55% vs 17%, p=0.01) and linear extensor erythema (64% vs 20%, p=0.02) than anti-NXP2 autoantibody positive patients without anti-Ro52 autoantibodies. Those with both anti-NXP2 and anti-Ro52 autoantibodies also had more frequent gastroesophageal regurgitation (55% vs 17%, p=0.04). Within the anti-MDA5 subgroup, however, anti-Ro52 autoantibodies were associated with less frequent linear extensor erythema (11% vs 50%, p=0.04). Patients with anti-Ro52 autoantibodies also had a higher mean early cardiac score, defined by the presence of cardiac symptoms at diagnosis divided by the number of symptoms assessed (p=0.05). There were no other significant differences in the prevalence of the muscle, lung, joint, cutaneous, gastrointestinal, or constitutional manifestations between patients with and without anti-Ro52 autoantibodies in univariate or multivariate analysis, or in examining these features in anti-Ro52 autoantibody positive patients in the presence of another MSA.

Disease severity among patients with anti-Ro52 autoantibodies

Several other differences in outcomes and medications received between patients positive and negative for anti-Ro52 autoantibodies suggested that anti-Ro52 autoantibodies are associated with more severe disease (Table 5). The disease course in patients with anti-Ro52 autoantibodies was more often chronic continuous (78% vs 52%, p=0.05) and less often monocyclic (3% vs 25%, p=0.02). Anti-Ro52 positive patients were more often American College of Rheumatology (ACR) functional class 4 (11% vs 4%, p=0.008) at the last clinical evaluation and had a higher mean ACR functional class score at that assessment (1.7 vs 1.4, p=0.007). Anti-Ro52 autoantibodies were also associated with an increased total number of medications received (mean 4.8 vs 3.8, p=0.05). Anti-Ro52 autoantibody

positive patients more often received intravenous pulse steroids (79% vs 52%, p=0.03). Anti-Ro52 autoantibody positive patients less often achieved clinical remission (5% vs 27% p=0.05). Lastly, on univariate analysis, but not multivariable analysis, patients with anti-Ro52 autoantibodies less often experienced a complete clinical response (17% vs 32%, p=0.04) and had more medication treatment trials per year (mean 3.5 vs 2.2, p=0.004).

Those with both anti-NXP2 and anti-Ro52 autoantibodies also more often had a severe (class IV) ACR functional class (27% vs 3%, p=0.03) and more frequent wheelchair use (60% vs 20%, p=0.03) as compared to patients positive for anti-NXP2 who were anti-Ro52 autoantibody negative. There was no other association of co-existing MSAs and anti-Ro52 autoantibodies on clinical outcomes or medications received.

Anti-Ro52 autoantibody titers

Anti-Ro52 autoantibody titers did not significantly differ between JDM, JPM, and JCTM groups. Overall, we found that higher anti-Ro52 titers are associated with shorter follow-up time, more treatment trials per year, higher early total symptom score, more total number of medications used, higher total functional class, higher severity at onset, higher early pulmonary score, higher early constitutional symptoms score, and higher total functional class in patients with juvenile IIM (all $p<0.05$; data not shown). However, as the Spearman correlation coefficients were ≤ 0.2 for each association, the clinical significance of high autoantibody titers is modest.

Table 5. Disease outcomes and medications used in juvenile myositis patients with and without anti-Ro52 autoantibodies

	Total (N=371) % (n/N) or Mean (SD)	Anti-Ro52 autoantibody positive (N=53) % (n/N) or Mean (SD)	Anti-Ro52 autoantibody negative (N=318) % (n/N) or Mean (SD)	Univariate p-value	Multivariate p-value
Disease					
Monocyclic course	22% (65/297)	3% (1/37) [b]	25% (64/260) [c]	0.003	0.02
Polycyclic course	23% (68/297)	19% (7/37) [b]	23% (61/260) [c]	0.5	0.9
Chronic continuous course	55% (164/297)	78% (29/37) [b]	52% (135/260) [c]	0.002	0.05
Steinbrocker functional class at final assessment					
Mean functional class	1.4 (0.8)	1.7 (1.0)	1.4 (0.8)	0.007	0.007
Functional class 1	70% (257/367)	53% (28/53)	73% (229/314) [c]	0.003	0.09
Functional class 2	21% (77/367)	34% (18/53)	19% (59/314) [c]	0.01	0.3
Functional class 3	4% (13/367)	2% (1/53)	4% (12/314) [c]	0.7	0.2
Functional class 4	5% (20/367)	11% (6/53)	4% (14/314) [c]	0.05	0.008
Mortality	4% (13/371)	6% (3/53)	3% (10/318)	0.4	0.4
Hospitalized	58% (206/355)	66% (35/53)	57% (171/302) [c]	0.2	0.4
Mean number of hospitalizations	1.3 (1.9)	1.3 (1.4)	1.3 (2.0)	0.9	0.8
Wheelchair use	19% (68/360)	24% (12/50) [b]	18% (56/310) [c]	0.3	0.2
Response to treatment					
Complete clinical response	30% (91/304)	17% (7/42)	32% (84/262) [c]	0.04	0.4
Remission	24% (74/312)	5% (2/43)	27% (72/269) [c]	0.002	0.05
Total number of medications used	3.9 (2.1)	4.8 (2.5)	3.8 (2.0)	0.003	0.05
Treatment trials per year	2.3 (2.8)	3.5 (3.0)	2.2 (2.7)	0.004	0.1
Medications received					
Oral steroids	99% (309/312)	100% (43/43) [b]	99% (266/269) [c]	1.0	.
Intravenous pulsed steroids	56% (174/312)	79% (34/43) [b]	52% (140/269) [c]	< 0.001	0.03
Methotrexate	74% (230/312)	86% (37/43) [b]	72% (193/269) [c]	0.05	0.4
Intravenous immunoglobulin	36% (112/312)	49% (21/43) [b]	34% (91/269) [c]	0.06	0.08
Other DMARDs	23% (73/312)	35% (15/43) [b]	22% (58/269) [c]	0.06	0.3

Dichotomous variables were represented as percentage (count/total), continuous variables as mean (SD). For the univariate analysis, dichotomous variables were compared using chi-squared or Fisher´s exact tests, as appropriate while continuous variables were compared using Student´s t-test. Multivariate analysis used linear or logistic regression adjusted for length of follow-up, year of onset and autoantibodies.

Abbreviations: ACR: American College of Rheumatology, DMARDs: disease modifying anti-rheumatic agents

[a] Azathioprine, Chlorambucil, Chloroquine, Colchicine, Cyclophosphamide, Cyclosporine, Dapsone, Hydroxychloroquine, Intravenous Immunoglobulin, Lefluonmide, Methotrexate, Mycophenolate mofetil, Sodium thiosulfate, Quinacrine
[b] $N \neq 53$ due to missing data
[c] $N \neq 318$ due to missing data

DISCUSSION

Here, we utilized a large cohort of juvenile myositis patients to study the prevalence and clinical significance of anti-Ro52 autoantibodies in children with IIM. We found anti-Ro52 autoantibodies to be strongly associated with ILD and other pulmonary manifestations in juvenile myositis patients. We also found that children with anti-Ro52 autoantibodies have more severe disease, underwent more intense treatment regimens, and had lower rates of disease remission than those without anti-Ro52 autoantibodies. In children with myositis, anti-Ro52 autoantibodies were associated with anti-aminoacyl tRNA synthetase autoantibodies, as previously described in adults.[11] We also found that anti-Ro52 autoantibodies were associated with anti-MDA5 autoantibodies in pediatric myositis patients, which has not been reported previously.

Importantly, our analyses indicate that the presence of anti-Ro52 autoantibodies is strongly associated with ILD, even after adjusting for the presence of MSAs such as anti-MDA5 and anti-aminoacyl tRNA synthetase autoantibodies. Indeed, the association of Ro52 reactivity with ILD is not limited to the anti-MDA5 and anti-aminoacyl tRNA synthetase autoantibody subgroups, but extends to other MSA subgroups that are not classically associated with ILD, such as children with anti-p155/140 (TIF-1) autoantibodies. Interestingly, none of the 5 Ro52-positive MSA-negative patients had ILD; however, the small number of patients limits our ability to draw definite conclusions about whether anti-Ro52 autoantibodies are associated with ILD in this subgroup. Current practice encourages screening juvenile myositis patients for MSAs such as anti-MDA5 and anti-aminoacyl tRNA synthetase autoantibodies, as these autoantibodies confer risk for developing ILD and their presence is a determinant of clinical management and patient prognosis. In light of the current findings demonstrating that anti-Ro52 autoantibodies are an independent predictor of ILD, screening juvenile myositis patients for these autoantibodies may also be prudent.

In adult patients with IIM, anti-Ro52 autoantibodies have been associated with poorer response to immunosuppressive drugs and decreased survival.[13, 15] Similarly, in our juvenile cohort, anti-Ro52 autoantibodies are associated with more severe disease and poorer outcomes. As the severity of other clinical manifestations, including muscle, joint, skin, gastrointestinal, and systemic features were not associated with Ro52 reactivity, it seems likely that disease severity seen in the anti-Ro52 positive patients is a consequence of pulmonary disease. Additional studies are required to clarify this point. Nonetheless, our findings highlight the potential utility of anti-Ro52 autoantibodies as a predictor of disease severity and poor prognosis in juvenile myositis, which underscores the potential utility of screening juvenile IIM patients for anti-Ro52 autoantibodies.

Of particular significance is the novel association of anti-Ro52 autoantibodies and anti-MDA5 autoantibodies in our JIIM cohort. In adult IIM patients, anti-Ro52 autoantibodies often co-occur with anti-Jo1 autoantibodies, and in adult anti-Jo1 positive patients, Ro52 reactivity is associated with more severe ILD. Until now neither the association of anti-Ro52 autoantibodies with anti-MDA5 autoantibodies, nor the association of anti-Ro52 autoantibodies with ILD across other MSA groups, has been observed in adult patients with myositis. However, a small case series described co-existing anti-Ro52 autoantibodies in 6 of 13 anti-MDA5 positive patients, 5 of whom had rapidly progressive ILD [Huang 2018]. Interestingly, only 1 of 33 patients in our JIIM cohort with ILD had rapidly progressive ILD and this patient was positive for both anti-MDA5 and anti-Ro52 autoantibodies.

Although we have now established an association between anti-aminoacyl tRNA synthetase and anti-Ro52 autoantibodies not only in adults, but also in children, it remains unclear why these autoantibodies co-occur. It has been proposed that local autoantibody production induced by type I IFN[27] could be a driving force behind the production of both anti-Jo1 and anti-Ro52 autoantibodies, given the increase in B-cell activating factor (BAFF) receptors in the sera of IIM patients with these autoantibodies.[28] In the current study of juvenile IIM, we now also demonstrate an association

between anti-MDA5 and anti-Ro52 autoantibodies. Interestingly, both MDA5 and Ro52 are cytosolic, interferon (IFN)-induced proteins; perhaps concurrent over-expression of these proteins in juvenile IIM patients leads to the development of autoimmunity against both. However, we do not have adequate type I IFN measurements to further examine this hypothesis.

This current study has several limitations. First, this cohort of patients with juvenile myositis had some data collected retrospectively, resulting in some missing data, and was collected over more than 20 years, with potential chronology bias. However, we adjusted the variables of this study for the year of diagnosis and tested the distribution of missing values across groups and did not find evidence of a significant bias. Second, although imaging studies were available to confirm the diagnosis of ILD in more than 90% of patients who had ILD, pulmonary function testing data were not available for many of the patients, as a number of the children were of young age when such testing is unreliable in children. Thus, we were not able to study whether ILD patients with anti-Ro52 autoantibodies had more severe pulmonary dysfunction than those without these autoantibodies. Also, we cannot confirm the absence of ILD as many of the children without clinical suspicion of ILD did not have imaging and/or pulmonary function testing. This however, is a limitation of standard clinical care in pediatric patients who have challenges to undergo such testing.

Overall, this study shows that anti-Ro52 autoantibodies are present in 14% of patients with juvenile myositis and are strongly associated with ILD, more severe illness, and poorer outcomes, even when correcting for the co-existence of MSAs. In juvenile myositis patients, anti-Ro52 autoantibodies are associated not only with the presence of anti-synthetase autoantibodies, as previously reported in adult myositis patients, but also with anti-MDA5 autoantibodies, and the co-existence of these MSAs increases the likelihood of ILD and poor outcome. The current standard of care in patients with juvenile myositis who have reactivity to MSAs associated with pulmonary manifestations (such as anti-MDA5 and anti-aminoacyl tRNA synthetase autoantibodies) is to have a high index of suspicion

for the development of ILD and modify management accordingly. Our data suggest that testing for anti-Ro52 autoantibodies may also have a role in disease monitoring, management, and patient prognosis in juvenile myositis patients. Future studies will be required to determine whether anti-Ro52 autoantibodies are not only useful biomarkers, but whether they also play a pathological role in the development of ILD and other disease manifestations in myositis patients.

REFERENCES:

1. Pachman LM, Khojah AM. Advances in Juvenile Dermatomyositis: Myositis Specific Antibodies Aid in Understanding Disease Heterogeneity. J Pediatr. 2018 Apr; 195:16-27.

2. Rider LG, Shah M, Mamyrova G, Huber AM, Rice MM, Targoff IN, et al. The myositis autoantibody phenotypes of the juvenile idiopathic inflammatory myopathies. Medicine (Baltimore). 2013 Jul; 92(4):223-243.

3. Rider LG, Nistala K. The juvenile idiopathic inflammatory myopathies: pathogenesis, clinical and autoantibody phenotypes, and outcomes. J Intern Med. 2016 Jul; 280(1):24-38.

4. Sakurai N, Nagai K, Tsutsumi H, Ichimiya S. Anti-CADM-140 antibody-positive juvenile dermatomyositis with rapidly progressive interstitial lung disease and cardiac involvement. J Rheumatol. 2011 May; 38(5):963-964.

5. Tansley SL, Betteridge ZE, Gunawardena H, Jacques TS, Owens CM, Pilkington C, et al. Anti-MDA5 autoantibodies in juvenile dermatomyositis identify a distinct clinical phenotype: a prospective cohort study. Arthritis Res Ther. 2014 Jul 2; 16(4):R138.

6. Kurtzman DJB, Vleugels RA. Anti-melanoma differentiation-associated gene 5 (MDA5) dermatomyositis: A concise review with an emphasis on distinctive clinical features. J Am Acad Dermatol. 2018 Apr; 78(4):776-785.

7. Kobayashi N, Takezaki S, Kobayashi I, Iwata N, Mori M, Nagai K, et al. Clinical and laboratory features of fatal rapidly progressive interstitial lung disease associated with juvenile dermatomyositis. Rheumatology (Oxford). 2015 May; 54(5):784-791.

8. Miller FW, Love LA, Barbieri SA, Balow JE, Plotz PH. Lymphocyte activation markers in idiopathic myositis: changes with disease activity and differences among clinical and autoantibody subgroups. Clin Exp Immunol. 1990 Sep; 81(3):373-379.

9. Targoff IN, Mamyrova G, Trieu EP, Perurena O, Koneru B, O'Hanlon TP, et al. A novel autoantibody to a 155-kd protein is associated with dermatomyositis. Arthritis Rheum. 2006 Nov; 54(11):3682-3689.

10. Kobayashi I, Okura Y, Yamada M, Kawamura N, Kuwana M, Ariga T. Anti-melanoma differentiation-associated gene 5 antibody is a diagnostic and predictive marker for interstitial lung diseases associated with juvenile dermatomyositis. J Pediatr. 2011 Apr; 158(4):675-677.

11. Pinal-Fernandez I, Casal-Dominguez M, Huapaya JA, Albayda J, Paik JJ, Johnson C, et al. A longitudinal cohort study of the anti-synthetase syndrome: increased severity of interstitial lung disease in black patients and patients with anti-PL7 and anti-PL12 autoantibodies. Rheumatology (Oxford). 2017 Jun 1; 56(6):999-1007.

12. La Corte R, Lo Mo Naco A, Locaputo A, Dolzani F, Trotta F. In patients with antisynthetase syndrome the occurrence of anti-Ro/SSA antibodies causes a more severe interstitial lung disease. Autoimmunity. 2006 May; 39(3):249-253.

13. Vancsa A, Csipo I, Nemeth J, Devenyi K, Gergely L, Danko K. Characteristics of interstitial lung disease in SS-A positive/Jo-1 positive inflammatory myopathy patients. Rheumatol Int. 2009 Jul; 29(9):989-994.

14. Bauhammer J, Blank N, Max R, Lorenz HM, Wagner U, Krause D, et al. Rituximab in the Treatment of Jo1 Antibody-associated Antisynthetase Syndrome: Anti-Ro52 Positivity as a Marker for Severity and Treatment Response. J Rheumatol. 2016 Aug; 43(8):1566-1574.

15. Marie I, Hatron PY, Dominique S, Cherin P, Mouthon L, Menard JF, et al. Short-term and long-term outcome of anti-Jo1-positive patients with anti-Ro52 antibody. Semin Arthritis Rheum. 2012 Jun; 41(6):890-899.

16. Srivastava P, Dwivedi S, Misra R. Myositis-specific and myositis-associated autoantibodies in Indian patients with inflammatory myositis. Rheumatol Int. 2016 Jul; 36(7):935-943.

17. Bohan A, Peter JB. Polymyositis and dermatomyositis (second of two parts). N Engl J Med. 1975 Feb 20; 292(8):403-407.

18. Shah M, Mamyrova G, Targoff IN, Huber AM, Malley JD, Rice MM, et al. The clinical phenotypes of the juvenile idiopathic inflammatory myopathies. Medicine (Baltimore). 2013 Jan; 92(1):25-41.

19. Habers GE, Huber AM, Mamyrova G, Targoff IN, O'Hanlon TP, Adams S, et al. Brief Report: Association of Myositis Autoantibodies, Clinical Features, and Environmental Exposures at Illness Onset With Disease Course in Juvenile Myositis. Arthritis Rheumatol. 2016 Mar; 68(3):761-768.

20. Oddis CV, Rider LG, Reed AM, Ruperto N, Brunner HI, Koneru B, et al. International consensus guidelines for trials of therapies in the idiopathic inflammatory myopathies. Arthritis Rheum. 2005 Sep; 52(9):2607-2615.

21. Kishi T, Bayat N, Ward MM, Huber AM, Wu L, Mamyrova G, et al. Medications received by patients with juvenile dermatomyositis. Semin Arthritis Rheum. 2018 Mar 28.

22. McCann LJ, Juggins AD, Maillard SM, Wedderburn LR, Davidson JE, Murray KJ, et al. The Juvenile Dermatomyositis National Registry and Repository (UK and Ireland)--clinical characteristics of children recruited within the first 5 yr. Rheumatology (Oxford). 2006 Oct; 45(10):1255-1260.

23. Pachman LM, Abbott K, Sinacore JM, Amoruso L, Dyer A, Lipton R, et al. Duration of illness is an important variable for untreated children with juvenile dermatomyositis. J Pediatr. 2006 Feb; 148(2):247-253.

24. Pachman LM, Hayford JR, Chung A, Daugherty CA, Pallansch MA, Fink CW, et al. Juvenile dermatomyositis at diagnosis: clinical characteristics of 79 children. J Rheumatol. 1998 Jun; 25(6):1198-1204.

25. Constantin T, Ponyi A, Orban I, Molnar K, Derfalvi B, Dicso F, et al. National registry of patients with juvenile idiopathic inflammatory myopathies in Hungary--clinical characteristics and disease course of 44 patients with juvenile dermatomyositis. Autoimmunity. 2006 May; 39(3):223-232.

26. Kishi T, Rider LG, Pak K, Barillas-Arias L, Henrickson M, McCarthy PL, et al. Association of Anti-3-Hydroxy-3-Methylglutaryl-Coenzyme A Reductase Autoantibodies With DRB1*07:01 and Severe Myositis in Juvenile Myositis Patients. Arthritis Care Res (Hoboken). 2017 Jul; 69(7):1088-1094.

27. Ittah M, Miceli-Richard C, Eric Gottenberg J, Lavie F, Lazure T, Ba N, et al. B cell-activating factor of the tumor necrosis factor family (BAFF) is expressed under stimulation by interferon in salivary gland epithelial cells in primary Sjogren's syndrome. Arthritis Res Ther. 2006; 8(2):R51.

28. Krystufkova O, Barbasso Helmers S, Venalis P, Malmstrom V, Lindroos E, Vencovsky J, et al. Expression of BAFF receptors in muscle tissue of myositis patients with anti-Jo-1 or anti-Ro52/anti-Ro60 autoantibodies. Arthritis Res Ther. 2014 Oct 10; 16(5):454.

A.4 More prominent muscle involvement in those dermatomyositis patients with anti-Mi2 autoantibodies. Neurology 2019 (PMID: 31594859).

In this longitudinal cohort study, we aimed to define the clinical phenotype of dermatomyositis (DM) with anti-Mi2 autoantibodies. We analyzed the prevalence and severity of clinical features at disease onset and during follow-up in patients with anti-Mi2-positive DM compared to patients with anti-Mi2-negative DM, AS, and IMNM. We also assessed the longitudinal titers of anti-Mi2 autoantibody.

A total of 58 patients with anti-Mi2-positive DM, 143 patients with anti-Mi2-negative DM, 162 patients with AS, and 170 patients with IMNM were included. Among patients with anti-Mi2-positive DM, muscle weakness was present in 60% at disease onset and occurred in 98% during longitudinal follow-up; fewer patients with anti-Mi2-negative DM developed weakness (85%; $p = 0.008$). Patients with anti-Mi2-positive DM were weaker and had higher CK levels than patients with anti-Mi2-negative DM or patients with AS. Muscle biopsies from patients with anti-Mi2-positive DM had prominent necrosis. Anti-Mi2 autoantibody levels correlated with CK levels and strength ($p < 0.001$). With treatment, most patients with anti-Mi2-positive DM had improved strength and CK levels; among 10 with multiple serum samples collected over 4 or more years, anti-Mi2 autoantibody titers declined in all and normalized in 3, 2 of whom stopped immunosuppressant treatment and never relapsed. Patients with anti-Mi2-positive DM had less calcinosis (9% vs 28%; $p = 0.003$), interstitial lung disease (5% vs 16%; $p = 0.04$), and fever (7% vs 21%; $p = 0.02$) than did patients with anti-Mi2- negative DM.

In conclusion, patients with anti-Mi2-positive DM have more severe muscle disease than patients with anti-Mi2-negative DM or patients with AS. Anti-Mi2 autoantibody levels correlate with disease severity and may normalize in patients who enter remission.

More prominent muscle involvement in those dermatomyositis patients with anti-Mi2 autoantibodies

Iago Pinal-Fernandez M.D., Ph.D.[1,2,3]*, Christopher Mecoli M.D.[4]*, Maria Casal-Dominguez M.D., Ph.D.[1,2]*, Katherine Pak M.D.[1], Yuji Hosono M.D., Ph.D.[1], Julio Huapaya M.D.[5], Wilson Huang BA[1], Jemima Albayda M.D.[4], Eleni Tiniakou M.D.[4], Julie Paik M.D.[4], Cheilonda Johnson M.D.[6], Sonye K. Danoff M.D., Ph.D.[4], Andrea M. Corse M.D.[2], Lisa Christopher-Stine M.D., M.P.H.[2,4], and Andrew L. Mammen M.D., Ph.D.[1,2,4]

* These authors contributed equally to this project.

[1]Muscle Disease Unit, Laboratory of Muscle Stem Cells and Gene Regulation, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland.

[2]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland.

[3]Faculty of Health Sciences, Universitat Oberta de Catalunya, Barcelona, Spain.

[4]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland.

[5]Department of Medicine, Medstar Georgetown University Hospital, Washington DC.

[6]Department of Medicine, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania.

Address correspondence to: Andrew L. Mammen, M.D., Ph.D., or Iago Pinal-Fernandez, M.D., Ph.D. Muscle Disease Unit, Laboratory of Muscle Stem Cells and Gene Regulation, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, 50 South Drive, Room 1141, Building 50, MSC 8024, Bethesda, MD 20892. E-mail: andrew.mammen@nih.gov or iago.pinalfernandez@nih.gov. Phone: 301-451-1199. Fax: 301-594-0305.

MANUSCRIPT WORD COUNT: 2972

SEARCH TERMS: Autoimmune diseases, muscle disease, myositis, dermatomyositis, autoantibodies.

FUNDING: This research was supported by the Intramural Research Program of the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health. SKD, LCS and the Myositis Research Database are supported by the Huayi and Siuling Zhang Discovery Fund. We are also grateful to Dr. Peter Buck, whose generous support made this work possible. CM and JA are supported by the Jerome L. Greene Foundation. IPF´s research was supported by a Fellowship from the Myositis Association. ET is supported by the Jerome L. Greene Foundation and the Rheumatology Research Foundation Science Development Award.

CONFLICT OF INTERESTS: No potential conflict of interest was reported by the authors.

DISLCOSURES:

I Pinal-Fernandez reports no disclosures relevant to the manuscript.

C Mecoli reports no disclosures relevant to the manuscript.

M Casal-Dominguez reports no disclosures relevant to the manuscript.

K Pak reports no disclosures relevant to the manuscript.

Y Hosono reports no disclosures relevant to the manuscript.

J Huapaya reports no disclosures relevant to the manuscript.

W Huang reports no disclosures relevant to the manuscript.

J Albayda reports no disclosures relevant to the manuscript.

E Tiniakou reports no disclosures relevant to the manuscript.

J Paik reports no disclosures relevant to the manuscript.

C Johnson reports no disclosures relevant to the manuscript.

SK. Danoff reports no disclosures relevant to the manuscript.

AM. Corse reports no disclosures relevant to the manuscript.

L Christopher-Stine reports no disclosures relevant to the manuscript.

AL. Mammen reports no disclosures relevant to the manuscript.

ABSTRACT

Objective: To define the clinical phenotype of dermatomyositis (DM) patients with anti-Mi2 autoantibodies.

Methods: In this longitudinal cohort study, the prevalence and severity of clinical features at disease onset and during follow-up in anti-Mi2-positive DM patients were compared to patients with anti-Mi2-negative DM, the antisynthetase syndrome (AS), and immune-mediated necrotizing myopathy (IMNM). Longitudinal anti-Mi2 autoantibody titers were assessed.

Results: Fifty-eight anti-Mi2-positive DM, 143 anti-Mi2-negative DM, 162 AS, and 170 IMNM patients were included. Among anti-Mi2-positive DM patients, muscle weakness was present in 60% at disease onset and occurred in 98% during longitudinal follow-up; fewer anti-Mi2-negative DM patients developed weakness (85%; p=0.008). Anti-Mi2-positive DM patients were weaker and had higher creatine kinase (CK) levels than anti-Mi2-negative DM or AS patients. Muscle biopsies from anti-Mi2-positive DM patients had prominent necrosis. Anti-Mi2 autoantibody levels correlated with CK levels and strength (p<0.001). With treatment, most anti-Mi2-positive DM patients had improved strength and CK levels; among 10 with multiple serum samples collected over 4 or more years, anti-Mi2 autoantibody titers declined in all and normalized in 3, 2 of whom stopped immunosuppressant treatment and never relapsed. Anti-Mi2-positive DM patients had less calcinosis (9% vs. 28%; p=0.003), interstitial lung disease (5% vs. 16%; p=0.04), and fever (7% vs. 21%; p=0.02) than anti-Mi2-negative DM patients.

Conclusions: Anti-Mi2-positive DM patients have more severe muscle disease than anti-Mi2-negative DM or AS patients. Anti-Mi2 autoantibody levels correlate with disease severity and may normalize in patients who enter remission.

INTRODUCTION

The inflammatory myopathies (IIM) are a heterogeneous family of diseases that affect skeletal muscle and, in some cases, the skin, lungs, and/or joints. At least four well-defined types of IIM are now widely recognized, including dermatomyositis (DM), the antisynthetase syndrome (AS), immune-mediated necrotizing myopathy (IMNM), and inclusion body myositis (IBM).[1]

Myositis-specific autoantibodies (MSAs) are a common serological feature of IIM patients and each MSA is closely associated with a particular type of IIM.[1, 2] For example, autoantibodies recognizing one of the tRNA synthetases (e.g., Jo1, PL-7, or PL-12) are found in AS, a disease characterized by myositis, interstitial lung disease (ILD), arthritis, Raynaud's phenomenon, and/or rash. In contrast, autoantibodies recognizing HMGCR or SRP are associated with IMNM, which is usually dominated by skeletal muscle involvement. Most patients with DM have an autoantibody recognizing Mi2, TIF1γ, NXP2, or MDA5; in addition to skin and muscle, DM may also affect other organ systems.

Within each type of IIM, individual MSAs define unique subtypes, each with its own clinical features, prognosis, and response to treatment. For instance, in DM, malignancy, cutaneous calcinosis, and rapidly progressive ILD are strongly associated with autoantibodies recognizing TIF1γ, NXP2, and MDA5, respectively.[1, 2] Prior reports have characterized anti-Mi2-positive DM as having mild muscle disease[3, 4] along with typical DM skin manifestations, infrequent ILD, and low cancer risk.[3, 5-7] However, these studies were limited by small numbers of patients,[5, 6] a lack of longitudinal assessments,[3, 5-7] and/or the absence of currently recognized comparison groups such as AS or IMNM.[3, 5-7] In addition, the evolution of anti-Mi2 autoantibody levels over years of follow-up has not

been described. Thus, the purpose of the current study was to define the phenotype of anti-Mi2-positive DM by comparing the prevalence and severity of the clinical features of these patients to those with anti-Mi2-negative DM, AS, and IMNM at disease onset and during the course of the disease. Furthermore, our study aimed to analyze the trend of anti-Mi2 autoantibody levels in patients with long follow-up times.

MATERIALS AND METHODS

Study population and autoantibody testing. All patients enrolled in the Johns Hopkins Myositis Center Longitudinal Cohort study between 2002 and 2018 were included in the current study if they were positive for autoantibodies recognizing Mi2, NXP2, TIF1γ, MDA5, Jo1, PL-7, PL-12, SRP, or HMGCR by at least two immunologic techniques from among the following: ELISA, *in vitro* transcription and translation immunoprecipitation, line blotting (EUROLINE myositis profile), or immunoprecipitation from S35-labeled HeLa cell lysates.[8, 9] Patients were included in the DM group if they had autoantibodies recognizing Mi2, NXP2, TIF1γ or MDA5. Alternatively, patients were classified as having AS if they had autoantibodies against Jo-1, PL-7, or PL-12. Finally, patients were included in the IMNM group if they tested positive for anti-SRP or anti-HMGCR autoantibodies.

Strength was evaluated by the examining physician using the Medical Research Council scale. This scale was transformed to Kendall´s 0-10 scale for analysis purposes as previously described.[10] Serial strength measurements for each patient were made by the same physician but more than 10 different physicians contributed to the measurements. For the purposes of analyses, right- and left-side measurements for arm abduction and hip flexion strength were combined and the average was used for calculations (possible range 0–10). Serum creatine kinase (CK) levels were included for the longitudinal analysis if obtained within a period of 6 weeks before or after strength testing. Skin manifestations (i.e., heliotrope rash or Gottron's sign), weakness, symptoms of esophageal involvement, antisynthetase syndrome-associated clinical features (e.g. mechanics hands, Raynaud's phenomenon, arthritis, fever), and other clinical features were documented both retrospectively at the onset of the disease (by

asking patients about features present at the onset of disease) and prospectively at each visit. Interstitial lung disease was defined through a multidisciplinary approach as recommended by the American Thoracic Society.[11] All available muscle biopsies were interpreted at the Johns Hopkins Neuromuscular Pathology Laboratory by pathologists blinded to autoantibody status. The pathologists systematically reported on the presence or absence of perifascicular atrophy, perivascular inflammation, primary inflammation (i.e., the invasion of non-necrotic fibers by mononuclear cells), and necrotizing myopathy (i.e., prominent myofiber necrosis in the absence of perifascicular atrophy or primary inflammation).

Anti-Mi2 autoantibody titers. For the quantitative anti-Mi2 autoantibody ELISA, 96-well ELISA plates were coated overnight at 4°C with 100ng of Mi2b protein (Abcam, ab124864) diluted in PBS. Replicate wells were incubated with phosphate-buffered saline (PBS) alone. After washing the plates, human serum samples, diluted 1:400 in PBS with 0.05% Tween (PBS-T), were added to wells (1 hour, 37°C). After washing, HRP-labeled goat anti-human antibody (Jackson ImmunoResearch 109-036-088; 1:10,000) was added to each well (30 minutes, 37° C). Color development was performed using SureBlue™ peroxidase reagent (KPL) and absorbances at 450 nm were determined. For each sample, the background absorbance from the PBS-coated wells was subtracted from that of the corresponding Mi2-coated well. Test sample absorbances were normalized using linear regression to a range of serial dilutions from an arbitrary positive anti-Mi2 patient, a reference serum included in every ELISA. Sera from 49 healthy control subjects enrolled at the National Institutes of Health Clinical Center were tested using this

ELISA, revealing a mean absorbance of 0.08 with a standard deviation of 0.03. The cutoff for a negative anti-Mi2 autoantibody titer was subsequently defined as the mean absorbance value plus three standard deviations of the healthy control subjects (i.e., 0.17).

Standard protocol approvals and patient consents. This study was approved by the Johns Hopkins and National Institutes of Health Institutional Review Boards; written informed consent was obtained from each participant.

Statistical analysis. Dichotomous variables were expressed as percentages and absolute frequencies, and continuous features were reported as means and standard deviations (SD). Pairwise comparisons for categorical variables between groups were made using chi-square test or Fisher's exact test, as appropriate. Student´s t-test was used to compare continuous variables among groups and paired t-test was used to compare the level of weakness of different muscle groups. CK, a highly positively skewed variable, was expressed as median, first, and third quartile for descriptive purposes, and was transformed through a base-10 logarithm for regression analysis.

The standardized mortality (SMR) and cancer (SCR) incidence rates were calculated as previously described.[9]

To account for differing numbers of visits per patient, the evolution of the CK levels and muscle strength were studied using multilevel linear regression models with random slopes and random intercepts. The mean of hip flexor and arm abductor strength (range 0-10) was used as the strength outcome for regression analysis.

Locally weighted regression was applied to graphically analyze the evolution of the strength, CK levels, and pulmonary function tests. Kaplan Meier curves and Cox regression were used to studying the evolution of each of the clinical features over time.

The influence of non-modifiable risk factors (sex, race, duration of disease, and age at the onset of the first symptom), the corticosteroid dose and the administration of intravenous immunoglobulins (IVIG), rituximab, methotrexate, azathioprine, and mycophenolate were used as adjusting covariates. Other treatments administered to less than 10% of the cohort were not included in the analysis.

All statistical analyses were performed using Stata/MP 14.1. A 2-sided p-value of 0.05 or less was considered statistically significant with no adjustment for multiple comparisons.

Data Availability Statement. All data relevant to the study are either included in the article or will be shared at the request of other investigators.

RESULTS

Epidemiologic features of anti-Mi2 patients

From among 2475 patients enrolled in the Johns Hopkins Myositis Center longitudinal cohort (which includes patients with inclusion body myositis and other muscle diseases), 533 patients (22%) were included in this study. Among these, 58 (11%) had anti-Mi2-positive DM, 143 (27%) had anti-Mi2-negative DM, 162 (30%) had AS, and 170 (32%) had IMNM. The general features of anti-Mi-2-positive patients and the comparator groups are detailed in Table 1. Compared to anti-Mi2-negative DM patients, anti-Mi2-positive DM patients were less likely to be female (62% vs. 78%; p=0.02) or white (64% vs. 79%; p=0.02). The prevalence of co-existing anti-Ro52 autoantibodies was markedly less in anti-Mi2-positive DM patients than in AS patients (17% vs. 81%, p<0.001). The longitudinal analysis included information from 5013 patient visits and 492 (10%) of these were from anti-Mi2-positive DM patients. The rates of mortality and cancer were not significantly different between the different myositis groups. Furthermore, anti-Mi2-positive DM patients did not have higher mortality or cancer rates than a control reference population (SMR=0.9, 95%CI 0.1-3.2; SCR=2.4, 95%CI 0.8-5.5). The two anti-Mi2-positive patients who died during the study follow-up did so at an advanced age from causes unrelated to the myositis.

Muscle Involvement

At disease onset, weakness was present in 60% of patients with anti-Mi2-positive DM and in 46% of anti-Mi2-negative DM patients (p=0.07). Compared to those with IMNM,

fewer anti-Mi2-positive DM patients had weakness at disease onset (60% vs. 88%; p<0.001) (Table 2). During the follow-up period, weakness occurred more commonly in anti-Mi2-positive DM patients than in DM patients without this autoantibody (98% vs. 85%; p=0.008) (Table 3).

At the initial visit, anti-Mi2-positive DM patients had weaker neck flexors, arm abductors, elbow extensors, hip flexors, hip extensors, and knee flexors than those with anti-Mi2-negative DM or AS (all p<0.05, Table 5). However, anti-Mi2-positive DM patients had stronger hip flexors than did IMNM patients at their first visit. During the course of follow-up, anti-Mi2-positive DM patients had significantly weaker mean hip flexor and mean arm abduction strength than anti-Mi2-negative or AS patients (all p<0.01, Table 4). During this time, patients in the anti-Mi2-positive DM group had substantially stronger mean hip flexors than those in the IMNM group (p<0.001). However, the mean arm abduction strength of the anti-Mi2-positive DM patients was remarkably similar to those with IMNM. Also, muscle enzymes were significantly higher in anti-Mi2-positive than in anti-Mi2-negative DM patients (p<0.001) (Table 4). Distal weakness was minimal in anti-Mi2-positive DM and the control groups (Table 5).

Multilevel regression analysis confirmed that anti-Mi-2-positive DM patients were weaker than anti-Mi2-negative DM or AS patients (all p<0.004) and stronger than IMNM patients (p=0.003) independent of the time from onset, age, sex, race or treatments at any given time during follow-up. This analysis also demonstrated that, compared to anti-Mi2-positive DM, CK levels were lower in anti-Mi2-negative DM (p<0.001), higher in IMNM (p<0.001), and similar in AS (p=0.6). The regression analysis also showed that in

anti-Mi2-positive DM, CK levels were inversely associated ($p<0.001$) with strength, independent of the aforementioned variables.

Kaplan Meier curves and Cox regression confirmed that most patients with anti-Mi2-positive DM developed weakness within the first 2 years of disease and that this occurred faster than in anti-Mi2-negative DM or AS (all $p<0.001$) independent of the age at onset, race or gender.

Of those anti-Mi2-positive DM patients who underwent thigh muscle MRI (n=23), muscle edema, muscle atrophy, fatty replacement, and fascial edema were present in 78%, 26%, 39%, and 43% of patients, respectively (Table 6). Compared to anti-Mi2-positive patients, more IMNM patients had atrophy (68% vs. 26%; $p<0.001$) and fatty replacement (87% vs. 39%; $p<0.001$).

Most anti-Mi2-positive DM patients with weakness at the first visit regained full strength within the first year of treatment (Figure 1) using a combination of corticosteroids along with methotrexate, mycophenolate, or azathioprine. Significant flares of weakness (defined either as increasing CK levels or worsening weakness) after treatment introduction were exceptional in our cohort.

Overall, anti-Mi2 autoantibody titers showed a weak but significant association with increased CK levels ($b=0.34$, $p<0.001$) and decreased strength ($b=-1.13$, $p<0.001$). Ten patients had 5 or more serum samples available during the course of disease and autoantibody titers declined over time in each patient. Indeed, antibody titers normalized in 3 (30%) of these patients. Importantly, 3 (30%) of these patients with decreasing autoantibody titers (including two of the 3 that reached the autoantibody normality

threshold) could stop treatment and never relapsed. Moreover, the strength and CK levels mirrored the evolution of the autoantibody titers in most of the patients.

Muscle biopsies were available for review from 27 anti-Mi2-positive DM, 32 anti-Mi2-negative DM, 26 AS, and 81 IMNM patients. There were no significant differences in the prevalence of perifascicular atrophy or perivascular inflammation between anti-Mi2-positive DM patients and those with anti-Mi2-negative DM or AS (Table 6). As expected, perifascicular atrophy was not present in muscle biopsies from IMNM patients. Primary inflammation (i.e. the invasion of non-necrotic fibers by mononuclear cells), although only present in a minority of cases, was more common in anti-Mi2-positive than in anti-Mi2-negative DM patients (19% vs. 0%; p=0.02). Myofiber necrosis was the most prominent histologic feature in many anti-Mi2-positive DM muscle biopsies. Nonetheless, the histological diagnosis of necrotizing myopathy was relatively uncommon in those with anti-Mi2-positive DM (19%), anti-Mi2-negative DM (6%), or AS (23%) compared to those with IMNM (78%;p<0.001).

Extramuscular manifestations

At disease onset, the prevalence of extramuscular manifestations affecting the skin, lungs, esophagus, and joints was similar between DM patients with and without anti-Mi2 autoantibodies (Table 2). However, over the course of the disease, compared to anti-Mi2-negative patients, anti-Mi2-positive DM patients had less calcinosis (9% vs. 28%; p=0.003), interstitial lung disease (5% vs. 16%; p=0.04), and fever (7% vs. 21%; p=0.02). Similarly, compared to AS patients, anti-Mi2-positive DM patients had more frequent DM-specific rashes (i.e., heliotrope or Gottron's sign/papules) (93% vs. 57%; p<0.001) and

dysphagia (53% vs. 37%; p=0.03) but less frequent mechanic's hands (21% vs. 56%; p<0.001), ILD (5% vs. 80%; p<0.001), arthritis (21% vs. 58%; p<0.001) and fever (7% vs. 22%; p=0.009).

DISCUSSION

In this study, we have shown that anti-Mi2-positive DM patients are more likely to have weakness at disease onset and during the course of disease than anti-Mi2-negative DM patients. Furthermore, DM patients with anti-Mi2 autoantibodies have higher maximum CK levels and more severe proximal muscle weakness than DM patients without these autoantibodies. Taken together, the data presented here suggest that anti-Mi2-positive DM patients have more severe muscle disease than anti-Mi2-negative DM patients. This study also shows that anti-Mi2-positive DM patients have more severe muscle disease than AS patients. In fact, the severity of muscle disease in the upper extremities of anti-Mi2-positive DM was comparable to that of IMNM. These findings contradict prior reports concluding that anti-Mi2-positive DM patients have relatively mild myositis.[3, 4] Although the reasons for this discrepancy are not immediately apparent, it should be noted that the conclusions of the prior reports were not supported by quantitative data such as the comparative analyses of muscle strength that were performed in the current study.

In contrast to the frequent and severe muscle disease seen in anti-Mi2-positive DM, this study shows that extramuscular manifestations are generally less common in this group than in anti-Mi2-negative DM or AS. Specifically, calcinosis, ILD, and fever were less common than in anti-Mi2-negative DM whereas ILD, arthritis and fever were less common than in AS. As expected, many extramuscular manifestations were more common in anti-Mi2-positive DM than in IMNM, which predominantly affects skeletal muscle.

Unlike other DM myositis-specific autoantibodies like anti- TIF1γ [12] and, to a lesser extent, anti-NXP2,[13] where the association with cancer has been confirmed by several groups, the association between anti-Mi2 autoantibodies and cancer is still a matter of debate.[3, 6, 7] In our study we could not find statistical evidence that anti-Mi2 autoantibodies are associated with cancer. However, the standardized cancer rate showed a trend towards a positive association with cancer (95%CI 0.8-5.5). Based on this, we believe that the safest recommendation for anti-Mi2 patients at this moment is to continue being screened for cancer within three years of the onset of myositis symptoms.

Prior studies have shown that serum levels of some MSAs, including anti-SRP[14] and Jo-1[15], are closely associated with disease activity and may normalize during periods of remission. Levels of other MSAs, such as anti-HMGCR, may have an association with disease activity but remain present at high levels even when the disease appears to be relatively quiescent.[16, 17] In the current study, we demonstrate that anti-Mi2 autoantibody levels are associated with muscle enzyme levels and strength, confirming the findings of a prior report.[18] We also show that in 30% of patients followed for four or more years, anti-Mi2 autoantibody levels not only decline, but even normalize as the disease becomes less active. Although a prior study demonstrated a median drop in anti-Mi2 levels of 38.1% 44 weeks after receiving B-cell depletion therapy, normalization of autoantibody levels was not reported.[18] The comparatively short follow-up period of this study may have precluded recording more substantial declines. Indeed, patients in the current study only had normalization of anti-Mi2 titers after 4-10 years of treatment.

Our observations regarding longitudinal anti-Mi2 autoantibody levels have several implications. First, since anti-Mi2 autoantibody levels declined and/or normalized even in

the majority of patients who did not receive rituximab, declines in autoantibody levels are most likely not exclusively due to B cell depletion. Also, because anti-Mi2 autoantibodies can become negative over time, it may be important to perform MSA testing early in the course of DM to avoid false negative results. Moreover, taken together with those of a prior study,[18] our results suggest that anti-Mi2 autoantibody levels may be a clinically useful biomarker of disease activity. However, future studies will be needed to determine whether it's safe to completely withdraw immunosuppressive medications from patients whose anti-Mi2 autoantibody levels have normalized.

The strengths of our study include the relatively long duration of follow-up of anti-Mi2-positive DM and controls groups at the Johns Hopkins Myositis Center, enabling the study of clinical trajectory and long-term outcomes. Furthermore, the large size of this cohort has allowed for the robust comparison of multiple control groups including large numbers of AS and IMNM patients. However, due to the long-term existence of our cohort, many recently developed outcome measures were not collected for many patients. These include the Cutaneous Dermatomyositis Disease Area and Severity Index (CDASI), patient and physician global assessments, and patient-reported outcome measures. Further work is warranted to incorporate these measures longitudinally for better comparison across cohorts.

Despite its limitations, this study shows that anti-Mi2-positive DM patients have more severe muscle disease than anti-Mi2-negative DM or AS patients. This finding supports the growing body of evidence that DM is a heterogeneous disease and that MSAs such as anti-Mi2 can be used to classify DM patients into more homogeneous subgroups.

FIGURE LEGENDS

Figure 1. Evolution of strength and creatine kinase levels in anti-Mi2-positive patients. In these graphs, the dots represent individual data points for strength and CK levels; the lines are the output of locally weighted regression analyses using this data to graphically analyze the evolution of the strength and CK over time.

AUTHORS

Name	Location	Role	Contribution
Iago Pinal-Fernandez, MD, PhD	NIH, Bethesda	Author	Design of the study, acquisition of data, analysis and interpretation of data, drafting the manuscript for intellectual content.
Christopher Mecoli	Johns Hopkins, Baltimore	Author	Design of the study, acquisition of data, analysis and interpretation of data, and drafting the manuscript for intellectual content.
Maria Casal-Dominguez, MD, PhD	NIH, Bethesda	Author	Design of the study, acquisition of data, analysis and interpretation of data and revising the manuscript for intellectual content.
Katherine Pak, MD	NIH, Bethesda	Author	acquisition of data, design of the study, analysis and interpretation of data and revising the manuscript for intellectual content.
Yuji Hosono	NIH, Bethesda	Author	Acquisition of data, design of the study, analysis and interpretation of data and revising the manuscript for intellectual content.
Julio Huapaya, MD	Medstar Georgetown University Hospital, Washington DC	Author	Acquisition of data and revising the manuscript for intellectual content.

Wilson Huang	NIH, Bethesda	Author	Acquisition of data, design of the study and revising the manuscript for intellectual content.
Jemima Albayda, MD	Johns Hopkins, Baltimore	Author	Acquisition of data and revising the manuscript for intellectual content.
Eleni Tiniakou, MD	Johns Hopkins, Baltimore	Author	Acquisition of data and revising the manuscript for intellectual content.
Julie Paik, MD, MHS	Johns Hopkins, Baltimore	Author	Acquisition of data and revising the manuscript for intellectual content.
Cheilonda Johnson, MD, MHS	University of Pennsylvania, Philadelphia	Author	Acquisition of data and revising the manuscript for intellectual content.
Sonye K. Danoff, MD, PhD	Johns Hopkins, Baltimore	Author	Acquisition of data and revising the manuscript for intellectual content.
Andrea Corse, MD	Johns Hopkins, Baltimore	Author	Acquisition of data, analysis and interpretation of data and revising the manuscript for intellectual content.
Lisa Christopher-Stine, MD, MPH	Johns Hopkins, Baltimore	Author	Acquisition of data and revising the manuscript for intellectual content.
Andrew L. Mammen, MD, PhD	NIH, Bethesda	Author	Design of the study, analysis and interpretation of data, revising the manuscript for intellectual content.

REFERENCES

1. Selva-O'Callaghan A, Pinal-Fernandez I, Trallero-Araguas E, Milisenda JC, Grau-Junyent JM, Mammen AL. Classification and management of adult inflammatory myopathies. Lancet Neurol 2018;17:816-828.

2. Benveniste O, Stenzel W, Allenbach Y. Advances in serological diagnostics of inflammatory myopathies. Curr Opin Neurol 2016;29:662-673.

3. Hengstman GJ, Vree Egberts WT, Seelig HP, et al. Clinical characteristics of patients with myositis and autoantibodies to different fragments of the Mi-2 beta antigen. Ann Rheum Dis 2006;65:242-245.

4. Ikeda N, Takahashi K, Yamaguchi Y, Inasaka M, Kuwana M, Ikezawa Z. Analysis of dermatomyositis-specific autoantibodies and clinical characteristics in Japanese patients. J Dermatol 2011;38:973-979.

5. Komura K, Fujimoto M, Matsushita T, et al. Prevalence and clinical characteristics of anti-Mi-2 antibodies in Japanese patients with dermatomyositis. J Dermatol Sci 2005;40:215-217.

6. Hamaguchi Y, Kuwana M, Hoshino K, et al. Clinical correlations with dermatomyositis-specific autoantibodies in adult Japanese patients with dermatomyositis: a multicenter cross-sectional study. Arch Dermatol 2011;147:391-398.

7. Lega JC, Fabien N, Reynaud Q, et al. The clinical phenotype associated with myositis-specific and associated autoantibodies: a meta-analysis revisiting the so-called antisynthetase syndrome. Autoimmun Rev 2014;13:883-891.

8. Mammen AL, Chung T, Christopher-Stine L, et al. Autoantibodies against 3-hydroxy-3-methylglutaryl-coenzyme A reductase in patients with statin-associated autoimmune myopathy. Arthritis Rheum 2011;63:713-721.

9. Pinal-Fernandez I, Casal-Dominguez M, Huapaya JA, et al. A longitudinal cohort study of the anti-synthetase syndrome: increased severity of interstitial lung disease in black patients and patients with anti-PL7 and anti-PL12 autoantibodies. Rheumatology (Oxford) 2017;56:999-1007.

10. Rider LG, Werth VP, Huber AM, et al. Measures of adult and juvenile dermatomyositis, polymyositis, and inclusion body myositis: Physician and Patient/Parent Global Activity, Manual Muscle Testing (MMT), Health Assessment Questionnaire (HAQ)/Childhood Health Assessment Questionnaire (C-HAQ), Childhood Myositis Assessment Scale (CMAS), Myositis Disease Activity Assessment Tool (MDAAT), Disease Activity Score (DAS), Short Form 36 (SF-36), Child Health Questionnaire (CHQ), physician global damage, Myositis Damage Index (MDI), Quantitative Muscle Testing (QMT), Myositis Functional Index-2 (FI-2), Myositis Activities Profile (MAP), Inclusion Body Myositis Functional Rating Scale (IBMFRS), Cutaneous Dermatomyositis Disease Area and Severity Index (CDASI), Cutaneous Assessment Tool (CAT), Dermatomyositis Skin Severity Index (DSSI), Skindex, and Dermatology Life Quality Index (DLQI). Arthritis Care Res (Hoboken) 2011;63 Suppl 11:S118-157.

11. Travis WD, Costabel U, Hansell DM, et al. An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. Am J Respir Crit Care Med 2013;188:733-748.

12. Trallero-Araguas E, Rodrigo-Pendas JA, Selva-O'Callaghan A, et al. Usefulness of anti-p155 autoantibody for diagnosing cancer-associated dermatomyositis: a systematic review and meta-analysis. Arthritis Rheum 2012;64:523-532.

13. Albayda J, Pinal-Fernandez I, Huang W, et al. Antinuclear Matrix Protein 2 Autoantibodies and Edema, Muscle Disease, and Malignancy Risk in Dermatomyositis Patients. Arthritis Care Res (Hoboken) 2017;69:1771-1776.

14. Benveniste O, Drouot L, Jouen F, et al. Correlation of anti-signal recognition particle autoantibody levels with creatine kinase activity in patients with necrotizing myopathy. Arthritis Rheum 2011;63:1961-1971.

15. Stone KB, Oddis CV, Fertig N, et al. Anti-Jo-1 antibody levels correlate with disease activity in idiopathic inflammatory myopathy. Arthritis Rheum 2007;56:3125-3131.

16. Werner JL, Christopher-Stine L, Ghazarian SR, et al. Antibody levels correlate with creatine kinase levels and strength in anti-3-hydroxy-3-methylglutaryl-coenzyme A reductase-associated autoimmune myopathy. Arthritis Rheum 2012;64:4087-4093.

17. Tiniakou E, Pinal-Fernandez I, Lloyd TE, et al. More severe disease and slower recovery in younger patients with anti-3-hydroxy-3-methylglutaryl-coenzyme A reductase-associated autoimmune myopathy. Rheumatology (Oxford) 2017;56:787-794.

18. Aggarwal R, Oddis CV, Goudeau D, et al. Autoantibody levels in myositis patients correlate with clinical response during B cell depletion with rituximab. Rheumatology (Oxford) 2016;55:991-999.

Table 1: General features of anti-Mi2 patients.

	Anti-Mi2 (n=58)	Non-Mi2 DM (n=143)	AS (n=162)	IMNM (n=170)
Female sex	62% (36)	78% (111)*	72% (117)	63% (107)
Race				
White	64% (37)	79% (113)*	58% (94)	68% (115)
Black	16% (9)	13% (18)	31% (51)*	24% (41)
Other races	21% (12)	8% (12)*	10% (17)*	8% (14)**
Age of onset (years)	48.9 (15.0)	46.5 (15.6)	45.7 (13.1)	50.2 (15.5)
Time of follow-up (years)	4.3 (3.6)	4.4 (3.6)	4.6 (3.9)	3.9 (3.8)
Number of visits per participant	8.5 (6.8)	10.2 (7.3)	9.6 (6.8)	8.9 (8.9)
Cancer associated myositis	9% (5)	8% (12)	3% (4)	5% (7)
Death during follow-up	3% (2)	5% (7)	9% (14)	3% (5)
Anti-Ro52	17% (10)	22% (32)	81% (131)***	22% (37)
Treatments				
Corticosteroids	90% (52)	82% (117)	94% (153)	74% (125)*
Azathioprine	28% (16)	30% (43)	52% (84)**	25% (42)
Methotrexate	57% (33)	52% (74)	42% (68)	51% (87)
Mycophenolate	31% (18)	38% (55)	40% (65)	19% (33)
IVIG	45% (26)	49% (70)	36% (58)	42% (71)
Rituximab	19% (11)	15% (22)	23% (38)	24% (40)

* p<0.05, ** p<0.01, *** p<0.001

Dichotomous variables were expressed as percentage (count) and continuous variables as mean (SD). Bivariate comparisons of continuous variables were made using Student´s t-test while bivariate comparisons of dichotomous variables were made either using chi-squared test or Fisher´s exact test, as appropriate. Each one of the clinical groups was compared to the sample of anti-Mi2 patients.

Table 2: Clinical features of anti-Mi2 patients at the onset of the disease.

	Anti-Mi2 DM (n=58)	Non-Mi2 DM (n=143)	AS (n=162)	IMNM (n=170)
Muscle involvement				
Muscle weakness	60% (35)	46% (66)	57% (92)	88% (150)***
Skin involvement				
DM-specific skin involvement	64% (37)	71% (102)	11% (18)***	2% (4)***
Raynaud's phenomenon	3% (2)	5% (7)	14% (23)*	6% (11)
Mechanics hands	2% (1)	5% (7)	15% (24)**	1% (1)
Calcinosis	0% (0)	3% (4)	2% (4)	0% (0)
Subcutaneous edema	5% (3)	8% (11)	7% (11)	2% (3)
Lung involvement				
Interstitial lung disease	2% (1)	6% (8)	49% (80)***	1% (2)
Esophageal involvement				
Gastroesophageal reflux disease	0% (0)	0% (0)	7% (12)*	1% (1)
Dysphagia	7% (4)	11% (16)	9% (14)	10% (17)
Joint involvement				
Arthritis	7% (4)	6% (9)	20% (32)*	1% (1)*
Arthralgia	16% (9)	16% (23)	45% (73)***	5% (9)*
Systemic involvement				
Fever	3% (2)	8% (11)	9% (15)	2% (3)

* p<0.05, ** p<0.01, *** p<0.001

Chi-squared or Fisher´s exact tests were used to compare each one of the clinical groups with the anti-Mi2 patients. The clinical features at the onset of the disease are based on retrospective reports from the patients.

Table 3: Cumulative clinical features of anti-Mi2 patients.

	Anti-Mi2 DM (n=58)	Non-Mi2 DM (n=143)	AS (n=162)	IMNM (n=170)
Muscle involvement				
Muscle weakness	98% (57)	85% (122)**	93% (150)	97% (165)
Skin involvement				
DM-specific skin involvement	93% (54)	96% (137)	57% (92)***	8% (14)***
Raynaud's phenomenon	29% (17)	22% (31)	36% (59)	16% (27)*
Mechanics hands	21% (12)	30% (43)	56% (91)***	4% (7)***
Calcinosis	9% (5)	28% (40)**	7% (12)	1% (1)**
Subcutaneous edema	17% (10)	24% (35)	26% (42)	5% (9)*
Lung involvement				
Interstitial lung disease	5% (3)	16% (23)*	80% (130)***	6% (10)
Esophageal involvement				
Gastroesophageal reflux disease	34% (20)	29% (41)	27% (44)	24% (41)
Dysphagia	53% (31)	53% (76)	37% (60)*	45% (77)
Joint involvement				
Arthritis	21% (12)	24% (34)	58% (94)***	6% (10)***
Arthralgia	52% (30)	54% (77)	63% (102)	33% (56)*
Systemic involvement				
Fever	7% (4)	21% (30)*	22% (36)**	7% (12)

* p<0.05, ** p<0.01, *** p<0.001

Chi-squared or Fisher´s exact tests were used to compare each one of the clinical groups with the anti-Mi2 patients.

Table 4. Activity of the disease.

	Anti-Mi2 DM (n=58)	Non-Mi2 DM (n=143)	AS (n=162)	IMNM (n=170)
Mean hip flexor strength	8.0 (2.4)	9.0 (1.5)***	9.1 (1.4)***	6.4 (2.9)***
Hip flexors strength at last visit	8.2 (2.7)	9.3 (1.6)***	9.3 (1.4)***	6.6 (3.5)**
Mean arm abductor strength	8.4 (2.1)	9.2 (1.3)**	9.4 (1.1)***	8.4 (2.0)
Arm abductors strength at last visit	8.5 (2.6)	9.4 (1.5)**	9.5 (1.1)***	8.5 (2.4)
Median CK	363 (158-1325)	105 (60-174)***	330 (116-1020)	1373 (502-2826)***
Maximum CK	3908 (2230-7070)	242 (110-1200)***	1464 (394-6024)***	5271 (2181-9786)
Mean aldolase	14.5 (11.5)	7.7 (3.9)***	23.6 (40.6)	33.9 (47.0)**
Maximum aldolase	22.8 (24.4)	9.4 (6.2)***	50.1 (169.0)	55.7 (70.4)**

* p<0.05, ** p<0.01, *** p<0.001

Strength and FVC values were expressed as means (SD) and CK as medians (Q1-Q3). Bivariate comparisons were made using Student´s t-test for the strength and Wilcoxon rank-sum test for CK. Follow-up strength was defined as the mean strength of all the visits, excluding the first one. Each one of the clinical groups was compared to the sample of anti-Mi2 patients.

Table 5. Pattern of weakness at the first visit of anti-Mi2 patients.

	Anti-Mi2 DM (n=58)	Non-Mi2 DM (n=143)	AS (n=162)	IMNM (n=170)
Neck flexors	8.1 (2.6)	9.1 (1.9)*	9.6 (1.3)*	8.7 (1.9)
Neck extensors	9.8 (0.7)	9.8 (0.7)	9.8 (0.7)	9.8 (0.9)
Arm abductors	7.7 (2.8)	8.7 (2.1)**	9.3 (1.4)***	8.0 (2.1)
Elbow flexors	8.9 (1.9)	9.3 (1.2)	9.7 (0.6)***	8.9 (1.7)
Elbow extensors	8.5 (2.1)	9.0 (1.4)*	9.6 (0.9)***	8.7 (1.7)
Wrist flexors	9.8 (0.8)	9.7 (0.7)	9.9 (0.4)	9.7 (0.9)
Wrist extensors	9.6 (1.3)	9.7 (0.8)	9.9 (0.4)	9.8 (0.8)
Finger flexors	9.9 (0.6)	9.8 (0.6)	9.8 (0.6)	9.7 (0.9)
Finger extensors	9.8 (0.6)	9.6 (1.1)	9.9 (0.5)	9.8 (0.6)
Hip flexors	7.1 (2.9)	8.6 (2.1)***	8.8 (1.9)***	5.8 (3.1)**
Hip extensors	9.0 (2.3)	9.7 (1.0)**	9.8 (0.7)**	8.9 (2.3)
Knee flexors	9.5 (1.4)	9.8 (0.6)*	9.9 (0.4)	9.0 (1.8)
Knee extensors	9.6 (1.0)	9.7 (1.0)	9.8 (0.5)	9.2 (1.7)
Ankle flexors	10.0 (0.2)	9.7 (1.2)	9.9 (0.3)	9.7 (1.1)
Ankle extensors	9.9 (0.7)	9.9 (0.4)	10.0 (0.0)	9.7 (1.2)

* p<0.05, ** p<0.01, *** p<0.001

Strength values were expressed as means (SD) and bivariate comparisons were made using Student´s t-test.

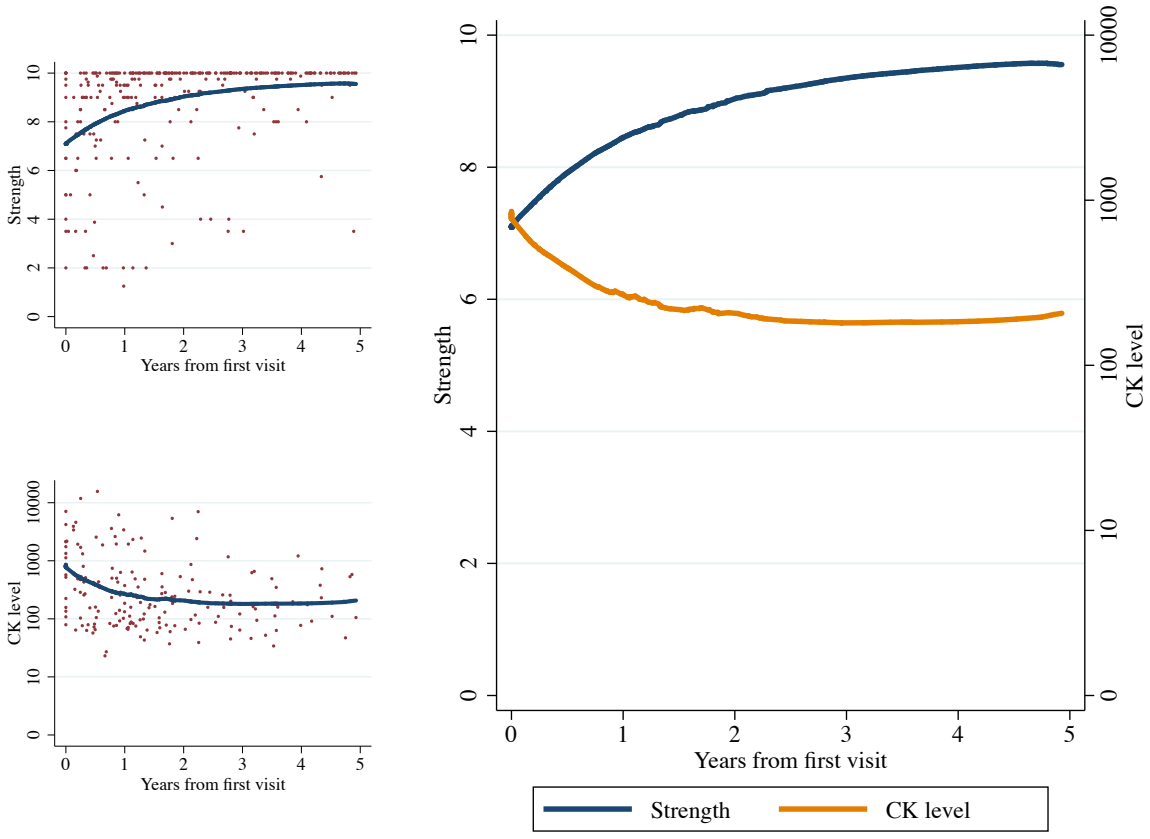
Table 6. Muscle biopsy and thigh magnetic resonance imaging results.

	Anti-Mi2 DM (n=27)	Non-Mi2 DM (n=32)	AS (n=26)	IMNM (n=81)
Muscle biopsy features				
Necrotizing myopathy	19% (5)	6% (2)	23% (6)	78% (63)***
Degenerating fibers	81% (22)	62% (20)	88% (21)	96% (77)*
Perifascicular atrophy	59% (16)	59% (19)	50% (13)	0% (0)***
Perivascular inflammation	48% (13)	69% (22)	65% (17)	26% (21)*
Primary inflammation	19% (5)	0% (0)*	31% (8)	17% (14)
Thigh MRI features				
Muscle edema	78% (18)	71% (48)	74% (40)	91% (79)
Atrophy	26% (6)	21% (14)	30% (16)	68% (59)***
Fatty replacement	39% (9)	50% (34)	56% (30)	87% (76)***
Fascial edema	43% (10)	59% (40)	67% (36)	36% (31)

* p<0.05, ** p<0.01, *** p<0.001

Chi-squared or Fisher´s exact tests were used to compare each one of the clinical groups with the anti-Mi2 patients.

Figure 1. Evolution of strength and creatine kinase levels in anti-Mi2-positive patients.



A.5 Validation of anti-Mi2 autoantibody testing by line blot. Autoimmun Rev 2020 (PMID: 31734399).

Immunoprecipitation is the gold standard for detecting anti-Mi2 autoantibodies. The objective of this study was to assess the performance of a commercially available anti-Mi2 autoantibody line blot test.

To do this we included all the patients from the Johns Hopkins Myositis Center tested for anti-Mi2α/β autoantibodies by immunoprecipitation (which tests for both anti-Mi2α and anti-Mi2β autoantibodies but does not distinguish between the two) and line blot (which tests separately for anti-Mi2α and anti-Mi2β autoantibodies). Patients who were anti-Mi2α and/or Mi2β -positive by IP and/or line blot had sera tested for anti-Mi2β autoantibodies by ELISA.

Among 666 patients, 35 (5%) were anti-Mi2α/β positive by IP. From among these, by line blot, 71% (n=25) were positive for both anti-Mi2α and anti-Mi2β, 23% (n=8) were exclusively positive for anti-Mi2α, and 6% (n=2) were exclusively positive for anti-Mi2β. False positive line blot results occurred in 5% (n=34) of patients, including 3% (n=1) for both anti-Mi2α and anti-Mi2β, 15% (n=5) for anti-Mi2α alone, and 82% (n=28) for anti-Mi2β alone. The signal intensity of anti-Mi2α and anti-Mi2β line blot tests correlated well with each other ($r=0.77$, $p<0.001$) and with anti-Mi2β ELISA titers (both $r>0.67$, both $p<0.001$).

In conclusion, a dual positive anti-Mi2α and anti-Mi2β line blot test reliably identifies anti-Mi2-positive patients. Exclusive anti-Mi2β line blot positive testing is most consistent with a false-positive result. Those who are only positive for anti-Mi2α by line blot require validation by another technique. The line blot signal intensity can be used to estimate anti-Mi2β autoantibody titers.

Validation of anti-Mi2 autoantibody testing by line blot

Iago Pinal-Fernandez[1,2,3]*, Katherine Pak[1]*, Maria Casal-Dominguez[1,2]*, Yuji Hosono[1], Christopher Mecoli[4], Lisa Christopher-Stine[2,4], and Andrew L. Mammen[1,2,4]

* These authors contributed equally to this project.

[1]Muscle Disease Unit, Laboratory of Muscle Stem Cells and Gene Regulation, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland.

[2]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland.

[3]Faculty of Health Sciences, Universitat Oberta de Catalunya, Barcelona, Spain.

[4]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland.

Address correspondence to: Andrew L. Mammen, M.D., Ph.D. or Iago Pinal-Fernandez, M.D., Ph.D. Muscle Disease Unit, Laboratory of Muscle Stem Cells and Gene Regulation, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, 50 South Drive, Room 1141, Building 50, MSC 8024, Bethesda, MD 20892. E-mail: andrew.mammen@nih.gov or iago.pinalfernandez@nih.gov. Phone: 301-451-1199. Fax: 301-594-0305.

MANUSCRIPT WORD COUNT: 1188 words

FUNDING: This research was supported by the Intramural Research Program of the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health. LCS and the Myositis Research Database are supported by the Huayi and Siuling Zhang Discovery Fund.

We are also grateful to Dr. Peter Buck, whose generous support made this work possible. IPF´s research was supported by a Fellowship from the Myositis Association.

CONFLICT OF INTEREST: No potential conflict of interest was reported by the authors.

KEYWORDS: Autoimmune diseases, muscle disease, myositis, dermatomyositis, autoantibodies.

Dermatomyositis (DM) is a subtype of inflammatory myopathy with characteristic skin and muscle involvement. A growing corpus of data suggests that DM-specific autoantibodies define subsets with unique phenotypes.[1, 2] Among the DM-specific autoantibodies, those recognizing two structurally similar proteins of the nucleosome remodeling and deacetylase (NuRD) complex, Mi2α and Mi2β, were the first to be described.[3] [4]

Immunoprecipitation (IP) is considered the gold standard for detecting anti-Mi2 autoantibodies. However, IP is expensive, labor-intensive, and takes weeks to complete. Anti-Mi2 line blots are simpler and faster to complete but the performance of commercial testing kits has not been adequately validated.[5-7] In this study, we analyzed the anti-Mi2 line blot's diagnostic value as well as its utility to estimate autoantibody titers.

All patients enrolled in the Johns Hopkins Myositis Center Longitudinal Cohort study between 2002 and 2018 with sera tested for anti-Mi2α/β autoantibodies by IP (which tests for both anti-Mi2α and anti-Mi2β autoantibodies but does not distinguish between the two) and line blot (which tests separately for anti-Mi2α and anti-Mi2β autoantibodies) were included in the study.

Anti-Mi2 IPs were performed at the Oklahoma Medical Research Foundation or the NIH Muscle Diseases Unit. EUROLine Myositis Profile 4 line blot strips (Euroimmun) processed in a EUROBlotOne device were used to detect anti-Mi2α and anti-Mi2β autoantibodies; a signal intensity of 15 units was set as the cutoff for positivity as recommended by the vendor. A quantitative anti-Mi2β ELISA was developed using anti-Mi2β antigen (Abcam, ab124864). The normality cut-off was set to 0.17 arbitrary absorbance units based on a sample of 49 control sera (mean + 3 standard deviations of the normalized absorbance).

Correlation between ELISA titers and line blot signal intensity was measured using Pearson´s r and linear regression. All statistical analyses were performed using Stata/MP 14.1. A 2-sided p-value

of 0.05 or less was considered statistically significant.

Of the 2475 patients enrolled in the Johns Hopkins Myositis Center Longitudinal Cohort (which includes patients with inclusion body myositis and other muscle diseases), 700 (28%) were tested for anti-Mi2α/β autoantibodies by IP and 666 of these had sera available and were subsequently tested for anti-Mi2α and anti-Mi2β autoantibodies by line blot. Of those patients tested by both IP and line-blot (n=666, 27% of the cohort), 35 (5%) were positive for anti-Mi2α/β autoantibodies by IP and all of these were positive for anti-Mi2α and/or anti-Mi2β autoantibodies by line blot: 71% (n=25) were positive for both anti-Mi2α and anti-Mi2β, 23% (n=8) were exclusively positive for anti-Mi2α, and 6% (n=2) were exclusively positive for anti-Mi2β autoantibodies. Of note, 34 of 631 (5%) patients that were anti-Mi2α/β negative by IP were positive (i.e., false positive) by line blot: 3% (n=1) were positive both for anti-Mi2α and anti-Mi2β, 15% (n=5) were exclusively positive for anti-Mi2α, and 82% (n=28) were exclusively positive for anti-Mi2β autoantibodies (Table 1). Of note, none of the anti-Mi2α false-positive samples had DM. In contrast, 61% (n=17) of the anti-Mi2β false-positives had DM, most of whom (65%, n=11) were positive for another myositis-specific autoantibody, including anti-Jo1, anti-MDA5, anti-Pm/Scl, anti-SAE or anti-TIF1γ.

It should be noted that in 38% of cases, serum samples tested by IP were collected on a different day than those tested by line blot. However, similar results were obtained when we performed the same analysis excluding these cases.

To determine if the anti-Mi2α and/or anti-Mi2β line blot signal intensities have quantitative value, we correlated these with results from an anti-Mi2β ELISA. The line blot signal intensities for anti-Mi2α and anti-Mi2β correlated will with anti-Mi2β ELISA titers ($r= 0.76$, $p<0.001$ and $r= 0.67$, $p<0.001$, respectively; Figure 1).

New immunologic techniques like the line blot have made testing for autoantibodies both quick and easy. However, for rare autoimmune diseases like myositis, companies may have insufficient numbers of positive and negative control samples to properly validate their assays.[5-7] Although a recent study assessed the utility of line blot compared to IP to detect myositis-specific autoantibodies[5], the small number of samples from patients with anti-Mi2 autoantibodies included in the study precluded drawing conclusions about the utility of the anti-Mi2α and anti-Mi2β line blot test.[5]

In this study, we have defined the diagnostic utility of the anti-Mi2α and anti-Mi2β line blot tests for anti-Mi2 autoantibodies. Our results show that only those subjects who are positive both for anti-Mi2α and anti-Mi2β autoantibodies by line blot are very likely to be anti-Mi2 positive by IP. Those who are positive for just for anti-Mi2β are usually false positive and those who are only positive for anti-Mi2α have only about a fifty percent chance of being true anti-Mi2 positives; further testing of the latter patients is required. These results, testing a large number of samples from patients with a rare disease, underscore the necessity for standardization of myositis autoantibody testing techniques for both clinical and research purposes.

We also demonstrate that that anti-Mi2 line blot signal intensities can be used to accurately estimate anti-Mi2 autoantibody titers. This may be clinically relevant, since anti-Mi2 autoantibody titers may be useful markers of disease activity. The fact that a readily available technique like line blot could be used to measure autoantibody titers has the potential of simplifying longitudinal autoantibody studies.

In summary, only those patients positive for both anti-Mi2α and anti-Mi2β by line blot can reliably be considered anti-Mi2 without further validation. Moreover, line blot signal intensity has quantitative value to estimate anti-Mi2 autoantibody titers.

REFERENCES

[1] O. Benveniste, W. Stenzel and Y. Allenbach. Advances in serological diagnostics of inflammatory myopathies. Curr Opin Neurol 2016;29:662-73. 10.1097/WCO.0000000000000376.

[2] A. Selva-O'Callaghan, I. Pinal-Fernandez, E. Trallero-Araguas, J.C. Milisenda, J.M. Grau-Junyent and A.L. Mammen. Classification and management of adult inflammatory myopathies. Lancet Neurol 2018;17:816-828. 10.1016/S1474-4422(18)30254-0.

[3] M. Reichlin and M. Mattioli. Description of a serological reaction characteristic of polymyositis. Clin Immunol Immunopathol 1976;5:12-20.

[4] H.P. Seelig, M. Renz, I.N. Targoff, Q. Ge and M.B. Frank. Two forms of the major antigenic protein of the dermatomyositis-specific Mi-2 autoantigen. Arthritis Rheum 1996;39:1769-71.

[5] F. Espinosa-Ortega, M. Holmqvist, H. Alexanderson, H. Storfors, T. Mimori, I.E. Lundberg, et al. Comparison of autoantibody specificities tested by a line blot assay and immunoprecipitation-based algorithm in patients with idiopathic inflammatory myopathies. Ann Rheum Dis 2019;78:858-860. 10.1136/annrheumdis-2018-214690.

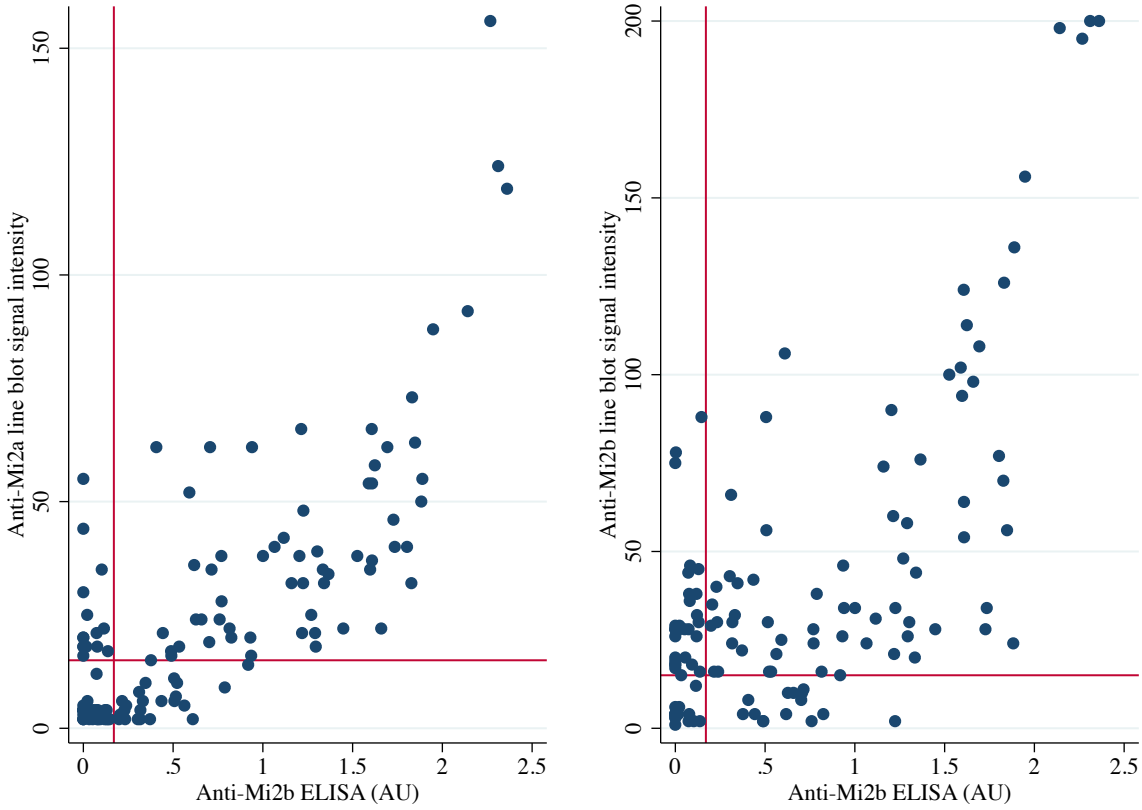
[6] J. Damoiseaux, J.B. Vulsteke, C.W. Tseng, A.C.M. Platteel, Y. Piette, O. Shovman, et al. Autoantibodies in idiopathic inflammatory myopathies: Clinical associations and laboratory evaluation by mono- and multispecific immunoassays. Autoimmun Rev 2019;18:293-305. 10.1016/j.autrev.2018.10.004.

[7] J. Ronnelid, S. Barbasso Helmers, H. Storfors, K. Grip, L. Ronnblom, K. Franck-Larsson, et al. Use of a commercial line blot assay as a screening test for autoantibodies in inflammatory myopathies. Autoimmun Rev 2009;9:58-61. 10.1016/j.autrev.2009.03.005.

Table 1. Immunoprecipitation results of anti-Mi2 positive patients by line blot.

	IP (% [n])	
	Anti-Mi2 +	**Anti-Mi2 -**
Euroimmune		
Mi2α+Mi2β	71% (25)	3% (1)
Mi2α only	23% (8)	15% (5)
Mi2β only	6% (2)	82% (28)
Total	35 (100%)	34 (100%)

Figure 1. Good correlation of the anti-Mi2β quantitative ELISA with the signal intensities of anti-Mi2α (left panel, r= 0.76, p<0.001) and anti-Mi2β (right panel, r= 0.67, p<0.001) detected by line blot. The red lines indicate the normal cutoff values for signal intensity of the line blot and for absorbance units (AU) of the ELISA.



A.6 In adults with myositis-specific autoantibodies, autoantibodies outperform the 2017 EULAR/ACR classification criteria to define phenotypes. Submitted.

The objective of this study was to evaluate the sensitivity of the 2017 EULAR/ACR criteria to classify inflammatory myopathy (IM) patients with MSAs and to compare the performance of autoantibodies with the EULAR/ACR classification to predict the clinical phenotype of MSAs-positive patients.

For this, we included 524 MSAs-positive IM patients from the Johns Hopkins Myositis Center. Each patient was classified using the EULAR/ACR classification criteria. The evolution of muscle strength and creatine kinase levels was studied using LOWESS and patient phenotypes were summarized using factor analysis of mixed data. We compared the ability of MSAs and the EULAR/ACR subgroups to predict the phenotype of patients by applying the AIC and the BIC to the linear regression models.

Nine percent of MSAs-positive patients did not meet EULAR/ACR criteria to be classified as inflammatory myopathy. Anti-HMGCR (20%), anti-SRP (9%), anti-MDA5 (11%), anti-PL12 (10%) and anti-PL7 (50%) patients had the highest failure rates. Around 10% of anti-SRP and anti-HMGCR patients were misclassified as IBM using EULAR/ACR criteria. LOWESS showed a characteristic evolution of the muscle involvement in each MSAs group and factor analysis of mixed data demonstrated that patients within each MSAs group had similar phenotypes. Application of both the AIC and BIC to the linear regression models revealed that MSAs better predict myositis phenotypes than the subgroups defined by the EULAR/ACR criteria.

In conclusion, MSAs outperform the 2017 EULAR/ACR classification to predict the clinical phenotypes of myositis patients. Thus, we propose using MSAs to build phenotypically homogeneous groups in myositis research.

In adults with myositis-specific autoantibodies, autoantibodies outperform the 2017 EULAR/ACR classification criteria to define phenotypes

Maria Casal-Dominguez M.D., Ph.D.[1,2,*], Iago Pinal-Fernandez M.D., Ph.D.[1,2,3,*], Katherine Pak M.D.[1], Wilson Huang M.D.[1], Jemima Albayda M.D. [2], Livia Casciola-Rosen Ph.D. [2], Julie J. Paik, M.D.[2], Eleni Tiniakou, M.D. [2], Christopher A. Mecoli, M.D., M.H.S. [2], Thomas E. Lloyd, M.D., Ph.D.[2], Sonye K. Danoff, M.D., Ph.D.[2,#], Lisa Christopher-Stine, M.D., M.P.H.[2,#], and Andrew L. Mammen M.D., Ph.D.[1,2,#]

*,# These authors contributed equally to this project.

[1]Muscle Disease Unit, Laboratory of Muscle Stem Cells and Gene Regulation, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland.

[2]Johns Hopkins University School of Medicine, Baltimore, Maryland.

[3]Faculty of Health Sciences and Faculty of Computer Science, Multimedia and Telecommunications, Universitat Oberta de Catalunya, Barcelona, Spain.

Address correspondence to: Andrew L. Mammen, M.D., Ph.D., or Iago Pinal-Fernandez, M.D., Ph.D. Muscle Disease Unit, Laboratory of Muscle Stem Cells and Gene Regulation, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, 50 South Drive, Room 1141, Building 50, MSC 8024, Bethesda, MD 20892. E-mail: andrew.mammen@nih.gov or iago.pinalfernandez@nih.gov. Phone: 301-451-1199. Fax: 301-594-0305.

MANUSCRIPT WORD COUNT: 2993

KEYWORDS: Autoimmune diseases, muscle disease, myositis, autoantibodies, classification.

FUNDING: This research was supported by the Intramural Research Program of the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health. The Rheumatic Diseases Research Core Center, where some of the autoantibodies were assayed, is supported by NIH P30-AR070254. We are grateful to Dr. Peter Buck, whose generous support made this work possible. IPF's research was supported by a Fellowship from the Myositis Association. ET was supported by Jerome L. Greene Foundation.

ACKNOWLEDGMENTS: This work is dedicated to the memory of Dr. Lori A. Love.

CONFLICT OF INTERESTS: No potential conflict of interest was reported by the authors.

KEY MESSAGES:

1.- Each myositis-specific autoantibody defines a unique group of myositis patients with characteristic muscle and extramuscular features.

2.- Myositis patients with myositis-specific autoantibodies are often misclassified by the 2017 EULAR/ACR classification criteria.

3.- Myositis-specific autoantibodies outperform the 2017 EULAR/ACR classification to predict the clinical phenotypes of myositis patients.

4.- Myositis-specific autoantibodies can be used to build phenotypically homogeneous groups for myositis research.

ABSTRACT

Objective: To evaluate the sensitivity of the 2017 EULAR/ACR criteria to classify inflammatory myopathy (IM) patients with myositis-specific autoantibodies (MSA) and to compare the performance of autoantibodies with the EULAR/ACR classification to predict clinical phenotype of MSA-positive patients.

Methods: This study included 524 MSA-positive IM patients from the Johns Hopkins Myositis Center. Each patient was classified using the EULAR/ACR classification criteria. The evolution of muscle strength and creatine kinase levels was studied using locally weighted polynomial regression (LOWESS) and patient phenotypes were summarized using factor analysis of mixed data (FAMD). We compared the ability of MSAs and the EULAR/ACR subgroups to predict the phenotype of patients by applying the Akaike information criterion (AIC) and the Bayesian information criteria (BIC) to the linear regression models.

Results: Nine percent of MSA-positive patients did not meet EULAR/ACR criteria to be classified as IM. Anti-HMGCR (20%), anti-SRP (9%), anti-MDA5 (11%), anti-PL12 (10%) and anti-PL7 (50%) patients had the highest failure rates. Around 10% of anti-SRP and anti-HMGCR patients were misclassified as IBM using EULAR/ACR criteria. LOWESS showed a characteristic evolution of the muscle involvement in each MSA group and FAMD demonstrated that patients within each MSA group had similar phenotypes. Application of both the AIC and BIC to the linear regression models revealed that MSAs better predict IM phenotypes than the subgroups defined by the EULAR/ACR criteria.

Conclusions: MSAs outperform the 2017 EULAR/ACR classification to predict the clinical phenotypes of IM patients. Thus, we propose using MSAs to build phenotypically homogeneous groups in myositis research.

INTRODUCTION

The inflammatory myopathies (IM) are a heterogeneous family of diseases that affect not only skeletal muscle, but often the skin, lungs, and/or joints.[1] Given their marked heterogeneity, classification criteria are needed to create well-defined, relatively homogeneous cohorts for clinical research and clinical trials.[2]

Importantly, myositis-specific autoantibodies (MSAs) are found in approximately 70% of IM patients without IBM, are not found in patients with other rheumatic or neuromuscular diseases, and are associated with unique clinical phenotypes.[3] Specifically, autoantibodies recognizing Mi2, TIF1γ, NXP2, or MDA5 are present in the group of IM patients with the hallmark cutaneous features of dermatomyositis (DM); autoantibodies recognizing the tRNA synthetases including Jo1, PL7, and PL12 are found in the group of IM patients with myositis, interstitial lung disease (ILD), and/or arthritis (i.e., the antisynthetase syndrome [AS]); and autoantibodies recognizing SRP or HMGCR are found in the group of patients with skeletal muscle-predominant IM characterized by necrotizing muscle biopsies (i.e., immune-mediated necrotizing myopathy [IMNM]). Furthermore, accumulating evidence suggests that individual MSAs define phenotypically distinct IM subgroups. For example, anti-MDA5 and anti-Mi2 are each DM-specific MSAs found in patients with characteristic DM rashes. However, whereas anti-MDA5-positive DM patients are likely to have severe ILD with minimal or no muscle involvement, anti-Mi2-positive DM patients rarely have ILD but often have severe myositis.

In 2017, new EULAR/ACR myositis criteria proposed classifying myositis based on a set of epidemiologic, clinical, and laboratory variables.[4] Importantly, among all known MSAs, only anti-Jo1 autoantibodies were included in the weighted score used to classify

patients as IM. Those adult patients classified as IM could be further subclassified in four different categories: polymyositis (PM)/immune-mediated necrotizing myositis (IMNM), inclusion body myositis (IBM), amyopathic dermatomyositis (ADM), and dermatomyositis (DM). However, because serologic tests for MSAs other than anti-Jo1 were not available when they were developed, the EULAR/ACR criteria did not utilize these for IM subclassification.[4]

In this study, we determined the characteristic phenotype and muscle involvement evolution of different MSA-positive IM patients. Moreover, we studied the sensitivity of the EULAR/ACR criteria to classify MSA-positive patients as IM, and compared the ability of the autoantibodies and the EULAR/ACR subgroups to predict the clinical phenotype of the patients. Based on our findings, we propose an alternative classification approach that uses MSAs to more precisely define homogeneous subsets of IM.

METHODS

Adult myositis patients and sera

Adult myositis patients enrolled in the Johns Hopkins Myositis Center Longitudinal Cohort study between 2002 to 2018 were included in the study if they were positive for autoantibodies recognizing Mi2, NXP2, TIF1γ, MDA5, Jo1, PL7, PL12, SRP, or HMGCR by at least two immunologic techniques from among the following: ELISA, immunoprecipitation of proteins generated by *in vitro* transcription and translation (IVTT-IP), line blotting (EUROLINE myositis profile), or immunoprecipitation from ^{35}S-methionine-labeled HeLa cell lysates.[5,6] Autoantibody groups with a prevalence of less than 2% (e.g. anti-OJ, anti-EJ) were excluded. The demographics, clinical, and laboratory features were collected prospectively at each visit as previously described (Supplementary Methods).[6]

Standard protocol approvals and patient consents. This study was approved by the Johns Hopkins Institutional Review Boards; written informed consent was obtained from each participant.

Analysis

Factor analysis of mixed data (Supplementary Methods) was used to summarize the phenotype of the patients based on the gender, race, age at onset, maximum creatine kinase (CK), presence of anti-Ro52 autoantibodies, and presence or absence during the course of the disease of muscle weakness, interstitial lung disease, arthritis, heliotrope or Gottron's rash, calcinosis, Raynaud's phenomenon, mechanic's hands, dysphagia, and

fevers. All the factors explaining more than 10% of the variance were retained for further analysis.

We compared the ability of MSA autoantibodies and EULAR/ACR subgroups to predict the resulting phenotypes using the Akaike information criterion (AIC)[7] and the Bayesian information criterion (BIC) of the linear regression (Supplementary Methods).[8] Both AIC and BIC are measures of quality of a statistical model, the lower the AIC or the BIC the better the variables predict the dependent variable (the factor-analysis-derived phenotype in this case). As a guideline, a model with 10 points lower AIC[9] or BIC[10] definitively fits the data better. Of note, both AIC and BIC penalize the number of parameters (it is harder to get low values with more parameters) and there were more categories in the autoantibodies than in the EULAR/ACR subgroup classification.

Dichotomous variables were expressed as absolute frequencies and percentages, and continuous variables were reported as means and standard deviations. Locally weighted polynomial regression (LOWESS) was used to study the evolution of the strength and the CK among autoantibodies (Supplementary Methods). Comparisons between groups were made using Chi-squared, Fisher's exact tests, or Student's t-test. FactoMineR v.2.1 was used to compute the factor analysis of mixed data and Stata 14 MP was used to perform the univariate, graphical, and regression analysis. A two-sided P value of <0.05 was considered significant with no correction for multiple comparisons.

RESULTS

Demographics and clinical characteristics of MSA-positive IM patients

The study included 524 myositis patients consecutively evaluated at the Johns Hopkins Myositis Center who were positive for a single MSA by at least two separate techniques. Among these, 24% had anti-Jo1 (n=127), 23% had anti-HMGCR (n=122), 12% had anti-TIF1γ (n=62), 11% had anti-Mi2 (n=59), 10% had anti-NXP2 (n=50), 8% had anti-SRP (n=44), 5% had anti-MDA5 (n=28), 4% had anti-PL12 (n=20), and 2% had anti-PL7 (n=12) autoantibodies.

The demographic and clinical features of these patients are provided in Table 1. Of note, the phenotypes associated with each MSA are consistent with prior reports. Thus, patients with autoantibodies against SRP and HMGCR are uniformly weak and rarely have arthritis, skin involvement, or ILD.[11,12] Among patients with autoantibodies against Mi2, NXP2, TIF1γ, and MDA5, almost all have DM-specific skin rashes, most are weak, and the presence of ILD and/or arthritis is uncommon except among those with anti-MDA5 autoantibodies.[13-16] Finally, the majority of patients with an antisynthetase autoantibody have weakness, ILD, skin involvement, and/or arthritis.[6]

Patients with different MSAs demonstrate distinct patterns of muscle involvement

Overall, muscle weakness with or without CK elevations was present in greater than 90% of the IM patients and was the most common clinical feature shared between each of the MSA-defined subgroups. To determine whether the severity of weakness and serum CK elevations varied among patients with different MSAs, we analyzed these features over time in each MSA-defined subgroup using locally weighted polynomial

regression. This revealed 5 distinct patterns of muscle involvement among the patients with the different MSAs (Figure 1).

First, patients with autoantibodies against Mi2 and NXP2 are initially moderately weak with moderately high CK levels; over time, they recover strength and normalize their CK levels. Second, those with autoantibodies against TIF1γ, MDA5, and PL12 have minimal weakness and normal CK levels throughout the course of the disease. Third, anti-Jo1-positive and anti-PL7-positive patients present with mild weakness and severely elevated CK levels, both of which gradually normalize over time. Fourth, anti-HMGCR-positive patients present with moderate weakness and severely elevated CK levels that improve but do not normalize, over time. And fifth, anti-SRP patients present with the most severe weakness and severely elevated CK levels; even after 3 years of follow-up, these patients continue to have moderately severe weakness and moderately elevated CK levels.

The same analysis performed on the subgroups defined by the 2017 EULAR/ACR classification criteria did not show apparent differences between patients classified as PM/IMNM and those classified as IBM. As expected, DM patients showed more severe muscle weakness and elevated CK levels than ADM patients and less than PM/IMNM and IBM (Supplementary Figure 1). This graphical analysis suggests that a significant amount of information on relevant myositis subgroups may be lost by using only the four adult classification subgroups in the 2017 EULAR/ACR IM criteria.

Factor analysis of mixed data shows that MSAs cluster IM patients into phenotypically similar subgroups

The above graphical analysis does not allow for a direct quantitative comparison between the autoantibodies and the 2017 EULAR/ACR classification subgroups. Moreover, although the graphical strength analysis suggests that different MSAs are associated with different patterns of muscle disease, it is not possible to compare the full phenotype associated with one MSA to that of another MSA by analyzing the strength, or any other single IM disease manifestation, in isolation. Therefore, we used factor analysis of mixed data (FAMD) to mathematically summarize the clinical phenotype of each patient. Using this technique, we were able to include both quantitative features (i.e., age at onset, maximum CK) and qualitative features (i.e., gender, race, and presence or absence of weakness, ILD, arthritis, heliotrope or Gottron's rash, calcinosis, Raynaud's phenomenon, mechanic's hands, dysphagia, and fevers) to model the overall phenotype of each IM patient included in this study.

Two principal component factors accounted for 18% (factor 1) and 11% (factor 2) of the clinical variability amongst all of the patients. The most heavily weighted factors in the first component (factor 1) were the presence of mechanic's hands, ILD, and arthritis. For factor 2, the most important parameters were maximum CK, race, and the presence of either heliotrope or Gottron's rashes (Supplementary Figure 2). Figure 2 depicts the value of both factors for each patient in two dimensions and demonstrates that those with the same MSA tend to cluster, indicating that they share a similar phenotype.

Although there was an overlap between the different MSAs, there were unique locations occupied by the 9 different MSAs in the FAMD graphs (Figure 2). Patients with autoantibodies against HMGCR and SRP clustered together in the lower right part of the graph, whereas those with anti-synthetase autoantibodies clustered in the upper middle

section of the graph. Patients with autoantibodies against Mi2, NXP2, and TIF1γ clustered lower than those with anti-synthetase autoantibodies, with increasing leftward shift comparing those with anti-Mi2 to anti-NXP2 to anti-TIF1γ autoantibodies. Anti-MDA5-positive patients occupied the upper left portion of the graph.

When the factor analysis distribution was performed using the 2017 EULAR/ACR subgroup classification of the patients, those classified as PM/IMNM and IBM shared a similar region of the graph. Likewise, patients classified as ADM overlapped with those classified as DM (Supplementary Figure 3). This graphical analysis also suggested that a significant amount of information on relevant myositis subgroups may be lost by using only the four adult classification subgroups in the 2017 EULAR/ACR IM criteria.

MSAs outperform the 2017 EULAR/ACR IM classification criteria to predict clinical phenotypes

Both MSAs and the 2017 EULAR/ACR IM classification subgroups were used to predict the factor-analysis-derived phenotypes of the IM patients using linear regression. In order to compare the relative quality of the models using the autoantibodies and the 2017 EULAR/ACR IM classification subgroups, we first used the Akaike information criterion to compare how well each model accounted for factor 1 and factor 2 from the FAMD. Both factor 1 and factor 2 were better modeled by MSAs than the EULAR/ACR classification subgroups with ΔAICs of 439 and 120, respectively. Similarly, using the Bayesian information criterion, MSAs outperformed the 2017 EULAR/ACR classification subgroups in modeling factor 1 and factor 2 with ΔBICs of 421 and 104, respectively (Supplementary Table 1). Taken together, these analyses provide quantitative evidence

that MSAs outperform the 2017 EULAR/ACR classification subgroups for predicting the clinical phenotype of IM patients.

2017 EULAR/ACR IM classification criteria fail to recognize certain autoantibody groups

In order to clarify the nature of discrepancies between the MSA-based and 2017 EULAR/ACR classification schemes, we explored how IM patients with each MSA were classified by the 2017 EULAR/ACR criteria (Table 2, Supplementary Tables 2-10). Overall, among MSA-positive IM patients, the 2017 EULAR/ACR criteria classified 51% as DM, 33% as PM/IMNM, and 4% as amyopathic DM. Surprisingly, 4% were classified by 2017 EULAR/ACR as IBM even though they all had MSAs, which are not found in IBM (Table 2). Patients misclassified as IBM were severely affected anti-HMGCR or anti-SRP patients who had either finger flexor weakness and were refractory to treatment (4 anti-HMGCR and 1 anti-SRP), and/or had muscle biopsies with rimmed vacuoles (11 of anti-HMGCR and 3 anti-SRP), which is characteristic of IBM but can also occur in other myopathic conditions.

Overall, the 2017 EULAR/ACR criteria misclassified 9% of MSA-positive IM patients as “not myositis”. Remarkably, this included 20% of anti-HMGCR-positive and 9% of anti-SRP-positive patients, all of whom also had proximal muscle weakness and/or elevated CK levels that, along with the autoantibodies, define this disease (Table 2). Also, 11% of anti-MDA5, 10% anti-PL12, and 50% anti-PL7 failed to meet the 2017 EULAR/ACR classification criteria (Table 2).

Among anti-Jo1-positive patients, the 2017 EULAR/ACR criteria classified 55% as DM and 43% as PM/IMNM. However, it is worth noting that there were no significant differences in the proportion of patients with ILD, muscle weakness, arthritis, calcinosis, or Raynaud's phenomenon or in the median proximal muscle strength or CK levels between those classified as DM and those classified as PM. Also, anti-HMGCR patients that did not fulfill the EULAR/ACR criteria and those that were classified as PM/IMNM had equivalent strength and CK levels. Similarly, anti-TIF1γ patients classified as DM and ADM had a similar prevalence of arthritis or skin involvement (Supplementary Tables 2-10).

Importantly, although the 2017 EULAR/ACR classification scheme requires a muscle biopsy for classification in patients with no skin rash, we applied the same criteria to all patients, regardless of the availability of a muscle biopsy. Notwithstanding this, we repeated the analysis excluding all the patients with no skin rash and no biopsy available and all conclusions of our analysis were still valid (Supplementary Tables 11-20, Supplementary Figure 4).

Table 1. Epidemiologic and clinical variables of the patients.

	Anti-SRP (n=44)	Anti-HMGCR (n=122)	Anti-Mi2 (n=59)	Anti-NXP2 (n=50)	Anti-TIF1g (n=62)	Anti-MDA5 (n=28)	Anti-Jo1 (n=127)	Anti-PL7 (n=12)	Anti-PL12 (n=20)	Total (n=524)
Female sex	70% (31)	60% (73)**	63% (37)	70% (35)	87% (54)**	68% (19)	72% (91)	58% (7)	80% (16)	69% (363)
Race										
White	50% (22)**	75% (91)	64% (38)	76% (38)	89% (55)***	64% (18)	66% (84)	25% (3)**	25% (5)***	68% (354)
Black	43% (19)***	17% (21)	15% (9)	14% (7)	6% (4)**	25% (7)	22% (28)	58% (7)**	75% (15)***	22% (117)
Other races	7% (3)	8% (10)	20% (12)**	10% (5)	5% (3)	11% (3)	12% (15)	17% (2)	0% (0)	10% (53)
Age of onset (years)	42.4 (14.6)**	54.2 (13.3)***	48.5 (15.1)	47.3 (16.8)	49.0 (14.5)	44.5 (11.0)	46.3 (12.9)	48.8 (12.4)	43.4 (12.5)	48.3 (14.3)
Cancer associated myositis	5% (2)	4% (5)	10% (6)	10% (5)	11% (7)	4% (1)	1% (1)**	8% (1)	10% (2)	6% (30)
Death during follow-up	2% (1)	3% (4)	3% (2)	4% (2)	6% (4)	4% (1)	6% (8)	17% (2)	20% (4)*	5% (28)
Anti-Ro52	41% (18)	15% (18)***	17% (10)***	16% (8)***	21% (13)**	36% (10)	82% (104)***	67% (8)	85% (17)***	39% (206)
Time of follow-up (years)	4.0 (3.1)	3.8 (3.9)	4.2 (3.6)	3.8 (3.4)	4.7 (3.3)	4.7 (4.3)	4.6 (3.9)	3.4 (3.3)	5.8 (4.0)	4.3 (3.7)
Number of visits per participant	11.4 (11.8)	8.2 (7.5)*	8.4 (6.7)	9.2 (7.2)	9.7 (6.4)	12.4 (7.6)*	9.1 (6.7)	10.1 (7.5)	13.6 (6.5)*	9.5 (7.6)
Treatments										
Corticosteroids	82% (36)	71% (87)***	88% (52)	86% (43)	74% (46)*	89% (25)	95% (121)***	83% (10)	95% (19)	84% (439)
Azathioprine	34% (15)	23% (28)**	27% (16)	26% (13)	21% (13)*	50% (14)	50% (64)***	42% (5)	70% (14)***	35% (182)
Methotrexate	59% (26)	50% (61)	58% (34)	60% (30)	55% (34)	25% (7)**	47% (60)	8% (1)**	30% (6)	49% (259)
Mycophenolate	32% (14)	14% (17)***	32% (19)	30% (15)	44% (27)*	39% (11)	39% (50)*	75% (9)**	25% (5)	32% (167)
Intravenous immunoglobulin	43% (19)	42% (51)	46% (27)	48% (24)	55% (34)*	32% (9)	37% (47)	25% (3)	35% (7)	42% (221)
Rituximab	50% (22)***	15% (18)*	19% (11)	14% (7)	16% (10)	18% (5)	22% (28)	33% (4)	30% (6)	21% (111)
ACR/EULAR classificatory criteria										
Upper extremity proximal weakness	95% (42)*	95% (116)***	92% (54)	92% (46)	73% (45)**	64% (18)**	80% (102)	75% (9)	75% (15)	85% (447)
Lower extremity proximal weakness	95% (42)	97% (118)*	98% (58)*	94% (47)	77% (48)***	71% (20)**	91% (116)	92% (11)	85% (17)	91% (477)
Neck flexor greater than neck extensor weakness	61% (27)*	54% (66)**	41% (24)	64% (32)**	47% (29)	18% (5)**	32% (41)**	17% (2)	20% (4)*	44% (230)
Proximal greater than distal lower extremity weakness	91% (40)**	90% (110)***	69% (41)	78% (39)	63% (39)*	50% (14)**	63% (80)**	50% (6)	75% (15)	73% (384)
Heliotrope rash	2% (1)***	4% (5)***	51% (30)*	56% (28)**	79% (49)***	79% (22)***	30% (38)	50% (6)	45% (9)	36% (188)
Gottron's papules	7% (3)***	4% (5)***	54% (32)**	44% (22)	77% (48)***	82% (23)***	33% (42)	17% (2)	45% (9)	35% (186)
Gottron's sign	5% (2)***	4% (5)***	51% (30)**	48% (24)*	69% (43)***	64% (18)***	31% (39)	17% (2)	40% (8)	33% (171)
Dysphagia or esophageal dysmotility	52% (23)	43% (53)	53% (31)	68% (34)**	50% (31)	36% (10)	37% (47)*	42% (5)	35% (7)	46% (241)
Elevated muscle enzymes	100% (44)**	100% (122)***	98% (58)**	92% (46)	55% (34)***	32% (9)***	89% (113)	83% (10)	65% (13)*	86% (449)
Muscle biopsy available	59% (26)**	59% (72)***	46% (27)	50% (25)	23% (14)**	14% (4)**	28% (35)**	17% (2)	15% (3)*	40% (208)
Endomysisial infiltration surrounding myofibers	8% (2)	24% (17)	37% (10)	4% (1)*	7% (1)	0% (0)	46% (16)***	0% (0)	0% (0)	23% (47)
Perifascicular atrophy	4% (1)***	3% (2)***	59% (16)**	60% (15)**	57% (8)	50% (2)	63% (22)***	50% (1)	0% (0)	32% (67)
Perimysial and/or perivascular infiltration	31% (8)*	38% (27)***	63% (17)	72% (18)*	79% (11)	25% (1)	77% (27)**	50% (1)	33% (1)	53% (111)
Rimmed vacuoles	12% (3)	15% (11)*	15% (4)	0% (0)	0% (0)	0% (0)	3% (1)	0% (0)	0% (0)	9% (19)
Other relevant clinical variables										
Interstitial lung disease	18% (8)*	2% (2)***	5% (3)***	6% (3)***	0% (0)***	71% (20)***	78% (99)***	100% (12)***	85% (17)***	31% (164)
Arthritis	7% (3)***	6% (7)***	20% (12)	14% (7)*	10% (6)***	71% (20)***	61% (77)***	33% (4)	65% (13)***	28% (149)
DM-specific skin involvement	9% (4)***	8% (10)***	92% (54)***	86% (43)***	100% (62)***	100% (28)***	56% (71)	50% (6)	70% (14)	56% (292)
Calcinosis	0% (0)**	1% (1)***	8% (5)	36% (18)***	13% (8)	46% (13)***	9% (12)	0% (0)	0% (0)	11% (57)
Raynaud's phenomenon	39% (17)*	7% (9)***	31% (18)	10% (5)**	19% (12)	54% (15)***	36% (46)**	33% (4)	30% (6)	25% (132)
Mechanic's hands	5% (2)***	4% (5)***	22% (13)	12% (6)**	26% (16)	75% (21)***	53% (67)***	67% (8)**	75% (15)***	29% (153)
Fever	9% (4)	7% (8)**	7% (4)	18% (9)	10% (6)	46% (13)***	19% (24)	17% (2)	50% (10)***	15% (80)

** p<0.05, ** p<0.01, *** p<0.001*

Chi-squared or Fisher´s exact tests were used to compare each group with the rest.

Table 2. ACR/EULAR criteria of autoantibody-positive adult myositis patients.

	Anti-SRP (n=44)	Anti-HMGCR (n=122)	Anti-Mi2 (n=59)	Anti-NXP2 (n=50)	Anti-TIF1g (n=62)	Anti-MDA5 (n=28)	Anti-Jo1 (n=127)	Anti-PL7 (n=12)	Anti-PL12 (n=20)	Total (n=524)
Negative	9% (4)	20% (25)***	5% (3)	4% (2)	2% (1)*	11% (3)	1% (1)***	50% (6)***	10% (2)	9% (47)
Subgroup										
PM/IMNM	73% (32)***	61% (74)***	3% (2)***	12% (6)***	0% (0)***	0% (0)***	43% (54)**	0% (0)*	20% (4)	33% (172)
IBM	9% (4)	11% (13)***	2% (1)	0% (0)	0% (0)	0% (0)	1% (1)	0% (0)	0% (0)	4% (19)
ADM	0% (0)	0% (0)*	0% (0)	2% (1)	18% (11)***	14% (4)*	1% (1)	8% (1)	5% (1)	4% (19)
DM	9% (4)***	8% (10)***	90% (53)***	82% (41)***	81% (50)***	75% (21)**	55% (70)	42% (5)	65% (13)	51% (267)
JM	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)
JDM	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)

** p<0.05, ** p<0.01, *** p<0.001*

Chi-squared or Fisher´s exact tests were used to compare each group with the rest. PM/IMNM: Polymyositis/immune-mediated necrotizing myositis; IBM: inclusion body myositis; ADM: amyopathic dermatomyositis; DM: dermatomyositis; JM: juvenile myositis other than DM; JDM: juvenile dermatomyositis.

Figure 1. Evolution of the strength (blue line) and creatine kinase (orange line) using locally weighted polynomial regression (LOWESS) of patients with different myositis-specific autoantibodies over time.

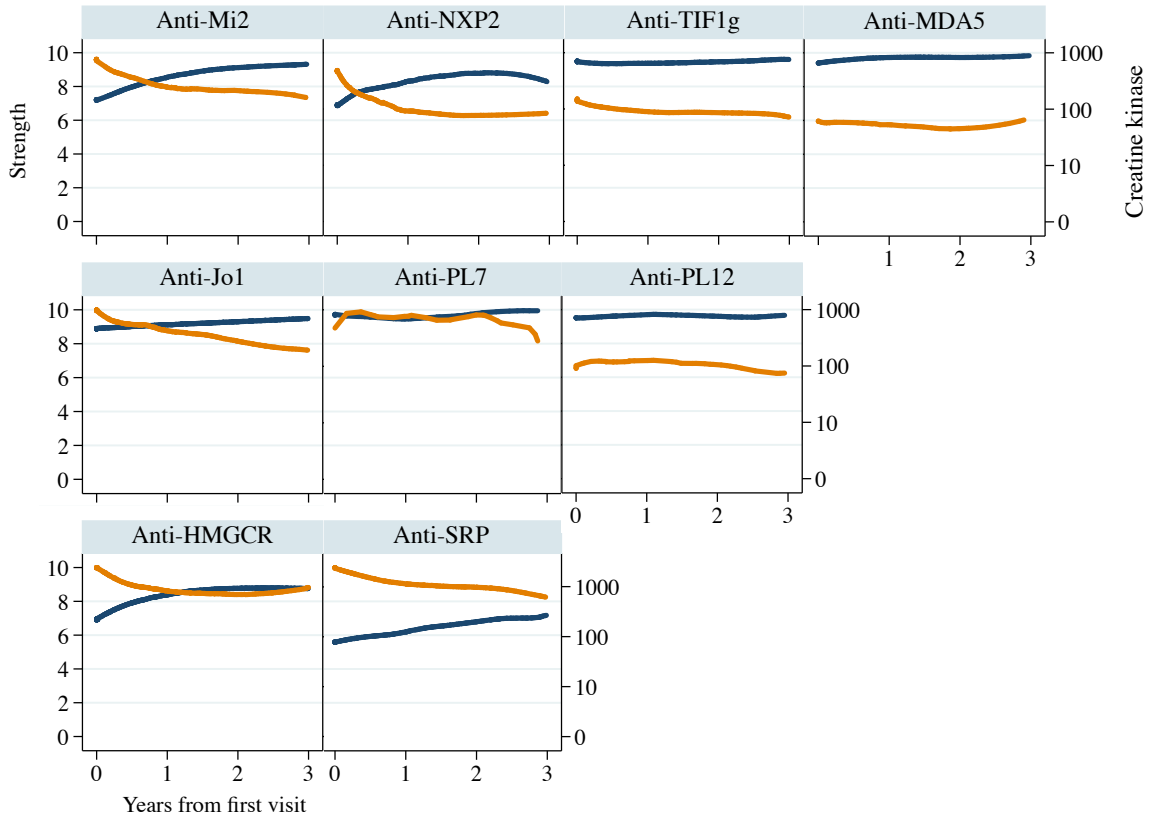
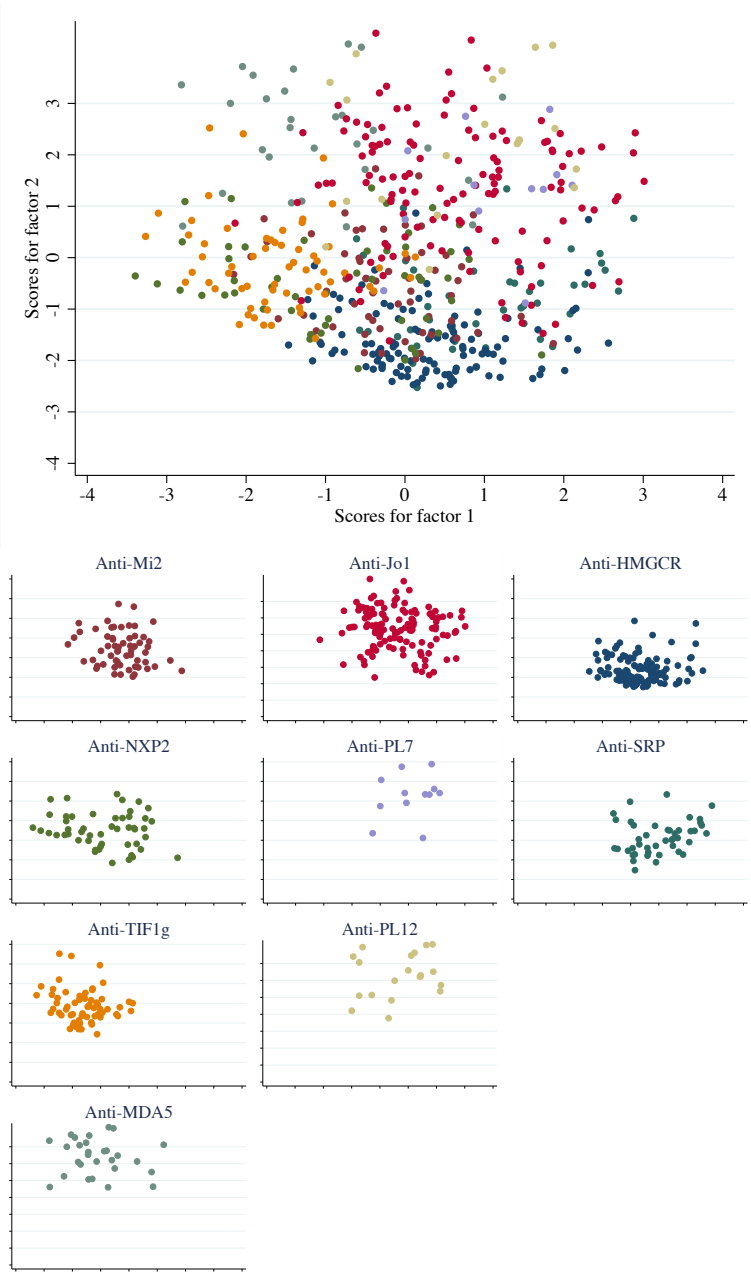


Figure 2. Factor analysis of mixed data summarizing the clinical phenotype of the autoantibody-positive adult myositis patients.



The factor analysis of mixed data included the gender, race, age at onset, median and maximum CK, presence of anti-Ro52 autoantibodies, and presence or absence during the course of the disease of: muscle weakness, interstitial lung disease, arthritis, heliotrope or Gottron's rash, calcinosis, Raynaud's phenomenon, mechanics hands, dysphagia and fevers. The only two factors explaining more than 10% of the variance (factor 1 and factor 2) were retained for further analysis.

DISCUSSION

The purpose of an IM classification scheme is to divide this heterogeneous collection of patients into phenotypically homogeneous groups. In the current study, which included over 500 MSA-positive IM patients, we show that MSAs are superior to the 2017 EULAR/ACR criteria to predict the phenotype of MSA-positive myositis patients. First, we provided graphical evidence that MSAs predict the course of muscle involvement as shown by the severity of weakness and muscle enzyme levels over the course of the disease. Second, we utilized FAMD to show that MSAs group individual IM patients into relatively homogeneous groups. Finally, we used the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), two closely related analytical tools based on information theory, to show that MSAs outperform the 2017 EULAR/ACR criteria to predict the phenotype of MSA-positive myositis patients. Taken together, these analyses validate the utility of using the MSA to classify patients with myositis in phenotypically homogeneous groups.

The utility of autoantibodies to classify patients with myositis into highly homogeneous clinical groups was first proposed by Love et al in 1991[17]. More recently, and in line with the new findings in the current manuscript, working groups of the European Neuromuscular Centre have proposed to classify patients with IMNM[18] or DM[19] based on the presence of MSAs and clinical features compatible with muscle or skin involvement.

Although the current study shows only that MSAs predict distinct clinical manifestations (e.g., weakness, rash, arthritis, and ILD), prior studies support the underlying hypothesis that MSAs define unique pathological states. For example, the

concept that each DM-specific MSA (i.e., anti-Mi2, -TIF1γ, -NXP2, and -MDA5) defines a unique DM subtype is supported by the observations that patients with different DM-specific MSAs have distinct histological features on muscle biopsy[20] and different gene expression profiles on transcriptomic analysis.[21] Similarly, evidence that anti-Jo1 autoantibodies define a single syndrome distinct from DM (i.e., AS) is supported by our observation that anti-Jo1-positive patients with and without rashes have otherwise indistinguishable clinical features including myositis, ILD, and arthritis. This is supported further by the observations that (i) the histological features of anti-Jo1 muscle biopsies are the same whether they come from a patient with or without a rash[20] and (ii) gene expression profiles from AS patients are homogeneous and easily distinguishable from those with DM.[21]

Based on our results, we propose using MSAs to inform patient selection for assembling IM cohorts consisting of the most phenotypically and clinically homogeneous groups. Such cohorts will optimize clinical research efforts as well as clinical trials. Since all our MSA-positive patients showed either muscle involvement (muscle weakness or CK elevation), interstitial lung disease, arthritis or characteristic skin involvement (either Gottron's papules, Gottron's sign or heliotrope), we propose that a MSA-positive patient showing any of these clinical features could be classified in their corresponding autoantibody category (Table 3). Thus, an anti-HMGCR patient with weakness would be classified as anti-HMGCR-related disease, an anti-Jo1 patient with ILD as anti-Jo1-related disease, and an anti-MDA5 patient with heliotrope as anti-MDA5-related disease. Our proposal is both sensitive and specific (both 100%) in our MSA-positive cohort. Validation of our findings in other cohorts will be an important next step.

Table 3. Proposal of myositis classification based on myositis-specific autoantibodies.

Myositis-specific autoantibody	+	Muscle weakness or Creatine kinase elevation or Interstitial lung disease or Arthritis or Gottron's sign or papules or Heliotrope

Using MSAs to subclassify IM has several limitations. First, IBM patients cannot be classified using this system because they do not have MSAs. Fortunately, since there are other well-accepted stand-alone schemes for classifying IBM patients (e.g., those of Griggs or Lloyd), this does not represent a practical limitation. Second, although ~30% of non-IBM IM patients do not have a known MSA, these antibody-negative groups may include a heterogeneous mixture of patients with different diseases that require better definition, as suggested by the ENMC working groups and others. Caution should, therefore, be used when including such populations of patients in research studies or clinical trials. However, for investigations where phenotypic homogeneity is of lesser importance (e.g. quality of life, evaluation of diagnostic techniques), the current 2017 EULAR/ACR criteria could be still used to decide if a patient meets criteria for myositis. Finally, in order for MSAs to be universally used for classification, “gold-standard” MSA detection methods will have to be defined.

In summary, we have used a large cohort of IM patients to demonstrate the utility MSAs to subclassify IM patients. Moreover, we have shown that MSAs perform better than the 2017 EULAR/ACR criteria to predict the clinical phenotype of MSA-positive IM

patients. Taken together, these analyses suggest that using MSAs to classify IM patients into clinically homogeneous groups may be advantageous in clinical trials, for clinical research, and, perhaps, in clinical practice.

REFERENCES

1. Selva-O'Callaghan A, Pinal-Fernandez I, Trallero-Araguas E, Milisenda JC, Grau-Junyent JM, Mammen AL. Classification and management of adult inflammatory myopathies. Lancet Neurol 2018;17:816-28.

2. Aggarwal R, Ringold S, Khanna D, et al. Distinctions between diagnostic and classification criteria? Arthritis Care Res (Hoboken) 2015;67:891-7.

3. McHugh NJ, Tansley SL. Autoantibodies in myositis. Nat Rev Rheumatol 2018;14:290-302.

4. Lundberg IE, Tjarnlund A, Bottai M, et al. 2017 European League Against Rheumatism/American College of Rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups. Ann Rheum Dis 2017;76:1955-64.

5. Mammen AL, Chung T, Christopher-Stine L, et al. Autoantibodies against 3-hydroxy-3-methylglutaryl-coenzyme A reductase in patients with statin-associated autoimmune myopathy. Arthritis Rheum 2011;63:713-21.

6. Pinal-Fernandez I, Casal-Dominguez M, Huapaya JA, et al. A longitudinal cohort study of the anti-synthetase syndrome: increased severity of interstitial lung disease in black patients and patients with anti-PL7 and anti-PL12 autoantibodies. Rheumatology (Oxford) 2017;56:999-1007.

7. Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control 1974;19:716-23.

8. Schwarz G. Estimating the Dimension of a Model. Ann Statist 1978;6:461-4.

9. Burnham KP, Anderson DR. Multimodel Inference:Understanding AIC and BIC in Model Selection. Sociological Methods & Research 2004;33:261-304.

10. Raftery AE. Bayesian model selection in social research. Sociol Methodol 1995;25:111-63.

11. Tiniakou E, Pinal-Fernandez I, Lloyd TE, et al. More severe disease and slower recovery in younger patients with anti-3-hydroxy-3-methylglutaryl-coenzyme A reductase-associated autoimmune myopathy. Rheumatology (Oxford) 2017;56:787-94.

12. Pinal-Fernandez I, Parks C, Werner JL, et al. Longitudinal Course of Disease in a Large Cohort of Myositis Patients With Autoantibodies Recognizing the Signal Recognition Particle. Arthritis Care Res (Hoboken) 2017;69:263-70.

13. Pinal-Fernandez I, Mecoli CA, Casal-Dominguez M, et al. More prominent muscle involvement in patients with dermatomyositis with anti-Mi2 autoantibodies. Neurology 2019;93:e1768-e77.

14. Albayda J, Pinal-Fernandez I, Huang W, et al. Antinuclear Matrix Protein 2 Autoantibodies and Edema, Muscle Disease, and Malignancy Risk in Dermatomyositis Patients. Arthritis Care Res (Hoboken) 2017;69:1771-6.

15. Fujimoto M, Hamaguchi Y, Kaji K, et al. Myositis-specific anti-155/140 autoantibodies target transcription intermediary factor 1 family proteins. Arthritis Rheum 2012;64:513-22.

16. Sato S, Hoshino K, Satoh T, et al. RNA helicase encoded by melanoma differentiation-associated gene 5 is a major autoantigen in patients with clinically amyopathic dermatomyositis: Association with rapidly progressive interstitial lung disease. Arthritis Rheum 2009;60:2193-200.

17. Love LA, Leff RL, Fraser DD, et al. A new approach to the classification of idiopathic inflammatory myopathy: myositis-specific autoantibodies define useful homogeneous patient groups. Medicine (Baltimore) 1991;70:360-74.

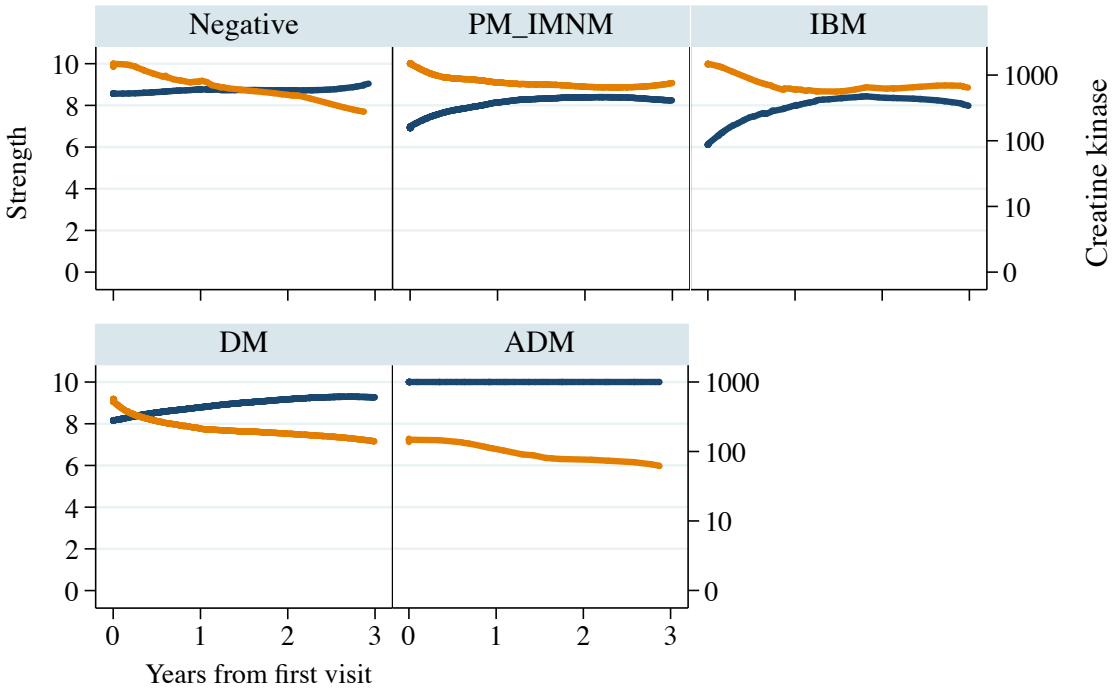
18. Allenbach Y, Mammen AL, Stenzel W, Benveniste O, Immune-Mediated Necrotizing Myopathies Working G. 224th ENMC International Workshop:: Clinico-sero-pathological classification of immune-mediated necrotizing myopathies Zandvoort, The Netherlands, 14-16 October 2016. Neuromuscul Disord 2017.

19. Mammen AL, Allenbach Y, Stenzel W, Benveniste O, Group EtWS. 239th ENMC International Workshop: Classification of dermatomyositis, Amsterdam, the Netherlands, 14-16 December 2018. Neuromuscul Disord 2019.

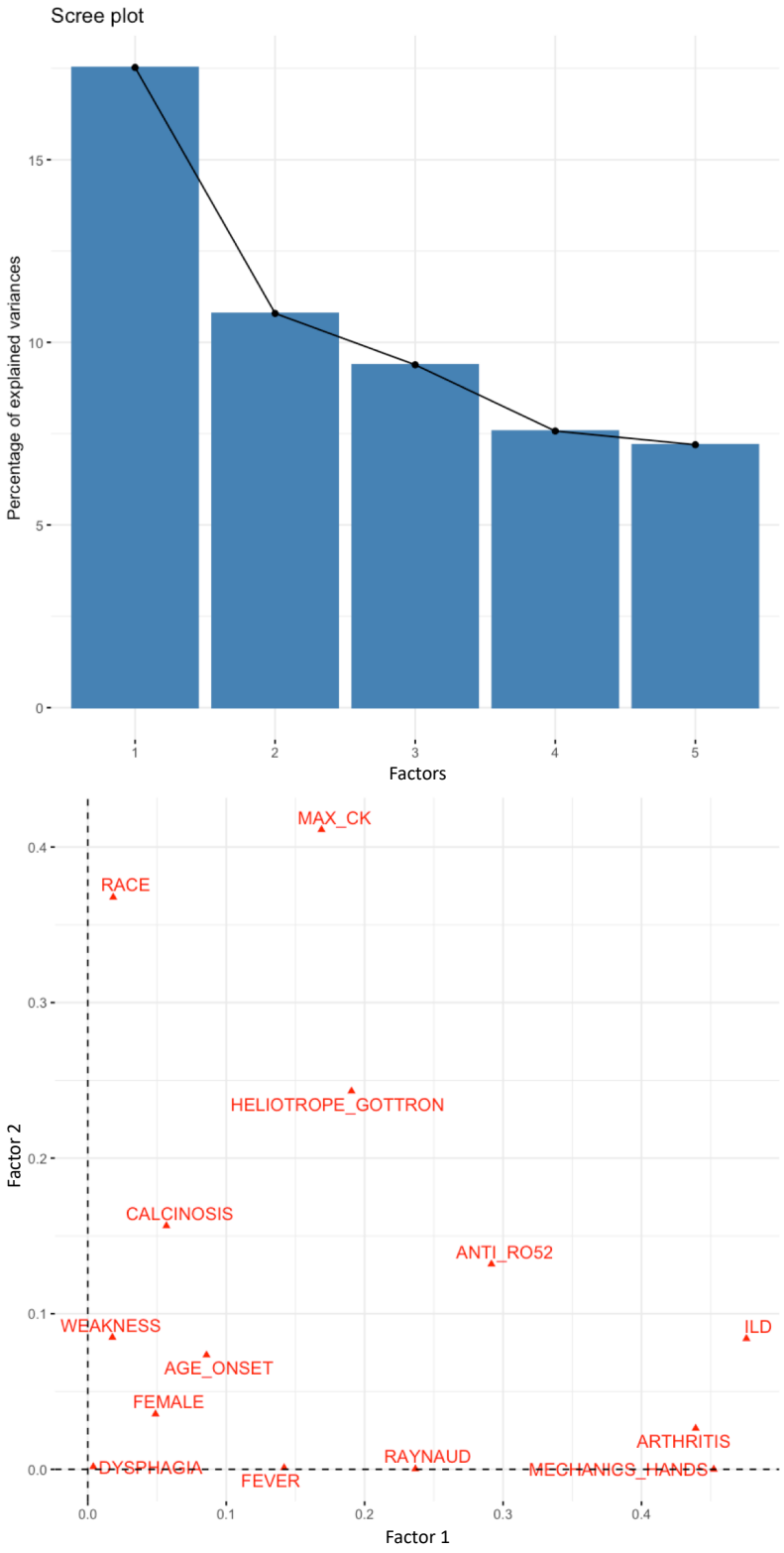
20. Pinal-Fernandez I, Casciola-Rosen LA, Christopher-Stine L, Corse AM, Mammen AL. The Prevalence of Individual Histopathologic Features Varies according to Autoantibody Status in Muscle Biopsies from Patients with Dermatomyositis. J Rheumatol 2015;42:1448-54.

21. Pinal-Fernandez I, Casal-Dominguez M, Derfoul A, et al. Identification of distinctive interferon gene signatures in different types of myositis. Neurology 2019;93:e1193-e204.

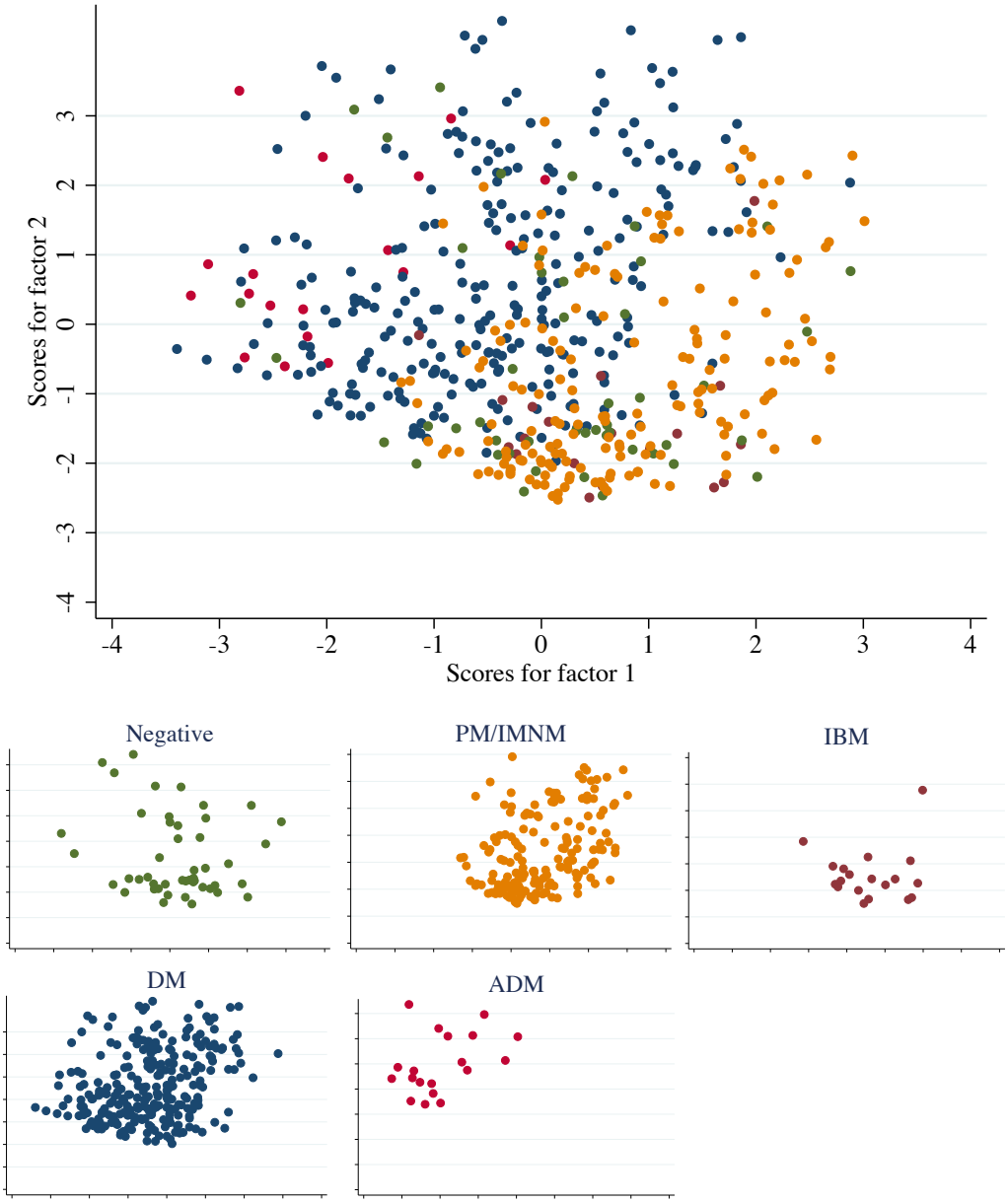
Supplementary Figure 1. Evolution of the strength (blue line) and creatine kinase (orange line) using locally weighted scatterplot smoothing of patients of the different 2017 EULAR/ACR myositis classification subgroups. PM_IMNM: polymyositis/immune-mediated necrotizing myositis; IBM: inclusion body myositis; DM: dermatomyositis; ADM: amyopathic dermatomyositis.



Supplementary Figure 2. a) Scree plot showing the percentage of explained variance for each factor. b) Scatterplot showing the weight of each one of the variables for factor 1 and factor 2.



Supplementary Figure 3. Factor analysis of mixed data summarizing the clinical phenotype of the autoantibody-positive adult myositis patients. Different color show the distinct 2017 EULAR/ACR adult subgroups for myositis patients.



The factor analysis of mixed data included the gender, race, age at onset, median and maximum CK, presence of anti-Ro52 autoantibodies, and presence or absence during the course of the disease of: muscle weakness, interstitial lung disease, arthritis, heliotrope or Gottron's rash, calcinosis, Raynaud's phenomenon, mechanics hands, dysphagia and fevers. The only two factors explaining more than 10% of the variance (factor 1 and factor 2) were retained for further analysis.

Supplementary Table 1. Comparison of the predictive ability of ACR/EULAR subgroup classification with the autoantibodies to predict the clinical phenotype of the patients mathematically summarized through factor analysis of mixed data. AIC: Akaike information criteria; BIC: Bayesian information criteria

	AIC autoantibodies	AIC ACR/EULAR	ΔAIC	BIC autoantibodies	BIC ACR/EULAR	ΔBIC
Factor 1	1459	1898	439	1498	1919	421
Factor 2	1454	1574	120	1492	1596	104

Supplementary Table 2. Clinical features of patients with anti-SRP autoantibodies divided by ACR/EULAR classification group.

	DM (n=4)	IBM (n=4)	Negative (n=4)	PM_IMNM (n=32)	Total (n=44)
Interstitial lung disease	0% (0)	0% (0)	75% (3)*	16% (5)	18% (8)
Muscle weakness	100% (4)	100% (4)	50% (2)**	100% (32)	95% (42)
Arthritis	0% (0)	0% (0)	0% (0)	9% (3)	7% (3)
DM-specific skin involvement	100% (4)***	0% (0)	0% (0)	0% (0)**	9% (4)
Calcinosis	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)
Raynaud's phenomenon	50% (2)	25% (1)	50% (2)	38% (12)	39% (17)
Mechanic's hands	25% (1)	0% (0)	0% (0)	3% (1)	5% (2)
Dysphagia	50% (2)	50% (2)	0% (0)*	59% (19)	52% (23)
Fever	0% (0)	0% (0)	0% (0)	12% (4)	9% (4)
Mean hip flexor strength	7.5 (3.4)	3.6 (1.6)	6.3 (4.3)	5.7 (2.8)	5.7 (2.9)
Hip flexors strength at last visit	7.3 (4.6)	2.8 (1.5)	8.0 (4.0)	5.9 (3.6)	5.9 (3.6)
Mean arm abductor strength	7.3 (3.3)	7.1 (2.4)	8.6 (2.4)	7.7 (2.0)	7.7 (2.1)
Arm abductors strength at last visit	8.2 (3.5)	7.0 (3.5)	8.8 (2.5)	7.8 (2.7)	7.8 (2.8)
Median CK	710 (548-4018)	572 (305-925)	1330 (494-4960)	1260 (564-2392)	944 (517-2193)
Maximum CK	6132 (3798-14814)	5156 (2226-14276)	8815 (3801-13750)	5912 (3277-10479)	6012 (3126-10923)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 3. Clinical features of patients with anti-HMGCR autoantibodies divided by ACR/EULAR classification group.

	DM (n=10)	IBM (n=13)	Negative (n=25)	PM_IMNM (n=74)	Total (n=122)
Interstitial lung disease	0% (0)	0% (0)	4% (1)	1% (1)	2% (2)
Muscle weakness	100% (10)	92% (12)	92% (23)	100% (74)	98% (119)
Arthritis	20% (2)	0% (0)	4% (1)	5% (4)	6% (7)
DM-specific skin involvement	100% (10)***	0% (0)	0% (0)	0% (0)***	8% (10)
Calcinosis	0% (0)	0% (0)	0% (0)	1% (1)	1% (1)
Raynaud's phenomenon	20% (2)	0% (0)	0% (0)	9% (7)	7% (9)
Mechanic's hands	20% (2)	8% (1)	0% (0)	3% (2)	4% (5)
Dysphagia	20% (2)	38% (5)	24% (6)*	53% (39)**	43% (52)
Fever	10% (1)	15% (2)	8% (2)	4% (3)	7% (8)
Mean hip flexor strength	7.9 (1.3)	6.2 (3.1)	7.0 (3.0)	6.7 (2.7)	6.8 (2.7)
Hip flexors strength at last visit	7.9 (1.6)	6.7 (3.5)	6.8 (3.6)	7.0 (3.6)	7.0 (3.4)
Mean arm abductor strength	9.3 (0.7)	8.8 (0.9)	9.2 (1.0)	8.5 (1.9)*	8.7 (1.6)
Arm abductors strength at last visit	9.6 (0.7)	9.4 (0.9)	9.1 (1.2)	8.7 (2.4)	8.9 (2.0)
Median CK	1262 (792-1859)	2161 (1569-2682)	1620 (646-4352)	1483 (493-3062)	1556 (526-3062)
Maximum CK	5085 (2226-8000)	8780 (3280-11610)	3538 (1890-8990)	4886 (2160-9166)	4908 (2160-10000)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 4. Clinical features of patients with anti-Mi2 autoantibodies divided by ACR/EULAR classification group.

	DM (n=53)	IBM (n=1)	Negative (n=3)	PM_IMNM (n=2)	Total (n=59)
Interstitial lung disease	6% (3)	0% (0)	0% (0)	0% (0)	5% (3)
Muscle weakness	100% (53)	100% (1)	67% (2)	100% (2)	98% (58)
Arthritis	23% (12)	0% (0)	0% (0)	0% (0)	20% (12)
DM-specific skin involvement	100% (53)***	0% (0)	33% (1)*	0% (0)**	92% (54)
Calcinosis	9% (5)	0% (0)	0% (0)	0% (0)	8% (5)
Raynaud's phenomenon	32% (17)	0% (0)	33% (1)	0% (0)	31% (18)
Mechanic's hands	21% (11)	0% (0)	67% (2)	0% (0)	22% (13)
Dysphagia	55% (29)	100% (1)	0% (0)	50% (1)	53% (31)
Fever	8% (4)	0% (0)	0% (0)	0% (0)	7% (4)
Mean hip flexor strength	8.1 (2.3)	8.7 (.)	9.8 (0.3)	4.2 (1.1)*	8.0 (2.4)
Hip flexors strength at last visit	8.3 (2.5)	9.5 (.)	10.0 (0.0)	3.5 (2.1)**	8.3 (2.6)
Mean arm abductor strength	8.4 (2.2)	9.6 (.)	9.9 (0.2)	6.5 (2.1)	8.5 (2.1)
Arm abductors strength at last visit	8.6 (2.7)	10.0 (.)	10.0 (0.0)	6.5 (2.1)	8.6 (2.6)
Median CK	554 (158-1448)	343 (343-343)	201 (77-364)	148 (108-188)	362 (155-1325)
Maximum CK	3830 (2003-7070)	2781 (2781-2781)	4113 (3635-6250)	6540 (5080-8000)	3900 (2230-7070)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 5. Clinical features of patients with anti-NXP2 autoantibodies divided by ACR/EULAR classification group.

	ADM (n=1)	DM (n=41)	Negative (n=2)	PM_IMNM (n=6)	Total (n=50)
Interstitial lung disease	0% (0)	7% (3)	0% (0)	0% (0)	6% (3)
Muscle weakness	0% (0)	100% (41)**	0% (0)**	100% (6)	94% (47)
Arthritis	0% (0)	15% (6)	0% (0)	17% (1)	14% (7)
DM-specific skin involvement	100% (1)	100% (41)***	50% (1)	0% (0)***	86% (43)
Calcinosis	0% (0)	41% (17)	0% (0)	17% (1)	36% (18)
Raynaud's phenomenon	0% (0)	10% (4)	0% (0)	17% (1)	10% (5)
Mechanic's hands	0% (0)	12% (5)	50% (1)	0% (0)	12% (6)
Dysphagia	0% (0)	73% (30)	50% (1)	50% (3)	68% (34)
Fever	0% (0)	22% (9)	0% (0)	0% (0)	18% (9)
Mean hip flexor strength	10.0 (.)	8.1 (1.9)	10.0 (0.0)	7.4 (2.1)	8.2 (1.9)
Hip flexors strength at last visit	10.0 (.)	8.8 (2.2)	10.0 (0.0)	8.8 (1.3)	8.9 (2.0)
Mean arm abductor strength	10.0 (.)	8.6 (1.8)	10.0 (0.0)	8.0 (2.3)	8.6 (1.8)
Arm abductors strength at last visit	10.0 (.)	8.9 (2.3)	10.0 (0.0)	8.6 (1.5)	9.0 (2.2)
Median CK	70 (70-70)	161 (72-377)	92 (66-117)	316 (203-456)	168 (72-377)
Maximum CK	114 (114-114)	2100 (499-4913)	1514 (117-2910)	4136 (751-7348)	1900 (462-5000)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 6. Clinical features of patients with anti-TIF1g autoantibodies divided by ACR/EULAR classification group.

	ADM (n=11)	DM (n=50)	Negative (n=1)	Total (n=62)
Interstitial lung disease	0% (0)	0% (0)	0% (0)	0% (0)
Muscle weakness	0% (0)***	100% (50)***	0% (0)	81% (50)
Arthritis	18% (2)	8% (4)	0% (0)	10% (6)
DM-specific skin involvement	100% (11)	100% (50)	100% (1)	100% (62)
Calcinosis	18% (2)	12% (6)	0% (0)	13% (8)
Raynaud's phenomenon	18% (2)	20% (10)	0% (0)	19% (12)
Mechanic's hands	36% (4)	24% (12)	0% (0)	26% (16)
Dysphagia	18% (2)*	58% (29)*	0% (0)	50% (31)
Fever	0% (0)	12% (6)	0% (0)	10% (6)
Mean hip flexor strength	10.0 (0.0)*	9.2 (1.3)*	10.0 (.)	9.3 (1.2)
Hip flexors strength at last visit	10.0 (0.0)	9.3 (1.6)	10.0 (.)	9.5 (1.4)
Mean arm abductor strength	10.0 (0.0)*	9.4 (0.9)*	10.0 (.)	9.5 (0.9)
Arm abductors strength at last visit	10.0 (0.0)	9.5 (1.1)	10.0 (.)	9.6 (1.0)
Median CK	116 (74-129)	102 (65-124)	52 (52-52)	105 (65-128)
Maximum CK	134 (99-209)	236 (115-588)*	55 (55-55)	207 (113-455)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 7. Clinical features of patients with anti-MDA5 autoantibodies divided by ACR/EULAR classification group.

	ADM (n=4)	DM (n=21)	Negative (n=3)	Total (n=28)
Interstitial lung disease	100% (4)	67% (14)	67% (2)	71% (20)
Muscle weakness	0% (0)***	100% (21)***	33% (1)	79% (22)
Arthritis	75% (3)	71% (15)	67% (2)	71% (20)
DM-specific skin involvement	100% (4)	100% (21)	100% (3)	100% (28)
Calcinosis	25% (1)	52% (11)	33% (1)	46% (13)
Raynaud's phenomenon	25% (1)	52% (11)	100% (3)	54% (15)
Mechanic's hands	75% (3)	76% (16)	67% (2)	75% (21)
Dysphagia	25% (1)	43% (9)	0% (0)	36% (10)
Fever	25% (1)	57% (12)	0% (0)	46% (13)
Mean hip flexor strength	10.0 (0.0)	9.6 (0.5)	10.0 (0.0)	9.7 (0.5)
Hip flexors strength at last visit	10.0 (0.0)	9.6 (0.7)	10.0 (0.0)	9.7 (0.7)
Mean arm abductor strength	10.0 (0.0)	9.8 (0.3)	10.0 (0.1)	9.8 (0.3)
Arm abductors strength at last visit	10.0 (0.0)	9.9 (0.3)	10.0 (0.0)	9.9 (0.3)
Median CK	77 (55-177)	59 (40-105)	84 (61-136)	64 (41-120)
Maximum CK	86 (56-244)	143 (65-256)	146 (74-160)	126 (66-235)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 8. Clinical features of patients with anti-Jo1 autoantibodies divided by ACR/EULAR classification group.

	ADM (n=1)	DM (n=70)	IBM (n=1)	Negative (n=1)	PM_IMNM (n=54)	Total (n=127)
Interstitial lung disease	100% (1)	73% (51)	100% (1)	100% (1)	83% (45)	78% (99)
Muscle weakness	0% (0)	100% (70)**	100% (1)	0% (0)	89% (48)	94% (119)
Arthritis	100% (1)	67% (47)	0% (0)	0% (0)	54% (29)	61% (77)
DM-specific skin involvement	100% (1)	100% (70)***	0% (0)	0% (0)	0% (0)***	56% (71)
Calcinosis	0% (0)	11% (8)	0% (0)	0% (0)	7% (4)	9% (12)
Raynaud's phenomenon	100% (1)	40% (28)	100% (1)	100% (1)	28% (15)	36% (46)
Mechanic's hands	0% (0)	67% (47)***	100% (1)	100% (1)	33% (18)***	53% (67)
Dysphagia	0% (0)	33% (23)	0% (0)	0% (0)	44% (24)	37% (47)
Fever	100% (1)	21% (15)	0% (0)	0% (0)	15% (8)	19% (24)
Mean hip flexor strength	10.0 (.)	8.9 (1.6)	10.0 (.)	10.0 (.)	8.9 (1.5)	9.0 (1.5)
Hip flexors strength at last visit	10.0 (.)	9.2 (1.5)	10.0 (.)	10.0 (.)	9.0 (1.6)	9.2 (1.6)
Mean arm abductor strength	10.0 (.)	9.4 (1.1)	10.0 (.)	10.0 (.)	9.4 (1.3)	9.4 (1.1)
Arm abductors strength at last visit	10.0 (.)	9.4 (1.0)	10.0 (.)	10.0 (.)	9.5 (1.4)	9.5 (1.2)
Median CK	261 (261-261)	390 (130-1316)	460 (460-460)	92 (92-92)	409 (161-1000)	403 (134-1098)
Maximum CK	261 (261-261)	2400 (679-6160)	467 (467-467)	107 (107-107)	1736 (563-9000)	1920 (469-6686)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 9. Clinical features of patients with anti-PL7 autoantibodies divided by ACR/EULAR classification group.

	ADM (n=1)	DM (n=5)	Negative (n=6)	Total (n=12)
Interstitial lung disease	100% (1)	100% (5)	100% (6)	100% (12)
Muscle weakness	0% (0)	100% (5)	100% (6)	92% (11)
Arthritis	0% (0)	40% (2)	33% (2)	33% (4)
DM-specific skin involvement	100% (1)	100% (5)*	0% (0)**	50% (6)
Calcinosis	0% (0)	0% (0)	0% (0)	0% (0)
Raynaud's phenomenon	0% (0)	60% (3)	17% (1)	33% (4)
Mechanic's hands	100% (1)	100% (5)	33% (2)	67% (8)
Dysphagia	0% (0)	100% (5)**	0% (0)*	42% (5)
Fever	0% (0)	20% (1)	17% (1)	17% (2)
Mean hip flexor strength	10.0 (.)	9.5 (0.5)	9.7 (0.4)	9.6 (0.4)
Hip flexors strength at last visit	10.0 (.)	9.5 (0.7)	9.5 (0.6)	9.6 (0.5)
Mean arm abductor strength	10.0 (.)	9.6 (0.4)	9.7 (0.5)	9.7 (0.5)
Arm abductors strength at last visit	10.0 (.)	9.8 (0.4)	10.0 (0.0)	9.9 (0.2)
Median CK	80 (80-80)	826 (352-1120)	236 (75-424)	346 (106-973)
Maximum CK	683 (683-683)	6797 (5668-8827)*	294 (115-584)	2766 (294-7812)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 10. Clinical features of patients with anti-PL12 autoantibodies divided by ACR/EULAR classification group.

	ADM (n=1)	DM (n=13)	Negative (n=2)	PM_IMNM (n=4)	Total (n=20)
Interstitial lung disease	100% (1)	85% (11)	100% (2)	75% (3)	85% (17)
Muscle weakness	0% (0)	100% (13)*	0% (0)*	100% (4)	85% (17)
Arthritis	0% (0)	69% (9)	50% (1)	75% (3)	65% (13)
DM-specific skin involvement	100% (1)	100% (13)***	0% (0)	0% (0)**	70% (14)
Calcinosis	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)
Raynaud's phenomenon	0% (0)	38% (5)	50% (1)	0% (0)	30% (6)
Mechanic's hands	0% (0)	77% (10)	100% (2)	75% (3)	75% (15)
Dysphagia	0% (0)	38% (5)	0% (0)	50% (2)	35% (7)
Fever	100% (1)	62% (8)	0% (0)	25% (1)	50% (10)
Mean hip flexor strength	10.0 (.)	9.4 (0.6)	10.0 (0.0)	9.1 (0.8)	9.4 (0.6)
Hip flexors strength at last visit	10.0 (.)	9.4 (0.6)	10.0 (0.0)	10.0 (0.0)	9.6 (0.6)
Mean arm abductor strength	10.0 (.)	9.6 (0.8)	10.0 (0.0)	8.9 (1.1)	9.5 (0.8)
Arm abductors strength at last visit	10.0 (.)	9.7 (1.0)	10.0 (0.0)	9.2 (1.0)	9.6 (0.9)
Median CK	96 (96-96)	172 (72-256)	66 (31-101)	63 (48-217)	99 (59-246)
Maximum CK	1000 (1000-1000)	356 (171-1938)	76 (31-121)	314 (115-658)	318 (119-1172)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 11. Comparison of the predictive ability of ACR/EULAR subgroup classification with the autoantibodies to predict the clinical phenotype of the patients mathematically summarized through factor analysis of mixed data excluding those patients with no skin rash and no muscle biopsy available. AIC: Akaike information criteria; BIC: Bayesian information criteria

	AIC autoantibodies	AIC ACR/EULAR	ΔAIC	BIC autoantibodies	BIC ACR/EULAR	ΔBIC
Factor 1	1093	1443	350	1129	1463	334
Factor 2	1093	1215	122	1128	1235	107

Supplementary Table 12. Clinical features of patients with anti-SRP autoantibodies divided by ACR/EULAR classification group excluding those patients with no skin rash and no muscle biopsy available.

	DM (n=4)	IBM (n=4)	Negative (n=2)	PM_IMNM (n=16)	Total (n=26)
Interstitial lung disease	0% (0)	0% (0)	50% (1)	12% (2)	12% (3)
Muscle weakness	100% (4)	100% (4)	50% (1)	100% (16)	96% (25)
Arthritis	0% (0)	0% (0)	0% (0)	12% (2)	8% (2)
DM-specific skin involvement	100% (4)***	0% (0)	0% (0)	0% (0)*	15% (4)
Calcinosis	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)
Raynaud's phenomenon	50% (2)	25% (1)	0% (0)	44% (7)	38% (10)
Mechanic's hands	25% (1)	0% (0)	0% (0)	0% (0)	4% (1)
Dysphagia	50% (2)	50% (2)	0% (0)	69% (11)	58% (15)
Fever	0% (0)	0% (0)	0% (0)	12% (2)	8% (2)
Mean hip flexor strength	7.5 (3.4)	3.6 (1.6)	6.5 (4.9)	5.3 (2.6)	5.5 (2.8)
Hip flexors strength at last visit	7.3 (4.6)	2.8 (1.5)	10.0 (0.0)	5.6 (3.2)	5.7 (3.5)
Mean arm abductor strength	7.3 (3.3)	7.1 (2.4)	9.6 (0.5)	7.2 (2.4)	7.4 (2.4)
Arm abductors strength at last visit	8.2 (3.5)	7.0 (3.5)	10.0 (0.0)	7.0 (3.1)	7.5 (3.1)
Median CK	710 (548-4018)	572 (305-925)	1330 (900-1761)	1157 (564-2444)	837 (558-1761)
Maximum CK	6132 (3798-14814)	5156 (2226-14276)	8815 (7130-10500)	5786 (3277-9673)	6012 (3159-10500)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 13. Clinical features of patients with anti-HMGCR autoantibodies divided by ACR/EULAR classification group excluding those patients with no skin rash and no muscle biopsy available.

	DM (n=10)	IBM (n=12)	Negative (n=15)	PM_IMNM (n=41)	Total (n=78)
Interstitial lung disease	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)
Muscle weakness	100% (10)	92% (11)	93% (14)	100% (41)	97% (76)
Arthritis	20% (2)	0% (0)	0% (0)	5% (2)	5% (4)
DM-specific skin involvement	100% (10)***	0% (0)	0% (0)	0% (0)***	13% (10)
Calcinosis	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)
Raynaud's phenomenon	20% (2)	0% (0)	0% (0)	15% (6)	10% (8)
Mechanic's hands	20% (2)	8% (1)	0% (0)	5% (2)	6% (5)
Dysphagia	20% (2)	42% (5)	40% (6)	56% (23)	46% (36)
Fever	10% (1)	17% (2)	13% (2)	7% (3)	10% (8)
Mean hip flexor strength	7.9 (1.3)	6.7 (2.6)	6.8 (3.2)	6.7 (2.8)	6.9 (2.7)
Hip flexors strength at last visit	7.9 (1.6)	7.5 (2.8)	7.0 (3.7)	7.2 (3.6)	7.3 (3.3)
Mean arm abductor strength	9.3 (0.7)	8.8 (1.0)	9.2 (1.0)	8.4 (1.9)	8.7 (1.5)
Arm abductors strength at last visit	9.6 (0.7)	9.5 (0.9)	9.2 (1.1)	8.5 (2.6)	8.9 (2.1)
Median CK	1262 (792-1859)	2102 (1284-3072)	1999 (474-4352)	1550 (493-3771)	1594 (526-3463)
Maximum CK	5085 (2226-8000)	9390 (3398-12962)	6900 (1983-14000)	5880 (2405-10580)	5940 (2323-10725)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 14. Clinical features of patients with anti-Mi2 autoantibodies divided by ACR/EULAR classification group excluding those patients with no skin rash and no muscle biopsy available.

	DM (n=53)	IBM (n=1)	Negative (n=2)	PM_IMNM (n=1)	Total (n=57)
Interstitial lung disease	6% (3)	0% (0)	0% (0)	0% (0)	5% (3)
Muscle weakness	100% (53)	100% (1)	50% (1)*	100% (1)	98% (56)
Arthritis	23% (12)	0% (0)	0% (0)	0% (0)	21% (12)
DM-specific skin involvement	100% (53)***	0% (0)	50% (1)	0% (0)	95% (54)
Calcinosis	9% (5)	0% (0)	0% (0)	0% (0)	9% (5)
Raynaud's phenomenon	32% (17)	0% (0)	0% (0)	0% (0)	30% (17)
Mechanic's hands	21% (11)	0% (0)	50% (1)	0% (0)	21% (12)
Dysphagia	55% (29)	100% (1)	0% (0)	100% (1)	54% (31)
Fever	8% (4)	0% (0)	0% (0)	0% (0)	7% (4)
Mean hip flexor strength	8.1 (2.3)	8.7 (.)	9.9 (0.1)	5.0 (.)	8.1 (2.3)
Hip flexors strength at last visit	8.3 (2.5)	9.5 (.)	10.0 (0.0)	5.0 (.)	8.4 (2.5)
Mean arm abductor strength	8.4 (2.2)	9.6 (.)	10.0 (0.0)	8.0 (.)	8.5 (2.1)
Arm abductors strength at last visit	8.6 (2.7)	10.0 (.)	10.0 (0.0)	8.0 (.)	8.6 (2.6)
Median CK	554 (158-1448)	343 (343-343)	283 (201-364)	108 (108-108)	364 (158-1325)
Maximum CK	3830 (2003-7070)	2781 (2781-2781)	5182 (4113-6250)	5080 (5080-5080)	3900 (2230-6766)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 15. Clinical features of patients with anti-NXP2 autoantibodies divided by ACR/EULAR classification group excluding those patients with no skin rash and no muscle biopsy available.

	ADM (n=1)	DM (n=41)	Negative (n=1)	PM_IMNM (n=3)	Total (n=46)
Interstitial lung disease	0% (0)	7% (3)	0% (0)	0% (0)	7% (3)
Muscle weakness	0% (0)*	100% (41)**	0% (0)*	100% (3)	96% (44)
Arthritis	0% (0)	15% (6)	0% (0)	0% (0)	13% (6)
DM-specific skin involvement	100% (1)	100% (41)***	100% (1)	0% (0)***	93% (43)
Calcinosis	0% (0)	41% (17)	0% (0)	0% (0)	37% (17)
Raynaud's phenomenon	0% (0)	10% (4)	0% (0)	33% (1)	11% (5)
Mechanic's hands	0% (0)	12% (5)	100% (1)	0% (0)	13% (6)
Dysphagia	0% (0)	73% (30)	100% (1)	67% (2)	72% (33)
Fever	0% (0)	22% (9)	0% (0)	0% (0)	20% (9)
Mean hip flexor strength	10.0 (.)	8.1 (1.9)	10.0 (.)	6.5 (2.5)	8.1 (2.0)
Hip flexors strength at last visit	10.0 (.)	8.8 (2.2)	10.0 (.)	9.0 (1.4)	8.9 (2.1)
Mean arm abductor strength	10.0 (.)	8.6 (1.8)	10.0 (.)	8.2 (3.2)	8.6 (1.8)
Arm abductors strength at last visit	10.0 (.)	8.9 (2.3)	10.0 (.)	8.5 (2.1)	9.0 (2.2)
Median CK	70 (70-70)	161 (72-377)	117 (117-117)	203 (82-341)	156 (72-349)
Maximum CK	114 (114-114)	2100 (499-4913)	117 (117-117)	1271 (751-24000)	1652 (300-4913)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 16. Clinical features of patients with anti-TIF1g autoantibodies divided by ACR/EULAR classification group excluding those patients with no skin rash and no muscle biopsy available.

	ADM (n=11)	DM (n=50)	Negative (n=1)	Total (n=62)
Interstitial lung disease	0% (0)	0% (0)	0% (0)	0% (0)
Muscle weakness	0% (0)***	100% (50)***	0% (0)	81% (50)
Arthritis	18% (2)	8% (4)	0% (0)	10% (6)
DM-specific skin involvement	100% (11)	100% (50)	100% (1)	100% (62)
Calcinosis	18% (2)	12% (6)	0% (0)	13% (8)
Raynaud's phenomenon	18% (2)	20% (10)	0% (0)	19% (12)
Mechanic's hands	36% (4)	24% (12)	0% (0)	26% (16)
Dysphagia	18% (2)*	58% (29)*	0% (0)	50% (31)
Fever	0% (0)	12% (6)	0% (0)	10% (6)
Mean hip flexor strength	10.0 (0.0)*	9.2 (1.3)*	10.0 (.)	9.3 (1.2)
Hip flexors strength at last visit	10.0 (0.0)	9.3 (1.6)	10.0 (.)	9.5 (1.4)
Mean arm abductor strength	10.0 (0.0)*	9.4 (0.9)*	10.0 (.)	9.5 (0.9)
Arm abductors strength at last visit	10.0 (0.0)	9.5 (1.1)	10.0 (.)	9.6 (1.0)
Median CK	116 (74-129)	102 (65-124)	52 (52-52)	105 (65-128)
Maximum CK	134 (99-209)	236 (115-588)*	55 (55-55)	207 (113-455)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 17. Clinical features of patients with anti-MDA5 autoantibodies divided by ACR/EULAR classification group excluding those patients with no skin rash and no muscle biopsy available.

	ADM (n=4)	DM (n=21)	Negative (n=3)	Total (n=28)
Interstitial lung disease	100% (4)	67% (14)	67% (2)	71% (20)
Muscle weakness	0% (0)***	100% (21)***	33% (1)	79% (22)
Arthritis	75% (3)	71% (15)	67% (2)	71% (20)
DM-specific skin involvement	100% (4)	100% (21)	100% (3)	100% (28)
Calcinosis	25% (1)	52% (11)	33% (1)	46% (13)
Raynaud's phenomenon	25% (1)	52% (11)	100% (3)	54% (15)
Mechanic's hands	75% (3)	76% (16)	67% (2)	75% (21)
Dysphagia	25% (1)	43% (9)	0% (0)	36% (10)
Fever	25% (1)	57% (12)	0% (0)	46% (13)
Mean hip flexor strength	10.0 (0.0)	9.6 (0.5)	10.0 (0.0)	9.7 (0.5)
Hip flexors strength at last visit	10.0 (0.0)	9.6 (0.7)	10.0 (0.0)	9.7 (0.7)
Mean arm abductor strength	10.0 (0.0)	9.8 (0.3)	10.0 (0.1)	9.8 (0.3)
Arm abductors strength at last visit	10.0 (0.0)	9.9 (0.3)	10.0 (0.0)	9.9 (0.3)
Median CK	77 (55-177)	59 (40-105)	84 (61-136)	64 (41-120)
Maximum CK	86 (56-244)	143 (65-256)	146 (74-160)	126 (66-235)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 18. Clinical features of patients with anti-Jo1 autoantibodies divided by ACR/EULAR classification group excluding those patients with no skin rash and no muscle biopsy available.

	ADM (n=1)	DM (n=70)	PM_IMNM (n=12)	Total (n=83)
Interstitial lung disease	100% (1)	73% (51)	67% (8)	72% (60)
Muscle weakness	0% (0)*	100% (70)	100% (12)	99% (82)
Arthritis	100% (1)	67% (47)	75% (9)	69% (57)
DM-specific skin involvement	100% (1)	100% (70)***	0% (0)***	86% (71)
Calcinosis	0% (0)	11% (8)	17% (2)	12% (10)
Raynaud's phenomenon	100% (1)	40% (28)	42% (5)	41% (34)
Mechanic's hands	0% (0)	67% (47)	50% (6)	64% (53)
Dysphagia	0% (0)	33% (23)	50% (6)	35% (29)
Fever	100% (1)	21% (15)	17% (2)	22% (18)
Mean hip flexor strength	10.0 (.)	8.9 (1.6)	8.0 (2.3)	8.8 (1.7)
Hip flexors strength at last visit	10.0 (.)	9.2 (1.5)	8.6 (2.8)	9.1 (1.7)
Mean arm abductor strength	10.0 (.)	9.4 (1.1)	8.7 (2.3)	9.3 (1.3)
Arm abductors strength at last visit	10.0 (.)	9.4 (1.0)	8.6 (2.8)	9.3 (1.4)
Median CK	261 (261-261)	390 (130-1316)	558 (210-1556)	412 (131-1316)
Maximum CK	261 (261-261)	2400 (679-6160)	2702 (891-9940)	2150 (587-6686)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 19. Clinical features of patients with anti-PL7 autoantibodies divided by ACR/EULAR classification group excluding those patients with no skin rash and no muscle biopsy available.

	ADM (n=1)	DM (n=5)	Total (n=6)
Interstitial lung disease	100% (1)	100% (5)	100% (6)
Muscle weakness	0% (0)	100% (5)	83% (5)
Arthritis	0% (0)	40% (2)	33% (2)
DM-specific skin involvement	100% (1)	100% (5)	100% (6)
Calcinosis	0% (0)	0% (0)	0% (0)
Raynaud's phenomenon	0% (0)	60% (3)	50% (3)
Mechanic's hands	100% (1)	100% (5)	100% (6)
Dysphagia	0% (0)	100% (5)	83% (5)
Fever	0% (0)	20% (1)	17% (1)
Mean hip flexor strength	10.0 (.)	9.5 (0.5)	9.6 (0.5)
Hip flexors strength at last visit	10.0 (.)	9.5 (0.7)	9.7 (0.6)
Mean arm abductor strength	10.0 (.)	9.6 (0.4)	9.7 (0.4)
Arm abductors strength at last visit	10.0 (.)	9.8 (0.4)	9.8 (0.3)
Median CK	80 (80-80)	826 (352-1120)	589 (184-1120)
Maximum CK	683 (683-683)	6797 (5668-8827)	6232 (4848-8827)

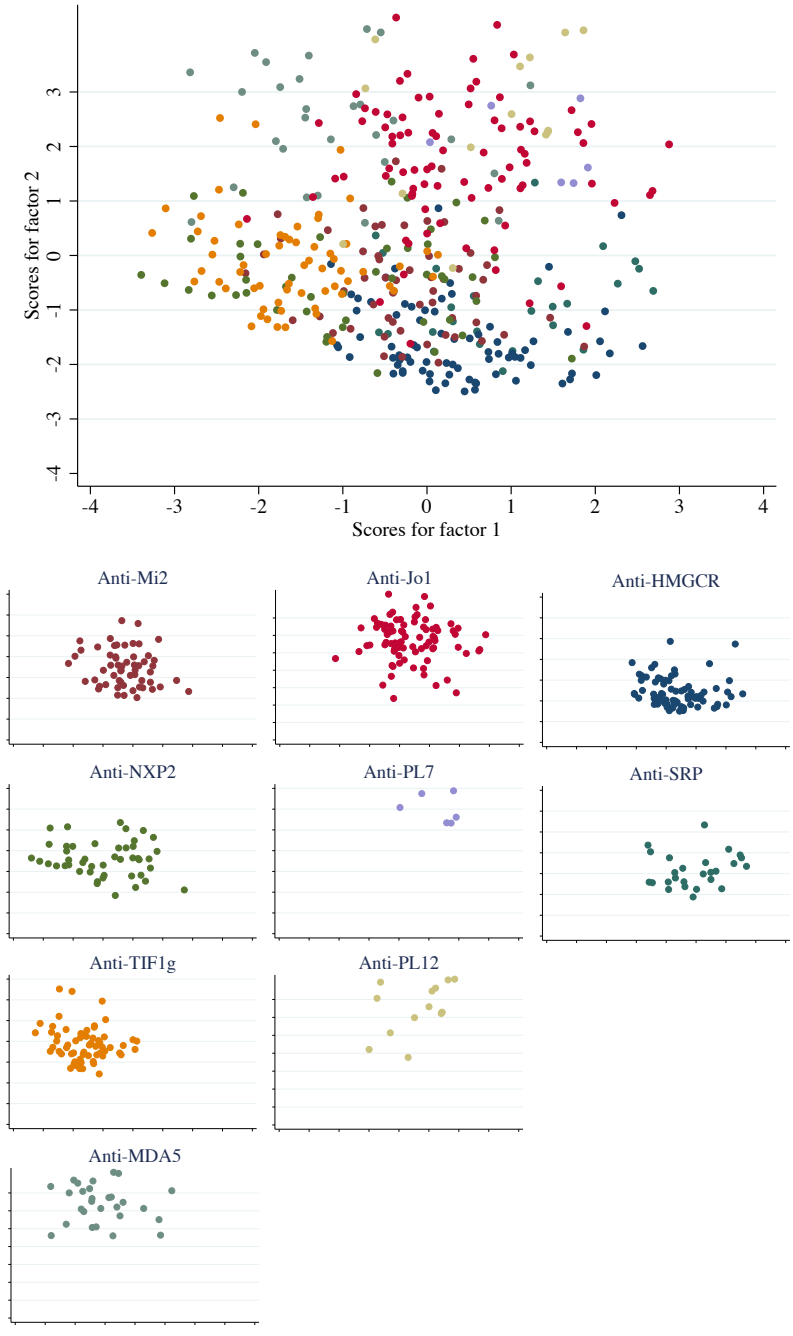
Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Table 20. Clinical features of patients with anti-PL12 autoantibodies divided by ACR/EULAR classification group excluding those patients with no skin rash and no muscle biopsy available.

	ADM (n=1)	DM (n=13)	Total (n=14)
Interstitial lung disease	100% (1)	85% (11)	86% (12)
Muscle weakness	0% (0)	100% (13)	93% (13)
Arthritis	0% (0)	69% (9)	64% (9)
DM-specific skin involvement	100% (1)	100% (13)	100% (14)
Calcinosis	0% (0)	0% (0)	0% (0)
Raynaud's phenomenon	0% (0)	38% (5)	36% (5)
Mechanic's hands	0% (0)	77% (10)	71% (10)
Dysphagia	0% (0)	38% (5)	36% (5)
Fever	100% (1)	62% (8)	64% (9)
Mean hip flexor strength	10.0 (.)	9.4 (0.6)	9.4 (0.6)
Hip flexors strength at last visit	10.0 (.)	9.4 (0.6)	9.5 (0.6)
Mean arm abductor strength	10.0 (.)	9.6 (0.8)	9.6 (0.8)
Arm abductors strength at last visit	10.0 (.)	9.7 (1.0)	9.7 (0.9)
Median CK	96 (96-96)	172 (72-256)	143 (72-256)
Maximum CK	1000 (1000-1000)	356 (171-1938)	377 (171-1938)

Dichotomous variables were expressed as a percentage (count) and continuous variables as mean (SD). Univariate comparisons of continuous variables were made using Student´s t-test while dichotomous variables were compared either using the chi-squared test or Fisher´s exact test, as appropriate.

Supplementary Figure 4. Factor analysis of mixed data summarizing the clinical phenotype of the autoantibody-positive adult myositis patients excluding those with no skin rash and no muscle biopsy available.



The factor analysis of mixed data included the gender, race, age at onset, median and maximum CK, presence of anti-Ro52 autoantibodies, and presence or absence during the course of the disease of: muscle weakness, interstitial lung disease, arthritis, heliotrope or Gottron's rash, calcinosis, Raynaud's phenomenon, mechanics hands, dysphagia and fevers. The only two factors explaining more than 10% of the variance (factor 1 and factor 2) were retained for further analysis.

Supplementary Methods

1. Clinical variables

At each visit, arm abduction and hip flexion strength were evaluated by the examining physician using the Medical Research Council scale. This scale was transformed to Kendall´s 0-10 scale for analysis purposes as previously described[1]. Serial strength measurements for each patient were made by the same physician. For the purposes of analyses, right- and left-side measurements for arm and hip strength were combined and the average was used for the calculations (possible range 0–10). All the available serum creatine kinase (CK) levels were included for analysis. Myositis-specific skin involvement (i.e., Gottron's papules, Gottron's sign, and heliotrope rash), calcinosis, symptoms of esophageal involvement, and antisynthetase syndrome-associated clinical features (e.g. mechanic's hands, Raynaud´s phenomenon, arthritis, fever) were documented both retrospectively at the onset of the disease and prospectively at each visit.

Interstitial lung disease was defined through a multidisciplinary approach as suggested by the American Thoracic Society[2].

All patients were screened for anti-Ro52 autoantibodies by line blot (EUROLINE myositis profile).

Cancer-associated myositis was defined as the cases of cancer occurring within three years of onset of the disease.[3]

2. Factor analysis of mixed data

As phenotype is a latent variable (i.e., a variable that cannot be directly observed but has to be inferred through a mathematical model), we used factor analysis of mixed data to model the phenotype of the patients. Factor analysis of mixed data is a principal component method dedicated to exploring data containing

both continuous and categorical variables. The continuous variables are scaled to unit variance and the categorical variables are transformed into a disjunctive data table and then scaled using the specific scaling of multiple correspondence analysis. This ensures that a balance between both continuous and categorical variables is maintained in determining the dimensions of variability. This method allows one to study the similarities between individuals taking into account mixed variables and to study the relationships between all the variables. It also provides graphical outputs such as the representation of the individuals, the correlation circle for the continuous variables and representations of the categories of the categorical variables, and also specific graphs to visualize the associations between both types of variables.

We used the package FactoMinerR v.2.1 to perform the factor analysis of mixed data and factoextra v. 1.0.6 to obtain the scree plot and the variable weight plot included in Supplementary Figure 2.

This was the code that was used to obtain the intermediate plot and export the factors to continue the analysis:

```
library("FactoMineR")
library("factoextra")
library("dplyr")

#Import the clinical variables, grouping variables (ATB_GROUP and EULAR_ACR_GROUP) and identifiers (ID)
setwd("project_dir")
df.total = read.csv("./famd.csv", header = TRUE)
df <- df.total[, !(names(df.total) %in% c("ID", "ATB_GROUP", "EULAR_ACR_GROUP"))]
res.famd <- FAMD(df, graph = FALSE)
eig.val <- get_eigenvalue(res.famd)

#Plot the scree plot
fviz_screeplot(res.famd)

#Plot the graph showing the weight of each variable for the two first factors
fviz_famd_var(res.famd, repel = TRUE)

#Obtain the values of the coordinates of the individuals on the factors
ind <- get_famd_ind(res.famd)
results <- ind$coord
results <- cbind(results, ID=df.total$ID)

#Output results to continue analysis
write.csv(results, "./factors.csv")
```

As input for the factor analysis of mixed data we selected a set of clinical, epidemiological, and laboratory parameters that were: 1) well documented in the

literature to be associated with the phenotype of patients with myositis, 2) systematically collected in our cohort, and 3) well defined. Thus, we included:

-Epidemiologic variables: Gender, race, age at onset.

-Clinical variables: presence or absence during the course of the disease of muscle weakness, interstitial lung disease, arthritis, heliotrope or Gottron's rash, calcinosis, Raynaud's phenomenon, mechanic's hands, dysphagia, and fevers.

-Laboratory values: Maximum creatine kinase (CK), presence of anti-Ro52 autoantibodies.

Anti-Ro52 autoantibodies were included because they are associated with the severity of the disease and with specific clinical features in patients with myositis.[4-7]

The detailed distribution of the muscle weakness, biopsy features, MRI patterns, or EMG findings were excluded from this section of the analysis because they were not available for all patients and restricting the sample size based on the availability of this data could bias the study.

3. Model comparison

We developed two distinct sets of linear regression models to predict each of the two selected components of the factor analysis. The first set used the different myositis-specific autoantibodies (Anti-SRP, HMGCR, Mi2, NXP2, TIF1g, MDA5, Jo1, PL7, PL12) as predicting variables and the second set used the adult 2017 EULAR/ACR categories: polymyositis/immune-mediated necrotizing myositis, inclusion body myositis, amyopathic dermatomyositis, dermatomyositis, and negative (i.e., not IM).

To compare the relative quality of the models using the myositis-specific autoantibodies and the 2017 EULAR/ACR categories we estimated the AIC and the BIC of each one of the linear regressions. Then, we subtracted the resulting estimates of the 2017 EULAR/ACR regression to the autoantibody regression both for factor 1 and factor 2. The model with the lowest AIC or BIC was considered superior and each one of the four comparisons (AIC for factors 1 and 2 and BIC for factors 1 and 2) was independent of each other.

Both the Akaike (AIC) and the Bayesian (BIC) information criteria are estimators of out-of-sample prediction error and thereby the relative quality of statistical models for a given set of data.[8, 9] Given a collection of models for the data, AIC and BIC estimate the quality of each model, relative to each of the other models. Thus, they provide a means for model selection.

The formula for AIC is:

$$\mathrm{AIC} = 2 - 2\ln(\hat{L})$$

Whereas BIC is formally defined as:

$$\mathrm{BIC} = \mathrm{k}\ln(\mathrm{n}) - 2\ln(\hat{L})$$

Where:

$\hat{L}$=the maximized value of the likelihood function of the model

n=the number of observations

k=the number of parameters estimated by the model

AIC and BIC are founded on information theory. When a statistical model is used to represent the process that generated the data, the representation will rarely be exact and so some information will be lost by using the model to represent the process. AIC and BIC estimate the relative amount of information lost by a given model: the less information a model loses, the higher the quality of that model.

In estimating the amount of information lost by a model, AIC and BIC deal with the trade-off between the goodness of fit of the model and the simplicity of the model. In other words, AIC and BIC deal with both the risk of overfitting and the risk of underfitting.

Given a set of candidate models for the data, the preferred model is the one with the minimum AIC or BIC value. Thus, AIC and BIC reward goodness of fit (as assessed by the likelihood function), but it also includes a penalty that is an increasing function of the number of estimated parameters. The penalty discourages overfitting, which is desired because increasing the number of parameters in the model almost always improves the goodness of the fit. In our particular case, this penalty would benefit the models built using the 2017 EULAR/ACR criteria (5 predicting variables) over those built using the myositis-specific autoantibodies (9 predicting variables).

The main difference between AIC and BIC is that the penalty for the number of parameters is larger for BIC. Thus, if ‘k ‘is the number of parameters and ‘n’ is the number of observations, the penalty for AIC is 2k, whereas for BIC is ln(n)k.

Different authors have suggested guidelines to interpret the magnitude of the differences in AIC and BIC between two models (DAIC and DBIC respectively). Thus, Burnham and Anderson[10] suggested that models having DAIC ≤ 2 have substantial support (evidence) to believe that are equivalent, those in which 4 ≤

DAIC ≤ 7 have considerably less support, and models having DAIC > 10 have essentially no support. Alternatively, Raftery[11] suggested that the evidence that two models are not equivalent would be weak with a DBIC between 0-2, positive with DBIC between 2-6, strong with DBIC between 6-10, and very strong with DBIC over 10. Based on this we selected a threshold of DAIC and DBIC to consider one model superior to others.

4. Graphical analysis of longitudinal data

A useful graphical method to represent the evolution of nonlinear parameters longitudinally is the locally weighted scatterplot smoothing (LOWESS). LOWESS fits simple models to localized segments of the data to build up a function that describes the deterministic part of the variation in the data, point by point.

The advantage of LOWESS over other methods (e.g. quadratic regression) is that it does not need to specify a function to fit the model to the data, making it simple and flexible for complex graphical representations. However, it requires a dense cloud of observations to be stable and it can be easily biased by outliers if the local density of data is low. Also, LOWESS does not return a simple mathematical function and, thus, is complicated to use for predictive purposes. Finally, LOWESS is relatively computationally intensive. Fortunately, with the range of observations that we used (in the hundreds) this was a negligible issue.

References

1. Rider LG, Werth VP, Huber AM, Alexanderson H, Rao AP, Ruperto N, et al. Measures of adult and juvenile dermatomyositis, polymyositis, and inclusion body myositis: Physician and Patient/Parent Global Activity, Manual Muscle Testing (MMT), Health Assessment Questionnaire (HAQ)/Childhood Health Assessment Questionnaire (C-HAQ), Childhood Myositis Assessment Scale (CMAS), Myositis Disease Activity Assessment Tool (MDAAT), Disease Activity Score (DAS), Short Form 36 (SF-36), Child Health Questionnaire (CHQ), physician global damage, Myositis Damage Index (MDI), Quantitative Muscle Testing (QMT), Myositis Functional Index-2 (FI-2), Myositis Activities Profile (MAP), Inclusion Body Myositis Functional Rating Scale (IBMFRS), Cutaneous Dermatomyositis Disease Area and Severity Index (CDASI), Cutaneous Assessment Tool (CAT), Dermatomyositis Skin Severity Index (DSSI), Skindex, and Dermatology Life Quality Index (DLQI). Arthritis Care Res (Hoboken). 2011 Nov; 63 Suppl 11:S118-157.

2. Travis WD, Costabel U, Hansell DM, King TE, Jr., Lynch DA, Nicholson AG, et al. An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. Am J Respir Crit Care Med. 2013 Sep 15; 188(6):733-748.

3. Troyanov Y, Targoff IN, Tremblay JL, Goulet JR, Raymond Y, Senecal JL. Novel classification of idiopathic inflammatory myopathies based on overlap syndrome features and autoantibodies: analysis of 100 French Canadian patients. Medicine (Baltimore). 2005 Jul; 84(4):231-249.

4. Pinal-Fernandez I, Casal-Dominguez M, Huapaya JA, Albayda J, Paik JJ, Johnson C, et al. A longitudinal cohort study of the anti-synthetase syndrome: increased severity of

interstitial lung disease in black patients and patients with anti-PL7 and anti-PL12 autoantibodies. Rheumatology (Oxford). 2017 Jun 1; 56(6):999-1007.

5. Vancsa A, Csipo I, Nemeth J, Devenyi K, Gergely L, Danko K. Characteristics of interstitial lung disease in SS-A positive/Jo-1 positive inflammatory myopathy patients. Rheumatol Int. 2009 Jul; 29(9):989-994.

6. Marie I, Hatron PY, Dominique S, Cherin P, Mouthon L, Menard JF, et al. Short-term and long-term outcome of anti-Jo1-positive patients with anti-Ro52 antibody. Semin Arthritis Rheum. 2012 Jun; 41(6):890-899.

7. Bauhammer J, Blank N, Max R, Lorenz HM, Wagner U, Krause D, et al. Rituximab in the Treatment of Jo1 Antibody-associated Antisynthetase Syndrome: Anti-Ro52 Positivity as a Marker for Severity and Treatment Response. J Rheumatol. 2016 Aug; 43(8):1566-1574.

8. Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974; 19(6):716-723.

9. Schwarz G. Estimating the Dimension of a Model. Ann Statist. 1978 1978/03; 6(2):461-464.

10. Burnham KP, Anderson DR. Multimodel Inference:Understanding AIC and BIC in Model Selection. Sociological Methods & Research. 2004; 33(2):261-304.

11. Raftery AE. Bayesian model selection in social research. Sociological Methodology 1995, Vol 25. 1995; 25:111-163.

A.7 Myositis autoantigen expression correlates with muscle regeneration but not autoantibody specificity. Arthritis Rheum 2019 (PMID: 30861336).

Although more than a dozen MSAs have been identified, most patients with myositis are positive for a single MSAs. The specific overexpression of a given myositis autoantigen in myositis muscle has been proposed as initiating and/or propagating autoimmunity against that particular autoantigen. The present study was undertaken to test this hypothesis.

In order to quantify autoantigen RNA expression, RNA sequencing was performed on muscle biopsy samples from control subjects, MSAs-positive patients with myositis, regenerating mouse muscles, and cultured human muscle cells.

Muscle biopsy samples were available from 20 control subjects and 106 patients with autoantibodies recognizing HMGCR (n = 40), SRP (n = 9), Jo-1 (n = 18), NXP2 (n = 12), Mi-2 (n = 11), TIF1γ (n = 11), or MDA5 (n = 5). The increased expression of a given autoantigen in myositis muscle was not associated with autoantibodies recognizing that autoantigen (all $q > 0.05$). In biopsy specimens from both myositis muscle and regenerating mouse muscles, autoantigen expression correlated directly with the expression of muscle regeneration markers and correlated inversely with the expression of genes encoding mature muscle proteins. Myositis autoantigens were also expressed at high levels in cultured human muscle cells.

In conclusion, most myositis autoantigens are highly expressed during muscle regeneration, which may relate to the propagation of autoimmunity. However, factors other than overexpression of specific autoantigens are likely to govern the development of unique autoantibodies in individual patients with myositis.

Myositis autoantigen expression correlates with muscle regeneration but not autoantibody specificity

Iago Pinal-Fernandez*, David R Amici*, Cassie A Parks, Assia Derfoul, Maria Casal-Dominguez, Katherine Pak, Richard Yeker, Paul Plotz, Jose C Milisenda, Josep M Grau-Junyent, Albert Selva-O'Callaghan, Julie Paik, Jemima Albayda, Andrea M Corse, Thomas E Lloyd, Lisa Christopher-Stine, Andrew L Mammen

*** These authors contributed equally to this project.**

Iago Pinal-Fernandez M.D., Ph.D., David R Amici B.S., Cassie A Parks B.S., Assia Derfoul Ph.D., Richard Yeker B.S., Katherine Pak M.D., Maria Casal-Dominguez M.D., Ph.D., Paul Plotz M.D., Ph.D., and Andrew L. Mammen, M.D., Ph.D.: National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD. Iago Pinal-Fernandez M.D., Ph.D., Maria Casal-Dominguez M.D., Ph.D., Julie J Paik M.D., Jemima Albayda M.D., Andrea Corse M.D., Thomas E. Lloyd M.D., Ph.D., Lisa Christopher-Stine M.D., M.P.H. and Andrew L. Mammen M.D., Ph.D.: Johns Hopkins University School of Medicine, Baltimore, MD. Jose C Milisenda M.D. and Josep M Grau-Junyent M.D., Ph.D.: Clinic Hospital and the University of Barcelona, Barcelona, Spain. Albert Selva-O'Callaghan M.D., Ph.D.: Vall d'Hebron Hospital and Autonomous University of Barcelona, Spain. Iago Pinal-Fernandez M.D., Ph.D.: Faculty of Health Sciences, Universitat Oberta de Catalunya, Barcelona, Spain.

Address correspondence to: Andrew L. Mammen, M.D., Ph.D., or Iago Pinal-Fernandez, M.D., Ph.D. Muscle Disease Unit, Laboratory of Muscle Stem Cells and Gene Expression, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, 50 South Drive, Room 1141, Building 50, MSC 8024, Bethesda, MD 20892. E-mail: andrew.mammen@nih.gov or iago.pinalfernandez@nih.gov. Phone: 301-451-1199. Fax: 301-594-0305.

FUNDING: This research was supported in part by the Intramural Research Program of the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health. The Myositis Research Database is supported by the the Huayi and Siuling Zhang Discovery Fund. IPF´s research was supported by a Fellowship from the Myositis Association.

COMPETING INTERESTS: None

Word count: 1756 words

Abstract

Objectives: Although more than a dozen myositis-specific autoantibodies (MSAs) have been identified, most myositis patients produce a single MSA. The specific overexpression of a given myositis autoantigen in myositis muscle has been proposed to initiate and/or propagate autoimmunity against that autoantigen. To test this hypothesis, we quantified autoantigen RNA expression in myositis muscle biopsies, regenerating mouse muscles, and cultured human muscle cells.

Methods: RNA-sequencing was performed on muscle biopsies from control subjects, muscle biopsies from MSA-positive myositis patients, regenerating mouse muscles, and cultured human muscle cells.

Results: Muscle biopsies were available from 20 control subjects and 106 patients with autoantibodies recognizing HMGCR (n=40), SRP (n=9), Jo1 (n=18), NXP2 (n=12), Mi2 (n=11), TIF1γ (n=11), or MDA5 (n=5). The increased expression of a given autoantigen in myositis muscle was not associated with autoantibodies recognizing that autoantigen (all q>0.05). In both myositis muscle biopsies and regenerating mouse muscles, myositis autoantigen expression correlated directly with the expression of markers of muscle regeneration and inversely with the expression of genes encoding mature muscle proteins. Myositis autoantigens were also expressed at high levels in cultured human muscle cells**.**

Conclusions: Most myositis autoantigens are highly expressed during muscle regeneration, which may relate to the propagation of autoimmunity. However, factors

other than specific autoantigen overexpression are likely to govern the development of unique autoantibodies in individual patients.

MESH Keywords: Myositis, Autoantibodies, Autoantigens, Skeletal Muscle, Regeneration

Key messages:

- Myositis autoantigen expression correlates directly with the expression of markers of muscle regeneration and inversely with the expression of genes encoding mature muscle proteins in human myositis muscle biopsies and regenerating mouse muscle.
- Myositis autoantigens are highly expressed during muscle differentiation in cultured human muscle cells**.**
- The expression of a given autoantigen in myositis muscle was not associated with autoantibodies recognizing that autoantigen.

Introduction

Myositis is a heterogeneous group of diseases that includes dermatomyositis (DM), immune-mediated necrotizing myopathy (IMNM), the antisynthetase syndrome and inclusion body myositis.(1) Each type of myositis is characterized by weakness, elevated muscle enzymes, and muscle biopsies featuring inflammation, necrotic myofibers, and/or regenerating muscle cells.(2) Most myositis patients also have only one of more than a dozen myositis-specific autoantibodies (MSAs), each of which is associated with a unique clinical phenotype.(3)

Regenerating muscle cells in myositis patient muscle biopsies have been shown to express high levels of several myositis autoantigens, including Mi2, TIF1γ, Jo1, HMGCR and SRP.(4-8) Given this observation, it has been proposed that increased expression of myositis autoantigens may initiate and/or maintain autoimmunity against these proteins. However, it has not been determined if autoantigens other than Mi2, TIF1γ and Jo1 are expressed at high levels in regenerating muscle, if autoantigen expression patterns differ between myositis subgroups (e.g., IMNM vs DM), or if a relationship exists between an autoantigen's expression level and the presence of the corresponding autoantibody. Here, we have addressed these issues by using high throughput next-generation sequencing to quantify the expression of myositis autoantigens and other genes in muscle biopsies from patients with defined MSAs.

Methods

Patient Samples and Autoantibody Testing

Muscle biopsies from patients enrolled in the longitudinal cohorts of the National Institutes of Health (Bethesda), the Johns Hopkins Myositis Center (Baltimore) the Clinic Hospital (Barcelona), and the Vall d'Hebron Hospital (Barcelona) were included in the study if the patients had one of the following MSAs: anti-HMGCR, -SRP, -Jo1, -NXP2, -Mi2, or - TIF1γ or -MDA5. Autoantibody testing was performed as previously described for anti-HMGCR (9) and by line blot for the others (EUROLINE Myositis Profile 4). Normal muscle biopsies were obtained from the Johns Hopkins Neuromuscular Pathology Laboratory (n=10) and the Skeletal Muscle Biobank of the University of Kentucky (n=10).

Cultured Human Skeletal Muscle Cells

Normal human skeletal muscle myoblasts (HSMM; Lonza) were cultured according to the provider's protocol. When 80% confluent, the cultures were induced to differentiate into myotubes by replacing the growth medium with differnation medium (DMEM, 2% horse serum, and L-glutamine). Two plates of cells were harvested before differentiaton and then daily for 6 days.

Mouse Muscle Injury

Muscle injury and regeneration were induced in mice using cardiotoxin as previously described (5). Briefly, C57BL/6 mice were unilaterally injured via

intramuscular (tibialis anterior) injection of 0.1 mL of 10 μM cardiotoxin (CTX). Muscle was harvested at days 3 (n=2), 5 (n=2), 7 (n=2), 10 (n=4), 14 (n=4), and 28 (n=3) post-injury. Contralateral (uninjured) tibialis anterior muscles were also collected (n=9). All the animal experiments were performed according to the NIH ethical guidelines.

RNA-sequencing

RNA-sequencing (RNA-seq) was performed as previously described (10). Briefly, muscle biopsies were homogenized in TRIzol using 1.4 mm ceramic bead low-binding tubes and the RNA was extracted using the regular TRIzol protocol. Concentration and quality of the resulting RNA was assessed using standard NanoDrop and TapeStation protocols, respectively. The median RIN value of the muscle biopsy samples RNA was 7 (interquartile range [IQR] 5.9-7.4) while for the cultured human skeletal muscle cells was 9.5 (IQR 9.1-9.6). Libraries were prepared using the NeoPrepTM system according to the TruSeqM Stranded mRNA Library Prep protocol (Illumina) and sequenced using the Illumina HiSeq 2500 or 3000. Reads were aligned using the STAR v.2.5 (11) and the abundance of each gene was quantified using StringTie v.1.3.3.(12) and the differential gene expression was performed using DESeq2 v.1.20.0.(13) The Benjamini-Hochberg correction was used to adjust for multiple comparisons and a corrected p-value (q-value) of 0.05 or less was considered statistically significant.

Data Analysis

Gene-expression (Transcripts Per Kilobase Million [TPM]) values were log-transformed (logTPM: log2[TPM+1]) or referenced to the normal biopsies (log2[fold

change]) and data were processed and visualized using the Python programming language, using the packages Numpy, Pandas and Seaborn. Spearman rho was used to quantify the correlation between genes of interest.

Results

Myositis autoantigen RNA expression correlates with muscle regeneration in myositis muscle biopsies

To assess the myositis autoantigen expression in patients with different MSAs, RNAseq was performed on 20 normal biopsies and muscle biopsy specimens from 9 anti-SRP, 40 anti-HMGCR, 11 anti-Mi2, 12 anti-NXP2, 11 anti- TIF1γ, 5 anti-MDA5 and 18 anti-Jo1 patients. There was no correlation between the autoantibody produced by a patient and the RNA expression level of the corresponding autoantigen (all $q>0.05$). For example, muscles from anti-Mi2-positive patients did not have increased CHD4 RNA (encoding the Mi2 autoantigen) expression compared to patients with other MSAs (Figure 1). Of note, the expression of IFIH1 (encoding the MDA5 autoantigen), which is an interferon inducible gene, was elevated in all DM autoantibody groups (i.e., anti-Mi2, anti-NXP2, anti-TIF1γ and anti-MDA5) compared to the anti-SRP- or anti-HMGCR-positive patients. In general, all myositis autoantigens except for NXP2 and TIF1γ were expressed at higher levels in biopsies from myositis patients than from control subjects (Supplementary Figure 1).

Since prior studies showed that TIF1γ (14) and Mi2 (5 6) are highly expressed in regenerating myofibers, we sought to determine whether there was a correlation between the expression of other myositis autoantigens and the expression of genes associated with regenerating muscle fibers. Indeed, among patients with all MSAs, we found a positive correlation between RNA expression of each myositis autoantigen and the expression of regeneration genes (i.e., myogenin [MYOG], MyoD, PAX7 and the perinatal [MYH3] and embryonic [MYH8] myosin heavy chains). In contrast, there was an inverse correlation between the RNA expression of genes encoding myositis

autoantigens and markers of mature muscle (i.e., skeletal muscle actin [ACTA1] and the adult skeletal muscle myosin heavy chains [MYH1 and MYH2]) (Figure 2a).

Myositis autoantigens correlate with muscle regeneration in regenerating mouse muscle

Myositis muscle may include mature myofibers as well as muscle cells at various stages of degeneration and regeneration. To define myositis autoantigen RNA expression during muscle regeneration, we utilized a mouse model in which muscle is injured with cardiotoxin and then allowed to regenerate; in this model, myoblast proliferation, myocyte differentiation, and myotube formation are synchronized. As in myositis muscle, myositis autoantigen RNA expression in the mouse muscles were positively correlated with markers of muscle differentiation and inversely with the expression of adult muscle genes (Figure 2b). Indeed, the RNA expression of all myositis autoantigens increased after cardiotoxin injection (Figure 3) and paralleled the expression of genes associated with muscle regeneration (MYOG, MyoD, PAX7, MYH3, and MYH8), which transiently increased after the muscle injury (Supplementary Figure 2). The expression levels of genes encoding mature muscle proteins (ACTA1, MYH1, and MYH2) transiently decreased after injury and subsequently increased following muscle repair (Supplementary Figure 2).

Myositis autoantigens are expressed at high levels of cultured human myoblasts

Myositis muscle biopsies may include infiltrating macrophages and T cells along with regenerating myofibers. Similarly, macrophages and T cells infiltrate regenerating mouse muscle where they remove cellular debris and promote muscle repair. Not

surprisingly, T cell- and macrophage-specific gene expression was correlated with the levels of myositis autoantigens in both myositis muscle and regenerating mouse muscles (Figure 2).

To confirm that myositis autoantigens are expressed at high levels in muscle cells rather than exclusively by inflammatory cells, we analyzed myositis autoantigen RNA expression in human myoblast cultures as they proliferated and then differentiated into myotubes. While genes specifically expressed by inflammatory cells were not present in the cultured muscle cells (Supplementary Figure 3), markers of muscle regeneration were expressed at levels equivalent to those measured in regenerating mouse muscles (Supplementary Figure 2). Moreover, all myositis autoantigens were expressed at levels equal to or higher than observed in regenerating mouse muscle (Figure 3). Taken together, these results demonstrate that proliferating myoblasts, differentiating myocytes, and newly formed myotubes contribute substantially to the expression levels of myositis antigens in myositis muscle biopsies.

Discussion

As several myositis autoantigens (i.e., Mi2, TIF1γ and Jo1) were previously shown to be expressed at high levels in regenerating muscle cells, it has been proposed that the overexpression of specific autoantigens in myositis muscle might drive the autoantigen-specific immune response.(6) In this study, we used RNAseq to systematically investigate autoantigen expression levels in muscle biopsies from myositis patients with each major MSA. We found that RNA levels of each myositis autoantigen are positively correlated with markers of muscle regeneration but that the levels of a given autoantigen are not associated with the presence of the corresponding autoantibody. Therefore, restricted autoantigen overexpression alone does not account for why myositis patients typically produce only a single MSA. Rather, it is likely that factors such as aberrant post-translational processing,(15) mislocalization of autoantigen, immunogenetic susceptibility,(16) and/or exposure to molecularly similar antigens (e.g. tumor antigens(17) determine which autoantigens will be targeted by the immune system in myositis patients.

We also showed that all myositis autoantigens are expressed at high levels not just in regenerating myositis muscle, but also in regenerating mouse muscles and in cultured human myoblasts. This indicates that elevated myositis autoantigen expression is a normal part of muscle regeneration/differentiation. Nonetheless, disease-related factors may also contribute to the myositis autoantigen overexpression. For example, *IFIH1* is expressed at low levels (<2) during all phases of cultured muscle cell differentiation compared to the expression levels of other myositis autoantigens (4-6). However, *IFIH1* is expressed at especially high levels in muscle biopsies from

patients with DM autoantibodies. Since interferon (IFN) levels are high in DM patients (18) and *IFIH1* is an IFN-inducible gene, we hypothesize that muscle regeneration and IFN both contribute to the high levels of *IFIH1* in DM muscle biopsies.

The primary limitation of this study is that we relied on RNA quantitation to assess gene expression levels. However, the utilization of high-throughput next-generation sequencing is also what allowed us to analyze the expression of many genes in each of many samples. For example, figure 2 summarizes the expression levels of 20 genes in over 106 myositis muscle biopsies and 26 mouse muscle specimens. Such an analysis would be impractical using immunoblotting techniques to quantify protein expression levels. Furthermore, we and others have previously shown that Mi2, TIF1γ, Jo1, HMGCR and SRP proteins are upregulated in regenerating cells of myositis muscle biopsies,(4-8) validating a correlation between RNA and protein levels for these autoantigens.

In summary, by utilizing RNAseq to quantitate autoantigen expression in a large number of myositis muscle biopsies from patients with defined autoantibodies, we have demonstrated that autoantigen expression is highly correlated with muscle regeneration but that expression of a given autoantigen is not associated with the presence of the corresponding autoantibody. Future studies will be required to determine why only one autoantigen is typically targeted by the immune system in a given myositis patient.

ACKNOWLEDGMENTS: Dr. Gustavo Gutierrez-Cruz, Dr. Stefania Dell'Orso and Faiza Naz from the NIAMS sequencing facility for all technical collaboration in RNAseq library construction and sequencing. The University of Kentucky Center for Muscle Biology provided comtrol human muscle samples.

REFERENCES

1. Selva-O'Callaghan A, Pinal-Fernandez I, Trallero-Araguas E, Milisenda JC, Grau-Junyent JM, Mammen AL. Classification and management of adult inflammatory myopathies. Lancet Neurol 2018;17:816-28.
2. Dalakas MC. Inflammatory Muscle Diseases. N Engl J Med 2015;373:393-4.
3. Mammen AL, Casciola-Rosen L, Christopher-Stine L, Lloyd TE, Wagner KR. Myositis-specific autoantibodies are specific for myositis compared to genetic muscle disease. Neurol Neuroimmunol Neuroinflamm 2015;2:e172.
4. Mohassel P, Rosen P, Casciola-Rosen L, Pak K, Mammen AL. Expression of the dermatomyositis autoantigen transcription intermediary factor 1γ in regenerating muscle. Arthritis Rheumatol 2015;67:266-72.
5. Mammen AL, Casciola-Rosen LA, Hall JC, Christopher-Stine L, Corse AM, Rosen A. Expression of the dermatomyositis autoantigen Mi-2 in regenerating muscle. Arthritis Rheum 2009;60:3784-93.
6. Casciola-Rosen L, Nagaraju K, Plotz P, Wang K, Levine S, Gabrielson E, et al. Enhanced autoantigen expression in regenerating muscle cells in idiopathic inflammatory myopathy. J Exp Med 2005;201:591-601.
7. Arouche-Delaperche L, Allenbach Y, Amelin D, Preusse C, Mouly V, Mauhin W, et al. Pathogenic role of anti-signal recognition protein and anti-3-Hydroxy-3-methylglutaryl-CoA reductase antibodies in necrotizing myopathies: Myofiber atrophy and impairment of muscle regeneration in necrotizing autoimmune myopathies. Ann Neurol 2017;81:538-48.
8. Allenbach Y, Arouche-Delaperche L, Preusse C, Radbruch H, Butler-Browne G, Champtiaux N, et al. Necrosis in anti-SRP(+) and anti-HMGCR(+)myopathies: Role of autoantibodies and complement. Neurology 2018;90:e507-e17.
9. Mammen AL, Chung T, Christopher-Stine L, Rosen P, Rosen A, Doering KR, et al. Autoantibodies against 3-hydroxy-3-methylglutaryl-coenzyme A reductase in patients with statin-associated autoimmune myopathy. Arthritis Rheum 2011;63:713-21.
10. Amici DR, Pinal-Fernandez I, Mázala DA, Lloyd TE, Corse AM, Christopher-Stine L, et al. Calcium dysregulation, functional calpainopathy, and endoplasmic reticulum stress in sporadic inclusion body myositis. Acta Neuropathol Commun 2017;5:24.
11. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15-21.
12. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 2015;33:290-5.
13. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
14. Mohassel P, Rosen P, Casciola-Rosen L, Pak K, Mammen AL. Expression of the dermatomyositis autoantigen transcription intermediary factor 1gamma in regenerating muscle. Arthritis Rheumatol 2015;67:266-72.

15. Levine SM, Raben N, Xie D, Askin FB, Tuder R, Mullins M, et al. Novel conformation of histidyl-transfer RNA synthetase in the lung: the target tissue in Jo-1 autoantibody-associated myositis. Arthritis Rheum 2007;56:2729-39.
16. Miller FW, Chen W, O'Hanlon TP, Cooper RG, Vencovsky J, Rider LG, et al. Genome-wide association study identifies HLA 8.1 ancestral haplotype alleles as major genetic risk factors for myositis phenotypes. Genes Immun 2015;16:470-80.
17. Pinal-Fernandez I, Ferrer-Fabregas B, Trallero-Araguas E, Balada E, Martinez MA, Milisenda JC, et al. Tumour TIF1 mutations and loss of heterozygosity related to cancer-associated myositis. Rheumatology (Oxford) 2018;57:388-96.
18. Greenberg SA, Pinkus JL, Pinkus GS, Burleson T, Sanoudou D, Tawil R, et al. Interferon-alpha/beta-mediated innate immune mechanisms in dermatomyositis. Ann Neurol 2005;57:664-78.

Figure 1: Mean RNA expression levels of the different myositis autoantigens (columns) in muscle biopsies from control patients and patients from the different autoantibody groups (rows). For example, SRP54 (first column) is expressed at uniformly high levels in all autoantibody groups and not higher in anti-SRP positive patients (cell with a white circle in the first column). Of note, the anti-MDA5 autoantigen (IFIH1) is highly expressed in patients with dermatomyositis (anti-Mi2, -NPX2, -TIF1γ, and -MDA5), moderately expressed in patients with anti-Jo1 autoantibodies and expressed at low levels in immune-mediated necrotizing myositis (anti-SRP and anti-HMGCR). White circles highlight the expression levels of each autoantigen in its corresponding autoantibody group.

TPM: Transcripts Per Kilobase Million.

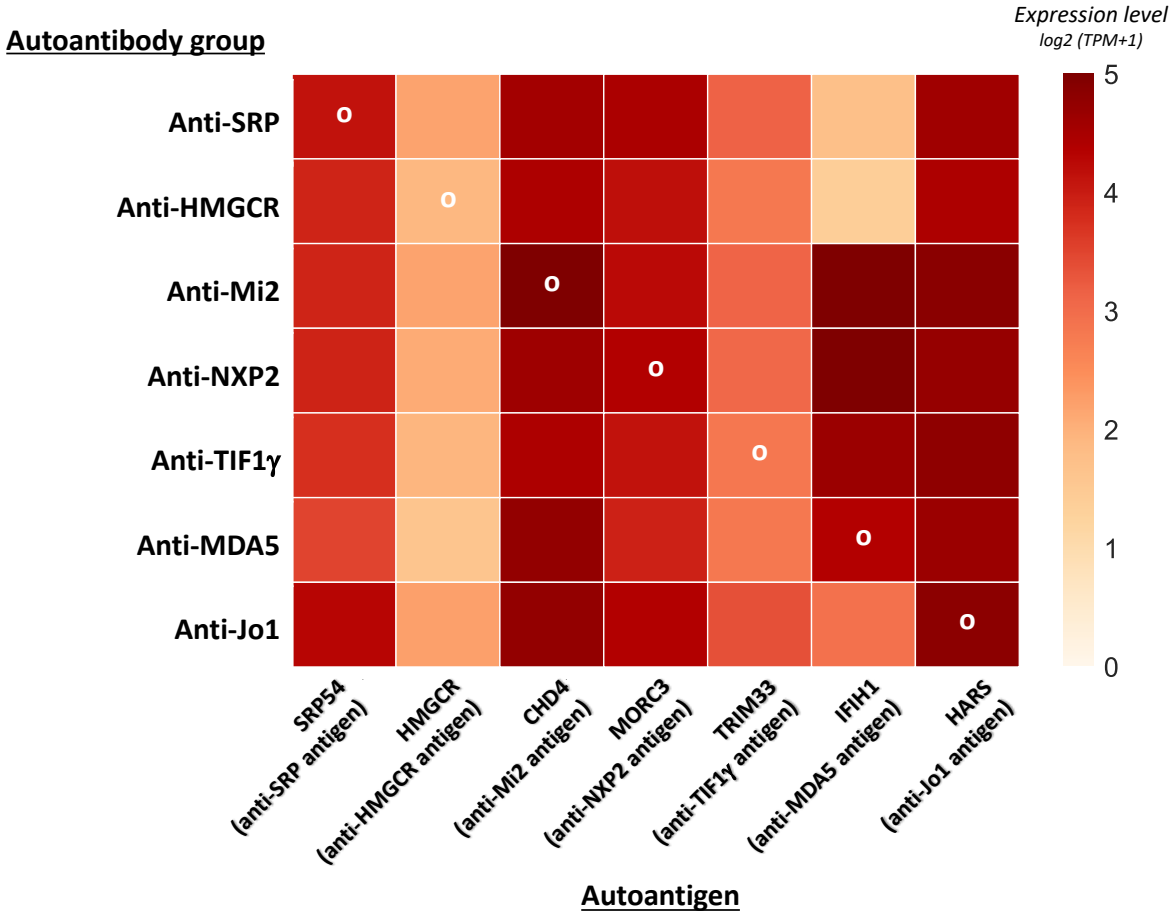


Figure 2: Matrices show correlation coefficients for the RNA expression of myositis autoantigens (rows) versus the expression of genes found in T cells (*CD3E*, *CD4*, and *CD8A*), macrophages (*CD14*), regenerating muscle cells (*NCAM1, MYOG, MYOD1, PAX7, MYH3, MYH8*), and mature muscle cells (*ACTA1*, *MYH1*, and *MYH2*)(columns) in (a) myositis muscle biopsies and (b) regenerating mouse muscles (all the time points were pooled for the analysis). Red denotes a positive correlation and blue denotes a negative correlation.

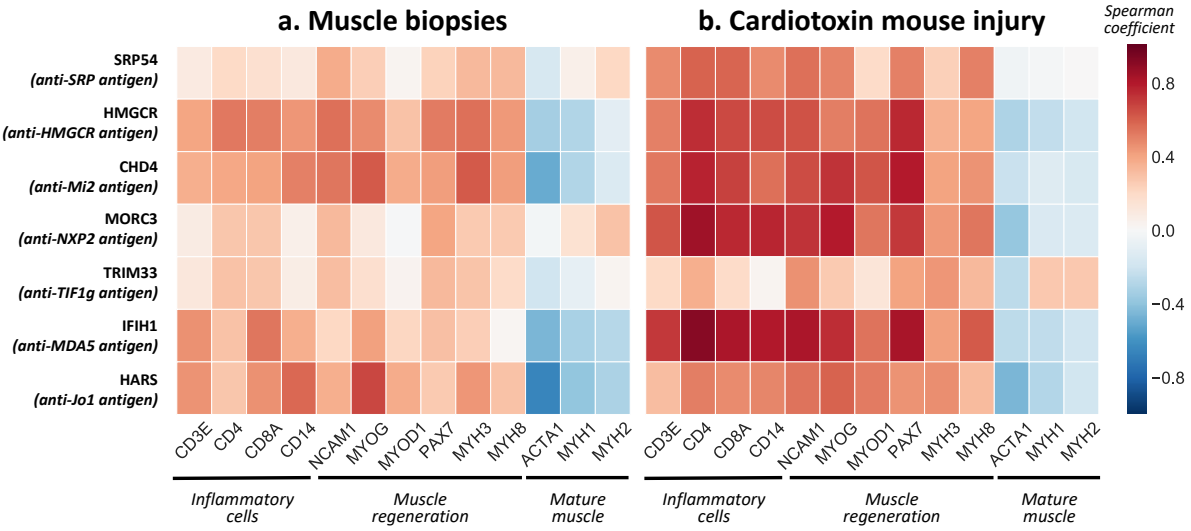
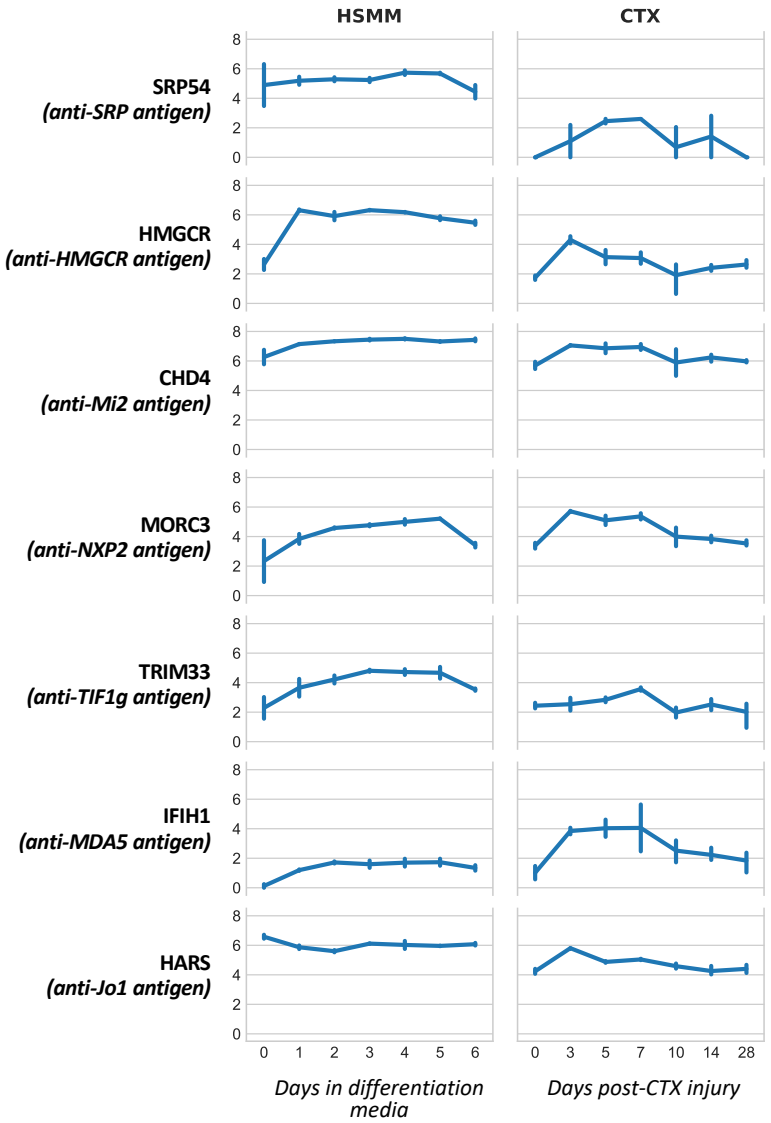
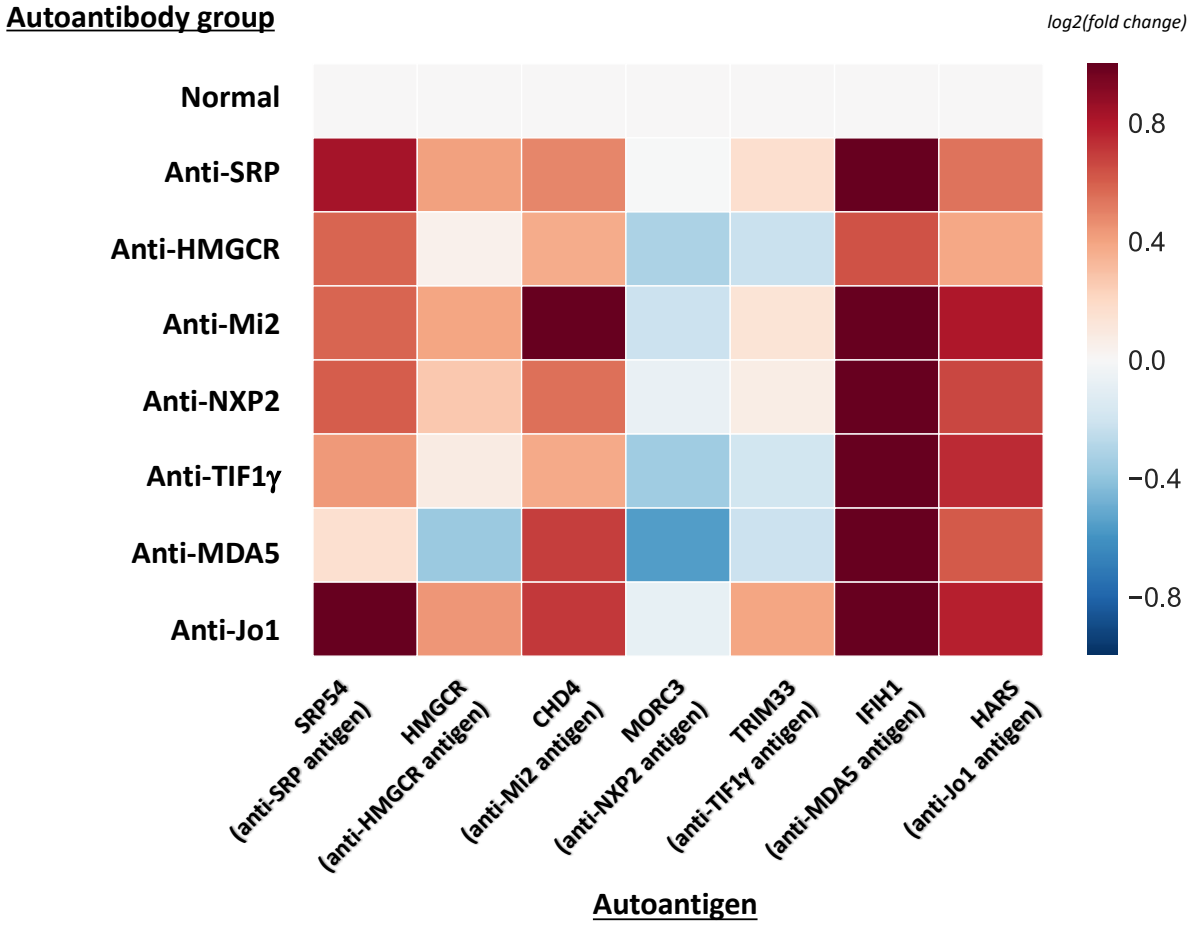


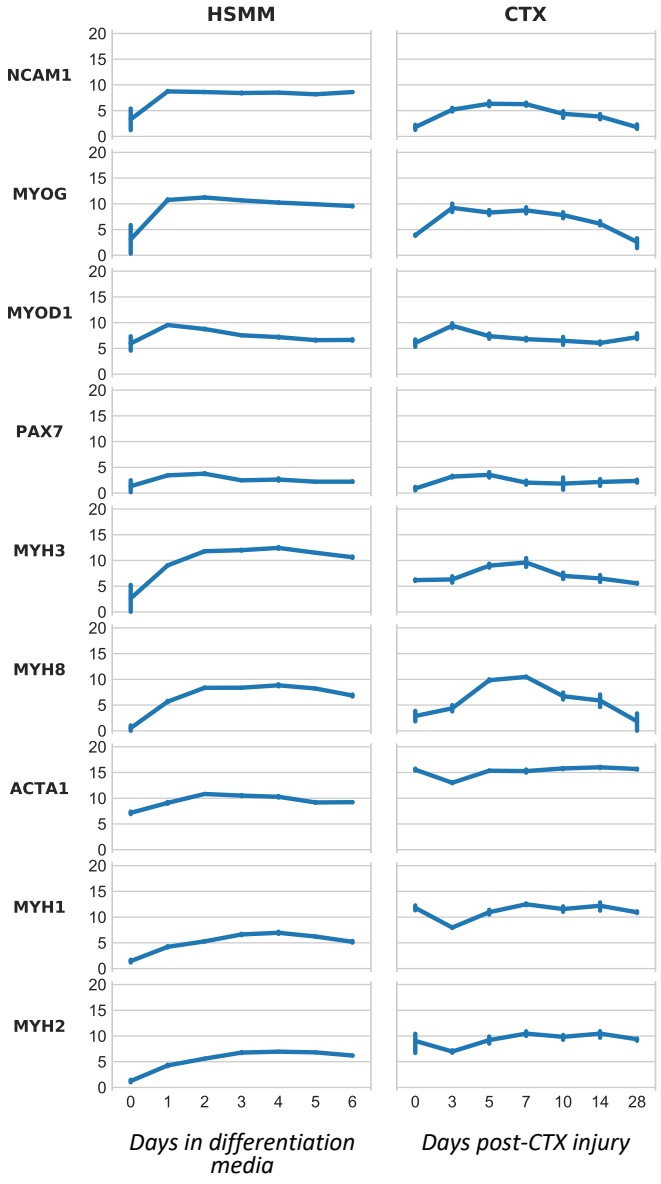
Figure 3: Evolution of the RNA expression levels (log_2[TPM+1]) of the different myositis autoantigens (rows) as human skeletal muscle myoblasts (HSMM) differentiate into myotubes and during the regeneration of mouse muscle following injury with cardiotoxin (CTX). In the HSMM model, proliferating myoblasts are placed in differentiation media at day 0 and allowed to differentiate into myotubes over the next 6 days. In the mouse injury model, the tibialis anterior muscle is injected with CTX at day 0 and the muscle is allowed to regenerate for as long as 28 days. Vertical lines indicate the 95% confidence interval for each value. CTX day 0 corresponds to the contralateral (uninjured) tibialis anterior muscle.



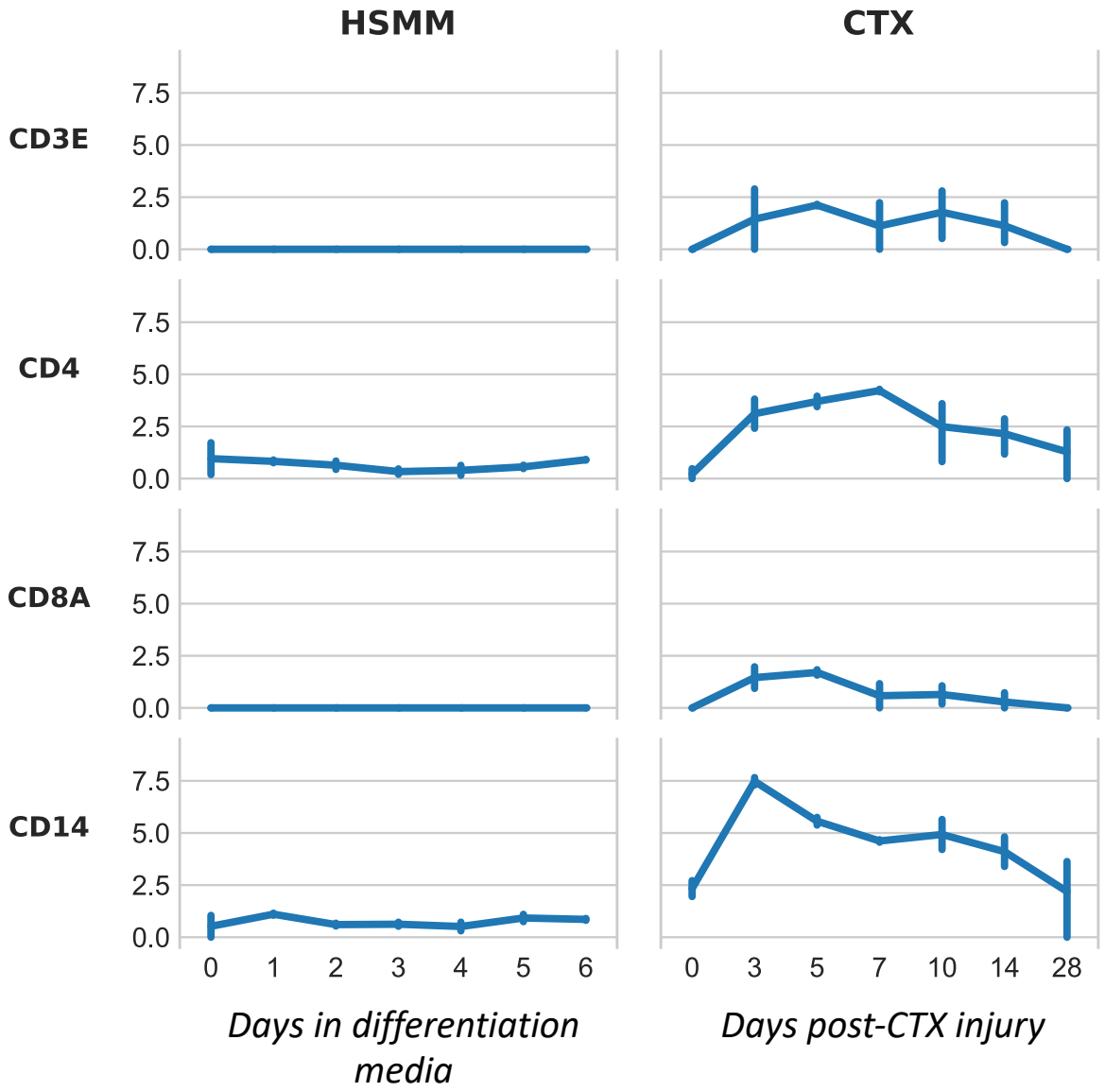
Supplementary Figure 1: Fold change of the RNA expression levels of the different myositis autoantigens (columns) in the different autoantibody groups (rows) compared to normal tissue (first row). Except for MORC3 (the anti-NXP2 antigen) and TRIM33 (anti-TIF1γ), the RNA expression of myositis autoantigens was higher in biopsies from patients with myositis than in normal tissue.



Supplementary Figure 2. Evolution of the RNA expression levels (log_2[TPM+1]) of genes expressed by differentiating myoblasts (*NCAM, MYOG, MYOD1, PAX7, MYH3*, and *MYH8*) and by mature myofibers (*ACTA1, MYH1*, and *MYH2*) (rows) during muscle differentiation in human skeletal muscle myoblasts (HSMM) and in the cardiotoxin mouse model (CTX). Markers of regenerating and mature muscle show similar expression levels in the cultured muscle cells and in the regenerating muscle tissue. Vertical lines indicate the 95% confidence interval of each individual value.



Supplementary Figure 3. Evolution of the RNA expression levels (log_2[TPM+1]) of inflammatory cell markers (rows) during muscle differentiation in human skeletal muscle myoblasts (HSMM) and in the cardiotoxin (CTX) mice model. Genes expressed by T cells (*CD3E, CD4*, and *CD8A*) and macrophages (*CD14*) are expressed at very low levels in HSMMs whereas they transiently increase after muscle injury in the CTX model. Vertical lines indicate the 95% confidence interval of each individual value.



A.8 Identification of distinctive interferon gene signatures in different types of myositis. Neurology 2019 (PMID: 31434690).

Activation of the type 1 IFN pathway is a prominent feature of DM muscle and may play a role in the pathogenesis of this disease. However, the relevance of the IFN1 pathway in patients with other types of myositis such as the AS, IMNM, and IBM is largely unknown. Moreover, the activation of the type 2 IFN pathway has not been comprehensively explored in myositis. In this cross-sectional study, our objective was to determine whether IFN1 and IFN2 pathways are differentially activated in different types of myositis by performing RNA sequencing on muscle biopsy samples from 119 patients with DM, IMNM, AS, or IBM and on 20 normal muscle biopsies.

The expression of IFN1-inducible genes was high in DM, moderate in AS, and low in IMNM and IBM. In contrast, the expression of IFN2-inducible genes was high in DM, IBM, and AS but low in IMNM. The expression of IFN-inducible genes correlated with the expression of genes associated with inflammation and muscle regeneration. Of note, ISG15 expression levels alone performed as well as composite scores relying on multiple genes to monitor activation of the IFN1 pathway in myositis muscle biopsies.

In conclusion, IFN1 and IFN2 pathways are differentially activated in different forms of myositis. This observation may have therapeutic implications because immunosuppressive medications may preferentially target each of these pathways.

Identification of distinctive interferon gene signatures in different types of inflammatory myopathy

Iago Pinal-Fernandez M.D., Ph.D., Maria Casal-Dominguez M.D., Assia Derfoul Ph.D., Katherine Pak M.D., Paul Plotz M.D., Frederick W Miller M.D., Ph.D., Jose C Milisenda M.D., Josep M Grau-Junyent M.D., Ph.D., Albert Selva-O'Callaghan M.D., Ph.D., Julie Paik M.D., Jemima Albayda M.D., Lisa Christopher-Stine M.D., M.P.H., Thomas E. Lloyd M.D., Ph.D., Andrea M Corse M.D., Andrew L Mammen M.D., Ph.D.

Iago Pinal-Fernandez M.D., Ph.D., Maria Casal-Dominguez M.D., Ph.D., Assia Derfoul Ph.D., Katherine Pak M.D., Paul Plotz M.D., Frederick W Miller, M.D., Ph.D., and Andrew L. Mammen, M.D., Ph.D.: National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD. Iago Pinal-Fernandez M.D., Ph.D., Maria Casal-Dominguez M.D., Ph.D., Julie J Paik M.D., Jemima Albayda M.D., Lisa Christopher-Stine M.D., M.P.H., Thomas E. Lloyd M.D., Ph.D., Andrea Corse M.D. and Andrew L. Mammen M.D., Ph.D.: Johns Hopkins University School of Medicine, Baltimore, MD. Jose C Milisenda M.D. and Josep M Grau-Junyent M.D., Ph.D.: Clinic Hospital and the University of Barcelona, Barcelona, Spain. Albert Selva-O'Callaghan M.D., Ph.D.: Vall d'Hebron Hospital and Autonomous University of Barcelona, Spain. Iago Pinal-Fernandez: Faculty of Health Sciences, Universitat Oberta de Catalunya, Barcelona, Spain.

Funding sources: This research was supported in part by the Intramural Research Program of the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health. The Johns Hopkins Myositis Center Research Database and Dr. LC-S are supported by the Huayi and Siuling Zhang Discovery Fund. IPF´s research was supported by a Fellowship from the Myositis Association.

Acknowledgments: Dr. Gustavo Gutierrez-Cruz, Dr. Stefania Dell'Orso and Faiza Naz from the NIAMS sequencing facility for all technical collaboration in RNAseq library construction and sequencing. The University of Kentucky Center for Muscle Biology provided comtrol human muscle samples.

Address correspondence to: Andrew L. Mammen, M.D., Ph.D., or Iago Pinal-Fernandez, M.D., Ph.D. Muscle Disease Unit, Laboratory of Muscle Stem Cells and Gene Regulation, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, 50 South Drive, Room 1141, Building 50, MSC 8024, Bethesda, MD 20892. E-mail: andrew.mammen@nih.gov or iago.pinalfernandez@nih.gov. Phone: 301-451-1199. Fax: 301-594-0305.

ABSTRACT

Objective: Activation of the type 1 interferon (IFN1) pathway is a prominent feature of dermatomyositis (DM) muscle and may play a role in the pathogenesis of this disease. However, the relevance of the IFN1 pathway in patients with other types of myositis, such as the antisynthetase syndrome (AS), immune-mediated necrotizing myopathy (IMNM), and inclusion body myositis (IBM), is largely unknown. Moreover, the activation of the type 2 interferon (IFN2) pathway has not been comprehensively explored in myositis. In this cross-sectional study, our objective was to analyze both IFN1 and IFN2 pathway activation in myositis by performing RNAseq on muscle biopsy samples from 119 patients with DM, IMNM, AS, or IBM as well as on 20 normal muscle biopsies.

Methods: The expression of IFN1- and IFN2-inducible genes was compared between the different groups.

Results: The expression of IFN1-inducible genes was high in DM, moderate in AS, and low in IMNM and IBM. In contrast, the expression of IFN2-inducible genes was high in DM, IBM, and AS but low in IMNM. The expression of IFN-inducible genes correlated with the expression of genes associated with inflammation and muscle regeneration. Of note, ISG15 expression levels alone performed as well as composite scores relying on multiple genes to monitor activation of the IFN1 pathway in myositis muscle biopsies.

Conclusions: IFN1 and IFN2 pathways are differentially activated in different forms of myositis. This observation may have therapeutic implications since immunosuppressive medications may preferentially target each of these pathways.

INTRODUCTION

Myositis is a heterogeneous family of systemic autoimmune diseases that includes the following groups: dermatomyositis (DM), immune-mediated necrotizing myopathy (IMNM), the antisynthetase syndrome (AS), and sporadic inclusion body myositis (IBM).[1 2] Importantly, myositis-specific autoantibodies (MSAs) help define additional myositis subgroups with unique clinical phenotypes.[1] For example, anti- TIF1γ and anti-MDA5 autoantibodies are each found in DM patients who have myositis and rash. However, whereas anti-TIF1 γ -positive patients have a high risk of cancer and a low risk of lung involvement, anti-MDA5-positive patients have a relatively low risk of cancer and a high risk of lung involvement. Additional MSAs associated with distinct clinical phenotypes include those found in patients with DM (anti-Mi2 and -NXP2), IMNM (anti-SRP and -HMGCR), and AS (anti-Jo1, -PL7, and PL-12).

The pathogenic mechanisms underlying the different types and subtypes of myositis are incompletely understood. However, the type 1 interferon (IFN) pathway has emerged as potentially relevant to DM pathogenesis.[3] Specifically, a marked overexpression of IFN1-inducible genes has been demonstrated in the muscle,[3] peripheral blood,[4, 5] and skin[6] of DM patients. Moreover, the expression levels of IFN1-inducible genes correlate with indicators of DM disease activity.[4, 5]

Three different families of ligands may activate the IFN pathway by binding to cell surface receptors: type 1 IFNs (IFN1; including IFN-alpha and IFN-beta), type 2 IFNs (IFN2; i.e., IFN-gamma), and type III IFNs (IFN3; i.e., IFN lambda).[7] These proteins bind to their corresponding surface receptors which, via the JAK/STAT signaling pathway, stimulate the expression of interferon-inducible genes.[8] Although there is considerable

overlap between the sets of genes induced by the different types of IFN[9, 10], a handful of genes are specifically stimulated by either IFN1 (e.g., *ISG15*,[11, 12] *IFI6*[13] and *MX1*[14]) or IFN2 (e.g., *GBP1, GBP2*,[10, 15] and *PSMB8*[16]).

Prior studies have established the preferential activation of the IFN1 pathway in DM muscle.[3] However, activation of the IFN1 pathway has not been compared between DM patients with different DM subtypes defined by the presence of different DM-autoantibodies. Furthermore, the IFN1 pathway activation was found to be low in IBM but has not been systematically explored in AS or IMNM.[3, 17, 18] Similarly, although IFN2 pathway activation has been implicated in IBM muscle,[19, 20] activation of IFN2 pathways in muscle biopsies from patients with IMNM, AS, and IBM has not been systematically analyzed. In this study, we assessed activation of both IFN1 and IFN2 pathways by analyzing gene expression data from RNAseq performed on a large number of muscle biopsies from patients with DM, IMNM, AS, and IBM, as well as normal comparator tissue.

MATERIALS AND METHODS

Patients, samples, and autoantibody testing

All the available muscle biopsies from patients enrolled in investigational review board-approved longitudinal cohorts of the National Institutes of Health (Bethesda), the Johns Hopkins Myositis Center (Baltimore) the Clinic Hospital (Barcelona), and the Vall d'Hebron Hospital (Barcelona) were included in the study if the patients fulfilled IBM criteria according to Lloyd,[21] or had one of the following MSAs: anti-NXP2, -Mi2, -TIF1γ, -MDA5, -HMGCR, -SRP, or -Jo1. Autoantibody testing was performed as previously described for anti-HMGCR [22] and by line blot for the others (EUROLINE Myositis Profile 4). Patients were classified as having the antisynthetase syndrome (AS) if they had autoantibodies against Jo-1, in the DM group if they had autoantibodies recognizing Mi2, NXP2, TIF1γ or MDA5 and in the IMNM group if they tested positive for anti-SRP or anti-HMGCR autoantibodies. Creatine kinase (CK) levels and strength assessments obtained closest to the time of muscle biopsy were used to assess the clinical activity of the disease. Muscle strength was evaluated by the examining physician using the Medical Research Council scale. This scale was transformed to Kendall´s 0-10 scale and the right- and left-side measurements for arm abduction and hip flexion strength were combined and the average was used for calculations (possible range 0–10) as previously described.[23] Normal muscle biopsies were obtained from the Johns Hopkins Neuromuscular Pathology Laboratory (n=10) and the Skeletal Muscle Biobank of the University of Kentucky (n=10).

Standard protocol approvals and patient consents.

This study was approved by the Institutional Review Boards at participating institutions and written informed consent was obtained from each participant.

RNA-sequencing

RNA-sequencing (RNA-seq) was performed as previously described [24]. Briefly, RNA was prepared using TRIzol. Libraries were prepared using the NeoPrepTM system according to the TruSeqM Stranded mRNA Library Prep protocol (Illumina) and sequenced using the Illumina HiSeq 2500 or 3000. Reads were aligned using the STAR v.2.5 [25], the abundance of each gene was quantified using StringTie v.1.3.3.[26] and the differential gene expression was performed using DESeq2 v.1.20.0.[27] The Benjamini-Hochberg correction was used to adjust for multiple comparisons and a corrected p-value (q-value) of 0.05 or less was considered statistically significant.

Interferon genes and pathways

Interferon pathway genes were collected from the Reactome biorepository (https://reactome.org/). General interferon-related genes and genes from the interferon type 1 and type 2 pathways were merged in a single list. The 13 genes included in the previously proposed interferon score in myositis were also added to the list.[3] The expression of the genes of this list was analyzed in the different autoantibody and clinical myositis subsets.

Data Analysis

Gene-expression (Fragments Per Kilobase of transcript per Million mapped reads [FPKM]) values were log-transformed (logFPKM: log2[FPKM+1]) for visualization purposes using the Python programming language and the packages Numpy, Pandas, and Seaborn. Correlation among continuous variables was measured using Spearman's rho.

Data Availability

Any anonymized data not published within the article will be shared by request from any qualified investigator.

RESULTS

Ranking IFN-inducible gene expression in myositis muscle biopsies

Muscle biopsy specimens were available from 119 myositis patients including 39 with DM (11 anti-Mi2-, 12 anti-NXP2-, 11 anti-TIF1γ, and 5 anti-MDA5-positive), 49 with IMNM (9 anti-SRP- and 40 anti-HMGCR-positive), 18 with anti-Jo1-positive AS, and 13 with IBM. Twenty normal muscle biopsy specimens were utilized as comparators. Expression levels of all genes were determined by RNAseq. The expression level of each gene from each major type of myositis (i.e., DM, IMNM, AS, and IBM) and each autoantibody group (i.e., anti-Mi2, -NXP2, -TIF1γ, -MDA5, -SRP, and -HMGCR) was compared to the expression level of the same gene in the comparator group. Differentially expressed genes were rank ordered by the degree of significance according to the adjusted p-value. From among the complete list of differentially expressed genes, IFN-inducible genes were identified; the top 10 upregulated IFN-inducible genes for each group are listed in Table 1.

Expression levels of IFN1-inducible genes

The most significantly upregulated interferon-inducible genes in DM muscle biopsies were *ISG15, IFI6, MX1, RSAD2, MX2, OAS1, IRF9, IFITM1, OAS3, and IFI35* (Table 1), all of which are preferentially induced by IFN1 (interferon α/β signaling of https://reactome.org/).[11-14] Importantly, among all differentially expressed genes in DM (not just IFN-induced genes), these ten IFN1-inducible genes were also among the most significantly upregulated (with all of them in the top 12 overall differentially-expressed genes) (Table 1).

The overexpression of IFN1-inducible genes was not restricted to DM muscle biopsies (Figure 1). However, the magnitude of this increase was markedly different among the different myositis types. Specifically, IFN1-inducible genes were expressed at markedly elevated levels in DM, at moderately increased levels in AS, and at minimally increased levels in IBM and IMNM (Figure 1). Using *ISG15* expression as an illustrative example, there was a 101-fold increase in DM (q-value $1.1 \bullet 10^{-91}$), an 8.7-fold increase in AS (q-value $1.8 \bullet 10^{-13}$), a 2.4-fold increase in IBM (q-value 0.01) and a 1.8-fold increase in IMNM (q-value 0.05) compared to comparator muscle biopsies (Figure 1). In DM, *ISG15* expression was 11 times higher than in AS (q-value $5.3 \bullet 10^{-27}$), 42 times higher than in IBM (q-value $1.6 \bullet 10^{-48}$) and 56 times higher than in IMNM (q-value $9.8 \bullet 10^{-109}$). Likewise, *ISG15* expression in AS was higher than in IBM and IMNM by 4 and 4.8 times, respectively (q-values 0.001 and $3.8 \bullet 10^{-11}$).

We next analyzed the expression levels of IFN1-inducible genes among autoantibody subgroups. Interestingly, *ISG15* and *IFI6* were the most significantly upregulated genes in all DM autoantibody groups (i.e., anti-Mi2, anti-NXP2, anti-TIF1γ and anti-MDA5) (Table 1). In each DM autoantibody subgroup, these two genes were upregulated by at least 60-fold compared to healthy comparators (all q-values < $1 \bullet 10^{-44}$) with no significant differences between the DM subgroups. Within IMNM, the expression of IFN1-inducible genes in those with anti-SRP autoantibodies was not significantly different compared to those with anti-HMGCR autoantibodies.

Expression levels of IFN2-inducible genes

The IFN2-specific genes *GBP1*, *GBP2*, and *PSMB8* were the three most significantly upregulated IFN-inducible genes in both AS and IBM. Also, in muscle biopsies from both AS and IBM patients, these three IFN2-inducible genes were within the top ten most upregulated genes overall (Table 1).

Compared to comparators, the expression of IFN2-inducible genes was increased by 7-14-fold in IBM, AS, and DM biopsies (all q-values < $1 \cdot 10^{-15}$) (Table 1 and Figure 2). There were no significant differences between AS or IBM and DM except that *GBP2* had slightly higher expression levels in IBM (fold-change 1.7, q-value 0.01) and AS (fold-change 1.8, q-value 0.02) compared to DM. In contrast, the magnitude of IFN2-inducible gene overexpression in IMNM compared to comparators was much lower (PSMB8 fold-change 2.5, q-value $7.6 \cdot 10^{-5}$). Compared to IMNM, IFN2-inducible genes were expressed at higher levels in DM (fold-change 5.6, q-value=$9.5 \cdot 10^{-25}$), AS (fold-change 5.2, q-value=$2.1 \cdot 10^{-12}$) and IBM (fold-change 3.7, q-value=$9.2 \cdot 10^{-7}$). There were no significant differences in the expression of IFN2-inducible genes between the different autoantibody subgroups within IMNM or DM.

Interestingly, the IFN2-inducible gene *IFI30* was one of the two most significantly upregulated interferon genes in both anti-SRP- and anti-HMGCR-positive IMNM patients. Compared to normal biopsies this gene showed a 7-fold-increase in IMNM (q-value = $4 \cdot 10^{-13}$), a 16-fold-increase in DM (q-value = $5.7 \cdot 10^{-32}$), a 15.8-fold-increase in AS (q-value = $1 \cdot 10^{-20}$) and a 7-fold-increase in IBM (q-value = $2.1 \cdot 10^{-9}$) (Table 1). Apart from *IFI30* gene expression, the relative magnitude of interferon-related genes among all differentially expressed genes in IMNM was modest compared to other types of

myositis. In fact, the first ranked interferon-inducible gene in IMNM was ranked 40th in the list of all differentially expressed genes. In contrast, the first ranked interferon-inducible gene was also first among all differentially expressed genes in DM and AS and the second among all differentially expressed genes in IBM (Table 1).

Expression levels of genes associated with inflammation and muscle regeneration

In each of the clinical and autoantibody subgroups studied, the expression of both IFN1- and IFN2-inducible genes was positively correlated with the expression of genes associated with inflammatory cells (T-cells [*CD3E, CD4, CD8A*] and macrophages [*CD14, CD68*]) and genes associated with muscle regeneration (*NCAM1, MYOG, MYOD1, PAX7, MYH3* and *MYH8*) (all q-values <0.05) (Figure 3). Conversely, interferon-inducible genes were inversely correlated with mature-muscle structural proteins (*ACTA1, MYH1, and MYH2*) (all q-values <0.05).

Strength measurements and CK levels obtained near the time of the muscle biopsy were available from 62 of the Hopkins patients (17 DM, 6 AS, 12 IBM, and 27 IMNM). Although there was a trend for DM, AS, and IMNM patients with higher levels of IFN-inducible genes to have higher CK levels and decreased strength, this was not statistically significant (Figure 4). However, IBM patients with higher levels of IFN-inducible genes had significantly higher CK levels (all p-value≤0.05) and a non-significant trend towards being stronger than those with lower levels of IFN-inducible genes. Since IBM patients often have relatively preserved muscle strength early in the course of the disease, we hypothesized that IFN-inducible gene expression might also

be highest early during the course of the disease. Indeed, we found that IBM patients with a shorter interval between onset of symptoms and muscle biopsy had higher expression levels of interferon-inducible genes (data not shown).

ISG15 gene expression compared to composite interferon scores to quantify the interferon signature

Several gene scoring systems have been proposed to measure the activation of the interferon pathway in myositis[28] and other autoimmune diseases[29]. Greenberg et al.[28] used a score combining 13 IFN1-inducible genes to study the relationship between IFN1 inducible gene expression and disease activity in blood of patients with dermatomyositis and polymyositis. We tried to test the utility of such score compared with simpler alternatives in myositis muscle biopsies.

First, we analyzed the correlations between the expression levels of the different IFN-inducible genes in all of the muscle biopsies included in the study. This revealed a high correlation between expression levels of each IFN-inducible gene with all the others (Figure 5). Second, since *ISG15* was, overall, the most highly expressed IFN1-inducible gene, we correlated the raw expression levels of *ISG15* with the previously proposed interferon score. This analysis revealed an almost perfect correlation between *ISG15* expression levels alone and the 13-gene composite score (Spearman rho: 0.94, p-value $1.5 \bullet 10^{-64}$, Figure 6a), suggesting it may be unnecessary to use a more complex scoring system to measure IFN1 pathway activation levels in myositis muscle.

Although expression levels of the IFN2-inducible genes *PSMB8, GBP1, and GBP2* were highly correlated with each other (Figure 5), the association between IFN2-

inducible genes with the 13 gene IFN1-gene score was restricted to DM. For example, *PSMB8* expression levels correlated well with the composite gene score in DM patients, but not in patients with AS or IBM (Figure 6b). This suggests that IFN1-inducible gene scores may be good indicators of IFN2 activation in DM, but not in AS or IBM. Rather, expression levels of *PSMB8, GBP1, and GBP2* should be considered as markers IFN2 pathway activation.

DISCUSSION

In this study, using RNAseq data from a large number of myositis and comparator muscle biopsies, we have established that the IFN1 pathway is activated, not only in DM patients as previously described,[3-6] but also in patients with AS, IMNM, and IBM. Quantitatively, the IFN1 pathway was most up-regulated in DM, with intermediate activation of the pathway in AS and lower levels of activation in IBM and IMNM. We also used RNAseq data to study activation of the IFN2 pathway, demonstrating robust activation in AS, IBM, and DM, but not in IMNM. We were also able to show that activation of the IFN pathway was associated with increased expression of inflammatory cell and muscle regeneration genes. Finally, we established that *ISG15* gene expression can be used as a surrogate marker of IFN1 pathway activation in myositis since it performs as well as a more complex composite score.

Interestingly, different collections of IFN-inducible genes were most prominently upregulated in the different groups. For example, the IFN1 genes *ISG15, IFI6, and MX1* were the most upregulated IFN-inducible genes in DM. In contrast, *IFI30, NCAM1, and SOCS3* were the most upregulated IFN-inducible genes in IMNM patients. Of note, the IFN2 genes *PSMB8, GBP2, and GBP1* were the most upregulated IFN-inducible genes in both AS and IBM patients, underscoring the prominence of the IFN2 pathway in these two diseases. Taken together, this suggests that the degree of activation of the IFN1 inflammatory pathways differs between DM, IBM, AS and IMNM, and most but not all types of myositis involve the IFN2 inflammatory pathways.

It is well-established that DM patients with different myositis autoantibodies have unique clinical manifestations. In fact, there are differences in muscle biopsy features

between DM patients with different autoantibodies.[30] For example, half of the muscle biopsies from anti-Mi2-positive DM patients include examples of lymphocytes surrounding and invading healthy muscle fibers; this histopathologic feature was never seen in DM patients with anti-NXP2 autoantibodies. Despite these histopathologic differences, the IFN gene signature was remarkably similar between DM patients with different myositis autoantibodies. Indeed, *ISG15* and *IFI6* were the top two IFN-inducible genes in each of the serologically defined DM subgroups and *MX1* and *MX2* present among the top ten IFN-inducible genes in each DM subgroup. These findings suggest that, at least with regard to activation of IFN pathways in the muscle, the different autoantibody subgroups of DM are more alike than different. Similarly, in IMNM patients with either anti-SRP or -HMGCR autoantibodies, *IFI30, NCAM1, VCAM1, ICAM1, SOC3, GBP2, and MT2A* were among the top ten IFN-inducible genes. We did not have a sufficient number of biopsies from patients with anti-PL7, anti-PL12, or other non-Jo1 antisynthetase autoantibodies to determine whether these serologic subgroups of the anti-synthetase syndrome share a similar IFN gene signature pattern.

Some investigators have shown that immunostaining muscle biopsies for specific IFN-inducible proteins can be used to distinguish between different types of myositis. For example, DM but not AS muscle biopsies stain positive for MxA (*MX1*)[31] or RIG-I (DDX58),[32] both IFN1 inducible genes. Our RNAseq data, which shows higher expression levels of these genes in DM than in AS (MX1 fold-change 4.7 and RIG-1 fold-change 3.3, both q-values<$5 \cdot 10^{-9}$), is consistent with this observation. Also, ISG15 overexpression was reported to be useful to diagnose patients with DM and perifascicular atrophy.[17] Accordingly, we found a marked preferential overexpression of

ISG15 in DM patients (ISG15 fold-change compared to comparator biopsies 101, q-value $1.1 \bullet 10^{-91}$).

As it was mentioned earlier, in this study we determined that *ISG15* expression levels alone can be used to reliably quantitate the activation of the IFN1 pathway in myositis muscle biopsies. In fact, measuring *ISG15* levels was equivalent to a composite score derived from measuring expression levels of 13 different IFN genes, which is concordant with previous data showing the marked specificity of ISG15 muscle transcript measurements for DM with perifascicular atrophy.[17] Also, we noted that either *ISG15* expression levels or the previously proposed composite IFN1 scores were useful for assessing activation of the IFN2 pathway in DM, but not in IBM or AS. Rather, directly measuring the expression levels of IFN2-inducible genes such as *PSMB8, GBP1 or GBP2* may be required.

This study has several limitations. For example, some less common autoantibody groups (e.g., non-anti-Jo1 AS patients) could not be included due to an insufficient number of biopsies. In addition, we only had relevant CK and strength information for muscle biopsies obtained at Johns Hopkins, which may have limited our ability to show significant associations between IFN pathway activation and markers of clinical disease activity, such as strength and CK levels.

In conclusion, this study demonstrates that DM muscle biopsies are characterized by high levels of both IFN1- and IFN2-inducible genes. In contrast, biopsies from patients with AS and IBM reveal gene expression patterns consistent with prominent IFN2 activation. Finally, RNAseq analysis reveals IMNM biopsies show relatively low activation of the interferon pathway. These findings are consistent with

recent case series suggesting the efficacy of JAK/STAT inhibitors in patients with DM.[33-37] They also suggest that these agents may be effective in patients with AS or IBM. However, the relatively modest activation of IFN pathways in IMNM does not provide compelling evidence to support the use of JAK/STAT inhibitors in this patient population.

APPENDIX 1: AUTHORS

Name	Location	Role	Contribution
Iago Pinal-Fernandez, MD, PhD	NIH, Bethesda	Author	Designed and conceptualized study; analyzed the data; drafted the manuscript for intellectual content
Maria Casal-Dominguez, MD, PhD	NIH, Bethesda	Author	Designed and conceptualized study; analyzed the data; drafted the manuscript for intellectual content
Assia Derfoul, PhD	NIH, Bethesda	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Katherine Pak, MD	NIH, Bethesda	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Paul Plotz, MD	NIH, Bethesda	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Frederick W Miller, MD, PhD	NIH, Bethesda	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Jose C Milisenda, MD	Clinic Hospital, Barcelona	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Josep M Grau-Junyent, MD, PhD	Clinic Hospital, Barcelona	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Albert Selva-O'Callaghan, MD, PhD	Vall d'Hebron Hospital, Barcelona	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Julie Paik, MD	Johns Hopkins, Baltimore	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Jemima Albayda, MD	Johns Hopkins, Baltimore	Author	Major role in the acquisition of data; revised the manuscript for intellectual content

Lisa Christopher-Stine, MD, MPH	Johns Hopkins, Baltimore	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Thomas E. Lloyd, MD, PhD	Johns Hopkins, Baltimore	Author	Major role in the acquisition of data; revised the manuscript for intellectual content
Andrea M Corse, MD	Johns Hopkins, Baltimore	Author	Major role in the acquisition of data; interpreted the data, revised the manuscript for intellectual content
Andrew L. Mammen, MD, PhD	NIH, Bethesda	Author	Designed and conceptualized study; interpreted the data; major role in the acquisition of data; drafted the manuscript for intellectual content

REFERENCES

1. Selva-O'Callaghan A, Pinal-Fernandez I, Trallero-Araguas E, Milisenda JC, Grau-Junyent JM, Mammen AL. Classification and management of adult inflammatory myopathies. Lancet Neurol 2018;17:816-828.
2. Mariampillai K, Granger B, Amelin D, et al. Development of a New Classification System for Idiopathic Inflammatory Myopathies Based on Clinical Manifestations and Myositis-Specific Autoantibodies. JAMA Neurol 2018;75:1528-1537.
3. Greenberg SA, Pinkus JL, Pinkus GS, et al. Interferon-alpha/beta-mediated innate immune mechanisms in dermatomyositis. Ann Neurol 2005;57:664-678.
4. Walsh RJ, Kong SW, Yao Y, et al. Type I interferon-inducible gene expression in blood is present and reflects disease activity in dermatomyositis and polymyositis. Arthritis Rheum 2007;56:3784-3792.
5. Baechler EC, Bauer JW, Slattery CA, et al. An interferon signature in the peripheral blood of dermatomyositis patients is associated with disease activity. Mol Med 2007;13:59-68.
6. Wong D, Kea B, Pesich R, et al. Interferon and biologic signatures in dermatomyositis skin: specificity and heterogeneity across diseases. PLoS One 2012;7:e29161.
7. Ivashkiv LB, Donlin LT. Regulation of type I interferon responses. Nat Rev Immunol 2014;14:36-49.
8. O'Shea JJ, Schwartz DM, Villarino AV, Gadina M, McInnes IB, Laurence A. The JAK-STAT pathway: impact on human disease and therapeutic intervention. Annu Rev Med 2015;66:311-328.
9. Zhou Z, Hamming OJ, Ank N, Paludan SR, Nielsen AL, Hartmann R. Type III interferon (IFN) induces a type I IFN-like response in a restricted subset of cells through signaling pathways involving both the Jak-STAT pathway and the mitogen-activated protein kinases. J Virol 2007;81:7749-7758.
10. Hall JC, Casciola-Rosen L, Berger AE, et al. Precise probes of type II interferon activity define the origin of interferon signatures in target tissues in rheumatic diseases. Proc Natl Acad Sci U S A 2012;109:17609-17614.
11. Zhang D, Zhang DE. Interferon-stimulated gene 15 and the protein ISGylation system. J Interferon Cytokine Res 2011;31:119-130.
12. D'Cunha J, Ramanujam S, Wagner RJ, Witt PL, Knight E, Jr., Borden EC. In vitro and in vivo secretion of human ISG15, an IFN-induced immunomodulatory cytokine. J Immunol 1996;157:4100-4108.
13. Kelly JM, Porter AC, Chernajovsky Y, Gilbert CS, Stark GR, Kerr IM. Characterization of a human gene inducible by alpha- and beta-interferons and its expression in mouse cells. EMBO J 1986;5:1601-1606.
14. Holzinger D, Jorns C, Stertz S, et al. Induction of MxA gene expression by influenza A virus requires type I or type III interferon signaling. J Virol 2007;81:7776-7785.
15. Hall JC, Baer AN, Shah AA, et al. Molecular Subsetting of Interferon Pathways in Sjogren's Syndrome. Arthritis Rheumatol 2015;67:2437-2446.
16. Hisamatsu H, Shimbara N, Saito Y, et al. Newly identified pair of proteasomal subunits regulated reciprocally by interferon gamma. J Exp Med 1996;183:1807-1816.

17. Salajegheh M, Kong SW, Pinkus JL, et al. Interferon-stimulated gene 15 (ISG15) conjugates proteins in dermatomyositis muscle with perifascicular atrophy. Ann Neurol 2010;67:53-63.
18. Greenberg SA, Sanoudou D, Haslett JN, et al. Molecular profiles of inflammatory myopathies. Neurology 2002;59:1170-1182.
19. Ivanidze J, Hoffmann R, Lochmuller H, Engel AG, Hohlfeld R, Dornmair K. Inclusion body myositis: laser microdissection reveals differential up-regulation of IFN-gamma signaling cascade in attacked versus nonattacked myofibers. Am J Pathol 2011;179:1347-1359.
20. Allenbach Y, Chaara W, Rosenzwajg M, et al. Th1 response and systemic treg deficiency in inclusion body myositis. PLoS One 2014;9:e88788.
21. Lloyd TE, Mammen AL, Amato AA, Weiss MD, Needham M, Greenberg SA. Evaluation and construction of diagnostic criteria for inclusion body myositis. Neurology 2014;83:426-433.
22. Mammen AL, Chung T, Christopher-Stine L, et al. Autoantibodies against 3-hydroxy-3-methylglutaryl-coenzyme A reductase in patients with statin-associated autoimmune myopathy. Arthritis Rheum 2011;63:713-721.
23. Rider LG, Werth VP, Huber AM, et al. Measures of adult and juvenile dermatomyositis, polymyositis, and inclusion body myositis: Physician and Patient/Parent Global Activity, Manual Muscle Testing (MMT), Health Assessment Questionnaire (HAQ)/Childhood Health Assessment Questionnaire (C-HAQ), Childhood Myositis Assessment Scale (CMAS), Myositis Disease Activity Assessment Tool (MDAAT), Disease Activity Score (DAS), Short Form 36 (SF-36), Child Health Questionnaire (CHQ), physician global damage, Myositis Damage Index (MDI), Quantitative Muscle Testing (QMT), Myositis Functional Index-2 (FI-2), Myositis Activities Profile (MAP), Inclusion Body Myositis Functional Rating Scale (IBMFRS), Cutaneous Dermatomyositis Disease Area and Severity Index (CDASI), Cutaneous Assessment Tool (CAT), Dermatomyositis Skin Severity Index (DSSI), Skindex, and Dermatology Life Quality Index (DLQI). Arthritis Care Res (Hoboken) 2011;63 Suppl 11:S118-157.
24. Amici DR, Pinal-Fernandez I, Mázala DA, et al. Calcium dysregulation, functional calpainopathy, and endoplasmic reticulum stress in sporadic inclusion body myositis. Acta Neuropathol Commun 2017;5:24.
25. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15-21.
26. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 2015;33:290-295.
27. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
28. Greenberg SA, Higgs BW, Morehouse C, et al. Relationship between disease activity and type 1 interferon- and other cytokine-inducible gene expression in blood in dermatomyositis and polymyositis. Genes Immun 2012;13:207-213.
29. El-Sherbiny YM, Psarras A, Md Yusof MY, et al. A novel two-score system for interferon status segregates autoimmune diseases and correlates with clinical features. Sci Rep 2018;8:5793.

30. Pinal-Fernandez I, Casciola-Rosen LA, Christopher-Stine L, Corse AM, Mammen AL. The Prevalence of Individual Histopathologic Features Varies according to Autoantibody Status in Muscle Biopsies from Patients with Dermatomyositis. J Rheumatol 2015;42:1448-1454.
31. Uruha A, Nishikawa A, Tsuburaya RS, et al. Sarcoplasmic MxA expression: A valuable marker of dermatomyositis. Neurology 2017;88:493-500.
32. Suarez-Calvet X, Gallardo E, Pinal-Fernandez I, et al. RIG-I expression in perifascicular myofibers is a reliable biomarker of dermatomyositis. Arthritis Res Ther 2017;19:174.
33. Hornung T, Janzen V, Heidgen FJ, Wolf D, Bieber T, Wenzel J. Remission of recalcitrant dermatomyositis treated with ruxolitinib. N Engl J Med 2014;371:2537-2538.
34. Kurasawa K, Arai S, Namiki Y, et al. Tofacitinib for refractory interstitial lung diseases in anti-melanoma differentiation-associated 5 gene antibody-positive dermatomyositis. Rheumatology (Oxford) 2018.
35. Kurtzman DJ, Wright NA, Lin J, et al. Tofacitinib Citrate for Refractory Cutaneous Dermatomyositis: An Alternative Treatment. JAMA Dermatol 2016;152:944-945.
36. Ladislau L, Suarez-Calvet X, Toquet S, et al. JAK inhibitor improves type I interferon induced damage: proof of concept in dermatomyositis. Brain 2018;141:1609-1621.
37. Paik JJ, Christopher-Stine L. A case of refractory dermatomyositis responsive to tofacitinib. Semin Arthritis Rheum 2017;46:e19.

Figure 1. Expression of type 1 interferon-inducible genes in myositis muscle biopsies. Relative (a) and raw (95% confidence interval) (b) expression levels of the type 1 interferon-inducible genes among the different clinical and serological groups. NT: normal biopsies, IBM: inclusion body myositis, [SRP, HMGCR]: immune-mediated necrotizing myositis autoantibody groups, [Mi2, NXP2, TIF1, MDA5]: dermatomyositis autoantibody groups, [Jo1]: antisynthetase autoantibody group.

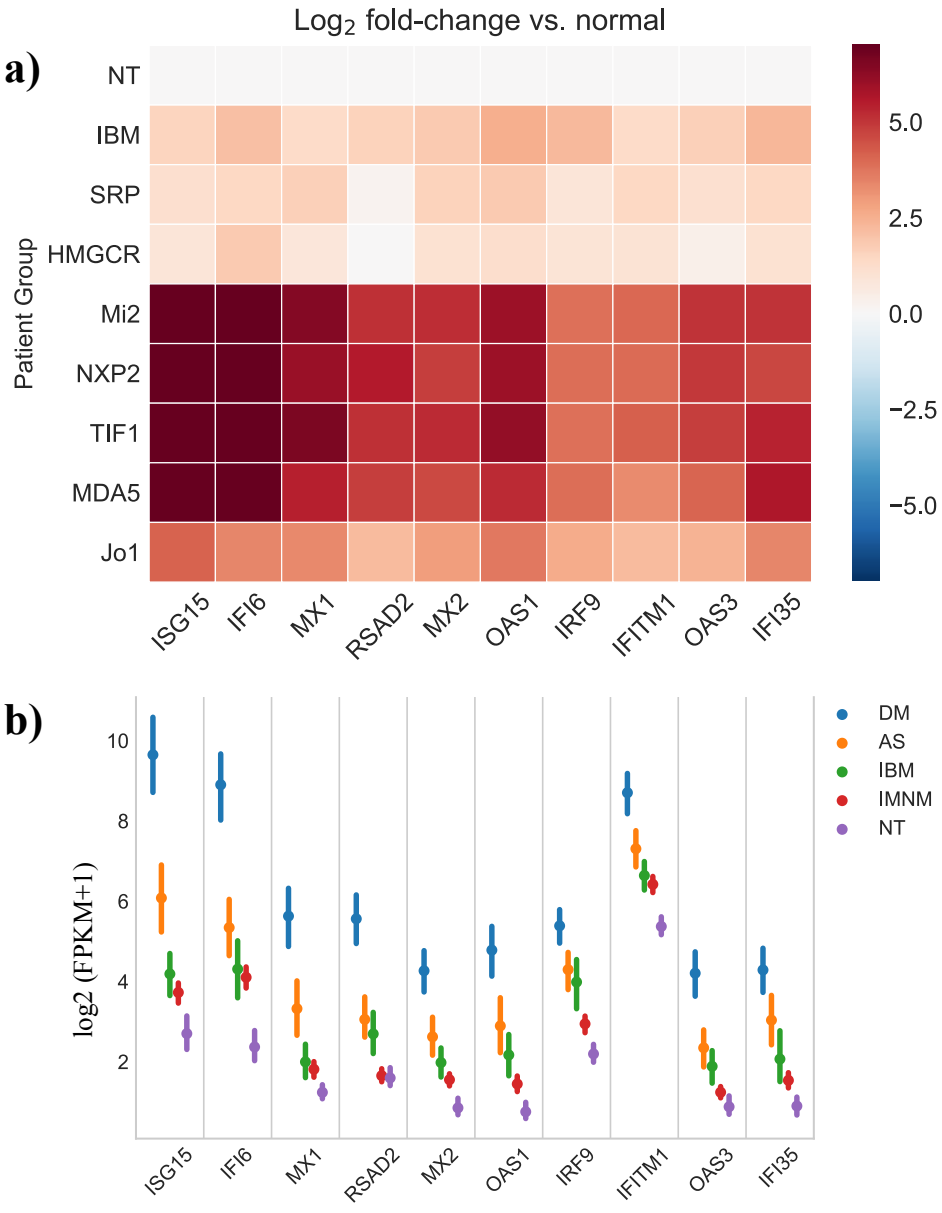


Figure 2. Expression of type 2 interferon-inducible genes in myositis muscle biopsies. Relative (a) and raw (95% confidence interval) (b) expression levels of type 2 interferon-inducible genes among the different clinical and serological groups. NT: normal biopsies, IBM: inclusion body myositis, [SRP, HMGCR]: immune-mediated necrotizing myositis autoantibody groups, [Mi2, NXP2, TIF1, MDA5]: dermatomyositis autoantibody groups, [Jo1]: antisynthetase autoantibody group.

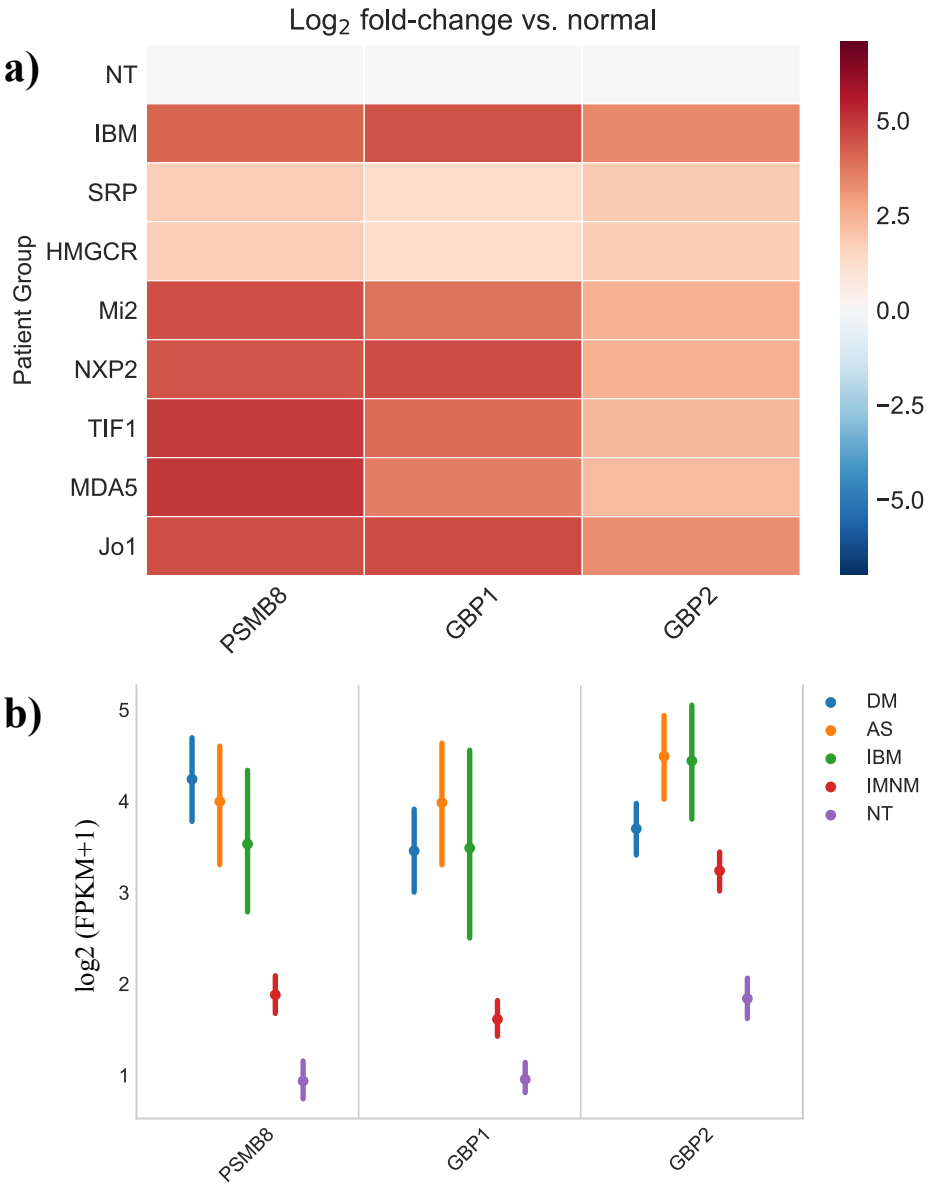


Figure 3. Correlation of IFN-inducible gene expression with expression of inflammatory cell and muscle regeneration genes. Correlation of type 1 and type 2 interferon-inducible-genes with the expression of genes related to T-cells (CD3E, CD4, and CD8A), macrophages (CD14 and CD68), muscle regeneration (NCAM1, MYOG, MYOD1, PAX7) and adult muscle structural proteins (ACTA1, MYH1, MYH2).

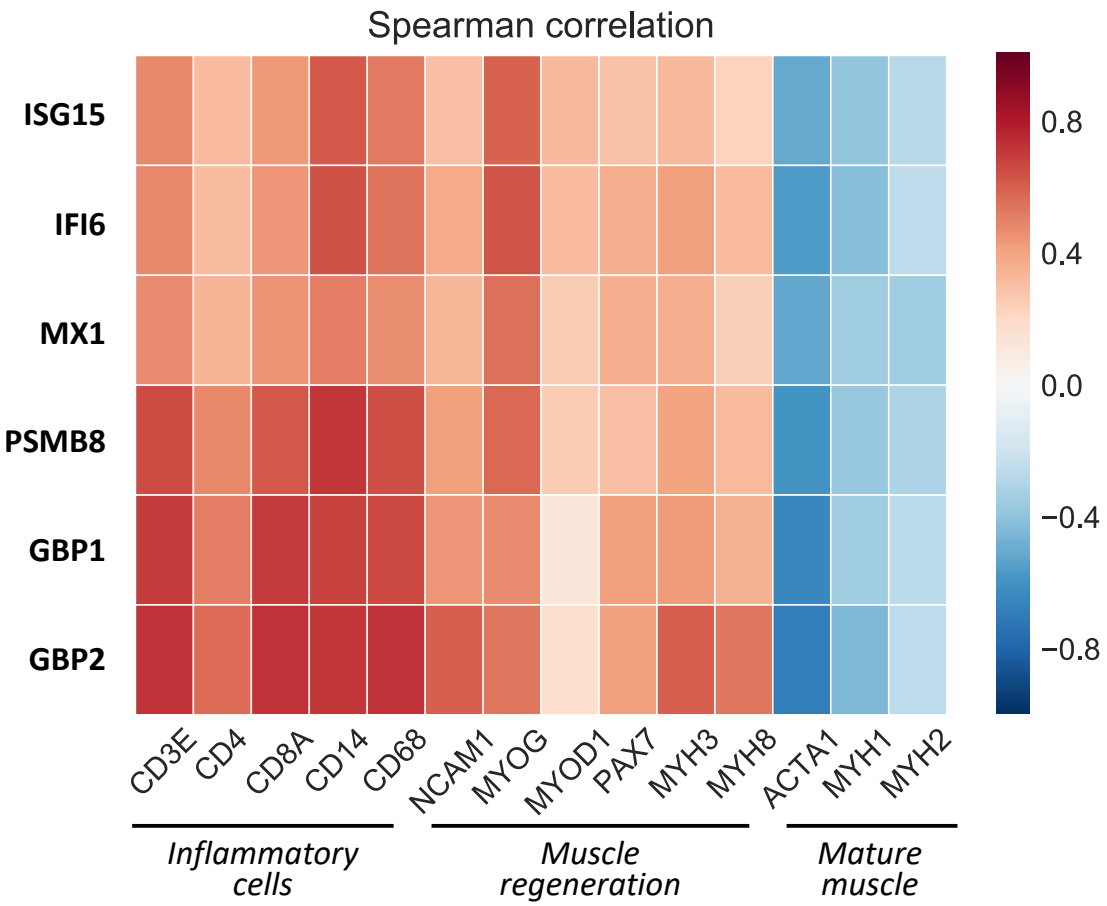


Figure 4. Correlation of type 1 and type 2 interferon-inducible genes with the CK and strength in different types of myositis. DM: dermatomyositis, IBM: inclusion body myositis, AS: antisynthetase syndrome, IMNM: immune-mediated necrotizing myositis.

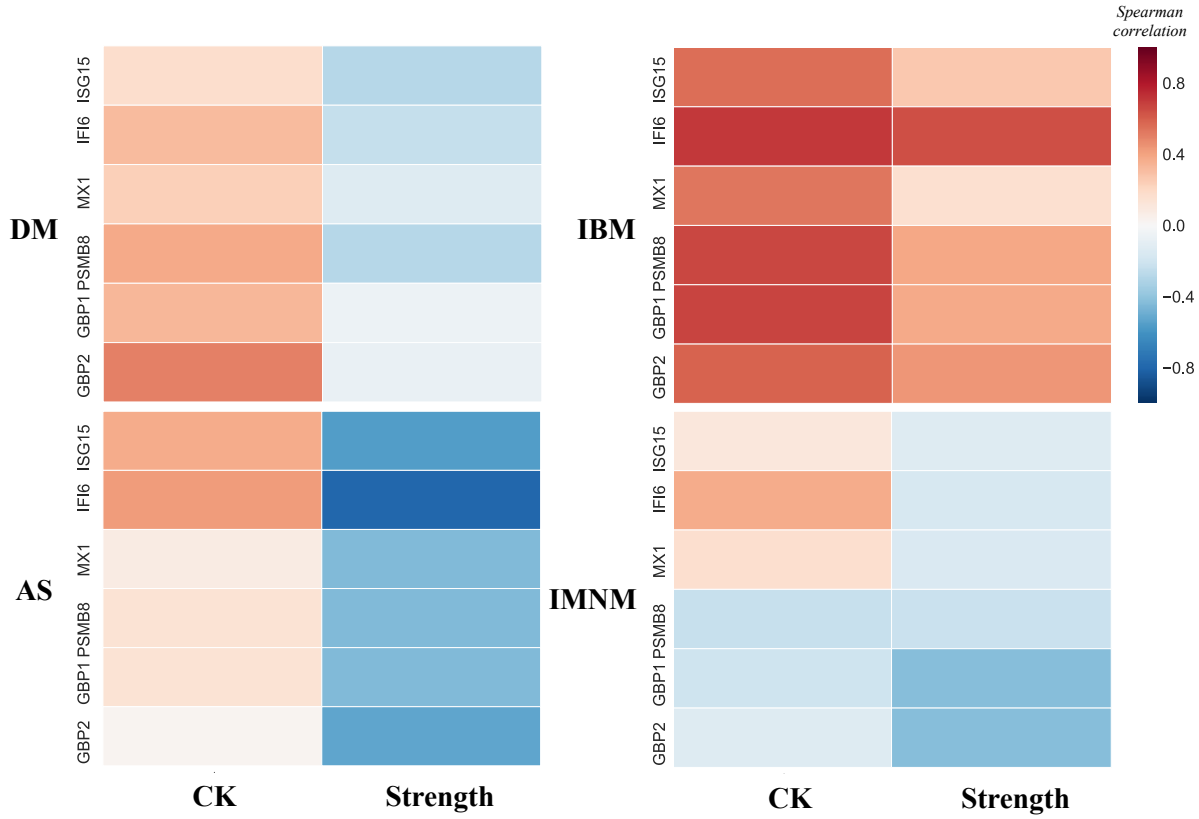


Figure 5. Spearman correlation of the different type 1 and type 2 interferon-inducible genes in all the biopsies included in the study.

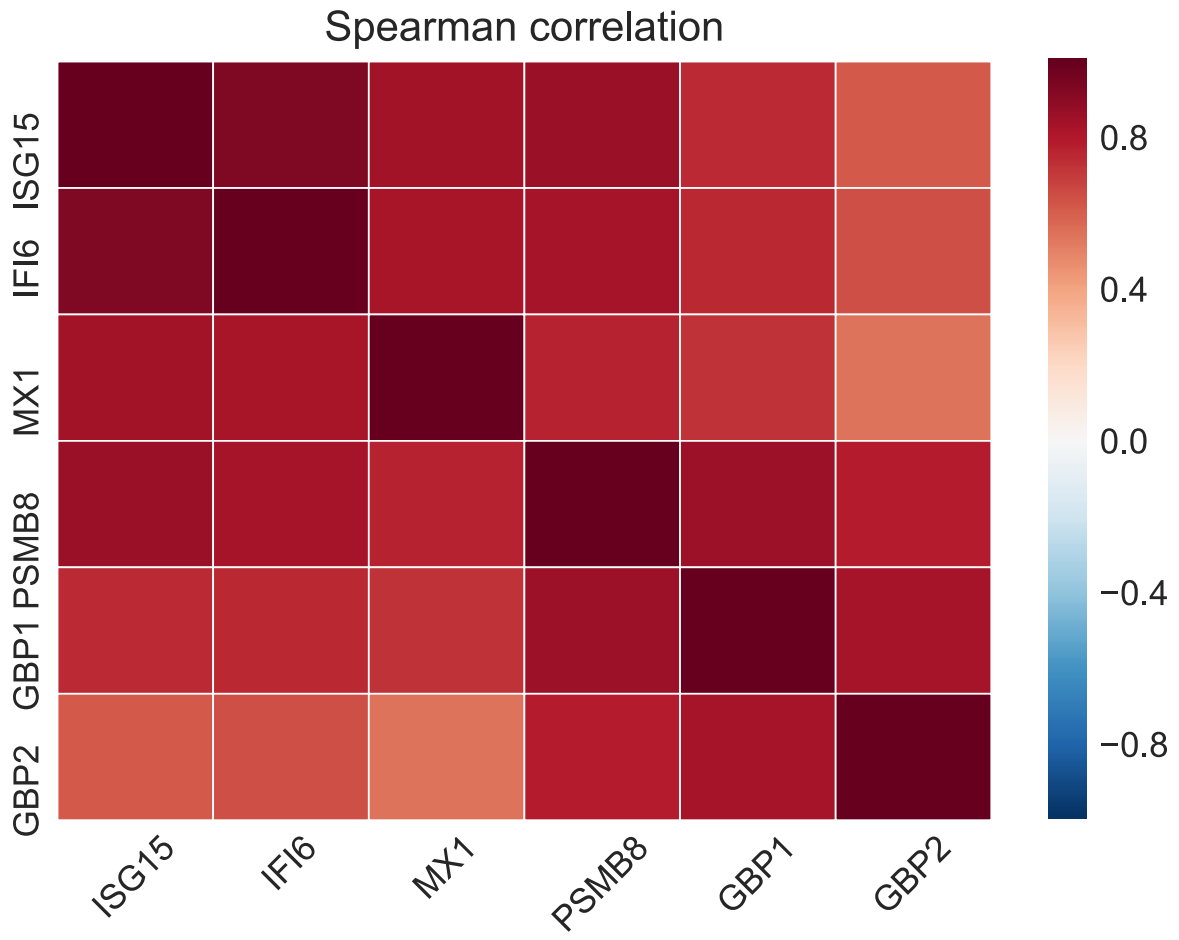


Figure 6. ISG15 or PSMB8 expression versus composite type 1 interferon-inducible gene scores. Correlation of the expression level (log_2[FPKM+1]) of ISG15 (a) and PSMB8 (b) with the previously proposed 13-gene type 1 interferon score.

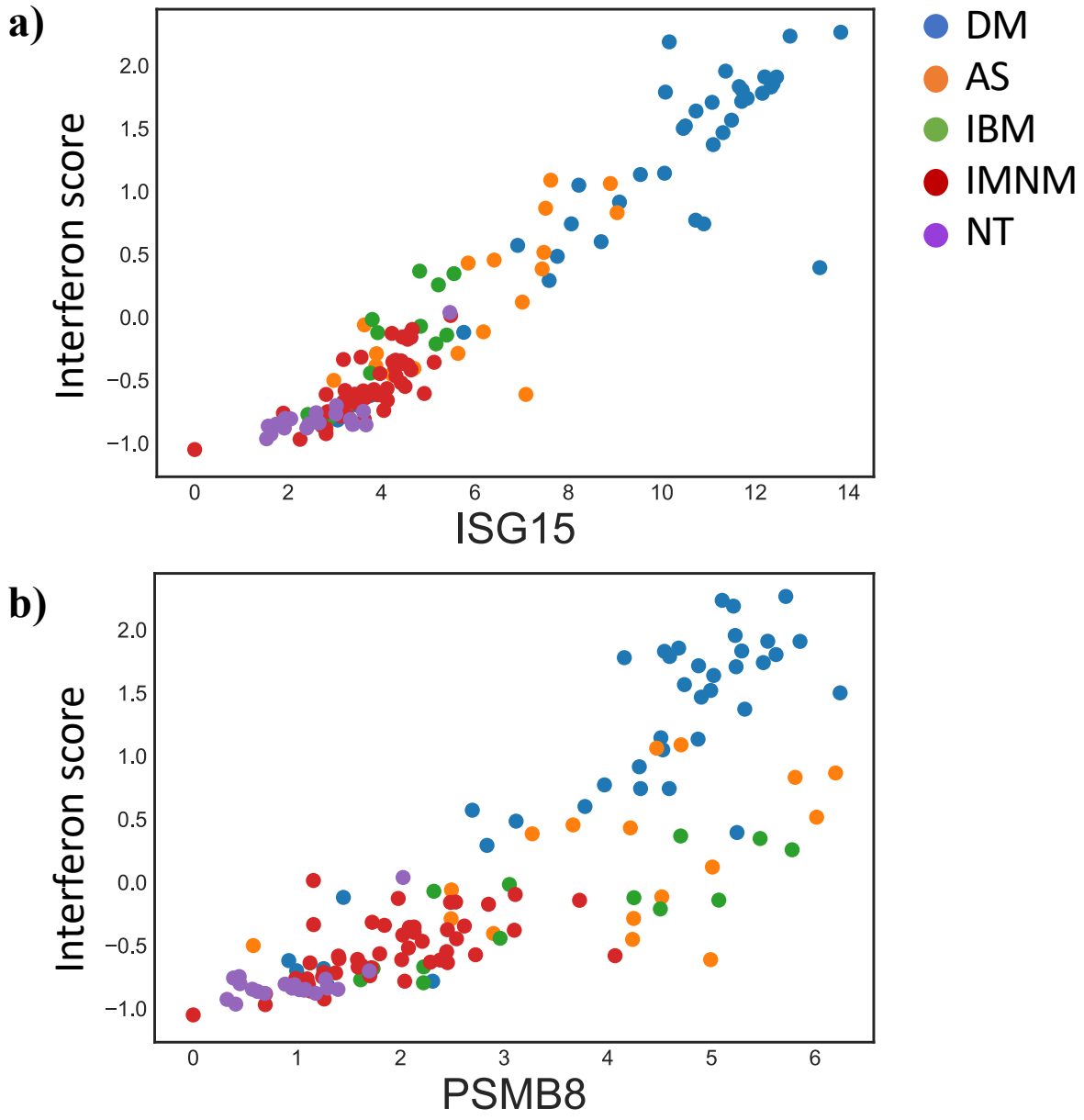


Table 1. Expression levels of the top 10 most significantly expressed genes of the interferon pathway in the different clinical and serologic myositis subgroups.

DM (all)				Mi2				NXP2				TIF1				MDA5			
Gene	Order	FC	qval	Gene	Order	FC	qval	Gene	Order	FC	qval	Gene	Order	FC	qval	Gene	Order	FC	qval
ISG15	1	101	1.1E-91	**IFI6**	1	62	2.6E-43	**ISG15**	1	110	1.5E-55	**ISG15**	1	84	1.5E-47	**ISG15**	1	163	1.2E-51
IFI6	2	67	2.7E-80	**ISG15**	2	67	5.9E-43	**IFI6**	2	70	7.8E-48	**IFI6**	2	62	1.7E-43	**IFI6**	2	72	2.6E-37
MX1	3	29	2.6E-56	**MX1**	3	32	1.9E-33	**RSAD2**	3	23	2.1E-33	**MX1**	3	41	9.0E-39	**IFI35**	4	30	2.3E-24
RSAD2	4	18	1.2E-49	**OAS1**	4	25	9.7E-28	**MX1**	4	29	8.6E-33	**MX2**	4	22	6.9E-32	**PSMB8**	9	19	3.3E-20
MX2	5	17	2.5E-49	**MX2**	5	18	1.2E-27	**IFIT2**	5	22	5.6E-31	**OAS1**	5	29	3.8E-31	**IFI27**	10	18	2.2E-19
OAS1	6	23	4.0E-48	**RSAD2**	7	17	8.1E-26	**OAS1**	7	26	4.7E-30	**IFITM1**	6	12	2.7E-28	**MX2**	14	16	1.6E-18
IRF9	7	10	2.5E-43	**OAS3**	9	16	3.6E-25	**OAS3**	8	16	1.3E-26	**RSAD2**	7	18	2.0E-27	**IRF7**	15	15	1.7E-18
IFITM1	8	10	8.6E-43	**IFITM1**	11	9	1.5E-22	**HERC5**	9	25	3.4E-26	**IFI44**	9	18	4.1E-27	**RSAD2**	16	16	1.7E-18
OAS3	10	14	2.3E-41	**UBE2L6**	12	11	3.7E-22	**MX2**	10	15	4.8E-26	**ISG20**	11	27	2.2E-26	**MX1**	20	18	8.6E-18
IFI35	11	18	2.3E-41	**IRF9**	13	9	3.8E-22	**IRF9**	11	10	7.7E-26	**UBE2L6**	13	13	6.4E-26	**IRF9**	25	10	4.3E-17

AS (Jo1)				IBM				IMNM (all)				SRP				HMGCR			
Gene	Order	FC	qval	Gene	Order	FC	qval	Gene	Order	FC	qval	Gene	Order	FC	qval	Gene	Order	FC	qval
PSMB8	1	13	2.6E-25	**GBP2**	2	7	1.3E-18	**IFI30**	40	7	4.0E-13	**IFI30**	52	7	1.9E-07	**NCAM1**	31	4	9.6E-14
GBP2	3	7	1.1E-22	**PSMB8**	4	9	1.7E-16	**NCAM1**	42	4	5.9E-13	**NCAM1**	67	4	1.1E-06	**IFI30**	39	7	1.3E-12
GBP1	9	12	9.8E-21	**GBP1**	5	11	3.5E-16	**SOCS3**	89	6	3.1E-10	**VCAM1**	159	4	3.1E-05	**SOCS3**	50	8	8.6E-12
IFI30	10	16	1.0E-20	**GBP5**	8	17	4.0E-15	**TRIM38**	148	3	7.4E-09	**ICAM1**	212	3	7.6E-05	**TRIM38**	67	3	7.5E-11
IRF1	14	8	5.0E-20	**GBP4**	17	6	3.4E-13	**VCAM1**	158	4	1.1E-08	**SOCS3**	220	5	8.2E-05	**GBP2**	143	3	6.8E-09
ISG20	35	11	2.1E-16	**STAT1**	24	6	1.5E-12	**GBP2**	191	3	2.4E-08	**GBP2**	289	3	2.2E-04	**VCAM1**	203	4	5.6E-08
ICAM1	36	7	2.2E-16	**B2M**	33	5	6.2E-12	**ICAM1**	197	3	3.0E-08	**TRIM38**	385	2	4.3E-04	**ICAM1**	227	3	8.3E-08
UBE2L6	41	6	7.9E-16	**CIITA**	47	6	1.7E-11	**MT2A**	312	3	4.6E-07	**MT2A**	407	3	5.8E-04	**MT2A**	301	4	4.2E-07
TRIM38	46	4	1.3E-15	**TRIM38**	52	4	2.3E-11	**TRIM62**	418	2	1.8E-06	**CD44**	425	3	6.3E-04	**TRIM8**	391	1	1.5E-06
GBP5	63	13	1.7E-14	**GBP6**	56	15	5.1E-11	**IRF5**	464	3	2.7E-06	**IRF5**	428	3	6.4E-04	**TRIM62**	450	2	2.8E-06

In each panel, the first column shows the gene name, the second column the rank of the gene relative to the whole list of differentially expressed genes (including non-interferon genes), the third column shows the fold-change and the fourth column shows the Benjamini-Hochberg adjusted p-value (q-value). [Mi2, NXP2, TIF1, MDA5]: dermatomyositis autoantibody groups, [Jo1]: antisynthetase (AS) autoantibody group, IBM: inclusion body myositis, IMNM: immune-mediated necrotizing myositis, [SRP, HMGCR]: immune-mediated necrotizing myositis autoantibody groups.

A.9 Machine learning algorithms reveal unique gene expression profiles in muscle biopsies from patients with different types of myositis. Ann Rheum Dis 2020 (PMID: 32546599).

The purpose of this study was to define unique gene expression profiles in muscle biopsies from patients with MSAs-positive DM, AS and IMNM as well as IBM.

RNA-seq was performed on muscle biopsies from 119 myositis patients with IBM or defined MSAss and 20 controls. Machine learning algorithms were trained on transcriptomic data and recursive feature elimination was used to determine which genes were most useful for classifying muscle biopsies into each type and MSA-defined subtype of myositis.

The support vector machine learning algorithm classified the muscle biopsies with >90% accuracy. Recursive feature elimination identified genes that are most useful to the machine learning algorithm and that are only overexpressed in one type of myositis. For example, CAMK1G (calcium/calmodulin-dependent protein kinase IG), EGR4 (early growth response protein 4) and CXCL8 (interleukin 8) are highly expressed in AS but not in DM or other types of myositis. Using the same computational approach, we also identified genes that are uniquely overexpressed in different MSAs-defined subtypes. These included apolipoprotein A4 (APOA4), which is only expressed in HMGCR myopathy, and MADCAM1 (mucosal vascular addressin cell adhesion molecule 1), which is only expressed in anti-Mi2-positive DM.

In conclusion, unique gene expression profiles in muscle biopsies from patients with MSAs-defined subtypes of myositis and IBM suggest that different pathological mechanisms underly muscle damage in each of these diseases.

Machine learning algorithms reveal unique gene expression profiles in muscle biopsies from patients with different types of myositis

Iago Pinal-Fernandez M.D., Ph.D., Maria Casal-Dominguez M.D., Ph.D., Assia Derfoul Ph.D., Katherine Pak M.D., Frederick W Miller M.D., Ph.D., Jose C Milisenda M.D., Josep M Grau-Junyent M.D., Ph.D., Albert Selva-O'Callaghan M.D., Ph.D., Carme Carrion-Ribas Ph.D., Julie J. Paik M.D., MHS, Jemima Albayda M.D., Lisa Christopher-Stine M.D., M.P.H., Thomas E. Lloyd M.D., Ph.D., Andrea M Corse M.D., Andrew L Mammen M.D., Ph.D.

Iago Pinal-Fernandez M.D., Ph.D., Maria Casal-Dominguez M.D., Ph.D., Assia Derfoul Ph.D., Katherine Pak M.D., and Andrew L. Mammen, M.D., Ph.D.: National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD. Iago Pinal-Fernandez M.D., Ph.D., Maria Casal-Dominguez M.D., Ph.D., Julie J Paik M.D., Jemima Albayda M.D., Lisa Christopher-Stine M.D., M.P.H., Thomas E. Lloyd M.D., Ph.D., Andrea Corse M.D. and Andrew L. Mammen M.D., Ph.D.: Johns Hopkins University School of Medicine, Baltimore, MD. Iago Pinal-Fernandez M.D., Ph.D., and Carme Carrion-Ribas Ph.D.: Faculty of Health Sciences, Universitat Oberta de Catalunya, Barcelona, Spain. Frederick W Miller, M.D., Ph.D., National Institute of Environmental Health Sciences, National Institutes of Health, Bethesda, MD. Jose C Milisenda M.D. and Josep M Grau-Junyent M.D., Ph.D.: Clinic Hospital and the University of Barcelona, Barcelona, Spain. Albert Selva-O'Callaghan M.D., Ph.D.: Vall d'Hebron Hospital and Autonomous University of Barcelona, Spain.

Address correspondence to: Andrew L. Mammen, M.D., Ph.D., or Iago Pinal-Fernandez, M.D., Ph.D. Muscle Disease Unit, Laboratory of Muscle Stem Cells and Gene Regulation, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, 50 South Drive, Room 1141, Building 50, MSC 8024, Bethesda, MD 20892. E-mail: andrew.mammen@nih.gov or iago.pinalfernandez@nih.gov. Phone: 301-451-1199. Fax: 301-594-0305.

Word count: 2999

KEY MESSAGES

What is already known about this subject?

- Different types of myositis are likely to have unique pathological mechanisms.

What does this study add?

- Machine learning algorithms can be trained on transcriptomic data to classify muscle biopsies from patients with DM, AS, IMNM, and IBM.
- Recursive feature elimination can be used to determine which genes are most important for the machine learning algorithms to classify the muscle biopsies.
- Only antisynthetase syndrome muscle biopsies express high levels of CAMKG, EGR4, and CXCL8 (interleukin 8).
- APOA4, a gene involved in cholesterol metabolism, is uniquely over-expressed in anti-HMGCR myopathy, which can be triggered by statins.
- MADCAM1, which recruits lymphocytes to target tissues, is uniquely over-expressed in muscle biopsies from those with anti-Mi2-positive dermatomyositis.

How might this impact on clinical practice?

Gene expression profiling of muscle biopsies from individual myositis patients may identify specific pathologic pathways that could be used to tailor therapies.

ABSTRACT

Objectives: Myositis is a heterogeneous family of diseases that includes dermatomyositis (DM), antisynthetase syndrome (AS), immune-mediated necrotizing myopathy (IMNM), inclusion body myositis (IBM), polymyositis, and overlap myositis. Additional subtypes of myositis can be defined by the presence of myositis-specific autoantibodies (MSAs). The purpose of this study was to define unique gene expression profiles in muscle biopsies from patients with DM, AS, IMNM, IBM, and the MSA-defined subtypes of myositis.

Methods: RNAseq was performed on muscle biopsies from 119 myositis patients with IBM or defined MSAs and 20 controls. Machine learning algorithms were trained on transcriptomic data and recursive feature elimination was used to determine which genes were most useful for classifying muscle biopsies into each type and MSA-defined subtype of myositis.

Results: The support vector machine learning algorithm classified the muscle biopsies with >90% accuracy. Recursive feature elimination identified genes most useful to the machine learning algorithm and that are only overexpressed in one type of myositis. For example, CAMK1G, EGR4, and CXCL8 are highly expressed in AS but not in DM or other types of myositis. Using the same computational approach, we also identified genes that are uniquely overexpressed in different MSA-defined subtypes. These included APOA4, which is only expressed in anti-HMGCR myopathy, and MADCAM1, which is only expressed in anti-Mi2-positive DM.

Conclusions: Unique gene expression profiles in muscle biopsies from patients with DM, AS, IMNM, IBM and different MSA-defined subtypes of myositis suggest that different pathological mechanisms underly muscle damage in each of these diseases.

MESH Keywords: Myositis, Autoantibodies, Autoantigens, Skeletal Muscle, Interferons

INTRODUCTION

The idiopathic inflammatory myopathies (IIM) are a heterogeneous family of diseases that includes six major types: dermatomyositis (DM), antisynthetase syndrome (AS), immune-mediated necrotizing myopathy (IMNM), inclusion body myositis (IBM), polymyositis, and overlap myositis [1]. Furthermore, 50-80% of IIM patients have myositis-specific autoantibodies (MSAs) that define phenotypically distinct IIM subtypes[2 3].

Muscle biopsies from patients with each major type of myositis have distinctive pathological features. For example, perifascicular myofiber atrophy and/or necrosis is a characteristic feature of both DM and AS, IMNM biopsies have abundant scattered necrotic myofibers, and IBM muscle biopsies usually include myofibers with cytoplasmic vacuoles[4]. However, histologic features that can reliably distinguish between DM and AS have not been identified. Similarly, histologic features cannot reliably be used to distinguish between different MSA-defined subtypes of DM or IMNM. Thus, it remains unclear whether different pathological pathways lead to muscle damage in the different myositis types and MSA-defined subtypes.

The advent of gene chip microarray and next-generation sequencing technologies has facilitated the use of myositis muscle biopsy gene expression profiles to identify pathological pathways. For example, microarray analysis led to the discoveries that type I and type II IFN-inducible genes are upregulated in muscle biopsies from patients with DM[5] and IBM[6 7], respectively. However, disease-specific gene expression profiles have not been fully described in patients with IMNM, AS, or any of the autoantibody-defined subtypes of DM. Furthermore, little attention has been

given to genes that are differentially expressed between patients with different types and subtypes of myositis.[8-11] In the current study, we trained machine learning algorithms to classify muscle biopsies using transcriptomic data from normal, IBM, and MSA-positive muscle biopsies; biopsies from the 20-50% of myositis patiens who are MSA-negative were not included in this study. We then used recursive feature elimination to identify novel disease-specific gene expression patterns that may be pathologically relevant in DM, AS, IMNM, IBM, and MSA-defined subtypes of myositis.

MATERIALS AND METHODS

Patients, samples, and autoantibody testing

Muscle biopsies obtained from subjects enrolled in IRB-approved longitudinal cohorts from the NIH (IRB number 91-AR-0196), the Johns Hopkins Myositis Center (IRB number NA_00007454), the Clinic Hospital (Barcelona; IRB number HCB/2015/0479), and the Vall d'Hebron Hospital (Barcelona; IRB number PR (AG) 68/2008) were included in the study if the patients fulfilled IBM criteria according to Lloyd,[12] or had one of the following MSAs: anti-NXP2, -Mi2, -TIF1γ, -MDA5, -HMGCR, -SRP, or -Jo1. Autoantibody testing was performed as previously described for anti-HMGCR and by line blot for the others (EUROLINE Myositis Profile 4). Patients were classified as having the antisynthetase syndrome (AS) if they had autoantibodies against Jo-1 and fulfilled Connor's AS criteria,[13] in the DM group if they had autoantibodies recognizing Mi2, NXP2, TIF1γ or MDA5 and in the IMNM group if they tested positive for anti-SRP or anti-HMGCR autoantibodies. Normal muscle biopsies were obtained from the Johns Hopkins Neuromuscular Pathology Laboratory (n=10) and the Skeletal Muscle Biobank of the University of Kentucky (n=10).

Standard protocol approvals and patient consents.

This study was approved by the Institutional Review Boards at participating institutions and written informed consent was obtained from each participant.

Human muscle biopsy processing, human skeletal muscle cell culture, and mouse muscle injury

See Supplementary Methods.

RNA-sequencing

RNA-sequencing (RNA-seq) was performed as previously described.[14] Briefly, RNA was prepared using TRIzol. Libraries were prepared using the NeoPrep™ system according to the TruSeq Stranded mRNA Library Prep protocol (Illumina) and sequenced using the Illumina HiSeq 2500 or 3000. Reads were aligned using the STAR v.2.5 25, the abundance of each gene was quantified using StringTie v.1.3.3.26 and the differential gene expression was performed using DESeq2 v.1.20 (Supplementary Methods). The Benjamini-Hochberg correction was used to adjust for multiple comparisons and a corrected p-value (q-value) of 0.05 or less was considered statistically significant.

Pathway analysis

We used Ingenuity Pathway Analysis v.01-07 and genes with a q-value below 0.05 and an expression ratio greater than 2 in each group compared to the rest of the biopsies were included in the analysis. Immunologic pathways with a z-score over 2 were selected.

RNAseq-based classification

To find the ability of RNAseq data to classify different types of myositis we first tested several classification models. Next, we performed stratified cross-validation to estimate the accuracy of each model. All steps were performed using Python v.3.6.3. Numpy v.1.13.3 and Pandas v.0.20.3 were used for data wrangling and basic statistical calculations, respectively (Supplementary Methods).

Those genes with significantly differential expression levels in one group compared to the rest of the biopsies were included in each model. The sample was split into a training set containing 2/3 of the observations and a test set containing the remaining 1/3. The training set was used to build the classificatory models and the testing set to evaluate the accuracy of the model. The machine learning models were developed using the package Scikit-learn v.0.19.1. Models were built using 2/3 random resamples of the data and tested in the remaining 1/3. The accuracy of classifying each of the myositis subsets was determined based on the mean and 95% CI of one thousand resampling cycles (Supplementary Methods).

Recursive feature elimination was applied to the whole dataset to rank each gene according to how useful it was for the model to differentiate the different patient groups. The RFE technique was applied through its implementation in Scikit-learn v.0.19.1 (Supplementary Methods).

Statement of patient and public involvement

Neither patients nor the public were involved in the design, conduct, reporting, or dissemination of this research.

Data availability statement

Deidentified RNAseq data will be made available upon request to Dr. Andrew Mammen (andrew.mammen@nih.gov)

RESULTS

Machine learning models accurately classify muscle biopsies

Muscle biopsy specimens were available from 119 myositis patients including 39 with DM (11 anti-Mi2-, 12 anti-NXP2-, 11 anti-TIF1γ-, and 5 anti-MDA5-positive), 49 with IMNM (9 anti-SRP- and 40 anti-HMGCR-positive), 18 with anti-Jo1-positive AS, and 13 with IBM. Twenty normal muscle biopsy specimens were utilized as comparators. Expression levels of all genes were determined for each sample by RNAseq. Details regarding the patients and their muscle biopsy features are found in Supplementary Table 1. Expression levels of genes associated with immune cells, regenerating myofibers, and mature skeletal muscle are found in Supplementary Figure 1.

First, we identified those genes with statistically significant differential expression in controls and each major type of myositis compared to the rest of the groups. A total of 10,141 differentially expressed genes were identified and the top 10 for each group are listed in Table 1. For example, the interferon-inducible gene ISG15 is the top differentially expressed gene in both DM and normal muscle biopsies; it is expressed at 43-fold higher levels in DM compared to the rest of the groups and at 17-fold lower levels in normal biopsies compared to the rest of the groups.

Table 1. Genes differentially expressed in muscle biopsies from each major type of myositis and controls compared to the rest of the samples.

NT			DM (All)			Anti-Mi2			Anti-NXP2			Anti-MDA5			Anti-TIF1g		
gene	FC	qval	gene	FC	qval	gene	FC	qval	gene	FC	qval	gene	FC	qval	gene	FC	qval
ISG15	-17	4.0E-40	**ISG15**	43	1.3E-139	**SCRT1**	14	1.4E-20	**ISG15**	12	4.2E-22	**ZFHX2**	11	1.3E-22	**MX1**	8	8.0E-16
IFI6	-15	6.6E-39	**IFI6**	25	2.9E-107	**KCNJ4**	9	2.0E-16	**IFI6**	8	3.0E-16	**ISG15**	18	1.3E-22	**ISG15**	9	8.1E-16
PSMB8	-9	9.2E-32	**RSAD2**	12	1.8E-78	**COL11A2**	5	2.0E-16	**RSAD2**	6	1.0E-15	**DNAH1**	10	4.5E-16	**IFI6**	7	3.0E-13
SECTM1	-10	2.3E-30	**MX1**	14	1.5E-75	**CHRM4**	11	1.0E-15	**KLHDC7B**	10	2.4E-14	**USP5**	5	2.8E-15	**HERC6**	5	1.8E-12
ACTC1	-11	6.0E-30	**CMPK2**	9	1.2E-65	**MADCAM1**	8	1.2E-13	**IFIT2**	5	7.7E-13	**RRP7A**	6	3.0E-15	**SUSD2**	5	3.6E-11
IFI30	-12	5.6E-29	**MX2**	8	3.4E-55	**IFI6**	7	1.2E-13	**MX1**	5	8.6E-11	**AGPAT2**	9	1.1E-14	**DHX58**	4	7.8E-11
SIGLEC1	-7	3.8E-28	**IFI27**	8	3.5E-55	**SPIB**	7	4.7E-13	**HERC5**	5	3.9E-10	**POU5F1P4**	9	1.3E-14	**MX2**	5	1.2E-10
MX1	-9	3.0E-27	**OAS3**	8	2.4E-54	**ISG15**	7	1.9E-12	**NDUFB2-AS1**	4	3.9E-10	**HOXB-AS1**	11	2.4E-14	**IFI44**	5	1.4E-10
OAS1	-9	3.3E-27	**HERC6**	7	7.6E-53	**MX1**	6	5.3E-12	**LOC101928053**	5	3.9E-10	**ACOT9**	5	3.0E-14	**RSAD2**	5	1.6E-10
MX2	-7	5.9E-27	**OAS1**	9	8.2E-52	**COX6B2**	6	9.1E-12	**OAS3**	4	6.1E-10	**FRA10AC1**	5	4.8E-14	**HELZ2**	4	1.8E-10

AS (Jo1)			IBM			IMNM (All)			Anti-HMGCR			Anti-SRP		
gene	FC	qval	gene	FC	qval	gene	FC	qval	gene	FC	qval	gene	FC	qval
EGR4	6	1.2E-09	**MYH4**	14	7.3E-19	**ISG15**	-12	3.0E-49	**ISG15**	-14	4.6E-51	**ISG15**	-6	4.5E-08
BRE-AS1	4	3.2E-09	**ISG15**	-6	3.9E-12	**RSAD2**	-7	2.1E-38	**RSAD2**	-7	2.4E-38	**RSAD2**	-4	1.9E-07
RNF165	4	3.2E-09	**CRYBG3**	7	1.8E-09	**KLHDC7B**	-10	1.5E-31	**KLHDC7B**	-15	2.4E-32	**IFI6**	-5	1.2E-06
CAMK1G	6	8.2E-08	**AHNAK**	2	3.5E-09	**IFI6**	-7	6.5E-30	**MX1**	-6	4.5E-31	**IRF9**	-3	9.9E-05
SAA1	5	1.6E-07	**FCRL6**	7	3.5E-09	**CMPK2**	-4	9.0E-25	**IFI44L**	-6	1.1E-30	**STAT1**	-3	2.2E-04
ALPL	3	4.2E-07	**GBP6**	7	6.4E-09	**MX1**	-5	1.2E-24	**CMPK2**	-5	6.8E-29	**IFI27**	-3	2.9E-04
IL1RL1	5	4.6E-07	**KIAA1147**	2	9.8E-09	**IFI27**	-5	5.8E-24	**OAS3**	-5	4.6E-28	**ZNFX1**	-2	2.9E-04
SPP1	5	5.9E-07	**PPM1L**	3	1.2E-08	**ZBP1**	-8	1.8E-23	**IFI44**	-5	1.4E-27	**DDX58**	-3	4.4E-04
MIR6087	4	1.5E-06	**LOC100128494**	5	3.5E-08	**DDX58**	-4	7.8E-23	**OAS2**	-4	5.6E-26	**NDUFS2**	3	4.4E-04
PBDC1	2	1.9E-06	**KIAA0754**	5	3.5E-08	**IFI44L**	-5	2.2E-22	**IFI6**	-6	2.1E-25	**IFIT2**	-3	4.4E-04

NT: normal muscle tissue; DM: dermatomyositis; AS: antisynthetase syndrome; IBM: inclusion body myositis; IMNM: immune-mediated necrotizing myositis; FC: fold-change; qval: adjusted p-value. The name and location of the genes is indicated in Supplementary Table 2.

To determine whether machine learning programs could use transcriptomic data to accurately classify patients into each major type of myositis or the control group, all differentially expressed genes were included in each of 10 machine learning models (Supplementary Methods). From among the models tested, the linear support vector machine (SVM) model performed the best with accuracies of 91% or greater to identify normal DM, AS, IMNM and IBM muscle biopsies. (Table 2).

Table 2. A comparison of machine learning models to classify muscle biopsies based on gene expression data. Accuracy and 95% confidence interval in the 1000 test sets of the different machine learning models to classify muscle biopsies into normal muscle tissue (NT), dermatomyositis (DM), antisynthetase syndrome (AS), inclusion body myositis (IBM) or immune-mediated necrotizing myopathy (IMNM).

	NT	DM	AS	IBM	IMNM
Linear SVM	94.7 [87.2-100.0]	92.0 [85.1-97.9]	91.0 [85.1-95.7]	95.0 [91.5-100.0]	92.0 [85.1-97.9]
AdaBoost	91.5 [83.0-97.9]	89.6 [80.9-95.7]	89.1 [83.0-93.6]	91.9 [80.9-97.9]	85.8 [76.6-93.6]
Gaussian Process	94.2 [87.2-100.0]	82.9 [74.5-91.5]	87.2 [80.9-91.5]	91.0 [85.1-95.7]	79.6 [68.1-89.4]
Nearest Neighbors	91.5 [85.1-97.9]	87.8 [80.9-95.7]	87.2 [83.0-89.4]	90.6 [89.4-93.6]	77.4 [66.0-87.2]
Random Forest	89.7 [83.0-95.7]	85.6 [76.6-93.6]	85.7 [78.7-91.5]	90.4 [87.2-93.6]	78.3 [68.1-87.2]
Neural Network	89.1 [72.3-97.9]	83.5 [44.7-95.7]	87.4 [74.4-93.6]	91.1 [89.4-97.9]	71.6 [36.2-95.7]
Decision Tree	87.8 [76.6-95.7]	86.5 [76.6-93.6]	85.0 [74.5-91.5]	85.7 [76.6-93.6]	76.1 [57.4-89.4]
RBF SVM	85.1 [85.1-85.1]	82.6 [76.6-87.2]	87.2 [87.2-87.2]	89.4 [89.4-89.4]	64.0 [63.8-66.0]
Gaussian Naïve Bayes	85.1 [85.1-85.1]	80.2 [70.2-89.4]	86.4 [83.0-89.4]	89.3 [87.2-91.5]	66.1 [53.2-78.7]
QDA	86.5 [78.7-93.6]	63.5 [48.9-76.6]	75.5 [61.7-87.2]	80.4 [68.1-89.4]	63.1 [46.8-76.6]

SVM: support vector machines; RBF: radial basis function; AdaBoost: adaptative boosting; QDA: quadratic discriminant analysis. The models are sorted based on the average accuracy of all the groups.

Identifying genes with unique expression patterns in DM, AS, INMM, and IBM

We expected that for each major type of myositis, those genes contributing most to the accuracy of the machine learning classification model would be involved in disease-specific pathological processes. To identify which among the thousands of

differentially expressed genes used by the linear SVM model are most useful to classify a biopsy into each type of myositis, we used the recursive feature elimination technique.[15] This method systematically eliminates genes with the weakest role in the model, leaving those that are most important to classify muscle biopsies into the correct group. Table 3 lists the 10 genes whose expression levels have the greatest utility to identify samples as belonging to each type of myositis or control group. Figure 1 shows the expression levels of the 3 most important genes from each group.

Table 3. The top 10 most useful genes to differentiate biopsy samples using the recursive feature elimination technique on the support vector machine model.

NT	DM	AS	IBM	IMNM
ACTC1	MX1	CAMK1G	MYH4	STAT1
LOC151121	TUBA1A	EGR4	H19	MYH8
SAA1	ISG15	CXCL8	JCHAIN	PSMB9
SOCS3	MCU	PROK2	CFAP126	KLF10
ANKRD1	HIST2H2AA3	NT5C3A	NT5C1A	MYBPH
NREP	IFI6	CXCL9	CCL13	ISG15
CCDC3	RARRES3	CAPN6	S100A9	MIR23A
PLEKHO1	CYB5R3	RAB13	COQ10A	COL3A1
SAA2	IGFN1	ANKRD28	DBNDD1	IGLL5
MYBPH	CDKN1A	C2ORF40	ZNF106	HIST1H2BD

NT: normal muscle tissue; DM: dermatomyositis; AS: antisynthetase syndrome; IBM: inclusion body myositis; IMNM: immune-mediated necrotizing myopathy; The name and location of the genes is indicated in Supplementary Table 2.

We first sought to validate this approach by determining whether it would identify key genes already known to play roles in DM pathogenesis. As genes upregulated by type I IFN are known to be expressed at high levels in DM muscle[5 16], we expected that expression levels of type I IFN-inducible genes should be important for the linear SVM model. Indeed, high expression levels of type 1 IFN-inducible genes MX1 and

ISG15 were among the 3 most important features used to identify DM muscle biopsies (Table 3).

When applied to the AS group, recursive feature elimination identified CAMK1G (calcium/calmodulin-dependent protein kinase IG), EGR4 (early growth response protein 4), and CXCL8 (interleukin 8) as the 3 most important genes (Table 3). Each of these genes is expressed at markedly higher levels in AS than in the other groups (Figure 1).

High expression levels of MYH4 (myosin heavy chain 4) and JCHAIN (the joining chain of multimeric IgA and IgM) were among the 3 most important features used by the linear SVM model to identify samples as belonging to the IBM group (Table 3 and Figure 1). In addition, the low expression level of H19 (a noncoding RNA) in IBM compared to DM, AS, and IMNM (Figure 1) appeared to be important for IBM classification.

Expression levels of STAT1 (signal inducer and activator of transcription 1), MYH8 (myosin heavy chain 8), and PSMB9 (proteasome subunit beta 9) were the top features used to classify a muscle biopsy as IMNM (Table 3). Based on the patterns of expression (Figure 1), the model seems to rely both on the low expression of IFN-inducible genes STAT1 and PSMB9 (expressed at high levels in DM, AS, and IBM) as well as the high expression of MYH8 (expressed at low levels in normal muscle) to classify biopsies as IMNM.

The expression levels of ACTC1 (actin alpha cardiac muscle 1), LOC151121 (a non-coding gene), and SAA1 (serum amyloid A1) were the top features used to classify normal muscle biopsies (Table 3). Interestingly, normal muscle biopsies were

characterized by low levels of ACTC1, which encodes a structural protein expressed during muscle regeneration[17] (Figure 1). Similarly, the SAA1 gene, which encodes the acute phase reactant serum amyloid A1, was expressed at low levels in normal muscles and high levels in all of the myositis groups. In contrast, LOC151121 was expressed at high levels in normal muscle but at low levels in all the myositis groups (Figure 1).

Identifying genes with unique expression patterns in the different subtypes of IMNM and DM

Using the same methodology, we next identified those genes that were most useful to classify biopsies according to the different autoantibody-defined subtypes within IMNM and DM. This revealed that APOA4 (apolipoprotein A4) was selectively expressed in IMNM patients with anti-HMGCR autoantibodies (Figure 2). Similarly, MADCAM1 (mucosal vascular addressin cell adhesion molecule 1) was exclusively detectable in DM patients with anti-Mi2 autoantibodies (Figure 2).

Pathway analysis

To gain further insight into the biological processes that distinguish each group compared to the others, we performed pathway analyses. For each analysis, we included the set of genes differentially expressed by at least two-fold in the type of myositis (or control) compared to the rest of the biopsies. Pathways annotated as related to the “cellular immune response”, “cytokine signaling”, and “humoral immune response” (i.e., immunologic pathways) were included in each analysis.

As expected, “interferon signaling” was the top over-represented immunologic pathway in DM (Figure 3). The AS and IBM biopsies shared the same top 3 over-

represented pathways that were not included DM, IMNM, or control biopsies. These included the T cell pathways “ICOS-ICOSL signaling in T helper cells”, “CD28 signaling in T helper cells”, and the “Th1 pathway”. No immunologic pathways were over-represented in IMNM biopsies. Rather, IMNM biopsies, like control biopsies, were notable for the under-representation of pathways that were important in DM, AS, and/or IBM.

Muscle regeneration genes are among the top differentially expressed genes in IMNM and are also overexpressed in other types of myositis

To classify biopsies as IMNM, linear SVM relied on the relative underexpression of genes expressed at high levels in other forms of myositis (e.g., STAT1 and PSMB8)[16] rather than on genes that were uniquely overexpressed in IMNM. To further investigate pathological processes important for IMNM, we considered the known functions of the top 10 overexpressed genes in biopsies from these patients (Table 4). Interestingly, several of these are known to play a role in skeletal muscle differentiation and/or muscle repair. For example, ACTC1 encodes alpha-actin which is expressed in early adult skeletal muscle.[17] Similarly, TNC encodes an extracellular matrix protein that is expressed only in actively remodeling musculoskeletal tissue.[18]

Table 4. The top 10 up-regulated genes in each type of myositis compared to normal biopsies.

DM (All)			Anti-Mi2			Anti-NXP2			Anti-MDA5			Anti-TIF1g		
gene	FC	qval	gene	FC	qval	gene	FC	qval	gene	FC	qval	gene	FC	qval
ISG15	101	1.06E-91	**IFI6**	62	2.6E-43	**ISG15**	110	1.5E-55	**ISG15**	163	1.2E-51	**ISG15**	84	1.5E-47
IFI6	67	2.73E-80	**ISG15**	67	5.9E-43	**IFI6**	70	7.8E-48	**IFI6**	72	2.6E-37	**IFI6**	62	1.7E-43
MX1	29	2.58E-56	**MX1**	32	1.9E-33	**RSAD2**	23	2.1E-33	**ZFHX2**	18	9.9E-25	**MX1**	41	9.0E-39
RSAD2	18	1.25E-49	**OAS1**	25	9.7E-28	**MX1**	29	8.6E-33	**IFI35**	30	2.3E-24	**MX2**	22	6.9E-32
MX2	17	2.48E-49	**MX2**	18	1.2E-27	**IFIT2**	22	5.6E-31	**ACP5**	35	2.2E-23	**OAS1**	29	3.8E-31
OAS1	23	4.00E-48	**LY6E**	16	4.5E-26	**KLHDC7B**	65	1.1E-30	**SECTM1**	26	1.0E-22	**IFITM1**	12	2.7E-28
IRF9	10	2.54E-43	**RSAD2**	17	8.1E-26	**OAS1**	26	4.7E-30	**ZBP1**	42	1.6E-22	**RSAD2**	18	2.0E-27
IFITM1	10	8.62E-43	**CMPK2**	15	2.2E-25	**OAS3**	16	1.3E-26	**CLEC4GP1**	19	5.4E-21	**TYMP**	30	2.0E-27
CMPK2	13	4.24E-42	**OAS3**	16	3.6E-25	**HERC5**	25	3.4E-26	**PSMB8**	19	3.3E-20	**IFI44**	18	4.1E-27
OAS3	14	2.27E-41	**KLHDC7B**	39	3.0E-23	**MX2**	15	4.8E-26	**IFI27**	18	2.2E-19	**DHX58**	13	6.2E-27

AS (Jo1)			IBM			IMNM (All)			Anti-HMGCR			Anti-SRP		
gene	FC	qval	gene	FC	qval	gene	FC	qval	gene	FC	qval	gene	FC	qval
PSMB8	13	2.64E-25	**GBP2**	7	1.27E-18	**SERPINA3**	22	1.63E-28	**ACTC1**	18	3.6E-31	**SERPINA3**	24	1.8E-17
ACTC1	18	2.33E-23	**BIRC3**	7	4.43E-18	**ACTC1**	15	1.08E-27	**SERPINA3**	20	5.8E-25	**ACTC1**	12	9.8E-13
GBP2	7	1.11E-22	**PSMB8**	9	1.70E-16	**CHRNA1**	8	6.40E-21	**MYH3**	17	8.1E-21	**HP**	18	1.6E-12
SAA1	37	1.11E-22	**GBP1**	11	3.50E-16	**IFITM10**	9	3.20E-20	**CHRNA1**	7	2.0E-19	**CHRNA1**	9	2.7E-12
SIK1	10	3.05E-21	**CCL13**	20	4.96E-16	**TNC**	11	5.38E-20	**IFITM10**	9	2.6E-19	**TNC**	11	6.7E-12
NNMT	14	9.65E-21	**ITGAL**	9	2.47E-15	**KRT80**	13	1.35E-19	**TNNT2**	16	1.7E-18	**DCLK1**	7	6.7E-12
MYH3	23	9.83E-21	**GBP5**	17	3.95E-15	**TNNT2**	15	1.35E-19	**ANKRD1**	10	2.4E-18	**KRT80**	13	6.7E-12
GADD45A	8	9.83E-21	**HLA-DQA1**	13	2.41E-14	**MYH3**	13	2.66E-19	**CSPG4**	5	2.4E-18	**IFITM10**	10	1.0E-11
GBP1	12	9.83E-21	**CD8A**	16	3.45E-14	**ANKRD1**	10	2.04E-18	**TNC**	10.06	2.6E-18	**RUNX1**	6.98	3.0E-11
IFI30	16	1.02E-20	**HLA-DOA**	10	3.58E-14	**DCLK1**	6	2.45E-18	**KRT80**	11.97	9.9E-18	**TNNT2**	13.81	3.3E-11

DM: dermatomyositis; AS: antisynthetase syndrome; IBM: inclusion body myositis; IMNM: inclusion body myositis; FC: fold-change; qval: adjusted p-value. The name and location of the genes is indicated in Supplementary Table 2.

To determine whether the other most overexpressed genes in IMNM play a role in muscle regeneration, we analyzed their expression levels in cultured human myoblasts as they differentiated to form myotubes. Each gene was expressed at low levels in myoblasts and at high levels in differentiating myotubes (Supplementary Fig 2). Similarly, these genes were expressed at low levels in healthy mouse muscle, but at high levels in regenerating mouse muscles following a muscle injury (Supplementary Fig 3). This pattern suggests that these genes are expressed as part of the muscle regeneration process induced by necrosis in IMNM muscle. Since regeneration is also a common feature of muscle biopsies from those with DM, AS, and IBM, we expected that muscle biopsies from each of these types of myositis should also have high levels of the genes overexpressed in IMNM. Indeed, even though they were not among the top 10 overexpressed genes in the other groups, each of these genes was highly expressed in the other types of myositis muscle but not control muscle (Supplementary Fig 4).

We next considered the known functions of the top 10 upregulated genes in DM, AS, and IBM compared to control muscle (Table 4). Consistent with prior studies, many of the top 10 differentially expressed genes in muscle biopsies from DM patients are inducible by interferon type I (e.g., ISG15[19 20], IFI6[21], and MX1[22]) (Table 4). Similarly, several of the most overexpressed genes in AS and IBM muscle biopsies are interferon type II inducible genes (e.g., PSMB8[23], GBP2, and GBP1[24 25]) (Table 4).

DISCUSSION

In this study, we showed that machine learning algorithms trained on transcriptomics data could accurately classify myositis muscle biopsies from DM, AS, IMNM, and IBM patients. This demonstrates that these IIM types have unique gene expression profiles. Indeed, by applying recursive feature elimination to the machine learning algorithms we identified novel gene markers (e.g., CAMK1G, EGR, and CXCL8) that are uniquely expressed in AS but not DM, even though these two diseases can be histologically indistinguishable. Moreover, we also identified genes (e.g., ACTC1 and SSA1) that are overexpressed in all types of myositis studied here but not in normal muscle. Finally, we confirmed previous observations related to the pathogenesis of myositis, including the role of interferon pathways in DM,[8 16] the prominence of muscle regeneration in IMNM,[26] and the presence of plasma cells in IBM (as evidenced by overexpression of JCHAIN, a plasma cell marker).[27 28]

We applied the same computational approach to identify genes that are uniquely upregulated in patients with different MSA-defined IIM subtypes. For example, although anti-SRP and anti-HMGCR myopathy muscle biopsies are histologically identical, we identified the APOA4 gene as being exclusively upregulated in the latter subtype of IMNM. Since statin exposure is a risk factor for developing anti-HMGCR myopathy but not other types of myositis[29], it is of interest that APOA4, which contributes to reverse cholesterol transport by facilitating the movement of cholesterol from the periphery to the liver for excretion[30], is only upregulated in anti-HMGCR myopathy muscle biopsies.

We also found that different MSA-defined DM subtypes had different gene expression profiles. For example, MADCAM1 was uniquely expressed in muscle biopsies from DM patients with anti-Mi2 autoantibodies. Of note, MADCAM1 is expressed on endothelial surfaces in the intestine where it mediates the migration of lymphocytes into the gut by binding to $\alpha_4\beta_7$ integrin found on the surface of CD4+ and CD8+ T-cells[31]. Since MADCAM1 recruits inflammatory cells to the gut in patients with colitis, we hypothesize that it could play a similar role in anti-Mi2-positive DM patients, who have more lymphocytic invasion of muscle fibers than DM patients with other autoantibodies[32]. This could have therapeutic implications since drugs that target the MADCAM1/$\alpha_4\beta_7$ pathway have already been developed.

This study was not designed to directly compare the performance of machine learning algorithms utilizing muscle biopsy transcriptomic data with the analysis of histologic features to diagnose different types of myositis. Still, the current study suggests that machine learning algorithms would fare favorably in such a comparison. For example, only 72% of biopsies from the included DM patients had perifascicular atrophy[32], the key feature required for histologic diagnosis of DM[33]. Nonetheless, the SVM algorithm diagnosed DM based on the muscle biopsy transcriptome with an accuracy of 92%. This raises the possibility that, with the availability of gene expression profile data collected from a large number of patients with different types of myopathy, machine learning algorithms could be diagnostically useful.

This study was limited in that we did not include muscle biopsies from all types of myositis. Indeed, we excluded biopsies from patients with polymyositis, overlap myositis, and MSA-negative forms of myositis. Furthermore, our analysis was restricted

to gene expression data and did not include analyses of the corresponding proteins. Nonetheless, by applying machine learning algorithms to muscle biopsy transcriptomic data, we have demonstrated that DM, AS, IMNM, and IBM can be distinguished based on their unique gene expression patterns. Furthermore, by applying recursive feature elimination to these classification models, we not only confirmed known pathological pathways in IIM, such as the role of type I interferon in DM, we also identified novel genes that are uniquely upregulated in other types and MSA-defined subtypes of myositis. We expect this computational approach could be useful for analyzing transcriptomic data from other autoimmune conditions in which there are different types and subtypes of the disease.

Figure 1. Expression levels of those genes most helpful to classify muscle biopsies into each type of myositis. The expression levels of the top 3 genes used by the support vector machine model to classify muscle biopsies from normal tissue (NT), dermatomyositis (DM), immune-mediated necrotizing myositis (IMNM), antisynthetase syndrome (AS) or inclusion body myositis (IBM).

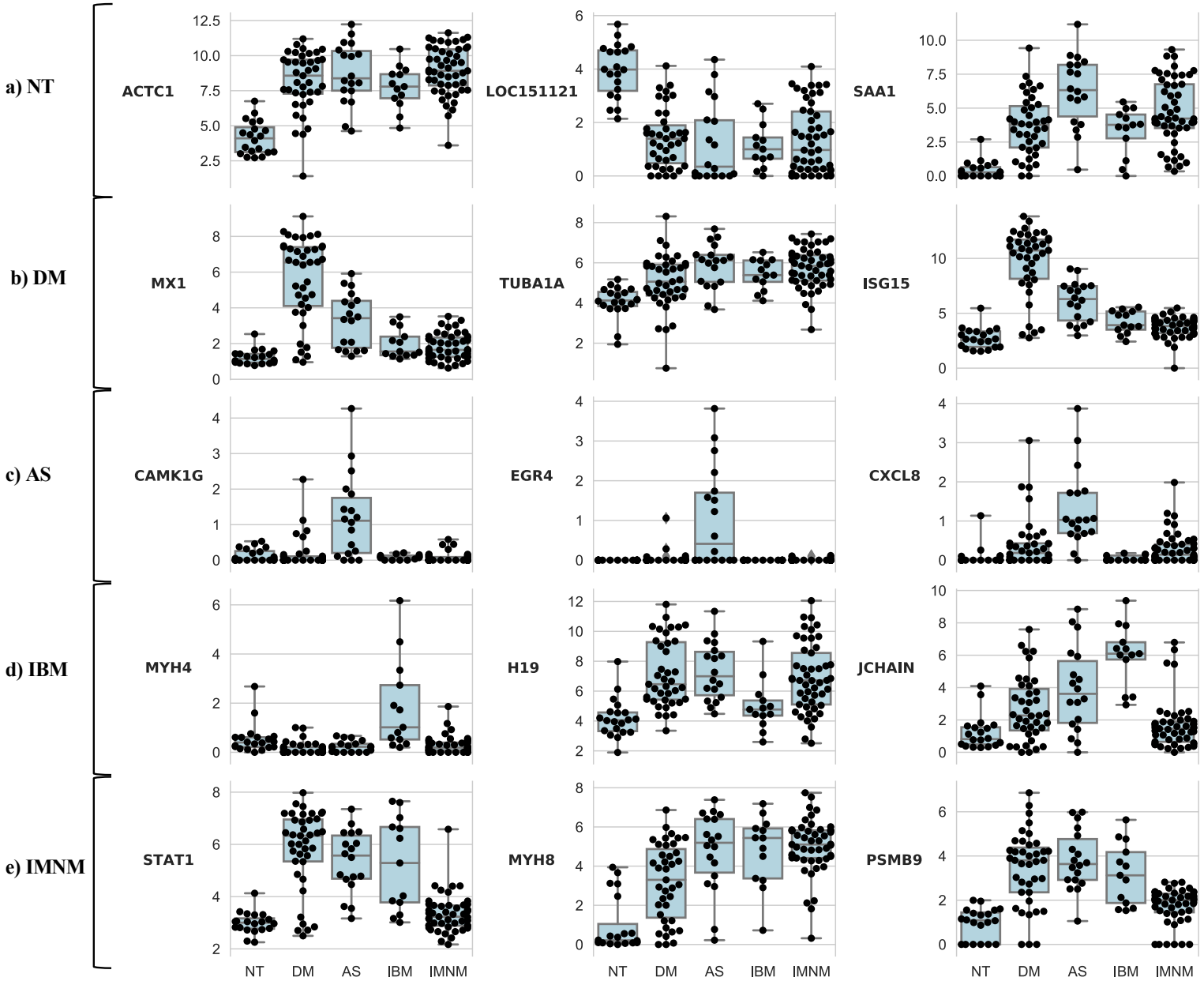
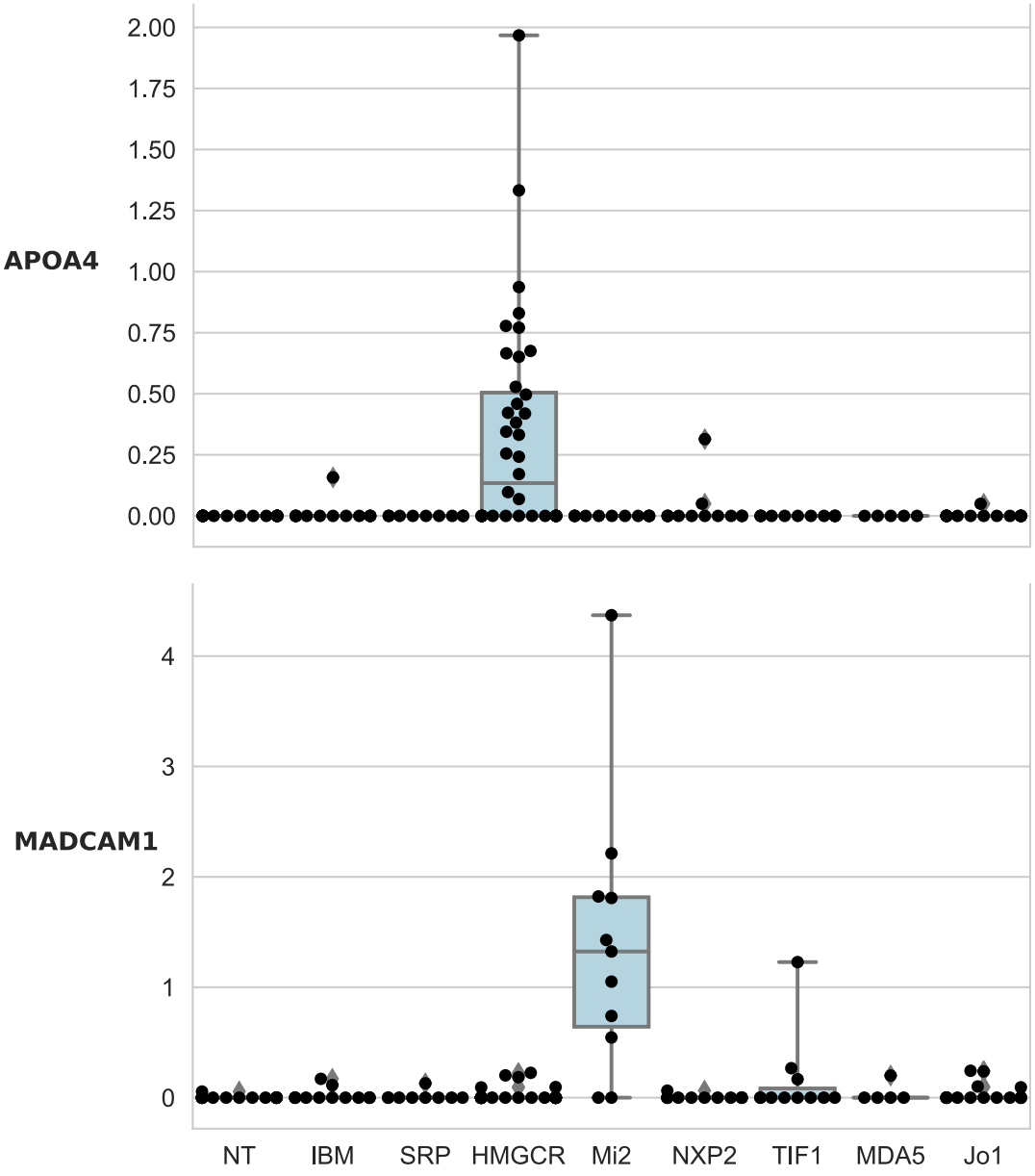
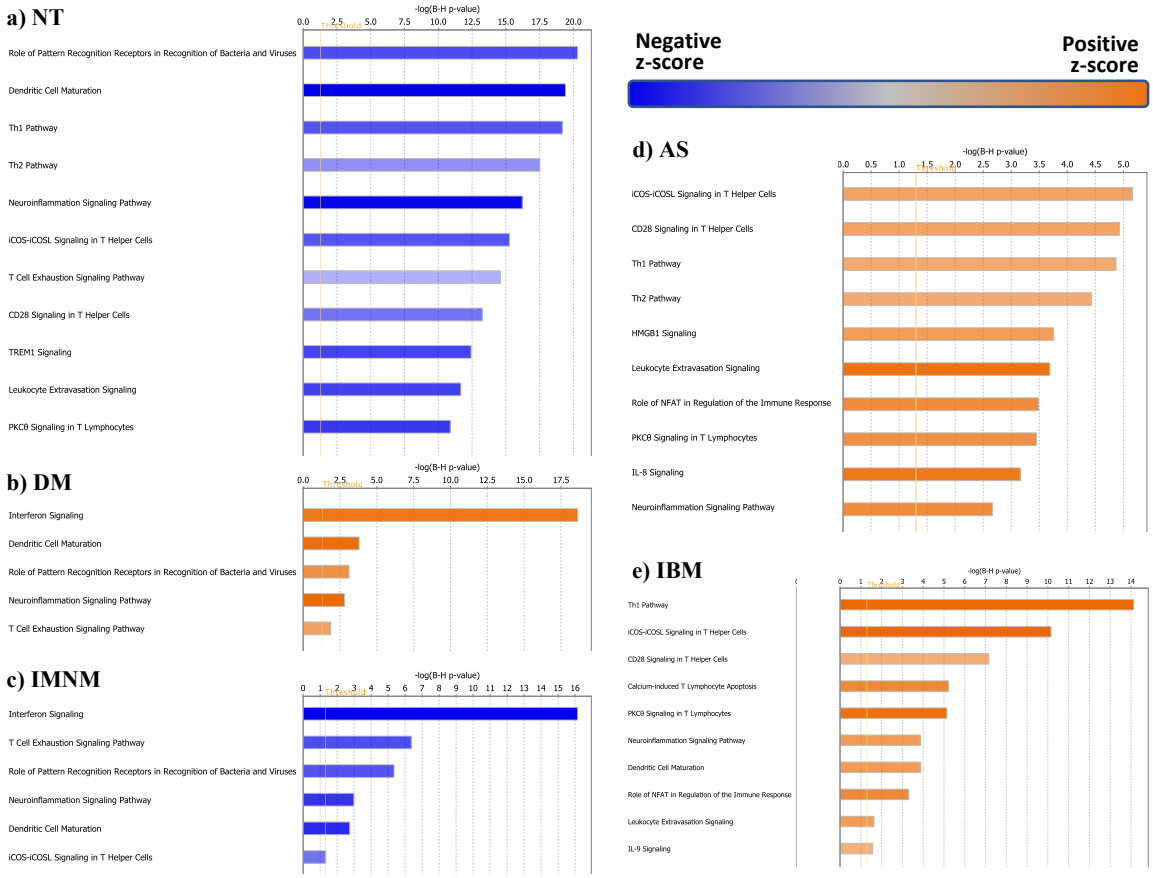


Figure 2. Genes selectively upregulated in different autoantibody-defined subtypes of myositis. APOA4 and MADCAM1 are selectively overexpressed (log2[FPKM + 1]) in anti-HMGCR IMNM (q-value compared to SRP: 0.0009) and anti-Mi2 DM (q-value compared to other DM antibodies: 2.9E-9), respectively.



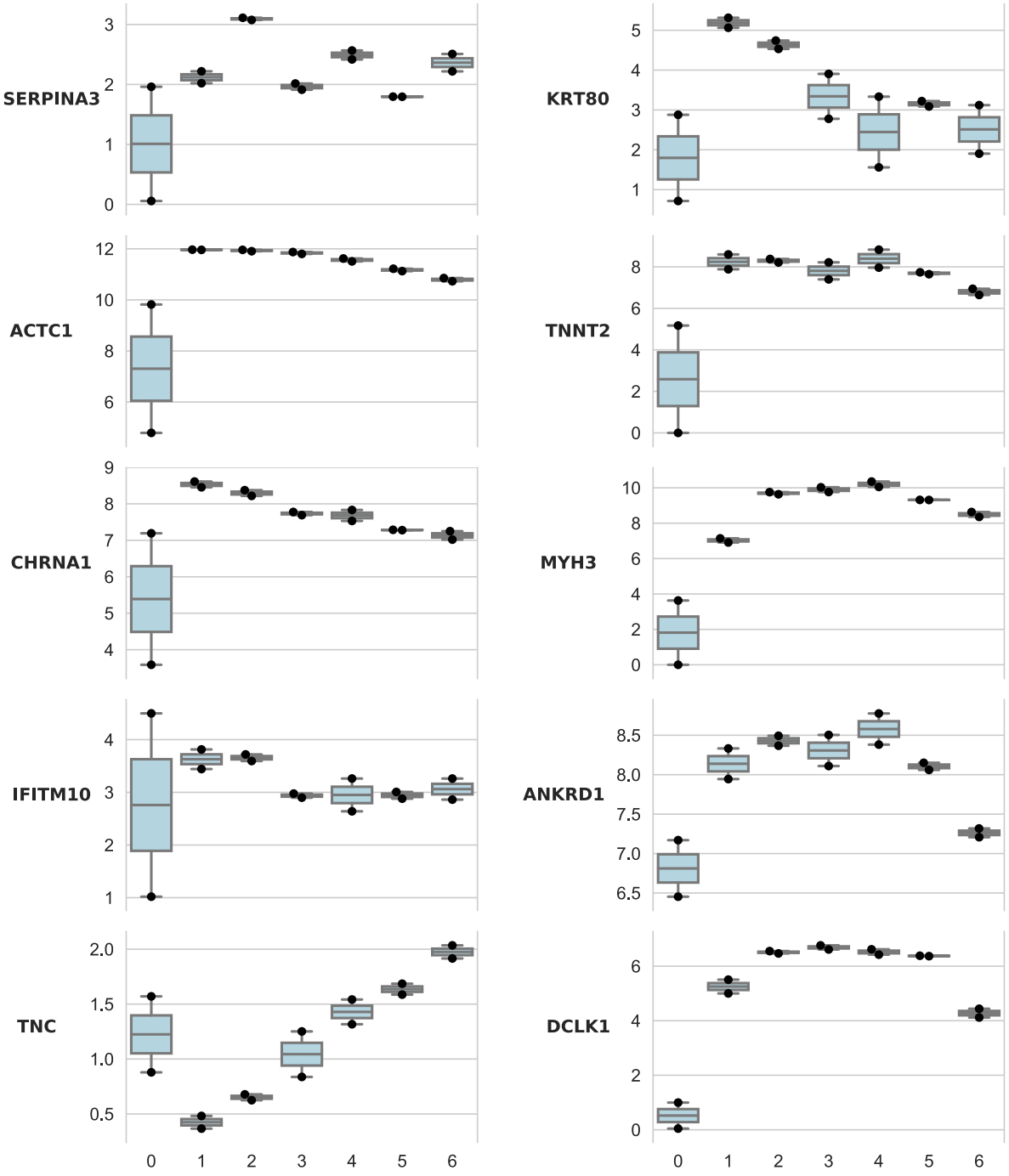
Normal tissue: NT; inclusion body myositis: IBM; anti-SRP IMNM: SRP; anti-HMGCR IMNM: HMGCR; anti-Mi2 DM: Mi2; anti-NXP2 DM: NXP2; anti-TIF1γ DM: TIF1; anti-MDA5 DM: MDA5; anti-Jo1 AS: Jo1.

Figure 3. Pathway analysis in myositis and normal muscle biopsies. The top 10 pathways of the different muscle biopsy groups are shown. NT, normal tissue; DM, dermatomyositis; IMNM, immune-mediated necrotizing myopathy; AS, antisynthetase syndrome; IBM, inclusion body myositis.

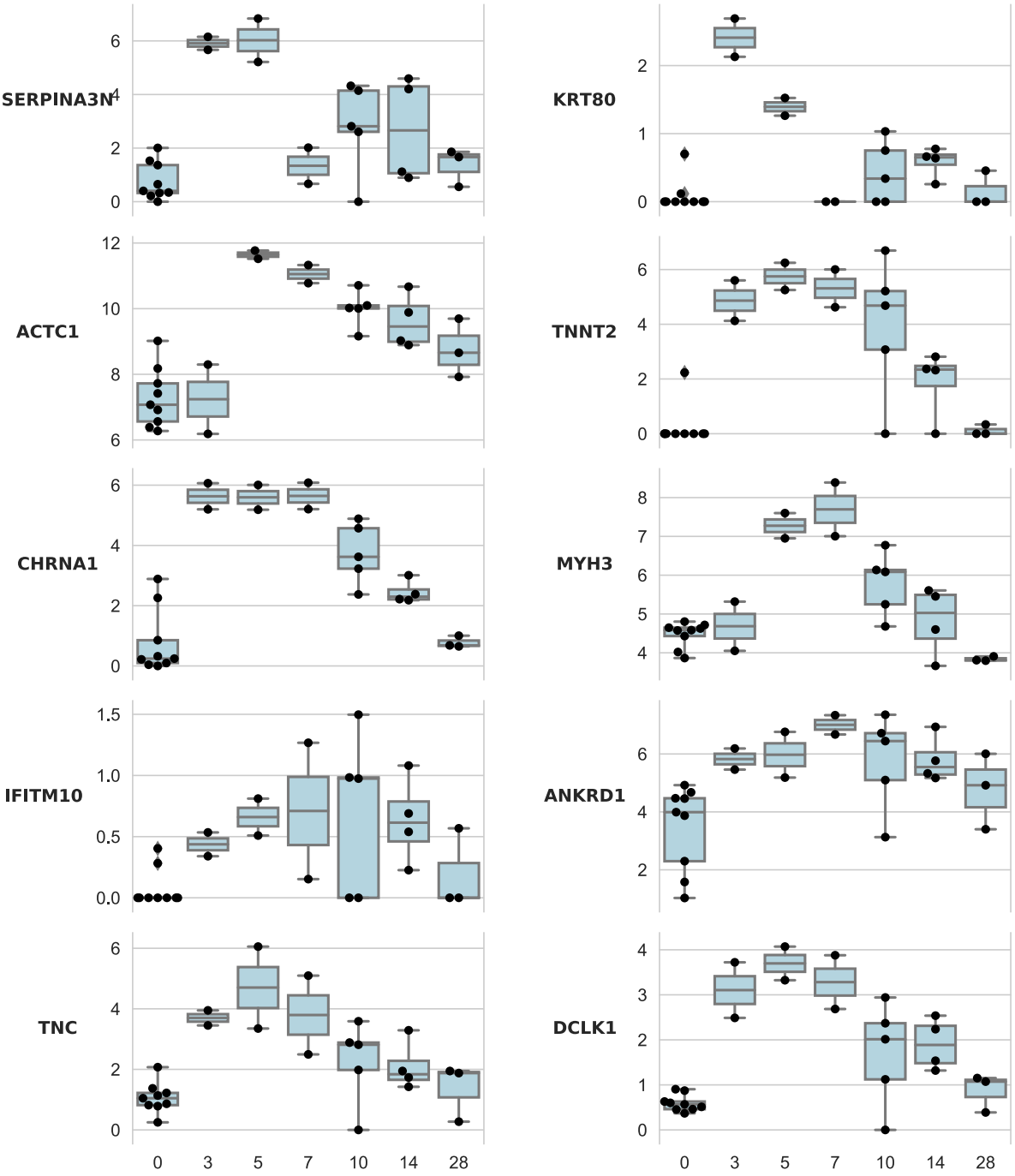


Supplementary Figure 1. RNA expression levels of genes associated with infiltrating immune cells, muscle regeneration, and adult skeletal muscle in myositis and healthy control muscle biopsies. The expression levels (log2[FPKM + 1]) of gene markers for T-cells (CD3E, CD4, and CD8A) and macrophages (CD14 and CD68), as well as muscle regeneration (NCAM1, MYOG, MYH3, MYH8) and adult muscle structural proteins (ACTA1, TTN, MYH1, MYH2) are shown for each of the major muscle biopsy groups. NT, normal tissue; DM, dermatomyositis; IMNM, immune-mediated necrotizing myopathy; AS, antisynthetase syndrome; IBM, inclusion body myositis.

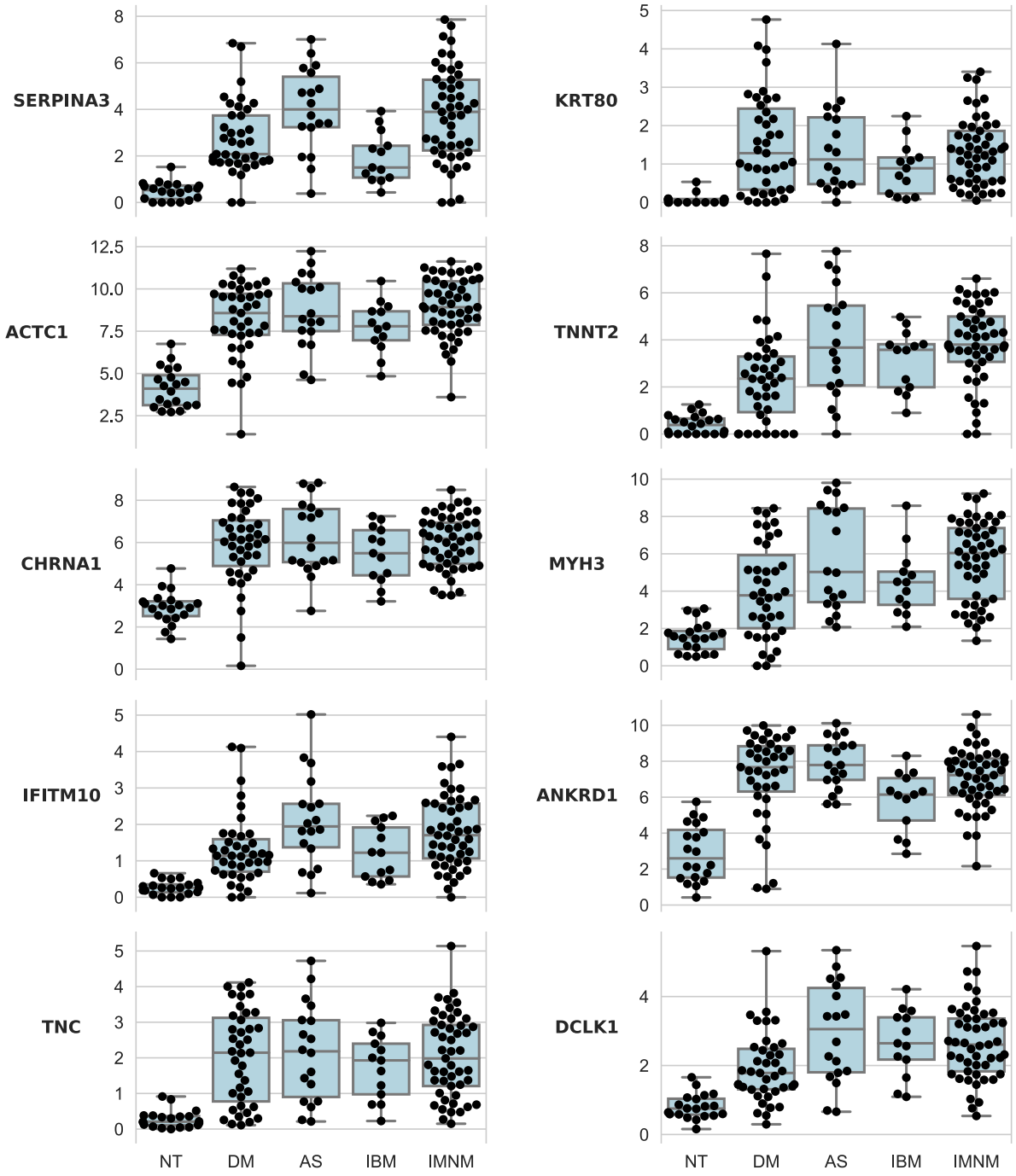
Supplementary Figure 2. Genes expressed at high levels in IMNM are expressed at low levels in proliferating human myoblasts and at high levels as they differentiate into myotubes. Evolution of the RNA expression levels (log2[FPKM + 1]) of the 10 most significantly differentially expressed genes in IMNM during the differentiation of human skeletal muscle myoblasts (HSMM) into myotubes. Proliferating myoblasts are placed in differentiation media on day 0 and allowed to differentiate into myotubes over the next 6 days.



Supplementary Figure 3. Genes expressed at high levels in IMNM are expressed at high levels during mouse muscle regeneration. Evolution of the expression levels (log2[FPKM + 1]) of the 10 most significantly differentially expressed genes in IMNM during mouse muscle regeneration. The tibialis anterior muscle was injected with cardiotoxin (CTX) on day 0 and allowed to regenerate for as long as 28 days. CTX day 0 corresponds to the contralateral (uninjured) tibialis anterior muscle.



Supplementary Figure 4. Muscle biopsies from all myositis groups express high levels of the most overexpressed genes in IMNM. RNA expression levels (log2[FPKM + 1]) of the 10 most upregulated genes in IMNM in all of the muscle biopsy groups. NT: normal tissue; DM: dermatomyositis: AS: antisynthetase syndrome; IBM: inclusion body myositis; IBM: inclusion body myositis.



Supplementary Table 1. Epidemiologic and muscle biopsy features of the study subjects.

	NT (n=20)	DM (n=39)	AS (n=18)	IBM (n=13)	IMNM (n=49)	Total (n=139)
Age at muscle biopsy	49.6 (13.4)	56.7 (15.5)	43.8 (13.7)	65.3 (10.8)	54.4 (14.5)	54.0 (15.1)
Years from symptom onset to biopsy	-	1.2 (2.5)	3.3 (4.9)	4.7 (5.0)	3.1 (4.2)	2.3 (3.8)
Female	70% (14)	59% (23)	78% (14)	46% (6)	57% (28)	61% (85)
Race						
White	75% (15)	72% (28)	50% (9)	85% (11)	63% (31)	68% (94)
Black	15% (3)	10% (4)	33% (6)	15% (2)	22% (11)	19% (26)
Other races	10% (2)	18% (7)	17% (3)	0% (0)	14% (7)	14% (19)
Treatments						
Any immunosuppressant	0% (0)	64% (25)	76% (13)	38% (5)	45% (22)	47% (65)
Corticosteroids	0% (0)	64% (25)	71% (12)	23% (3)	43% (21)	44% (61)
Azathioprine	0% (0)	8% (3)	18% (3)	0% (0)	6% (3)	7% (9)
Methotrexate	0% (0)	18% (7)	24% (4)	15% (2)	24% (12)	18% (25)
Mycophenolate	0% (0)	13% (5)	6% (1)	8% (1)	6% (3)	7% (10)
IVIG	0% (0)	5% (2)	6% (1)	0% (0)	6% (3)	4% (6)
Rituximab	0% (0)	3% (1)	0% (0)	0% (0)	2% (1)	1% (2)
Tacrolimus	0% (0)	3% (1)	0% (0)	0% (0)	2% (1)	1% (2)
Cyclosporine	0% (0)	3% (1)	6% (1)	0% (0)	0% (0)	1% (2)
Cyclophosphamide	0% (0)	0% (0)	6% (1)	0% (0)	0% (0)	1% (1)
Biopsy features						
Perifascicular atrophy	0% (0)	72% (28)	31% (5)	0% (0)	4% (2)	26% (35)
Perivascular inflammation	0% (0)	69% (27)	44% (7)	23% (3)	29% (14)	37% (51)
Myofiber necrosis	0% (0)	79% (31)	62% (10)	85% (11)	90% (44)	70% (96)
Myofiber regeneration	0% (0)	64% (25)	62% (10)	69% (9)	84% (41)	62% (85)
Primary inflammation (invasion)	0% (0)	0% (0)	25% (4)	62% (8)	12% (6)	13% (18)
Rimmed vacuoles	0% (0)	0% (0)	0% (0)	85% (11)	4% (2)	10% (13)
None of the above	100% (20)	13% (5)	17% (3)	0% (0)	8% (4)	23% (32)

NT: normal muscle tissue; IBM: inclusion body myositis; AS: antisynthetase syndrome; IMNM: immune-mediated necrotizing myositis; DM: dermatomyositis.

Supplementary Table 2. Name and location of the genes identified in the study.

Gene symbol	Gene name	Location
ACOT9	acyl-CoA thioesterase 9	Xp22.11
ACP5	acid phosphatase 5, tartrate resistant	19p13.2
ACTC1	actin alpha cardiac muscle 1	15q14
AGPAT2	1-acylglycerol-3-phosphate O-acyltransferase 2	9q34.3
AHNAK	AHNAK nucleoprotein	11q12.3
ALPL	alkaline phosphatase, biomineralization associated	1p36.12
ANKRD1	ankyrin repeat domain 1	10q23.31
ANKRD28	ankyrin repeat domain 28	3p25.1
BIRC3	baculoviral IAP repeat containing 3	11q22.2
BRE-AS1	BABAM2 antisense RNA 1	2p23.2
C2ORF40	ECRG4 augurin precursor	2q12.2
CAMK1G	calcium/calmodulin dependent protein kinase IG	1q32.2
CAPN6	calpain 6	Xq23
CCDC3	coiled-coil domain containing 3	10p13
CCL13	C-C motif chemokine ligand 13	17q12
CD8A	CD8a molecule	2p11.2
CDKN1A	cyclin dependent kinase inhibitor 1A	6p21.2
CFAP126	cilia and flagella associated protein 126	1q23.3
CHRM4	cholinergic receptor muscarinic 4	11p11.2
CHRNA1	cholinergic receptor nicotinic alpha 1 subunit	2q31.1
CLEC4GP1	C-type lectin domain family 4 member G pseudogene 1	19p13.2
CMPK2	cytidine/uridine monophosphate kinase 2	2p25.2
COL11A2	collagen type XI alpha 2 chain	6p21.32
COL3A1	collagen type III alpha 1 chain	2q32.2
COQ10A	coenzyme Q10A	12q13.3
COX6B2	cytochrome c oxidase subunit 6B2	19q13.42
CRYBG3	crystallin beta-gamma domain containing 3	3q11.2
CSPG4	chondroitin sulfate proteoglycan 4	15q24.2
CXCL8	C-X-C motif chemokine ligand 8	4q13.3
CXCL9	C-X-C motif chemokine ligand 9	4q21.1
CYB5R3	cytochrome b5 reductase 3	22q13.2
DBNDD1	dysbindin domain containing 1	16q24.3
DCLK1	doublecortin like kinase 1	13q13.3
DDX58	DExD/H-box helicase 58	9p21.1
DHX58	DExH-box helicase 58	17q21.2
DNAH1	dynein axonemal heavy chain 1	3p21.1
EGR4	early growth response 4	2p13.2
FCRL6	Fc receptor like 6	1q23.2
FRA10AC1	FRA10A associated CGG repeat 1	10q23.33
GADD45A	growth arrest and DNA damage inducible alpha	1p31.3
GBP1	guanylate binding protein 1	1p22.2
GBP2	guanylate binding protein 2	1p22.2
GBP5	guanylate binding protein 5	1p22.2
GBP6	guanylate binding protein family member 6	1p22.2
H19	H19 imprinted maternally expressed transcript	11p15.5
HELZ2	helicase with zinc finger 2	20q13.33
HERC5	HECT and RLD domain containing E3 ubiquitin protein ligase 5	4q22.1
HERC6	HECT and RLD domain containing E3 ubiquitin protein ligase family member 6	4q22.1
HIST1H2BD	H2B clustered histone 5	6p22.2
HIST2H2AA3	H2A clustered histone 18	1q21.2
HLA-DOA	major histocompatibility complex, class II, DO alpha	6p21.32
HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1	6p21.32
HOXB-AS1	HOXB cluster antisense RNA 1	17q21.32
HP	haptoglobin	16q22.2
IFI27	interferon alpha inducible protein 27	14q32.12
IFI30	IFI30 lysosomal thiol reductase	19p13.11
IFI35	interferon induced protein 35	17q21.31
IFI44	interferon induced protein 44	1p31.1
IFI44L	interferon induced protein 44 like	1p31.1
IFI6	interferon alpha inducible protein 6	1p35.3
IFIT2	interferon induced protein with tetratricopeptide repeats 2	10q23.31
IFITM1	interferon induced transmembrane protein 1	11p15.5
IFITM10	interferon induced transmembrane protein 10	11p15.5
IGFN1	immunoglobulin like and fibronectin type III domain containing 1	1q32.1
IGLL5	immunoglobulin lambda like polypeptide 5	22q11.22
IL1RL1	interleukin 1 receptor like 1	2q12.1
IRF9	interferon regulatory factor 9	14q12
ISG15	ISG15 ubiquitin like modifier	1p36.33
ITGAL	integrin subunit alpha L	16p11.2
JCHAIN	joining chain of multimeric IgA and IgM	4q13.3
KCNJ4	potassium inwardly rectifying channel subfamily J member 4	22q13.1
KIAA0754	KIAA0754	1p34.2

KIAA1147	DENN domain containing 11	7q34
KLF10	Kruppel like factor 10	8q22.3
KLHDC7B	kelch domain containing 7B	22q13.33
KRT80	keratin 80	12q13.13
LOC100128494	uncharacterized LOC100128494	11q13.4
LOC101928053	uncharacterized LOC101928053	11p15.4
LOC151121	long intergenic non-protein coding RNA 1854	2q21.1
LY6E	lymphocyte antigen 6 family member E	8q24.3
MADCAM1	mucosal vascular addressin cell adhesion molecule 1	19p13.3
MCU	mitochondrial calcium uniporter	10q22.1
MIR23A	microRNA 23a	19p13.12
MIR6087	microRNA 6087	Xq22.3
MX1	MX dynamin like GTPase 1	21q22.3
MX2	MX dynamin like GTPase 2	21q22.3
MYBPH	myosin binding protein H	1q32.1
MYH3	myosin heavy chain 3	17p13.1
MYH4	myosin heavy chain 4	17p13.1
MYH8	myosin heavy chain 8	17p13.1
NDUFB2-AS1	NDUFB2 antisense RNA 1	7q34
NDUFS2	NADH:ubiquinone oxidoreductase core subunit S2	1q23.3
NNMT	nicotinamide N-methyltransferase	11q23.2
NREP	neuronal regeneration related protein	5q22.1
NT5C1A	5'-nucleotidase, cytosolic IA	1p34.2
NT5C3A	5'-nucleotidase, cytosolic IIIA	7p14.3
OAS1	2'-5'-oligoadenylate synthetase 1	12q24.13
OAS2	2'-5'-oligoadenylate synthetase 2	12q24.13
OAS3	2'-5'-oligoadenylate synthetase 3	12q24.13
PBDC1	polysaccharide biosynthesis domain containing 1	Xq13.3
PLEKHO1	pleckstrin homology domain containing O1	1q21.2
POU5F1P4	POU class 5 homeobox 1 pseudogene 4	1q22
PPM1L	protein phosphatase, Mg2+/Mn2+ dependent 1L	3q25.33-q26.1
PROK2	prokineticin 2	3p13
PSMB8	proteasome 20S subunit beta 8	6p21.32
PSMB9	proteasome 20S subunit beta 9	6p21.32
RAB13	RAB13, member RAS oncogene family	1q21.3
RARRES3	phospholipase A and acyltransferase 4	11q12.3
RNF165	ring finger protein 165	18q21.1
RRP7A	ribosomal RNA processing 7 homolog A	22q13.2
RSAD2	radical S-adenosyl methionine domain containing 2	2p25.2
RUNX1	RUNX family transcription factor 1	21q22.12
S100A9	S100 calcium binding protein A9	1q21.3
SAA1	serum amyloid A1	11p15.1
SAA2	serum amyloid A2	11p15.1
SCRT1	scratch family transcriptional repressor 1	8q24.3
SECTM1	secreted and transmembrane 1	17q25.3
SERPINA3	serpin family A member 3	14q32.13
SIGLEC1	sialic acid binding Ig like lectin 1	20p13
SIK1	salt inducible kinase 1	21q22.3
SOCS3	suppressor of cytokine signaling 3	17q25.3
SPIB	Spi-B transcription factor	19q13.33
SPP1	secreted phosphoprotein 1	4q22.1
STAT1	signal transducer and activator of transcription 1	2q32.2
SUSD2	sushi domain containing 2	22q11.23
TNC	tenascin C	9q33.1
TNNT2	troponin T2, cardiac type	1q32.1
TUBA1A	tubulin alpha 1a	12q13.12
TYMP	thymidine phosphorylase	22q13.33
USP5	ubiquitin specific peptidase 5	12p13
ZBP1	Z-DNA binding protein 1	20q13.31
ZFHX2	zinc finger homeobox 2	14q11.2
ZNF106	zinc finger protein 106	15q15.1
ZNFX1	zinc finger NFX1-type containing 1	20q13.13

Supplementary Table 3. A comparison of machine learning models to classify muscle biopsies based on gene expression data using the the differentially expressed genes contained in the training set of each cycle. Accuracy and 95% confidence interval in the 100 test sets of the different machine learning models to classify muscle biopsies into normal muscle tissue (NT), dermatomyositis (DM), antisynthetase syndrome (AS), inclusion body myositis (IBM) or immune-mediated necrotizing myopathy (IMNM).

	NT	DM	AS	IBM	IMNM
Linear SVM	94.5 [91.8-96.9]	92.3 [89.8-93.9]	91.3 [89.8-93.9]	94.9 [93.9-95.9]	92.7 [89.8-95.9]
AdaBoost	92.2 [89.8-93.9]	89.4 [85.7-91.8]	88.1 [85.7-89.8]	92.6 [91.8-93.9]	86.4 [83.7-89.8]
Nearest Neighbors	91.7 [89.8-93.9]	89 [85.7-91.8]	87.9 [87.8-87.8]	91.2 [89.8-91.8]	78.9 [75.5-83.7]
Neural Net	91.7 [87.8-95.9]	86.4 [85.7-91.8]	89.1 [87.8-91.8]	91.6 [89.8-92.9]	78.2 [65.3-91.8]
Gaussian Process	93.3 [91.8-95.9]	84.1 [81.6-87.8]	87.8 [85.7-89.8]	90.9 [89.8-91.8]	79.4 [75.5-83.7]
Random Forest	90.6 [89.8-91.8]	85.2 [81.6-87.8]	86 [83.7-87.8]	91.2 [89.8-91.8]	81.2 [77.6-85.7]
RBF SVM	85.7 [85.7-85.7]	85.5 [83.7-87.8]	87.8 [87.8-87.8]	89.8 [89.8-89.8]	75.1 [71.4-77.6]
Decision Tree	87.5 [85.7-90.8]	85.6 [81.6-89.8]	83.2 [79.6-87.8]	86.1 [83.7-89.8]	77.1 [70.4-84.7]
Gaussian Naive Bayes	86 [85.7-85.7]	81.2 [77.6-85.7]	84.8 [83.7-87.8]	90.3 [89.8-91.8]	69.6 [65.3-73.5]
QDA	84.2 [81.6-87.8]	62.3 [59.2-66.3]	71.7 [67.3-75.5]	80 [75.5-83.7]	59.4 [53.1-65.3]

SVM: support vector machines; RBF: radial basis function; AdaBoost: adaptative boosting; QDA: quadratic discriminant analysis. The models are sorted based on the average accuracy of all the groups.

REFERENCES

1. Mariampillai K, Granger B, Amelin D, et al. Development of a New Classification System for Idiopathic Inflammatory Myopathies Based on Clinical Manifestations and Myositis-Specific Autoantibodies. *JAMA Neurol* 2018;75:1528-37.
2. Love LA, Leff RL, Fraser DD, et al. A new approach to the classification of idiopathic inflammatory myopathy: myositis-specific autoantibodies define useful homogeneous patient groups. *Medicine; analytical reviews of general medicine, neurology, psychiatry, dermatology, and pediatries* 1991;70:360-74.
3. Betteridge Z, McHugh N. Myositis-specific autoantibodies: an important tool to support diagnosis of myositis. *Journal of internal medicine* 2016;280:8-23.
4. Selva-O'Callaghan A, Pinal-Fernandez I, Trallero-Araguas E, et al. Classification and management of adult inflammatory myopathies. *Lancet Neurol* 2018;17:816-28.
5. Greenberg SA, Pinkus JL, Pinkus GS, et al. Interferon-alpha/beta-mediated innate immune mechanisms in dermatomyositis. *Annals of Neurology* 2005;57:664-78.
6. Ivanidze J, Hoffmann R, Lochmuller H, et al. Inclusion body myositis: laser microdissection reveals differential up-regulation of IFN-gamma signaling cascade in attacked versus nonattacked myofibers. *Am J Pathol* 2011;179:1347-59.
7. Allenbach Y, Chaara W, Rosenzwajg M, et al. Th1 response and systemic treg deficiency in inclusion body myositis. *PloS one* 2014;9:e88788.
8. Greenberg SA, Pinkus JL, Pinkus GS, et al. Interferon-alpha/beta-mediated innate immune mechanisms in dermatomyositis. *Ann Neurol* 2005;57:664-78.
9. Hamann PD, Roux BT, Heward JA, et al. Transcriptional profiling identifies differential expression of long non-coding RNAs in Jo-1 associated and inclusion body myositis. *Sci Rep* 2017;7:8024.
10. Raju R, Dalakas MC. Gene expression profile in the muscles of patients with inflammatory myopathies: effect of therapy with IVIg and biological validation of clinically relevant genes. *Brain* 2005;128:1887-96.
11. Greenberg SA, Sanoudou D, Haslett JN, et al. Molecular profiles of inflammatory myopathies. *Neurology* 2002;59:1170-82.
12. Lloyd TE, Mammen AL, Amato AA, et al. Evaluation and construction of diagnostic criteria for inclusion body myositis. *Neurology* 2014;83:426-33.
13. Connors GR, Christopher-Stine L, Oddis CV, et al. Interstitial lung disease associated with the idiopathic inflammatory myopathies: what progress has been made in the past 35 years? *Chest* 2010;138:1464-74.
14. Amici DR, Pinal-Fernandez I, Mazala DA, et al. Calcium dysregulation, functional calpainopathy, and endoplasmic reticulum stress in sporadic inclusion body myositis. *Acta Neuropathol Commun* 2017;5:24.
15. Guyon I, Weston J, Barnhill S, et al. Gene selection for cancer classification using support vector machines. *Machine Learning* 2002;46:389-422.

16. Pinal-Fernandez I, Casal-Dominguez M, Derfoul A, et al. Identification of distinctive interferon gene signatures in different types of myositis. *Neurology* 2019;93:e1193-e204.
17. Gunning P, Ponte P, Kedes L, et al. Chromosomal location of the co-expressed human skeletal and cardiac actin genes. *Proc Natl Acad Sci U S A* 1984;81:1813-7.
18. Fluck M, Mund SI, Schittny JC, et al. Mechano-regulated tenascin-C orchestrates muscle repair. *Proc Natl Acad Sci U S A* 2008;105:13662-7.
19. Zhang D, Zhang DE. Interferon-stimulated gene 15 and the protein ISGylation system. *J Interferon Cytokine Res* 2011;31:119-30.
20. D'Cunha J, Ramanujam S, Wagner RJ, et al. In vitro and in vivo secretion of human ISG15, an IFN-induced immunomodulatory cytokine. *J Immunol* 1996;157:4100-8.
21. Kelly JM, Porter AC, Chernajovsky Y, et al. Characterization of a human gene inducible by alpha- and beta-interferons and its expression in mouse cells. *EMBO J* 1986;5:1601-6.
22. Holzinger D, Jorns C, Stertz S, et al. Induction of MxA gene expression by influenza A virus requires type I or type III interferon signaling. *J Virol* 2007;81:7776-85.
23. Hisamatsu H, Shimbara N, Saito Y, et al. Newly identified pair of proteasomal subunits regulated reciprocally by interferon gamma. *J Exp Med* 1996;183:1807-16.
24. Hall JC, Casciola-Rosen L, Berger AE, et al. Precise probes of type II interferon activity define the origin of interferon signatures in target tissues in rheumatic diseases. *Proc Natl Acad Sci U S A* 2012;109:17609-14.
25. Hall JC, Baer AN, Shah AA, et al. Molecular Subsetting of Interferon Pathways in Sjogren's Syndrome. *Arthritis Rheumatol* 2015;67:2437-46.
26. Watanabe Y, Uruha A, Suzuki S, et al. Clinical features and prognosis in anti-SRP and anti-HMGCR necrotising myopathy. *J Neurol Neurosurg Psychiatry* 2016;87:1038-44.
27. Greenberg SA, Bradshaw EM, Pinkus JL, et al. Plasma cells in muscle in inclusion body myositis and polymyositis. *Neurology* 2005;65:1782-87.
28. Lim JH, Cho SJ, Park SK, et al. Stage-specific expression of two neighboring Crlz1 and IgJ genes during B cell development is regulated by their chromatin accessibility and histone acetylation. *J Immunol* 2006;177:5420-9.
29. Mammen AL. Statin-Associated Autoimmune Myopathy. *N Engl J Med* 2016;374:664-9.
30. Qu J, Ko CW, Tso P, et al. Apolipoprotein A-IV: A Multifunctional Protein Involved in Protection against Atherosclerosis and Diabetes. *Cells* 2019;8.
31. Berlin C, Berg EL, Briskin MJ, et al. Alpha 4 beta 7 integrin mediates lymphocyte binding to the mucosal vascular addressin MAdCAM-1. *Cell* 1993;74:185-95.
32. Pinal-Fernandez I, Casciola-Rosen LA, Christopher-Stine L, et al. The Prevalence of Individual Histopathologic Features Varies according to Autoantibody Status in Muscle Biopsies from Patients with Dermatomyositis. *J Rheumatol* 2015;42:1448-54.
33. Hoogendijk JE, Amato AA, Lecky BR, et al. 119th ENMC international workshop: trial design in adult idiopathic inflammatory myopathies, with the exception of inclusion body myositis, 10-12 October 2003, Naarden, The Netherlands. *Neuromuscular disorders : NMD* 2004;14:337-45.

Competing interests: None

Acknowledgments: The authors thank Dr. Gustavo Gutierrez-Cruz, Dr. Stefania Dell'Orso and Faiza Naz from the NIAMS sequencing facility for all their technical collaboration in making the RNAseq libraries and sequencing them, and the University of Kentucky Center for Muscle Biology for providing normal human muscle samples for the study.

Funding: This research was supported in part by the Intramural Research Program of the National Institute of Arthritis and Musculoskeletal and Skin Diseases and the National Institute of Environmental Health Sciences of the National Institutes of Health. The Myositis Research Database and Dr. LC-S are supported by the Huayi and Siuling Zhang Discovery Fund. IPF´s research was supported by a Fellowship from the Myositis Association. The authors also thank Dr. Peter Buck for support.

Supplementary Methods

1. Cultured human skeletal muscle cells

Normal human skeletal muscle myoblasts (HSMM; Lonza) were cultured according to the manufacturer's protocol. When 80% confluent, the cultures were induced to differentiate into myotubes by replacing the growth medium with differentiation medium (DMEM, 2% horse serum, and L-glutamine). Two plates of cells were collected for RNA extraction at 7 separate time points: immediately before differentiation and then daily for 6 days.

2. Mouse Muscle Injury

Muscle injury and regeneration were induced in mice using cardiotoxin (CTX) as previously described.[1] Briefly, 6 week-old C57BL/6 mice were unilaterally injured by intramuscular injection of 0.1 mL of 10 μM CTX into the tibialis anterior (TA) muscle. Injured TA muscles were harvested at days 3 (n=2), 5 (n=2), 7 (n=2), 10 (n=4), 14 (n=4), and 28 (n=3) post-injury. Contralateral (uninjured) TA muscles were also collected (n=9). Muscle tissue was snap-frozen and stored at -80 degrees Celsius.

3. Human muscle biopsy processing

a) Freezing of muscle tissues

Open muscle biopsies were placed in an aluminum foil envelope. 2-methylbutane (isopentane) was pre-chilled using liquid nitrogen and the aluminum foil envelopes were submerged in the isopentane for 15 seconds. After this, the samples were placed in cryovials at -80º for long-term storage. Samples collected at other institutions were shipped in dry ice to the NIH Muscle Disease Unit.

b) RNA extraction protocol

Required reagents

1. TRIzol (ThermoFisher # 15596026)
2. Chloroform
3. 100% isopropanol
4. 75% ethanol
5. 1.4 ceramic bead homogenizing tubes (VWR # 10032-358)
6. RNase-free water
7. 1.5 mL low-binding tubes
8. GlycoBlue (ThermoFisher #AM9515)

Step 0: Set-up reagents and equipment

1. Spray down centrifuge, then set at 4º.
2. Set heat block at 55ºC.
3. Set Precellys 24 homogenizer at 3x15'' and 6500rpm.
4. Obtain dry ice and ice.
5. Thaw GlycoBlue.

Step 1: Homogenization

1. Get TRIzol from 4º.
2. Get homogenization tubes with beads.
3. Label tubes and set in dry ice.
4. Get samples from -80ºC and place in dry ice.
5. Cut 1-2mm piece of the muscle biopsy with a #11 surgical blade.
6. Fill tubes in ice with 1mL TRIzol and immediately homogenize 3 x 15 minutes.

Step 2: Phase separation

1. Get chloroform, 100% isopropanol, 75% ethanol, and RNase-free water.
2. Wait for 5 minutes at room temperature.
3. Add 0.2 mL of chloroform.
4. Shake 15 seconds by hand.
5. Incubate 2-3 minutes at room temperature.
6. Centrifuge the sample at 12,000g for 15 minutes-4ºC.
7. Label new 1.5mL tubes for all 4 samples.

8. Remove the aqueous phase samples, being careful not to touch interphase.
9. Place aqueous phases into new tubes.
10. Store interphase and organic phases -80ºC.

Step 3: RNA Isolation

1. Add 1.5uL of GlycoBlue, flicker and short spin.
2. Add 0.5 mL of 100% isopropanol.
3. Incubate at room temperature for 10 minutes.
4. Centrifuge the sample at 12,000g for 10 minutes at 4ºC.

Step 4: RNA wash

1. Remove supernatant (leave 0.5 mm).
2. Wash pellet with 1mL of 75% ethanol.
3. Vortex sample briefly.
4. Centrifuge the sample at 7,500g for 5 minutes at 4ºC.
5. Discard the wash (with a pipette).
6. Dry RNA pellet for 5 minutes.

Step 5: RNA resuspension

1. Resuspend RNA pellet in 40µL RNase-free water.
2. Incubate on heat block 10 minutes.
3. Measure the quantity of RNA with NanoDrop.

Step 6: Evaluate RNA quality with 4200 TapeStation

1. Allow High Sensitivity RNA Sample buffer (5067- 5580) to equilibrate at room temperature for 30 minutes.
2. Thaw High Sensitivity RNA ladder (5067- 5581) and total RNA samples on ice.
3. Launch the Agilent 4200 TapeStation Controller Software and select RNA assay mode under “settings”.
4. Flick the High Sensitivity RNA ScreenTape device (5067- 5579) and load it into the 4200 TapeStation instrument.
5. Place loading tips (5067- 5598) into the Agilent 4200 TapeStation instrument.
6. Vortex reagents and spin down before use.
7. Prepare diluted Ladder solution by adding 10 µL RNase free water to the High Sensitivity RNA
8. Ladder vial and mix thoroughly. Pipette 1 µL High Sensitivity RNA Sample Buffer and 2 µL diluted High Sensitivity RNA Ladder at position A1 in a tube strip (401428).
9. For each sample, pipette 1 µL High Sensitivity RNA Sample Buffer and 2 µL RNA sample in a tube strip.
10. Cap tube strips with ladder or sample.
11. Mix liquids in sample and ladder vials using the IKA vortex at 2000 rpm for 1 min.
12. Spin down to position the sample and ladder at the bottom of the well plate and tube strip.

13. Samples and ladder denaturation:
    a. Heat samples and ladder to 72 °C (162 °F) for 3 min.
    b. Place samples and ladder on ice for 2 min.
    c. Spin down to position the samples and ladder at the bottom of the well plate and tube strip.
14. Sample Analysis:
    a. Load samples into the Agilent 4200 TapeStation instrument. Carefully remove caps of tube strips.
    b. Place the ladder in position A1 on tube strip holder in the 4200 TapeStation instrument.
    c. Select the required sample positions on the 4200 TapeStation Controller Software.
    d. Click “Start”.
    e. The Agilent Tapestation Analysis Software opens after the run and displays results.

c) Sample size requirements, RNA yield, and quality

- The frozen muscle biopsy specimens were placed on glass pre-cooled with dry ice and a 1-2mm section of muscle was removed with a #11 surgical blade.

- An average of 11ug (SD 12ug) of RNA was recovered from each muscle biopsy specimen. 65ng of the samples were used to prepare the RNA library using the

NeoPrepTM system according to the TruSeq Stranded mRNA Library Prep protocol (Illumina) and sequenced using the Illumina HiSeq 2500 or 3000.

- The integrity of the RNA was verified using a standard quality metric denominated RNA integrity number (RIN) value, showing a median value of 7 (interquartile range [IQR] 5.9–7.4) for the muscle biopsy samples.

4. Differential gene expression

We performed the differential expression between different subgroups using DESeq2 v.1.20.[2] DESeq2 performs an internal normalization where the geometric mean is calculated for each gene across all samples. The counts for a gene in each sample is then divided by this mean. The median of these ratios in a sample is the size factor for that sample. This procedure corrects for library size and RNA composition bias, which can arise for example when only a small number of genes are very highly expressed in one experiment condition but not in the other.

Additionally, DESeq2 automatically detects count outliers using Cooks' distance and removes these genes from the analysis. DESeq2 v.1.20 also performs independent filtering which maximizes the number of genes which will have a Benjamini and Hochberg-adjusted p–value less than a critical value set by default to 0.1; removing these genes with low counts improves the detection power by making the multiple testing adjustment of the p-values less severe. To speed up the computations we prefiltered genes with a total count across conditions below 10. Since these genes would have been excluded from the analysis afterward anyways, this did not influence the calculations at all.

DESeq2 uses shrinkage estimation for dispersions and fold changes. A dispersion value is estimated for each gene through a model fit procedure. Using these estimations, the package fits a negative binomial generalized linear model for each gene and uses the Wald test for significance testing. The Wald test P values

from the subset of genes that pass the independent filtering step are adjusted for multiple testing using the procedure of Benjamini and Hochberg.[3]

Loading the data, prefiltering and fitting the model was done with the following code:

```
library(DESeq2)
if (packageVersion("DESeq2") != "1.20.0") {
  stop("DESeq2 version is not 1.20.0, please use version 1.20.0")
}

#Set working directory
project_folder <- "/working_directory_path"
setwd(project_folder)

#Import count data and the sample covariates
countData <- as.matrix(read.csv("./anonymized_gene_counts.csv", row.names="gene_id"))
sample_covariates <- t(read.csv("./anonymized_gene_counts.csv", nrows=1, header = FALSE)[-1])
colnames(sample_covariates) <- "GROUP"

#Load data in DESeq2
dds <- DESeqDataSetFromMatrix(countData = countData, colData = sample_covariates, design= ~ GROUP)

#Prefiltering
dds <- dds[ rowSums(counts(dds)) >= 10, ]

#Fitting the model
deseq_fitted <- DESeq(dds, betaPrior=T)
```

To ensure the stability of the central tendency and dispersion values of each biological group between different sections of the study, the normalization process included the totality of the samples even if that specific comparison did not include some of those samples.

For example, the comparison between anti-Jo1 and normal biopsies was performed using the following code:

```
#Calculating the differential expression between anti-Jo1 and normal biopsies
deseq_results <- results(deseq_fitted, contrast=c("GROUP", "Jo1", "NT"))
write.csv(deseq_results[with(deseq_results, order(padj, pvalue)),], file=("./jo1_vs_nt.csv"))
```

We assigned equal weights to each autoantibody subgroups within DM and IMNM to avoid giving more importance to differentially expressed transcriptomic features of autoantibody subgroups with a higher number of biopsies at this stage of the analysis.

For example, to compare DM and IMNM we used the following code:

```
#Calculating the differential expression between DM and IMNM biopsies
resultsNames(deseq_fitted)
#'Intercept' 'GROUPHMGCR' 'GROUPIBM' 'GROUPJo1' 'GROUPMDA5' 'GROUPMi2' 'GROUPNT' 'GROUPNXP2'
'GROUPSRP' 'GROUPTIF1'
deseq_results <- results(deseq_fitted, contrast=c(0,-1/2,0,0,1/4,1/4,0,1/4,-1/2,1/4))
write.csv(deseq_results[with(deseq_results, order(padj, pvalue)),], file=("./dm_vs_imnm.csv"))
```

5. RNAseq-based classification

a) Import data and normalize FPKM

After importing the FPKM levels of all genes, we performed a logarithmic transformation and then normalized the data, maintaining the relative expression of the gene levels.

The rationale behind log-transforming the RNA expression values was to make variation similar across different orders of magnitude.[4]

As per the gene normalization, given that genes with low expression levels are more prone to technical bias in RNAseq,[5] the different genes were normalized so the relative expression of the gene levels would be preserved.

The following code was used to perform this step of the analysis:

```
import pandas as pd
import numpy as np

#Import data
df = pd.read_csv('anonymized_gene_fpkm.csv')

#Generate lists of samples
nt = df['GROUP'] == 'NT'
dm = df['GROUP'].isin(['Mi2', 'NXP2', 'TIF1', 'MDA5'])
asys = df['GROUP'] == 'Jo1'
imnm = df['GROUP'].isin(['HMGCR', 'SRP'])
ibm = df['GROUP'] == 'IBM'
groups = {'nt':nt, 'dm':dm, 'as':asys, 'ibm':ibm, 'imnm':imnm}

df = df.set_index('GROUP')

#Log-transform FPKM values
df = np.log2(df+1)

#Normalize FPKM maintaining relative expression of gene levels
df = (df-np.mean(df.values))/np.std(df.values)
```

b) Filter genes with low signal-to-noise ratio

To filter genes with low signal-to-noise ratio, we selected all the genes that were significantly different (with a cutoff of q-value <0.05) in each group compared to the rest. This was performed using the following lines of code:

```
#Filter genes with q-value>0.05
diff_expression_dfs = {}
sig_genes = []

for group in groups.keys():
       diff_expression = pd.read_csv('./' + group + '_vs_all.csv')
       diff_expression = diff_expression[diff_expression['padj'] < 0.05]
       sig_genes = sig_genes + list(diff_expression['gene'])
sig_genes = list(set(sig_genes))
sig_genes = [gene.replace('-', '_').upper() for gene in sig_genes]
df = df[sig_genes]
df.shape
```

c) Stratified cross-validation

We performed a stratified 3-fold-cross-validation of the samples included in the study. Stratification was used in order to ensure that there were enough samples in each cycle to build the models.

d) Model training

With the objective of showing that the information contained in the RNAseq has classificatory value in myositis, we tested a set of classificatory machine learning models using the default parameters (for example, for AdaBoost we used the default algorithm AdaBoost-SAMME)[6]. The rationale behind performing this screening was that there were no comprehensive studies in the field to guide our selection of the best model for this application.

The cross-validation and training of the models were performed with the following lines of code:

```
import logging
from sklearn.model_selection import StratifiedKFold
from sklearn import metrics
from sklearn.svm import LinearSVC, SVC
from sklearn.ensemble import RandomForestClassifier, AdaBoostClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.naive_bayes import GaussianNB
from sklearn.tree import DecisionTreeClassifier
from sklearn.neural_network import MLPClassifier
from sklearn.gaussian_process import GaussianProcessClassifier
from sklearn.discriminant_analysis import QuadraticDiscriminantAnalysis

logging.basicConfig(level=logging.INFO, filename="models_log.log")

np.random.seed(1)

classifiers = {
    "Linear SVM": LinearSVC(),
    "RBF SVM": SVC(),
    "Random Forest": RandomForestClassifier(),
    "Nearest Neighbors": KNeighborsClassifier(),
    "Gaussian Process": GaussianProcessClassifier(),
    "Decision Tree": DecisionTreeClassifier(),
    "Neural Net": MLPClassifier(),
    "AdaBoost": AdaBoostClassifier(),
    "Gaussian Naive Bayes": GaussianNB(),
    "QDA": QuadraticDiscriminantAnalysis()
    }

for model, clf in zip(classifiers.keys(), classifiers.values()):
    results=[]

    print(model)
    logging.info(model)

    for j in range(1000):
        results_set = []

        for i in groups.values():
            #Diagnostic value
            skf = StratifiedKFold(n_splits=3,shuffle=True)
            skf.get_n_splits(df, i)

            train_index, test_index = next(skf.split(df, nt))

            try:
                clf.fit(df.iloc[train_index], i.iloc[train_index])
                y_pred=clf.predict(df.iloc[test_index])
            except:
                pass

            results_set.append(metrics.accuracy_score(i.iloc[test_index], y_pred))

        results.append(np.array(results_set))
        logging.info(["{0:0.2f}".format(item) for item in results_set])

    mean = np.mean(results, axis=0)
    ci_lower = np.percentile(results, 2.5, axis=0)
    ci_upper = np.percentile(results, 97.5, axis=0)

    for n, i in enumerate(groups.keys()):
        print("   {0}: {2:0.1f} [{1:0.1f}-{3:0.1f}]".format(i, ci_lower[n]*100,
mean[n]*100, ci_upper[n]*100))
```

e) Comparing the effect of including the gene-selection in the internal validation.

We decided to use all the genes that were significantly different (with a cutoff of q-value <0.05) in each group compared to the rest using all the samples. An alternative would have been to include the differentially expressed genes contained in the training set of each cycle. However, this approach was excessively computationally expensive. Nonetheless, to demonstrate the equivalency of these approaches, we modified our pipeline to train 100 cross-validation cycles using only the differentially expressed genes resulting from each training set. The performance of the models was equivalent using both methods (compare Table 2 with Supplementary Table 3).

f) Gene ranking based on the linear support vector machine classification

Once we determined that the linear SVM outperformed the rest of the models, we trained the whole dataset using this algorithm and then applied the recursive feature elimination technique [7] one gene at a time to sort the importance of the genes for each group of subjects.

The following lines of code were used to rank the genes for the different groups of subjects.

```
from sklearn.feature_selection import RFE

for name, dataset in zip(groups.keys(), groups.values()):
    print("Top 10 variables " + name)

    rfe = RFE(LinearSVC(), 1)
    rfe = rfe.fit(df, dataset)

    results = pd.DataFrame(sorted(zip(rfe.ranking_, df.columns)), columns = ['position', 'gene'])

    results.to_excel('../svm_gene_importance_' + name + '.xlsx', index=False)

    print(results.head())
```

References

1. Mammen AL, Casciola-Rosen LA, Hall JC, Christopher-Stine L, Corse AM, Rosen A. Expression of the dermatomyositis autoantigen Mi-2 in regenerating muscle. Arthritis Rheum. 2009 Dec; 60(12):3784-3793.
2. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15(12):550.
3. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. Behav Brain Res. 2001 Nov 1; 125(1-2):279-284.
4. Zwiener I, Frisch B, Binder H. Transforming RNA-Seq data to improve the performance of prognostic gene signatures. PLoS One. 2014; 9(1):e85150.
5. Bhargava V, Head SR, Ordoukhanian P, Mercola M, Subramaniam S. Technical variations in low-input RNA-seq methodologies. Sci Rep. 2014 Jan 14; 4:3678.
6. Zhu J, Zou H, Rosset S, Hastie T. Multi-class AdaBoost.
7. Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. Mach Learn. 2002; 46(1-3):389-422.