ORIGINAL COMMUNICATION



Comprehensive transcriptomic analysis shows disturbed calcium homeostasis and deregulation of T lymphocyte apoptosis in inclusion body myositis

Mridul Johari^{1,2} · Anna Vihola^{1,2,3} · Johanna Palmio⁴ · Manu Jokela^{3,5} · Per Harald Jonson^{1,2} · Jaakko Sarparanta^{1,2} · Sanna Huovinen⁶ · Marco Savarese^{1,2} · Peter Hackman^{1,2} · Bjarne Udd^{1,2,4,7}

Received: 15 October 2021 / Revised: 9 February 2022 / Accepted: 13 February 2022 / Published online: 2 March 2022 © The Author(s) 2022

Abstract

Objective Inclusion body myositis (IBM) has an unclear molecular etiology exhibiting both characteristic inflammatory T-cell activity and rimmed-vacuolar degeneration of muscle fibers. Using in-depth gene expression and splicing studies, we aimed at understanding the different components of the molecular pathomechanisms in IBM.

Methods We performed RNA-seq on RNA extracted from skeletal muscle biopsies of clinically and histopathologically defined IBM (n=24), tibial muscular dystrophy (n=6), and histopathologically normal group (n=9). In a comprehensive transcriptomics analysis, we analyzed the differential gene expression, differential splicing and exon usage, downstream pathway analysis, and the interplay between coding and non-coding RNAs (micro RNAs and long non-coding RNAs).

Results We observe dysregulation of genes involved in calcium homeostasis, particularly affecting the T-cell activity and regulation, causing disturbed Ca^{2+} -induced apoptotic pathways of T cells in IBM muscles. Additionally, LCK/p56, which is an essential gene in regulating the fate of T-cell apoptosis, shows increased expression and altered splicing usage in IBM muscles.

Interpretation Our analysis provides a novel understanding of the molecular mechanisms in IBM by showing a detailed dysregulation of genes involved in calcium homeostasis and its effect on T-cell functioning in IBM muscles. Loss of T-cell regulation is hypothesized to be involved in the consistent observation of no response to immune therapies in IBM patients. Our results show that loss of apoptotic control of cytotoxic T cells could indeed be one component of their abnormal cytolytic activity in IBM muscles.

Keywords Inclusion body myositis · Calcium · Differential expression · Differential splicing · T cells

Mridul Johari mridul.johari@helsinki.fi

- ¹ Folkhälsan Research Center, Helsinki, Finland
- ² Department of Medical and Clinical Genetics, Medicum, University of Helsinki, Helsinki, Finland
- ³ Neuromuscular Research Center, Department of Genetics, Fimlab Laboratories, Tampere, Finland
- ⁴ Neuromuscular Research Center, Department of Neurology, Tampere University and University Hospital, Tampere, Finland
- ⁵ Division of Clinical Neurosciences, Department of Neurology, Turku University Hospital, Turku, Finland
- ⁶ Department of Pathology, Fimlab Laboratories, Tampere University Hospital, Tampere, Finland
- ⁷ Department of Neurology, Vaasa Central Hospital, Vaasa, Finland

Introduction

Inclusion body myositis (IBM) is a late-onset, acquired muscle disease with unclear etiology, and the poorly understood molecular pathogenesis is under debate due to several factors. The CD8 + T-cell infiltration and overexpression of class I MHC antigens in all muscle fibers indicate an autoimmune cascade and are the most consistent finding together with the degeneration of myofibers. However, IBM largely remains refractory to immunosuppressive drugs [1], and comprehensive clinical trials have generally been ineffective [2]. A partial clinical and histopathological overlap with other rimmed-vacuolar (RV) myopathies [3] including accumulations of similar proteins in the RVs [4] support a degenerative pathophysiology. Accumulation/aggregation of these misfolded proteins suggests that IBM could be a protein aggregate disease with the immune-mediated cytotoxic inflammation as a resulting secondary feature [5]. However, there is a significant variance in nature and the number of accumulated proteins observed in the IBM muscle biopsies [6]. Similar aggregates observed in HIV-associated IBM [7] suggest that protein aggregation can still be a downstream effect of immune dysfunction. Additionally, the occurrence of rare familial cases [8] and a strong association with immune MHC locus 8.1 ancestral haplotype [9, 10] support a possible genetic predisposition for IBM.

Analysis of tissue-specific mRNAs and subsequent RNAseq based transcriptomics studies focused on understanding the expression of genes, participating pathways, and networks can increase our understanding of underlying disease pathomechanisms. Prior studies have investigated the differential gene expression in IBM muscles for both the inflammatory and the degenerative pathology [11–17]. However, no study has attempted a comprehensive analysis of RNAseq data combining differential gene expression, differential exon, and splicing usage along with an in-depth analysis of the relation between dysregulation of coding and regulatory RNAs in IBM muscles.

Our study used RNA extracted from muscle biopsies of IBM patients, of non-myositis RV-myopathy disease group, and a histopathologically normal group. We first studied the differential expression of coding, long non-coding RNAs (lncRNAs), and micro RNAs (miRNAs) and then evaluated their possible interplay. Additionally, we studied the transcriptome-wide differential exon and splicing usage. We observed a significant association with genes involved in various calcium-related pathways and identified disturbed calcium regulation specific to T cells in IBM muscles, highlighting the relevance of calcium homeostasis for T-cell activity in IBM muscles. In particular, we identified calcium-induced T lymphocyte apoptosis to be disturbed in IBM muscles.

Materials and methods

Patients and skeletal muscle biopsies

Muscle biopsies (predominantly Tibialis anterior or Vastus lateralis) from 24 Finnish patients diagnosed with clinically and pathologically defined IBM according to the ENMC criteria [18] were included. The age of onset was 60 ± 11 years (median \pm SD), and the age at muscle biopsy was 70 ± 9 years. Additionally, muscle biopsies from six patients with genetically diagnosed Tibial muscular dystrophy (TMD, caused by the heterozygous FINmaj mutation the titin gene) [19] were included. In the TMD cohort, the age of onset was 49 ± 11 years, and age at biopsy 54 ± 14 years. Nine muscle biopsies from individuals that underwent leg amputation for reasons other than a muscle disease [20] were also included. These amputations were due to vascular complications but without ischemic changes in the used muscle tissue. Additionally, these nine biopsies did not show pathologically defined muscle degeneration or inflammation. Age at sampling for amputees was 70 ± 11 years. All muscle biopsies were snap-frozen and stored at -80 °C. Muscle biopsies were collected at the Tampere Neuromuscular Research Center, Tampere University Hospital, Finland.

RNA extraction, selection, and library preparation

Muscle tissue homogenization steps were performed using SpeedMill PLUS (Analytik Jena AG, Germany). RNA was extracted with Qiagen RNeasy Plus Universal Mini Kit (Qiagen, Hilden, Germany) and treated with Invitrogen TURBO DNAse buffer (ThermoFisher Scientific, MA, USA) according to the manufacturers' instructions. RNA was quantified and qualitatively assessed using High Sensitivity RNA ScreenTape (Agilent Technologies, CA, USA) on Agilent 4200 TapeStation system (Agilent Technologies).

Library preparations and sequencing were performed at Oxford Genomics Center, University of Oxford. For PolyA+RNA selection, the NEBNext Ultra II Directional RNA Library Prep kit (E7760) for Illumina (NEB, Beverly, MA, USA) was used to prepare strand-specific RNA-seq libraries. Libraries were multiplexed and sequenced on HiSeq4000: 75 bp paired-end sequencing (Illumina, CA, USA), and an average of ~47 million reads per sample were produced. Samples with enough RNA were used for library preparation for small RNA (<200 nt) selection (18 IBM, nine amputees, and four TMD). NEBNext Small RNA Library Prep Set (E7330) for Illumina was used per the manufacturer's instructions (NEB). Libraries were multiplexed and sequenced on HiSeq2500: 50 bp single-end sequencing (Illumina), and an average of ~ 10 million reads per sample were produced.

RNA-seq data pre-processing, QC, and alignment

Adapter sequences and low-quality bases were removed with fastp [21]. Trimmed sequences were then mapped with STAR 2.7.0d [22] (STAR, RRID: SCR_004463) with index generated from Gencode.v29 human reference (release date 05.2018, based on ENSEMBL GRCh38.p12) and comprehensive gene annotation (primary assembly) using the STAR two-pass method according to the guidelines from the ENCODE project for alignment of long RNA (> 200 nt) and small RNA (< 200 nt) data.

RNA-seq quantification and differential gene expression analysis

Uniquely mapped fragments were summarized and quantified (referred to as counts) by featureCounts [23] (feature-Counts, RRID: SCR_012919) using Gencode.v29 primary comprehensive gene annotation, which lists 58,780 RNAs including 19,969 protein-coding, 16,066 non-coding, and 22,745 other types of RNAs (primary gene expression analysis). Separate quantification of counts for lncRNA (IncRNA analysis) was done using long non-coding RNA gene annotation from Gencode.v29 (a subset of the primary annotation). Quantification of counts for miRNAs (miRNA analysis) in 31 samples was done using miRBase human miRNA annotation (Release 22.1 October 2018) [24]. Differential gene expression analysis was performed with DESeq2 [25] (v1.26.0) (DESeq2, RRID: SCR_015687) in Rstudio (v1.2.5019) (RStudio, RRID: SCR_000432) based on R (v3.6.3) (R Project for Statistical Computing, RRID: SCR_001905). Counts were normalized with variance stabilizing transformation function within DESeq2. A principal component analysis (PCA) was performed on the gene expression data of the IBM samples compared to amputee and TMD groups. Further, pairwise comparisons between cohorts were performed using the Wald test. Log₂ fold changes (LFC) were shrunk using 'ashr' adaptive shrinkage estimation [26], and results were generated with default independent filtering for increasing power. Only genes with LFC values larger than ± 1.5 and a Benjamini–Hochberg adjusted p value of ≤ 0.01 were considered further. Genes specifically dysregulated in IBM muscles were considered for downstream analysis.

Pathway analysis

Ingenuity Pathway Analysis (IPA, QIAGEN Inc.) (Ingenuity Pathway Analysis, RRID: SCR_008653) was used for pathway analysis and enrichment analysis of the obtained differential gene expression data. Using Ingenuity Pathways Knowledge Base (Ingenuity Pathways Knowledge Base, RRID: SCR_008117), IPA mapped and annotated genes to the pathways and predicted activation state based on the direction of changes comparing it with the change in the database.

Differential splicing analysis

To investigate differential usage of exons and splicing, independent of the differential gene expression analysis, we used QoRTS [27] java-based application (v1.3.6) (QoRTs, RRID: SCR_018665) to prepare counts from exons and splice junctions (known and novel) from the aligned data. Downstream analysis of this data was performed using JunctionSeq [28] (v1.16.0) in R. JunctionSeq results produce a q-value (based on FDR) on gene-level analysis, which considers that one or more exon/junction in this gene is differentially used. A conservative q-value threshold of 0.01 was used to select significant observations. IBM-specific differentially expressed genes and differentially spliced genes were compared (Fig. 1). Statistical over-enrichment analysis for Gene ontology terms in categories: Molecular function, biological process, and cellular component, was performed on results obtained from QoRTs/JunctionSeq using clusterProfiler [29] (clusterProfiler, RRID: SCR_016884). Gene sets were compared using UpSet plot [30].

Estimation of immune cell fraction and immunofluorescence images

We used Immunedeconv [31] in R along with CIBER-SORT [32] (CIBERSORT, RRID:SCR_016955) to estimate the fraction of immune cells in our transcriptome dataset. CIBERSORT can characterize the heterogeneity of immune cells in bulk RNA-seq. This validated deconvolution approach can differentiate among 22 immune cell types including seven T-cell types, naïve and memory B cells, plasma cells, natural killer (NK) cells and myeloid subsets.

To show that increased *LCK* expression in IBM muscle compared to controls originates from T cells, we prepared 6 µm frozen sections of four IBM biopsies that were available for further experiments. The slides were immunostained overnight using primary antibodies anti-LCK (RabMab, clone EPR20798-107, Abcam ab227975, RRID: AB_2905531) and anti-CD3 (mouse mAb, clone OKT3, eBioscienceTM/Thermo Fisher Scientific, RRID:AB_467057). Alexa Fluor 488/594 -conjugated secondary antibodies (Thermo Fisher Scientific) were used for detection, and Hoechst was used for nuclear counterstaining. Microscopic images were obtained using the Zeiss Axio Imager M2 system (Carl Zeiss AG), with 40×original magnification.

Results

Expression signature in IBM muscles

Figure 1a shows the summarized workflow of the methodology. The PCA shown in Fig. 1b explains the differences between the three cohorts. Pairwise comparisons were performed to reduce the potential confounding effects of groups, which identified 2288 and 302 genes specifically upor downregulated in the IBM cohort, respectively (Fig. 1c).

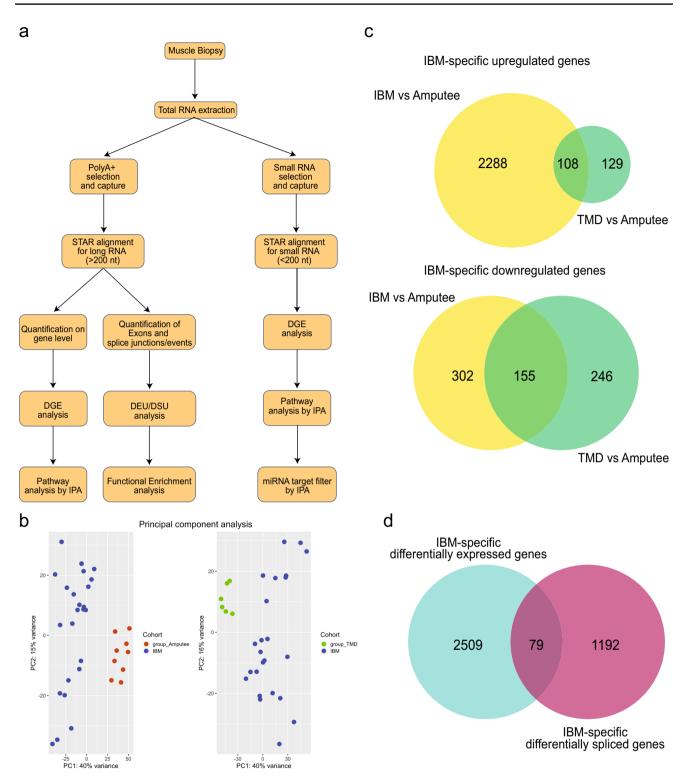


Fig. 1 a Workflow and methodology used in this study. **b** Principal component analysis of gene expression results showing the pairwise comparison between different groups: IBM, TMD and Amputees. **c** IBM-specific differentially expressed genes were determined by com-

paring IBM cases with amputee and TMD groups. **d** Comparison between IBM-specific differentially expressed genes (cyan) and IBM-specific differentially spliced genes (magenta)

Non-coding RNA analyses resulted in 497 lncRNAs upregulated, 106 lncRNAs downregulated, 140 miRNAs upregulated, and 126 miRNAs explicitly downregulated in the IBM cohort compared to other groups. These IBM-specific dysregulated RNAs were used for downstream pathway analysis using IPA workflow. The top 15 genes dysregulated specifically in IBM muscles, with their functional annotations and normalized expression in the different cohorts, are shown in Fig. 2.

Pathway analysis

We performed IPA workflow analysis on IBM-specific dysregulated genes to better understand the pathways and the upstream regulators associated with the observed expression dysregulation. Out of these, 2588 genes, 596 lncRNAs, and 257 miRNAs mapped to the Ingenuity database. From the primary gene expression analysis, IPA identified 91 pathways as significantly altered. Table 1 shows a summary of the IPA results with the top identified pathways.

The top upstream regulators in both miRNA and lncRNA analysis are shown in Tables 2 and 3, respectively. We identified an increased expression of the lncRNA *DNM3OS* (DNM3 antisense RNA) and *MIAT* (Myocardial infarction associated transcript) from these analyses. IPA suggested this dysregulation may be due to *JDP2* (Jun Dimerization Protein 2) and *TARDBP* (TAR DNA Binding Protein), acting as an upstream regulator of *DNM3OS* and *MIAT*, respectively (Table 4).

Dysregulation of calcium-related pathways in IBM muscles

IPA identified calcium-induced T lymphocyte apoptosis as one of the most significant pathways dysregulated in IBM muscles (Table 1). Our IBM-specific dataset contained 69 genes with significant dysregulation out of the 232 genes annotated in this pathway. A part of this pathway, including the major players, is shown in Fig. 3. Another pathway outside the top results identified that 29 genes (29/208, p=7.05E-03) significantly dysregulated in our dataset are also involved in calcium signaling. These results prompted us to investigate further for calcium-related issues in cellular signaling, and we found that IPA also detects dysregulation of the following processes, mobilization of Ca²⁺ (80 genes), the release of Ca²⁺ (51 genes), as significantly disturbed in IBM muscles (Table 4).

Altered exon usage and splicing pattern in IBM muscles

To explore IBM-specific exon usage, we performed an independent transcriptome-wide differential splicing analysis in our three cohorts. We obtained a list of 1,271 differentially spliced genes in IBM from our differential splicing analysis. These transcripts either showed IBM-specific increased usage of a known junction or a known exon or contained a novel exon-exon junction resulting in an alternative isoform. To understand the diverse portfolio of mature mRNAs created from pre-mRNAs, we used gene ontology over-enrichment analysis on these 1271 differentially spliced genes and identified the first splicing signature specific to IBM muscles. To understand the different classes over-represented in these genes, we performed statistical over-enrichment analysis using clusterProfiler for all three GO categories as seen in Fig. 4a-c. Our analysis showed an enrichment of genes involved in the structure and organization of actin filament assembly in IBM muscles and, interestingly, proteins involved in mRNA processing and metabolism.

We then compared the list of differentially spliced genes with differentially expressed genes in our analysis and found an overlap of 79 genes (Fig. 1d). Next, we wanted to observe the overlap between six different sets of genes, namely IBM-specific differentially spliced genes, calcium-induced T Lymphocyte apoptosis, Mobilization of Ca2+, Flux of Ca2+, Quantity of Ca2+, and Release of Ca2+ (Fig. 4d). We observed 10 genes to be associated with calcium-related processes; HLA-DPA1, HLA-DPB1, and HLA-DQB1 are associated with calcium-induced T Lymphocyte apoptosis, ANXA1 is associated with mobilization, flux, and release of Ca²⁺, CCL4 is associated with mobilization, flux, and quantity of Ca²⁺, GRK3 and RARRES2 are associated with mobilization, SH3KBP1 with flux, and ITGAM with the quantity of Ca²⁺. In particular, one specific differentially spliced gene, LCK, is part of all six sets.

Figure 5a shows the gene expression of *LCK* in three cohorts, with expression in IBM muscles being significantly higher than the others ($log_2FC = +2.86$, padj = 3.50E-11, ranking = 355/2590). Additionally, Fig. 5b shows the differential splicing pattern observed in *LCK* in all three groups. The highlighted E016 corresponds to an alternative exon (chr1:32274818-32274992, GRCh38).

Immune cell profiling and immunofluorescence

Using Immunedeconv and CIBERSORT, IBM muscle transcriptome showed a strong signature of CD4 + /CD8 + T cells as compared to the Amputee and TMD cohort (Supplementary Fig. 1). Immuno-stainings using anti-LCK and anti-CD3 antibodies showed strong expression of LCK in

а	Symbol	Mean count	log2FC	pvalue	padj	GO terms
	KDM7A	1147.06	-1.69	4.07E-32	2.39E-28	Iron ion binding and oxidoreductase activity
	LGALS3BP*	2922.28	2.59	5.91E-31	2.89E-27	Scavenger receptor activity
	TLR3	128.42	2.29	3.98E-28	1.30E-24	Regulation of dendritic cell cytokine production
	ATP6V0A1 3393.79 -1.76 3.01E-27 8.83E-24 ATPase bindin		8.83E-24	ATPase binding and proton-transporting ATPase activity		
	IRF8*	199.90	3.89	4.66E-27	1.24E-23	DNA-binding transcription factor activity, RNA polymerase II-specific
	SLC7A7	216.91	3.16	2.08E-26	5.09E-23	Amino acid transmembrane transporter activity
	FGL2	1102.10	3.27 3.14E-26 6.58E-23 T cell activation via T cell receptor contact with antigen bound to MHC molecule on antigen presenting cell			
	B2M*			Positive regulation of T cell mediated cytotoxicity		
	MYL4*	610.22	5.44			Actin monomer binding, Calcium ion binding
	UBE2L6*	986.03	2.47	1.86E-25	3.04E-22	Ubiquitin-protein transferase activity
	SPPL2B	1161.90	-1.66	2.99E-25	4.39E-22	Protein homodimerization activity
	TUBA1A*	2951.15	2.50	4.36E-25	5.81E-22	GTPase activity, structural molecule activity
	TFAP4	234.41	-1.52	4.32E-25	5.81E-22	Transcription regulatory region sequence-specific DNA binding
	HLA-DRA	8658.66	3.25	7.29E-25	9.00E-22	Antigen processing and presentation of endogenous peptide antigen via MHC class II
	SOX11	96.02	4.03	1.97E-24	2.23E-21	DNA-binding transcription factor activity, RNA polymerase II-specific
b						
D	KDN	174		LGAL	S3BP	TLR3 ATP6V0A1 IRF8
300 100 50	•- •		3000 - 1000 - 300 -			
100	SLC7	47	1	FGL	.2	B2M MYL4 UBE2L6
Too Normalised counts 10 10 Normalised counts 10 10 10 10 10 10 10 10 10 10 10 10 10	0 - 0 -		3000 - 1000 - 300 - 100 -	Å	OF ROLLING	1e+05 - 3e+04 - 1e+04 - 100
500 300 100 50	o - Ç	28	10000 - 3000 - 1000 -	TUBA	1A	TFAP4 HLA-DRA SOX11 500 3000 1000 100 300 3000 1000 100 100 3000 100 100 100 1000 1000 100

Cohort 🛱 group_Amputee 🖨 group_TMD 🖨 IBM

the inflammatory T-cell infiltrates of IBM muscle biopsies (Supplementary Fig. 2).

Discussion

In this study, we aimed to identify a more detailed IBMspecific molecular signature, using different RNA-seq based methods that can help us explore the inflammatory

◄Fig. 2 a Top 15 differentially expressed genes specific to IBM muscles. Log₂ fold change (log2FC) of IBM versus amputees calculated by DEseq2 after shrinkage estimations. '+'/'-' sign denotes the direction of change, i.e., positive log2FC values indicate overexpressed genes in IBM muscles, and negative log2FC values indicate underexpressed genes in IBM muscles. The p value of significance and adjusted *p* value using the Benjamini–Hochberg corrections and associated GO terms are shown for each gene. Genes marked with * are also observed as significantly dysregulated in Hamann et al. [13]. b Normalized gene expression in the different cohorts is presented as boxplots. Median and quartile values are shown, with whiskers reaching up to 1.5 times the interquartile range. Individual expression levels are shown as jitter points. The raincloud plots illustrate the distribution of data in each cohort. The scaled *Y*-axis shows log normalized counts

and degenerative parts in depth. Antigen-driven T-cell cytotoxicity is the most reproducible and plausible part of the complex molecular pathomechanism in IBM. However, it remains unknown what antigen drives this IBM-specific immune cascade.

As part of the RV pathology, accumulated proteins or the unfolded protein response have been hypothesized to prompt an immune reaction [5]. A recent unbiased proteomics study dissected these RVs in IBM [33]. Interestingly, the protein encoded by one of our top differentially expressed genes, MYL4, is also detected in the RVs in IBM along with ANXA1, which is both differentially expressed and differentially spliced in IBM muscles. In our study design, we considered TMD, another RV muscle disease but without immune involvement, to understand if there are any RVspecific antigens in IBM muscles. Additionally, using agematched histopathologically normal muscles from amputees, we aimed to understand if general inflammatory signatures can be replicated and studied in more detail using additional methods such as non-coding RNAs and differential splicing studies. Consequently, our strong study design and robust methodology helped us replicate findings from previous studies [11–17] and identify essentially new calcium-related issues in IBM muscles and their link with the altered T-cell cytotoxicity in IBM muscle fibers.

We found that several genes contributing to calcium homeostasis are differentially expressed in IBM muscles resulting in dysregulation of several critical pathways, specifically, calcium-induced T lymphocyte apoptosis and related Nur77 signaling. Ca^{2+} is a universal second messenger in T cells, and it is known to regulate proliferation and differentiation of T cells and T-cell effector functions [34]. The complexity and duration of Ca^{2+} signals and resultant cytoskeletal rearrangements determine the fate of T cells in response to an antigen [35]. On one hand, a short-term increase in intracellular Ca^{2+} concentration results in the cytolytic activity of T cells; on the other hand, prolonged elevation results in proliferation, differentiation, and maturation of näive T cells into Th1, Th2, and Th17 subtypes and the production of cytokines [34].

Ca²⁺ signaling is known to optimize the interaction between T cells and antigen-presenting cells [35]. The binding of antigen/MHC complexes (CD8⁺-MHC class I/CD4⁺-MHC class II) to T-cell receptors (TCR) activates Src-family protein tyrosine kinases, e.g., LCK and FYN at the cytoplasmic side of the TCR/CD3 complex.

The genes in the NR4A family (NR4A1/Nur77, NR4A2/ Nurr1, NR4A3/Nor1) act as critical molecular switches in cell survival and inflammation. Human NR4A1 encodes for a homolog of a mouse protein called Nur77, a zinc transcription factor expressed as an early gene in T cells upon antigen-TCR interaction. In addition to being a transcriptional activator, Nur77 has an apoptotic role in T regulatory fate [36] and other non-genomic proapoptotic functions via mitochondrial interactions with Bcl-2 [37]. T cells deficient in Nur77 have been shown to have high proliferation, enhanced T-cell activation, and increased susceptibility for T-cell-mediated inflammatory diseases [38]. The expression of Nur77 is Ca^{2+} dependent and is controlled by the myocyte enhancer factor 2 (MEF2) transcription factor [39], whose DNA-binding and transcriptional activity is enhanced by Calcineurin. Another calcium-dependent transcription factor, CABIN1, acts as a transcriptional repressor of MEF2, thus keeping the Nur77 promoter silent in the absence of a TCR signal [40]. The interaction between CABIN1 and calcineurin is influenced by intracellular Ca²⁺ and PKC activation, resulting in hyperphosphorylation of CABIN1 and its subsequent transcription repressing activity. An increase in intracellular Ca²⁺ concentration activates the interaction of the calmodulin family of genes (CALM) with CABIN1, triggering the dissociation of MEF2 from Cabin and MEF2 to become transcriptionally active [41]. In the nucleus, NFAT interacts with MEF2 and enhances its transcriptional activity by recruiting the co-activator p300 for the transcription of Nur77.

In our dataset, 69 genes mapping to the calciuminduced T Lymphocyte apoptosis and 72 genes mapping the Nur77 signaling in T Lymphocytes are differentially expressed in IBM muscles. As seen in Fig. 3, several essential genes like ZAP70, LCK, different subunits of Protein Kinase C, and ATP2A1 which encodes for SERCA, are significantly changed in IBM muscles. Additionally, we also observed genes associated with the mobilization of Ca^{2+} , the release of Ca^{2+} , the quantity of Ca^{2+} , and the flux of Ca^{2+} as significantly dysregulated in IBM muscles, indicating a possible widespread disturbance with the handling of calcium entry and release in cells. In T cells, especially, this disturbance could dramatically impact their activation, differentiation, and most likely, the regulation of T-cell apoptosis will be disturbed. Table 1Top 10 dysregulatedcanonical pathways identifiedby IPA. The significance of theidentified pathway is shownwith a p value and the numberof differentially expressed genesobserved in the IBM-specificdataset compared to the numberof genes present in the databasefor each pathway

Ingenuity canonical pathways	<i>p</i> value	Number of genes in dataset/ number of genes in database
Dendritic cell maturation	5.72E-31	106/357
T-cell receptor signaling	2.30E-25	97/355
T-cell exhaustion signaling pathway	3.12E-25	94/338
Cdc42 signaling	7.63E-24	88/315
iCOS-iCOSL signaling in T helper cells	4.10E-23	81/280
CD28 signaling in T helper cells	1.13E-22	82/290
OX40 signaling pathway	4.03E-21	68/222
Calcium-induced T lymphocyte apoptosis	1.29E-20	69/232
Nur77 signaling in T lymphocytes	3.07E-20	72/253
Role of NFAT in regulation of the immune response	4.57E-19	87/360

Table 2From the miRNAanalysis, upstream bindingpartners are shown along withtheir target miRNA. A p valueand associated GO terms areshown

Upstream regulator	p value	GO terms and annotations
AGO2	7.89E-23	RNA polymerase II complex binding
SSB	4.74E-19	RNA binding
<i>TP53</i>	7.59E-09	Transcription regulatory region sequence-specific DNA binding
RNA polymerase III	4.09E-06	Synthesis of small RNA, RNA polymerase activity

Table 3 From the long non-coding RNA analysis, upstream binding partners are shown along with their target lncRNA. A p value and associated GO terms are shown

Upstream regulator	Target molecule in dataset	p value	GO terms and annotations
JDP2	DNM3OS	4.10E-03	DNA-binding transcription factor activity, RNA polymerase II-specific
miR-338-3p	NR2F1-AS1	4.10E-03	Negative regulation of gene expression; negative regulation of IL-6 pro- duction; negative regulation of cytokine production involved in inflam- matory response
miR-150-5p	MIAT	4.10E-03	mRNA binding involved in posttranscriptional gene silencing
TARDBP	MIAT	2.03E-02	RNA polymerase II cis-regulatory region sequence-specific DNA binding
mir-150	MIAT	2.63E-02	mRNA binding involved in posttranscriptional gene silencing
PGF	DNM3OS	3.43E-02	Protein binding, signal transduction
FUS	RMRP	3.43E-02	mRNA binding, mRNA stabilization
DDX58	EGOT	4.61E-02	Double-stranded RNA binding

Table 4 In cell signaling processes, different pathways associated with calcium homeostasis are shown along with their p value and a prediction state

Functional annotations	p value	Predicted activation state
Mobilization of Ca ²⁺	5.04E-26	Increased
Flux of Ca ²⁺	4.68E-13	Increased
Quantity of Ca2+	4.34E-09	Increased
Release of Ca ²⁺	5.00E-09	Increased

Apoptosis in T cells is necessary to resolve their inflammatory activity, and defective or delayed apoptosis may contribute to the pathogenesis of inflammatory diseases [42]. In this scenario, loss of apoptotic control could be one mechanism explaining the lack of immune-suppressive therapeutic effect in IBM [43].

The diversity of the skeletal muscle proteome is, among others, dependent on the diversity of exon usage in premRNAs [44]. From our transcriptome-wide splicing analysis within the differentially expressed genes, we identified 79

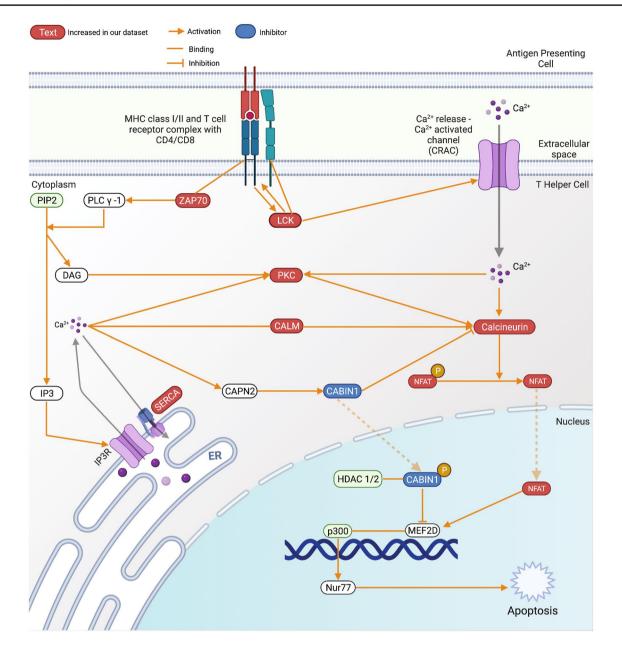


Fig. 3 The calcium-induced T lymphocyte apoptosis pathway with gene expression changes observed in IBM compared to groups. Created with BioRender.com

genes, out of which ten are associated with different calcium-related functions. Amongst these, LCK is a T lymphocyte-specific protein tyrosine kinase involved in downstream events of antigen–TCR interaction. LCK/p56 is essential in transducing signals leading to apoptotic cell death in mature T cells [45], and its activity is tightly regulated to protect against hyperactivation of T cells and autoimmunity, thus maintaining T-cell homeostasis [46]. Moreover, LCK also selectively influences the flux and release of calcium in cells [47]. In our analysis, LCK is both differentially expressed and differentially spliced in IBM muscles. Additionally, further evidence is provided by immunoanalysis of IBM muscles showing strong expression of the LCK protein in the T-cell infiltrates. Disturbed T-cell apoptosis and the dysregulation of LCK in IBM muscles provide novel insights into the molecular mechanisms of IBM. Considering the crucial regulatory activity of LCK, it might be a potential therapeutic target for IBM patients.

We also observe dysregulation of several non-coding RNAs in our study. Previously, Hamann and colleagues have discussed lncRNAs in the context of IBM [13]. The benefits of our study design, especially the homogenous molecular pathology and the larger sample size, let us dig deeper into the dysregulation of lncRNAs specific to IBM muscles. We

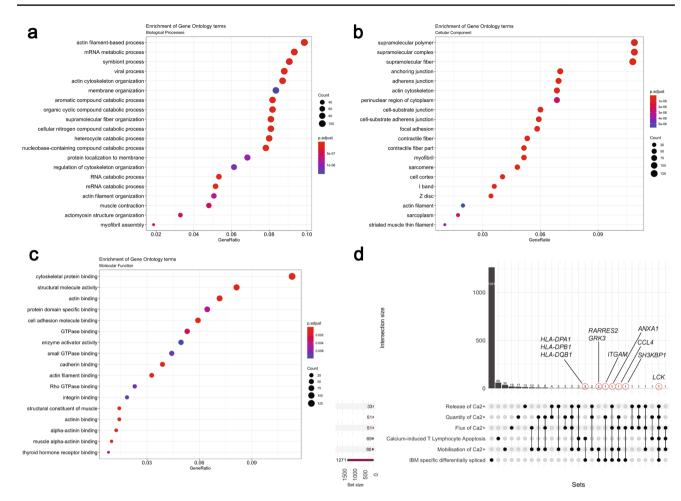


Fig. 4 Statistical over-representation tests were performed on a list of differentially spliced RNAs, using clusterProfiler for **a** Biological Processes, **b** Cellular component, and **c** Molecular function. **d** An UpSet plot is shown comparing six different sets, namely, IBM-specific differentially spliced (1271 genes), mobilization of Ca²⁺ (80 genes), calcium-induced T lymphocyte apoptosis (69 genes), the flux of Ca²⁺ (51 genes), quantity of Ca²⁺ (51 genes), and release of Ca²⁺

(33 genes). Dots and lines represent subsets of different lists. The horizontal bar graph (wine color) represents the size of each set, while the vertical histogram (black) represents the number of RNAs in each subset. The 10 RNAs that are both differentially expressed and differentially spliced are shown with a red circle with their gene names (black)

identified that JDP2 (DNA-binding transcription factor) and TARDBP/TDP-43 (DNA and RNA binding protein) might have altered regulator activity since their downstream noncoding partners (DNM3OS and MIAT, respectively) are significantly overexpressed IBM muscles. Additionally, both these proteins are specific to RNA polymerase II (RNA Pol II), facilitating transcription and pre-mRNA maturation. Alteration in RNA or DNA-binding proteins (expression or localization) associated with the activity of the spliceosome machinery can directly affect the downstream events. Since TDP-43 is accumulated in RVs, one possibility is that the unavailability of TDP-43 can affect its transcription and splicing activities. The normal expression of TARDBP we observe in IBM patients is expected and is in coherence with the previous reports [48]. In inherited muscle diseases, damaging variants in the disease-associated gene can result in mislocalization and accumulation of mutant protein in the muscle fibers. Previous studies have reported rare exonic variants in genes, including VCP and SQSTM1 in IBM [49, 50]. However, in our cohort of IBM patients, there were no rare exonic TARDBP, VCP, or SQSTM1 variants [9] that could suggest a possible association with abnormal protein turnover and accumulation/aggregation. Therefore, further evidence to suggest the pathogenic role of variants in such genes and their downstream effect on pathological protein accumulation in IBM is still missing. However, the potential downside of TDP-43 not being available for its traditional roles, such as effective splicing because of the aggregation, is noteworthy. Further evidence of possible dysregulation of splicing in IBM muscles comes from our differential splicing results where proteins involved in mRNA processing, transcription, and regulation are enriched, suggesting that

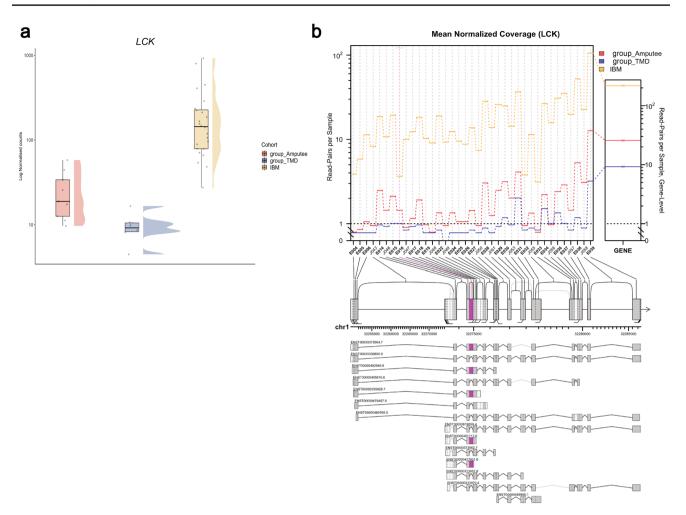


Fig. 5 a Normalized LCK expression in the different cohorts (as explained in Fig. 2b). **b** Altered isoform expression of *LCK* using JunctionSeq showing estimated normalized mean read-pair count for each exon and splice junctions in the different cohorts (left) as well as for the whole *LCK* gene (right). The significantly alternatively spliced

additional studies are needed to understand the possible impact of dysregulated mRNA processing in IBM muscles.

Previously, Pinal-Fernandez and colleagues observed that calcium-induced T lymphocyte apoptosis was a significant IBM-specific dysregulated pathway in their extensive analysis of different inflammatory myopathies but did not comment further on the possible importance [17]. Additionally, using a smaller sample size, Amici and colleagues identified calcium signaling as one of the significant disturbed pathways in IBM muscles and hypothesized its potential major role [14]. Furthermore, previous gene expression studies in IBM have analyzed data primarily from microarrays [11, 12, 15]. Only recently paired-end reads RNA-seq have been used in IBM studies [13, 14, 16, 17]. While microarray-based analyses are comparable for differential expression studies, RNA-seq based methodologies are superior for in-depth transcriptome analysis. In our study design, we used matched muscle biopsies and

feature, E016 (pink), corresponds to chr1:32274818–32274992 (GRCh38). The alternative *LCK* transcripts used in the JunctionSeq analysis are shown below with their corresponding ENSEMBL identifiers

state-of-the-art RNA-seq analysis tools. Our analyses show novel molecular events in IBM muscles which increase our understanding of IBM and provide valuable additions to improve the therapeutic interventions considering the disturbed calcium homeostasis, dysregulation of LCK, and associated deregulation of apoptotic control of T cells in IBM muscles.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00415-022-11029-7.

Acknowledgements The authors would like to thank the patients and their families. We thank the Oxford Genomics Centre at the Wellcome Centre for Human Genetics (funded by Wellcome Trust grant reference 203141/Z/16/Z) for the generation and initial processing of the sequencing data. We would like to thank CSC – IT Center for Science, Finland, for its computational resources. We thank Helena Luque, Merja Soininen, Talha Qureshi and Eini Penkkimäki for their technical assistance. We would like to thank Peter-Bram 't Hoen for his critical review of this manuscript. Author contributions MJ: conceptualization of the study, funding acquisition, data analysis and curation, methodology, project administration, visualization, writing the original draft, review, and editing of the manuscript. AV: data acquisition and analysis, methodology, writing the original draft. JP: patient samples and data collection, writing the original draft. MJ: patient samples and data collection, review, and editing of the manuscript. PHJ.: conceptualization of the methodology, review, and editing of the manuscript. SH: patient samples and data collection, review, and editing of the manuscript. MS: conceptualization of the study and methodology, review and editing of the manuscript. PH: funding acquisition, project administration, and review and editing of the manuscript. BU: conceptualization of the study funding acquisition, supervision, project administration, review, and editing of the manuscript.

Funding Open Access funding provided by the University of Helsinki including Helsinki University Central Hospital. This work was supported by the Folkhälsan Research Foundation, Doctoral program in Integrative Life Science (ILS) and Doctoral school in Health Sciences (DSHealth), University of Helsinki (MJ), the Päivikki and Sakari Sohlberg Foundation (MJ), the Biomedicum Helsinki Foundation (MJ), Finska läkaresällskapet (BU/MJ), the Finnish Medical Foundation (JP), the Paulo foundation (MS), the Jane and Aatos Erkko Foundation (PH) and the Sigrid Jusélius Foundation (BU).

Data availability Raw counts and normalized DESeq2 counts from polyA + RNAs and miRNAs are available in GEO as superseries GSE151758.

Declarations

Conflicts of interest The authors report no conflicts of interest.

Ethical approval The study was performed in line with the principles of the Declaration of Helsinki. Ethical approval for this study falls under HUS:195/13/03/00/11. Informed consent from the patients was obtained at the time of sample collection.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

- Ikenaga C, Kubota A, Kadoya M et al (2017) Clinicopathologic features of myositis patients with CD8-MHC-1 complex pathology. Neurology 89:1060–1068. https://doi.org/10.1212/WNL. 000000000004333
- Dalakas MC, Koffman B, Fujii M, Spector S, Sivakumar K, Cupler E (2001) A controlled study of intravenous immunoglobulin combined with prednisone in the treatment of IBM. Neurology 56:323–327. https://doi.org/10.1212/wnl.56.3.323

- Cai H, Yabe I, Sato K et al (2012) Clinical, pathological, and genetic mutation analysis of sporadic inclusion body myositis in Japanese people. J Neurol 259:1913–1922. https://doi.org/10. 1007/s00415-012-6439-0
- Askanas V, Engel WK (2005) Sporadic inclusion-body myositis: a proposed key pathogenetic role of the abnormalities of the ubiquitin-proteasome system, and protein misfolding and aggregation. Acta Myol 24:17–24
- Askanas V, Engel WK (1998) Sporadic inclusion-body myositis and hereditary inclusion-body myopathies: current concepts of diagnosis and pathogenesis. Curr Opin Rheumatol 10:530–542. https://doi.org/10.1097/00002281-199811000-00005
- Greenberg SA (2010) Theories of the pathogenesis of inclusion body myositis. Curr Rheumatol Rep 12:221–228. https://doi.org/ 10.1007/s11926-010-0102-5
- Hiniker A, Daniels BH, Margeta M (2016) T-cell-mediated inflammatory myopathies in HIV-positive individuals: a histologic study of 19 cases. J Neuropathol Exp Neurol 75:239–245. https:// doi.org/10.1093/jnen/nlv023
- Amato AA, Shebert RT (1998) Inclusion body myositis in twins. Neurology 51:598–600. https://doi.org/10.1212/wnl.51.2.598
- Johari M, Arumilli M, Palmio J et al (2017) Association study reveals novel risk loci for sporadic inclusion body myositis. Eur J Neurol 24:572–577. https://doi.org/10.1111/ene.13244
- Rothwell S, Chinoy H, Lamb JA et al (2019) Focused HLA analysis in Caucasians with myositis identifies significant associations with autoantibody subgroups. Ann Rheum Dis 78:996–1002. https://doi.org/10.1136/annrheumdis-2019-215046
- Greenberg SA, Sanoudou D, Haslett JN et al (2002) Molecular profiles of inflammatory myopathies. Neurology 59:1170–1182. https://doi.org/10.1212/wnl.59.8.1170
- Eisenberg I, Eran A, Nishino I et al (2007) Distinctive patterns of microRNA expression in primary muscular disorders. Proc Natl Acad Sci USA 104:17016–17021. https://doi.org/10.1073/pnas. 0708115104
- Hamann PD, Roux BT, Heward JA et al (2017) Transcriptional profiling identifies differential expression of long non-coding RNAs in Jo-1 associated and inclusion body myositis. Sci Rep 7:8024. https://doi.org/10.1038/s41598-017-08603-9
- Amici DR, Pinal-Fernandez I, Mazala DA et al (2017) Calcium dysregulation, functional calpainopathy, and endoplasmic reticulum stress in sporadic inclusion body myositis. Acta Neuropathol Commun 5:24. https://doi.org/10.1186/s40478-017-0427-7
- Greenberg SA, Pinkus JL, Kong SW, Baecher-Allan C, Amato AA, Dorfman DM (2019) Highly differentiated cytotoxic T cells in inclusion body myositis. Brain 142:2590–2604. https://doi.org/ 10.1093/brain/awz207
- Pinal-Fernandez I, Casal-Dominguez M, Derfoul A et al (2019) Identification of distinctive interferon gene signatures in different types of myositis. Neurology 93:e1193–e1204. https://doi.org/10. 1212/WNL.00000000008128
- Pinal-Fernandez I, Casal-Dominguez M, Derfoul A et al (2020) Machine learning algorithms reveal unique gene expression profiles in muscle biopsies from patients with different types of myositis. Ann Rheum Dis 79:1234–1242. https://doi.org/10.1136/ annrheumdis-2019-216599
- Rose MR, Group EIW (2013) 188th ENMC International Workshop: Inclusion Body Myositis, 2-4 December 2011, Naarden, The Netherlands. Neuromuscul Disord 23:1044–1055. https://doi.org/ 10.1016/j.nmd.2013.08.007
- Hackman P, Vihola A, Haravuori H et al (2002) Tibial muscular dystrophy is a titinopathy caused by mutations in TTN, the gene encoding the giant skeletal-muscle protein titin. Am J Hum Genet 71:492–500. https://doi.org/10.1086/342380
- Huovinen S, Penttila S, Somervuo P et al (2015) Differential isoform expression and selective muscle involvement in muscular

dystrophies. Am J Pathol 185:2833–2842. https://doi.org/10. 1016/j.ajpath.2015.06.018

- Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi. org/10.1093/bioinformatics/bty560
- Dobin A, Davis CA, Schlesinger F et al (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21. https://doi. org/10.1093/bioinformatics/bts635
- Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923–930. https://doi.org/10.1093/ bioinformatics/btt656
- Griffiths-Jones S (2004) The microRNA Registry. Nucleic Acids Res 32:D109-111. https://doi.org/10.1093/nar/gkh023
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. https://doi.org/10.1186/s13059-014-0550-8
- Stephens M (2016) False discovery rates: a new deal. Biostatistics 18:275–294. https://doi.org/10.1093/biostatistics/kxw041
- Hartley SW, Mullikin JC (2015) QoRTs: a comprehensive toolset for quality control and data processing of RNA-Seq experiments. BMC Bioinformatics 16:224. https://doi.org/10.1186/ s12859-015-0670-5
- Hartley SW, Mullikin JC (2016) Detection and visualization of differential splicing in RNA-Seq data with JunctionSeq. Nucleic Acids Res 44:e127. https://doi.org/10.1093/nar/gkw501
- Yu G, Wang LG, Han Y, He QY (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16:284–287. https://doi.org/10.1089/omi.2011.0118
- Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H (2014) UpSet: visualization of intersecting sets. IEEE Trans Vis Comput Graph 20:1983–1992. https://doi.org/10.1109/TVCG.2014.23462 48
- Sturm G, Finotello F, Petitprez F et al (2019) Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. Bioinformatics 35:i436–i445. https:// doi.org/10.1093/bioinformatics/btz363
- Newman AM, Liu CL, Green MR et al (2015) Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 12:453–457. https://doi.org/10.1038/nmeth.3337
- Güttsches A-K, Brady S, Krause K et al (2017) Proteomics of rimmed vacuoles define new risk allele in inclusion body myositis. Ann Neurol 81:227–239. https://doi.org/10.1002/ana.24847
- Joseph N, Reicher B, Barda-Saad M (2014) The calcium feedback loop and T cell activation: how cytoskeleton networks control intracellular calcium flux. Biochim Biophys Acta 1838:557–568. https://doi.org/10.1016/j.bbamem.2013.07.009
- Lewis RS (2001) Calcium signaling mechanisms in T lymphocytes. Annu Rev Immunol 19:497–521. https://doi.org/10.1146/ annurev.immunol.19.1.497
- Calnan BJ, Szychowski S, Chan FK, Cado D, Winoto A (1995) A role for the orphan steroid receptor Nur77 in apoptosis accompanying antigen-induced negative selection. Immunity 3:273–282. https://doi.org/10.1016/1074-7613(95)90113-2

- Thompson J, Winoto A (2008) During negative selection, Nur77 family proteins translocate to mitochondria where they associate with Bcl-2 and expose its proapoptotic BH3 domain. J Exp Med 205:1029–1036. https://doi.org/10.1084/jem.20080101
- Liebmann M, Hucke S, Koch K et al (2018) Nur77 serves as a molecular brake of the metabolic switch during T cell activation to restrict autoimmunity. Proc Natl Acad Sci 115:E8017–E8026. https://doi.org/10.1073/pnas.1721049115
- Youn HD, Chatila TA, Liu JO (2000) Integration of calcineurin and MEF2 signals by the coactivator p300 during T-cell apoptosis. EMBO J 19:4323–4331. https://doi.org/10.1093/emboj/19.16. 4323
- Youn HD, Liu JO (2000) Cabin1 represses MEF2-dependent Nur77 expression and T cell apoptosis by controlling association of histone deacetylases and acetylases with MEF2. Immunity 13:85–94. https://doi.org/10.1016/s1074-7613(00)00010-8
- Palmer E (2003) Negative selection-clearing out the bad apples from the T-cell repertoire. Nat Rev Immunol 3:383–391. https:// doi.org/10.1038/nri1085
- Raza K, Scheel-Toellner D, Lee CY et al (2006) Synovial fluid leukocyte apoptosis is inhibited in patients with very early rheumatoid arthritis. Arthritis Res Ther 8:R120. https://doi.org/10. 1186/ar2009
- Greenberg SA (2019) Inclusion body myositis: clinical features and pathogenesis. Nat Rev Rheumatol 15:257–272. https://doi. org/10.1038/s41584-019-0186-x
- Nakka K, Ghigna C, Gabellini D, Dilworth FJ (2018) Diversification of the muscle proteome through alternative splicing. Skeletal Muscle 8:8. https://doi.org/10.1186/s13395-018-0152-3
- Yu X-Z, Levin SD, Madrenas J, Anasetti C (2004) Lck is required for activation-induced T cell death after TCR ligation with partial agonists. J Immunol 172:1437–1443. https://doi.org/10.4049/ jimmunol.172.3.1437
- 46. McNeill L, Salmond RJ, Cooper JC et al (2007) The differential regulation of Lck kinase phosphorylation sites by CD45 is critical for T cell receptor signaling responses. Immunity 27:425–437. https://doi.org/10.1016/j.immuni.2007.07.015
- Thompson JL, Shuttleworth TJ (2012) A plasma membrane-targeted cytosolic domain of STIM1 selectively activates ARC channels, an arachidonate-regulated store-independent Orai channel. Channels (Austin) 6:370–378. https://doi.org/10.4161/chan.21947
- Cortese A, Plagnol V, Brady S et al (2014) Widespread RNA metabolism impairment in sporadic inclusion body myositis TDP43-proteinopathy. Neurobiol Aging 35:1491–1498. https:// doi.org/10.1016/j.neurobiolaging.2013.12.029
- Weihl CC, Baloh RH, Lee Y et al (2015) Targeted sequencing and identification of genetic variants in sporadic inclusion body myositis. Neuromuscul Disord 25:289–296. https://doi.org/10. 1016/j.nmd.2014.12.009
- Gang Q, Bettencourt C, Machado PM et al (2016) Rare variants in SQSTM1 and VCP genes and risk of sporadic inclusion body myositis. Neurobiol Aging 47:218 e211-218 e219. https://doi.org/ 10.1016/j.neurobiolaging.2016.07.024