

Analysis of human urinary extracellular vesicles reveals disordered renal metabolism in myotonic dystrophy type 1

Received: 19 December 2023

Accepted: 17 January 2025

Published online: 05 March 2025



Preeti Kumari¹, Lauren M. Sullivan¹, Zhaozhi Li¹, E. Parker Conquest¹, Elizabeth Cornforth², Rojashree Jayakumar¹, Ningyan Hu¹, J. Alexander Sizemore¹, Brigham B. McKee¹, Robert R. Kitchen³, Paloma González-Pérez¹, Constance Linville⁴, Karla Castro⁵, Hilda Gutierrez⁶, Soleil Samaan⁶, Elise L. Townsend², Basil T. Darras⁷, Seward B. Rutkove⁶, Susan T. Iannaccone⁵, Paula R. Clemens^{8,9}, Araya Puwanant⁴, Sudeshna Das¹✉ & Thurman M. Wheeler¹✉

Chronic kidney disease (CKD) and the genetic disorder myotonic dystrophy type 1 (DM1) each are associated with progressive muscle wasting, whole-body insulin resistance, and impaired systemic metabolism. However, CKD is undocumented in DM1 and the molecular pathogenesis driving DM1 is unknown to involve the kidney. Here we use urinary extracellular vesicles (EVs), RNA sequencing, droplet digital PCR, and predictive modeling to identify downregulation of metabolism transcripts *Phosphoenolpyruvate carboxykinase-1*, *4-Hydroxyphenylpyruvate dioxygenase*, *Dihydropyrimidinase*, *Glutathione S-transferase alpha-1*, *Aminoacylase-1*, and *Electron transfer flavoprotein B* in DM1. Expression of these genes localizes to the kidney, especially the proximal tubule, and correlates with muscle strength and function. In DM1 autopsy kidney tissue, characteristic ribonuclear inclusions are evident throughout the nephron. We show that urinary organic acids and acylglycines are elevated in DM1, and correspond to enzyme deficits of downregulated genes. Our study identifies a previously unrecognized site of DM1 molecular pathogenesis and highlights the potential of urinary EVs as biomarkers of renal and metabolic disturbance in these individuals.

Myotonic dystrophy type 1 (dystrophia myotonica; DM1) is an autosomal dominant multisystem disorder that features myotonia, progressive muscle atrophy, insulin resistance, cardiac conduction disturbance, and disordered metabolism^{1,2}. Prevalence is estimated as 1

in 3000–8000 worldwide³. DM1 is caused by a CTG repeat expansion (CTG^{exp}) in the 3' untranslated region of the *DM Protein Kinase (DMPK)* gene⁴. In the general population, the *DMPK* gene contains between 5 and 34 CTG repeats, while in DM1, the expansion is 50 or more.

¹Department of Neurology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. ²School of Health and Rehabilitation Sciences, Massachusetts General Hospital Institute of Health Professions, Boston, MA, USA. ³Cardiovascular Research Center, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. ⁴Department of Neurology, Wake Forest University School of Medicine, Winston-Salem, NC, USA. ⁵Department of Pediatrics, University of Texas Southwestern, Dallas, TX, USA. ⁶Department of Neurology, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA. ⁷Department of Neurology, Boston Children's Hospital and Harvard Medical School, Boston, MA, USA. ⁸Department of Neurology, University of Pittsburgh, Pittsburgh, PA, USA. ⁹Neurology Section, Veteran's Affairs Pittsburgh Health Care System, Pittsburgh, PA, USA.

✉ e-mail: sdas5@mg.harvard.edu; twheeler1@mg.harvard.edu

The size of the expansion is a predictor of clinical severity⁵. Symptoms in DM1 result from the accumulation of *DMPK*-CUG^{exp} RNA in nuclear inclusions of skeletal muscle, heart, brain, and other affected tissues. This pathogenic RNA sequesters proteins in the muscleblind-like (MBNL) family, resulting in partial loss of MBNL protein function⁶. CUG^{exp} transcripts also lead to increased steady-state levels of RNA-binding proteins in the CUGBP1, Elav-like factor (CELF) family⁷. The unbalanced MBNL and CELF1 activity disrupts regulatory control of alternative splicing and transcript abundance⁸. A second form of myotonic dystrophy, DM2, results from a CCTG repeat expansion in intron 1 of the *CNBP* gene⁹. DM1 and DM2 share an RNA gain-of-function pathogenesis involving partial loss of MBNL1 protein function and some clinical features, including myotonia and muscle weakness. However, symptoms in DM2 are generally milder and, in contrast to DM1, CELF1 protein levels are not elevated in affected tissues¹⁰.

Chronic kidney disease (CKD) is defined as abnormalities of kidney structure or function lasting for more than 3 months¹¹. The glomerular filtration rate (GFR) is considered the best overall indicator of kidney function. In healthy young men and women, GFR is ~125 ml/min/1.73 m². A threshold GFR of <60 ml/min/1.73 m² for more than 3 months indicates CKD¹¹. In DM1, CKD has never been documented as a clinical feature and there are no clinical care recommendations to monitor kidney function¹³. However, two recent studies have reported reduced GFR in DM1. A retrospective analysis that used cystatin C to estimate GFR found lower values in adults with DM1 (mean ~101 ml/min/1.73 m²; *N* = 141) than in adults with other muscular dystrophies (means 110–124 ml/min/1.73 m²)¹². A second study that measured GFR using plasma clearance of the X-ray contrast agent iohexol found a mean value of 74 ml/min/1.73 m² in DM1 (*N* = 94)¹³. GFR values were >90 ml/min/1.73 m² (defined as normal) in 15% of DM1 participants, 60–89 (mild impairment) in 68%, and 30–59 (moderate impairment) in 17%. None had a GFR 15–29 (severe impairment) or <15 (uremic). Due to the cross-sectional nature of the study, the duration of the GFR deficit is unknown, and, therefore, it's unclear how many would classify as CKD.

Extracellular vesicles (EVs) are membrane-encased particles released and taken up by cells as a form of extracellular communication¹⁴. EVs contain mRNA, noncoding RNAs, microRNAs—collectively termed extracellular RNAs (exRNAs)—that can serve as biomarkers of cancers and other disease states^{15,16}. Previous -omics studies have identified the urinary tract as the source 99.96% of urinary EVs, with the apical portion the major contributor^{17,18}. The remaining 0.04% of urinary EVs is believed to arise from cells infiltrating the kidney and/or contaminating skin cells. Urinary EVs carry exRNAs, proteins, and metabolites that reflect the molecular physiology and pathology of the parent cells, enabling their use as liquid biopsies of kidney, bladder, and prostate tissues¹⁵. In a previous study, our group found that urinary EVs contain alternative splice variant biomarkers of DM1, suggesting that the kidney and other urinary tract tissues may be involved in the DM1 molecular pathogenesis¹⁹. In this study, we use RNA sequencing of urinary EVs, droplet digital PCR (ddPCR), DM1 autopsy kidney tissue, predictive modeling, and urinary metabolites to test the hypothesis that DM1 pathogenesis involves the kidney.

Results

Differential gene expression in DM1 urinary EVs

To expand our understanding of the exRNA profile in DM1, we performed sequencing of urinary EVs isolated from individuals with DM1 or Duchenne muscular dystrophy (DMD) (Fig. 1a and Supplementary Fig. 1). Genes with mean transcripts per million (TPM) values of ≥ 5 were analyzed for differential gene expression using the edgeR package in R²⁰ (Supplementary Dataset). We identified 87 differentially expressed genes (DEGs) in DM1 vs. DMD (Fig. 1b). Next we applied principal component analysis (PCA), a statistical technique for dimensionality reduction that transforms the original variables into a

new set of orthogonal (uncorrelated) variables (i.e., the principal components) to capture maximum variance in the data, for convenient visualization of similarities and differences between samples and groups^{21,22}. PCA of the 87 DEGs produced a large separation between the two groups along the first principal component, PC1 (Fig. 1c). Using Gene Set Enrichment Analysis (GSEA) and the Gene Ontology Biologic Process datasets^{23,24}, the top 10 downregulated categories all involve metabolism of small molecules, organic acids, amino acids, or carbohydrates, while the upregulated genes include negative regulation of RNA and nucleobase metabolism, negative regulation of biosynthesis, and cytoskeletal organization (Fig. 1d). Specific DEGs include *Phosphoenolpyruvate Carboxykinase-1* (*PCK1*), *4-Hydroxyphenylpyruvate Dioxygenase* (*HPD*), *Dihydropyrimidinase* (*DPYS*), *Glutathione S-Transferase A 1* (*GSTA1*), *Aminoacylase-1* (*ACY1*), *Electron Transfer Flavoprotein B* (*ETFB*), *Glutaryl CoA Dehydrogenase* (*GCDH*), *Hydroxysteroid 17 Beta Dehydrogenase 14* (*HSD17B14*), *Phosphatidylinositol Glycan Anchor Biosynthesis Class T* (*PIGT*), *Bridging Integrator 1* (*BINI*), *Keratin 20* (*KRT20*), and *15-Hydroxyprostaglandin Dehydrogenase* (*HPGD*) (Fig. 1e).

Next, we used gene-specific primers to examine DEGs by RT-PCR and/or quantitative ddPCR. The transcript *FAM168A* was used as a reference gene due to its relatively low coefficient of variation (0.12) observed by RNASeq (Supplementary Fig. 2). In urinary EVs, *FAM168A*-normalized expression of *PCK1*, *HPD*, *DPYS*, *GSTA1*, *ACY1*, *ETFB*, *GCDH*, *HSD17B14*, *PIGT*, and *BINI* was reduced in the DM1 group relative to DM2, DMD, or unaffected (UA) healthy control individuals, while expression of *KRT20* and *HPGD* was elevated in the DM1 group, all matching the sequencing data (Fig. 2a and Supplementary Fig. 2). ddPCR quantification using a second reference gene, *General Transcription Factor 2B* (*GTF2B*) gave similar results and showed strong correlation with the *FAM168A*-normalized values (Supplementary Figs. 3 and 4). Gene expression patterns appeared similar in females and males, and across ages (Supplementary Figs. 5 and 6). Of the 12 ddPCR-confirmed DEGs, *HPD*, *ETFB*, and *HSD17B14* correlated with leukocyte CTG repeat length in males (*r* = −0.61 to −0.51, *P* = 0.036–0.046; Supplementary Fig. 7), and only *BINI* has a previously known interaction with DM1-related transcripts/proteins (Supplementary Fig. 8; String database [v. 12.0]²⁵). In contrast to urinary EVs, ddPCR gene expression patterns appeared similar in serum EVs of the DM1 and UA groups (Fig. 2b and Supplementary Fig. 9).

Localization of downregulated genes to the kidney

Within urinary tract tissues, four of the DEGs in DM1 urinary EVs are localized almost exclusively to the kidney: >99.8% of *PCK1* and *DPYS*, >99.6% of *GSTA1*, and >96.8% of *HPD* (Adult Genotype Tissue Expression [GTEx] Project: <https://www.gtexportal.org/home/aboutAdultGtex>) (Table 1). Using gene-specific primers, ddPCR, and commercially available total RNA from human tissues, we confirmed that expression of *PCK1*, *HPD*, *DPYS*, and *GSTA1* was limited to the kidney and *KRT20* the bladder, while the remaining DEGs were evident in multiple tissues (Supplementary Fig. 9). To evaluate the role of the kidney as a source of DEGs in DM1 urinary EVs, we performed ddPCR on DM1 renal autopsy tissue from individuals ages 50–55 with advanced DM1 (Muscle Impairment Rating Scale score of 4 [MIRS 4]²⁶; *N* = 3), and age 80 with minimal DM1 (MIRS 3, including normal strength of ankle dorsiflexion; *N* = 1) and a CTG repeat expansion <100. We found that expression of metabolism genes *PCK1*, *HPD*, *DPYS*, *GSTA1*, *ACY1*, and *ETFB* is reduced in DM1 kidney, with tissue from more severely affected individuals showing greater reduction than tissue from the individual with mild muscle impairment (Fig. 3a). Examination of the Kidney Tubule Expression Atlas^{27,28} revealed transcripts per million (TPM) values of >1000 in the proximal tubule for *Pck1*, *Acy1*, and *Etfb*, >2000 for *Gsta1*, and >4000 for *Hpd*. Within the nephron, *Pck1* (>99%), *Hpd* (97%), *Dpys* (100%), and *Gsta1* (94%) are localized to the proximal tubule (Fig. 3b and Supplementary

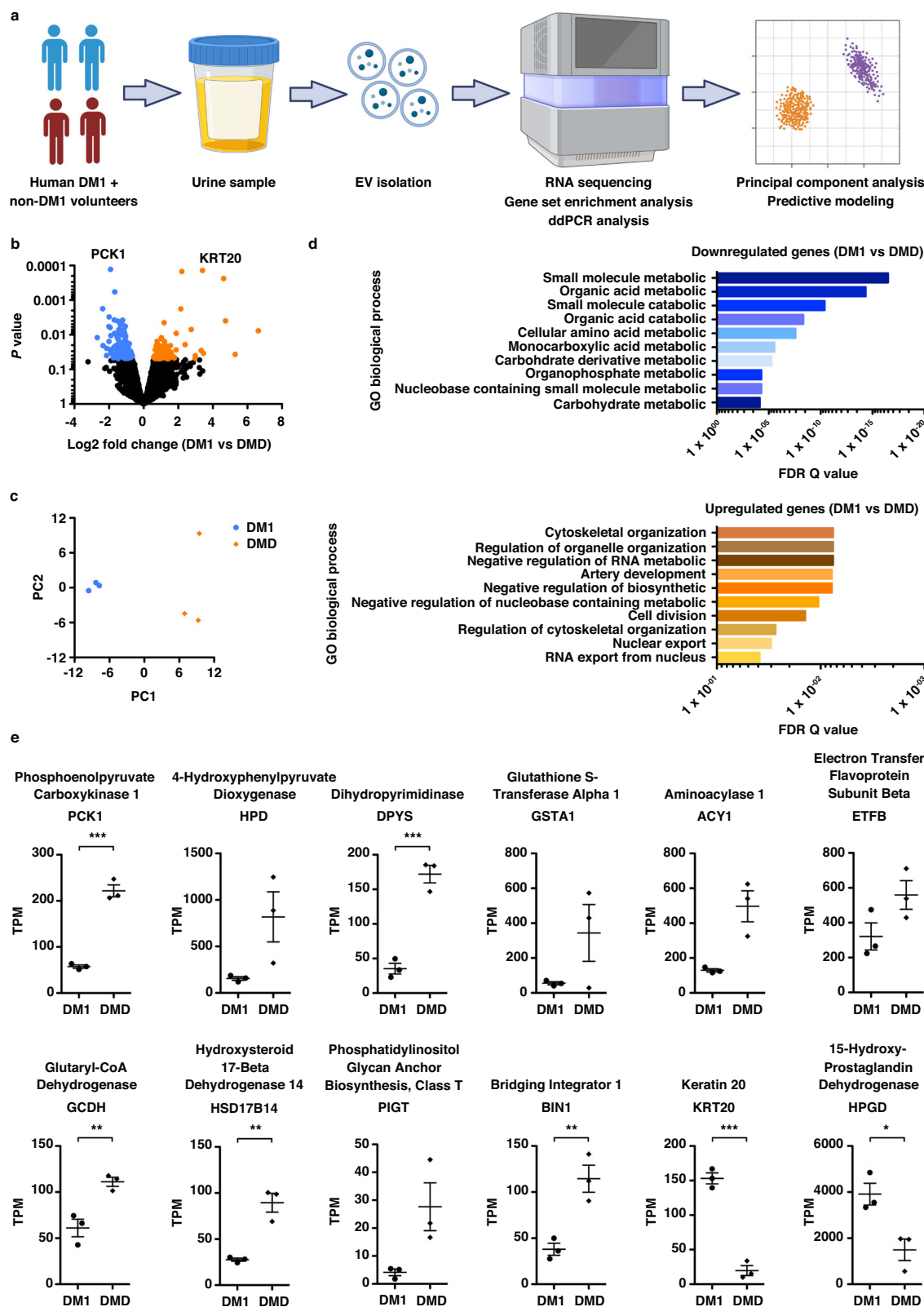


Fig. 10). The majority of *Acy1* (63%) and *Gcdh* mRNA (52%) also are found in the proximal tubule, while *Etfb* is evident in most regions of the nephron.

Urine contains a mixture of epithelial cells, red blood cells, and white blood cells that may be contributing to the urinary EV pool. Epithelial cell types found in the urine include (1) renal tubular, (2) transitional that arise from the renal pelvis, ureters, or bladder, and (3)

squamous that arise from the urethra during the process of micturition²⁹. To explore this possibility that a mixed population of urinary cells may be useful as gene expression biomarkers of DM1, we performed ddPCR to quantify gene expression. Genes that are downregulated in DM1 urinary EVs also tended to be expressed at lower levels in mixed urine cells of DM1 as compared to UA, although to a lesser degree and the differences statistically non-significant

Fig. 1 | Study overview and RNA sequencing analysis. **a** Study overview. EV extracellular vesicle, ddPCR droplet digital polymerase chain reaction. Created in BioRender (Kumari; <https://BioRender.com/s97f089>)⁸⁰. **b** Urine exRNA sequencing analysis in DM1 and DMD. Volcano plot of log₂ of the fold change (DM1 vs DMD) following edgeR analysis. Using a *P* value < 0.05, downregulated transcripts are highlighted in blue and upregulated transcripts are shown in orange. **c** Principal component analysis of differentially expressed transcripts using a log₂ fold change >1.0 or < -1.0 and a *P* value < 0.05. **d** Gene Set Enrichment Analysis (GSEA) using the

Molecular Signatures Database and the Gene Ontology Biologic Process gene sets comparing DM1 and DMD exRNA samples. The top 10 categories of downregulated (blue) and upregulated (orange) genes in DM1 by false discovery rate (FDR) *Q* value are shown. **e** Transcripts per million (TPM) values in DM1 and DMD (*N* = 3 each) for individual differentially expressed genes. ****P* = 0.0003 (*PCK1*), 0.0008 (*DPYS*), and 0.0002 (*KRT20*); ***P* = 0.0096 (*GCDH*), 0.0039 (*HSD17B14*), and 0.0089 (*BIN1*); **P* = 0.0219 (two-tailed *t* tests). Error bars indicate ± s.e.m. Source data are provided as a Source Data file.

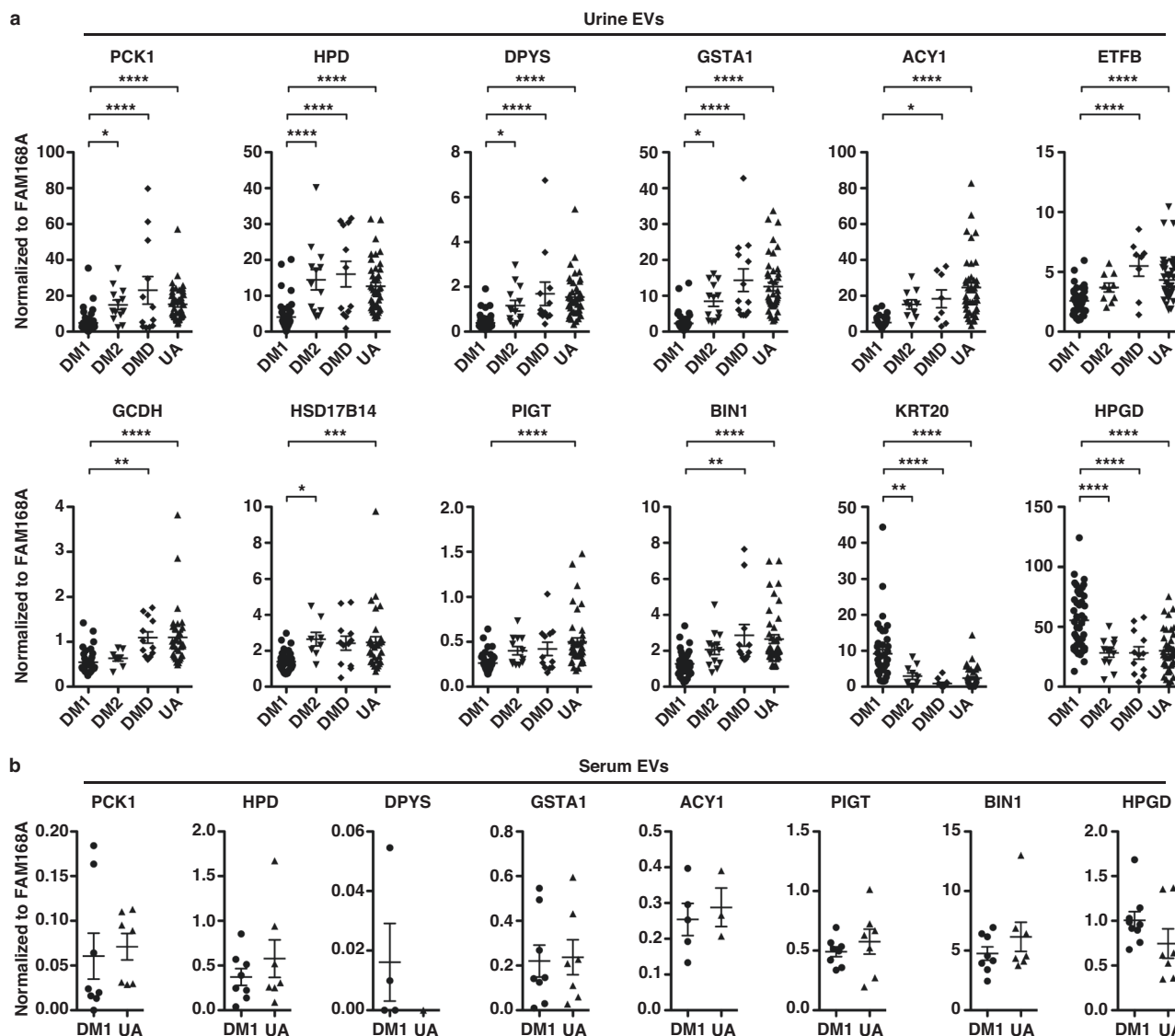


Fig. 2 | ddPCR quantification of gene expression in EVs. **a** We used ddPCR to quantify expression of *PCK1*, *HPD*, *DPYS*, *GSTA1*, *ACY1*, *ETFB*, *GCDH*, *HSD17B14*, *PIGT*, *BIN1*, *KRT20*, and *HPGD* normalized to *FAM168A* in urinary EVs of DM1 (*N* = 39–47), DM2 (*N* = 10–13), DMD (*N* = 8–12), and UA controls (*N* = 27–39).

*****P* < 0.0001; ****P* < 0.001; ***P* < 0.01; **P* < 0.05 (one-way ANOVA/Tukey). **b** ddPCR quantification of *PCK1*, *HPD*, *DPYS*, *GSTA1*, *ACY1*, *PIGT*, *BIN1*, and *HPGD* expression normalized to *FAM168A* in serum EVs of DM1 (*N* = 4–8) and UA controls (*N* = 3–7). Error bars indicate ± s.e.m. Source data are provided as a Source Data file.

(Supplementary Fig. 11). Of the transcripts upregulated in DM1 exRNA, *KRT20* also was significantly increased in mixed urine cells of DM1 vs. UA controls, while *HPGD* appeared similar in both groups. Taken together, these findings are consistent with a prior study that found alternative splice events in urine exRNA were more reliable indicators of DM1 molecular pathogenesis than those found in a mixed population of urine cells¹⁹.

Ultracentrifugation isolates a pellet consisting of EVs and non-EV-associated high molecular weight protein and protein complexes, some of which bind RNA³⁰. To determine whether the mRNAs that we are measuring are localized to EV cargo, we used size exclusion chromatography (SEC) to separate EVs from non-EV-associated proteins and RNAs. Using ddPCR, we found that the purified EV fraction is enriched for mRNAs as compared to the mixed EV/non-EV protein

Table 1 | Gene expression in human tissues

Gene	Transcripts per million (TPM), GTEx portal											% of total urinary tract	
	Kidney cortex	Kidney medulla	Bladder	Prostate	Liver	Whole blood	Muscle, skeletal	Heart, atrial appendage	Heart, LV	Anterior cingulate cerebral cortex	Frontal cerebral cortex	Bladder	Kidney
<i>PCK1</i>	147.6	63.45	0.41	0.28	530	0.067	0.22	0.51	0.1	0.28	0.36	0.19	99.81
<i>HPD</i>	168.7	22.98	6.09	2.7	1002	2.51	1.03	2.19	0.59	0.16	0.1	3.08	96.92
<i>DPYS</i>	39.09	12.04	0.09	2.42	132.8	0.03	0.02	0.03	0.01	0.7	1.3	0.18	99.82
<i>GSTA1</i>	101.2	32.41	0.5	6.4	424	0.15	0.03	0.09	0.04	0.09	0.25	0.37	99.63
<i>ACY1</i>	113	58.38	13.55	24.58	125.6	2.56	14.44	9.55	7.4	7.58	7.96	7.33	92.67
<i>ETFB</i>	20.97	24.95	24.44	20.29	60.99	10.96	34.43	39.74	46.42	27.52	28.55	34.74	65.26
<i>GCDH</i>	12.35	10.52	19.04	20.09	36.06	6.46	13.12	9.89	13.29	10.09	10.8	45.43	54.57
<i>HSD17B14</i>	35.28	18.45	28.07	23.61	27.27	0.35	0.69	5.39	1.18	27.43	25.1	34.32	65.68
<i>PIGT</i>	10.27	13	28.93	20.44	6.76	8.67	12.82	23.24	17.82	4.95	8.05	55.42	44.58
<i>BIN1</i>	55.11	44.03	125.8	77.59	17.88	22.64	953.2	7.77	7.34	101.6	134.4	55.93	44.07
<i>KRT20</i>	0.02	0.01	3.26	0.09	0.03	0.01	0.02	0.02	0.01	0	0	99.09	0.91
<i>HPGD</i>	7.37	5.53	32.91	6.79	33.56	4.89	0.81	0.25	0.24	0.28	0.29	71.84	28.16
<i>DMPK</i>	28.21	34.95	79.36	97.62	7.72	4.58	98.16	106.3	100.9	28.47	33.42	55.68	44.32
<i>CELF1</i>	16.7	28.68	41.46	36.6	24.23	18.45	27.36	15.27	12.16	18.75	24.11	47.74	52.26
<i>MBNL1</i>	12.1	15.16	71.96	42.01	13.68	26.99	74.52	26.74	19.76	13.04	18.7	72.53	27.47
<i>MBNL2</i>	15.36	19.93	44.73	37.1	26.88	1.93	27.91	20.74	15.46	28.44	47.49	55.90	44.10

LV left ventricle. Transcript abundance of urinary EV DEGs *PCK1*, *HPD*, *DPYS*, *GSTA1*, *ACY1*, *ETFB*, *GCDH*, *HSD17B14*, *PIGT*, *BIN1*, *KRT20*, and *HPGD* in adult human tissues (from the Adult Genotype Tissue Expression (GTEx) project; GTEx portal, <https://www.gtexportal.org/home/>). *DMPK*, the gene that causes DM1, and genes encoding RNA-binding proteins involved in DM1 molecular pathogenesis, *CELF1*, *MBNL1*, and *MBNL2*, also are included. The transcript abundance is displayed as transcripts per million and as percent (%) total of each transcript within the urinary tract.

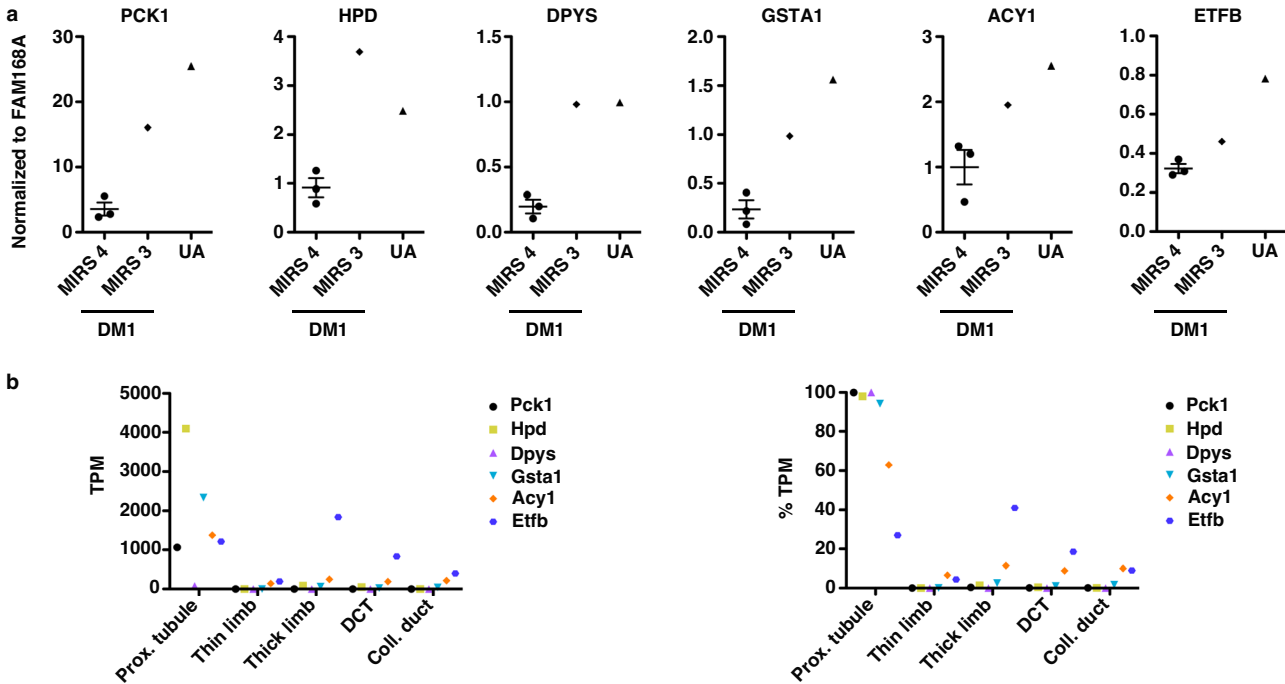


Fig. 3 | ddPCR quantification of gene expression in tissues and localization within the nephron. **a** ddPCR quantification of *PCK1*, *HPD*, *DPYS*, *GSTA1*, *ACY1*, and *ETFB* normalized to reference gene *FAM168A* in human DM1 autopsy kidney tissue from individuals ages 50–55 with advanced DM1 (Muscle Impairment Rating Scale score of 4 [MIRS 4]²⁶; *N* = 3) and age 80 with minimal DM1 (MIRS 3, including normal strength of ankle dorsiflexion; *N* = 1) and a CTG repeat expansion <100. Commercially available total RNA of human kidney served as a control. Error bars indicate ±

s.e.m. **b** Renal tubule localization of *Pck1* (black circles), *Hpd* (yellow squares), *Dpys* (purple triangles), *Gsta1* (turquoise inverted triangles), *Acy1* (orange diamonds), and *Etfb* (blue hexagons) expression by RNA sequencing of microdissected rat nephrons (Kidney Tubule Expression Atlas²⁷). TPM transcripts per million, Prox. tubule proximal tubule, DCT distal convoluted tubule, Coll. duct collecting duct. Source data are provided as a Source Data file.

pellet isolated by ultracentrifugation (UCF) from the same samples (Supplementary Fig. 12).

DM1 molecular pathogenesis in the kidney

In DM1 skeletal muscle, heart, and cerebral cortex, *DMPK*-CUG^{exp} RNA sequesters MBNL proteins to form nuclear inclusions, which leads to changes in alternative splicing and gene expression patterns as compared to non-DM1 controls^{31,32}. Due to the shared molecular pathogenesis across tissues, renal dysfunction previously reported in DM1¹³ and the gene expression changes that we observed in DM1 kidney may be related to nuclear accumulation of *DMPK*-CUG^{exp} transcripts within kidney tissue. According to the Adult GTEx Project (<https://www.gtexportal.org/home/aboutAdultGtex>), *DMPK* expression in the kidney cortex (28.21 transcripts per million [TPM]) and kidney medulla (34.95 TPM) is similar to that in the anterior cingulate cerebral cortex (28.47 TPM) and the frontal cerebral cortex (33.42 TPM) (Table 1). If *DMPK* TPM values of 28–33 are sufficient to induce DM1 molecular pathogenesis in the cerebral cortex, it is logical to conclude that *DMPK* TPM values of 28–35 also may be sufficient induce DM1 molecular pathogenesis in the kidney. To investigate this possibility further, we performed fluorescence in situ hybridization and immunofluorescence analysis of DM1 autopsy kidney specimens ($N = 4$). We show that CUG^{exp} RNA nuclear inclusions co-localize with MBNL proteins in the proximal tubule, loop of Henle, collecting duct, and in podocytes, which are located on the urinary space side of the glomerular basement membrane (Fig. 4).

Principal component analysis and predictive modeling validate DEGs as indicators of DM1

ddPCR expression of the downregulated genes, especially those that localize mostly to the proximal tubule, strongly correlate with each other, while the two upregulated bladder-enriched genes, *KRT20* and *HPGD*, correlate with one another (Fig. 5a). To confirm that the gene expression changes we observed in this study indicate DM1 disease status, we used PCA and predictive modeling. The PCA combined the ddPCR gene expression values of several transcripts to serve as a weighted sum for each individual. Using the four proximal tubule genes (*PCK1*, *HPD*, *DPYS*, and *GSTA1*), PCA separated the DM1 and UA groups along the first principal component, PC1 (Fig. 5b), confirming that the combined changes in expression of these four genes indicate DM1 disease status and identifying the proximal tubule as a site of pathogenesis in DM1. By including an additional four genes (*PIGT*, *BINI*, *KRT20*, and *HPGD*) for an eight-transcript PCA, the weighted sum of differential gene expression between groups is greater than with four transcripts, and further separated the groups along both PC1 and the second principal component, PC2. The six downregulated genes contribute fairly evenly to the separation along PC1, while the two upregulated genes are responsible for the separation along PC2 (Fig. 5c).

To generate a predictive model for gene expression, we pooled the eight-transcript ddPCR data from 49 participants ($N = 24$ DM1; 25 UA), then randomly assigned 76% ($N = 37$) to a training set and the remaining 24% ($N = 12$) to a held-out test set¹⁹, maintaining an equal proportion of each group in both sets. Using our training set with principal component regression³³ and a threshold of 0.5, such that values greater than the threshold were DM1 and below threshold UA, the model was 97.5% accurate in a fivefold cross-validation set (Fig. 5d). The predictive model was accurate in distinguishing all DM1 patients from UA participants in the held-out test set ($N = 6$ each group), and in 19 of the next 22 individuals enrolled after model implementation, for a total of 34 evaluated (Fig. 5e). Overall, the model correctly predicted 16 of 19 DM1 and all 15 UA participants (31/34 = 91%). Of the DM1 individuals that the model predicted as UA, all three have mild clinical features that overlap with the UA group (Muscle Impairment Rating Scale score of 2, where the minimum 1 is

no impairment and the maximum of 5 is severe proximal muscle weakness²⁶).

Molecular mechanism of transcript downregulation

RNA-binding proteins in the CELF and MBNL families act antagonistically to regulate mRNA expression and alternative splicing by targeting sequence-specific regions in hundreds of mRNAs⁸. These proteins bind independently but functionally compete to specify transcript repression (CELF) or localization/stabilization (MBNL)⁸. In DM1, steady-state levels of CELF1 protein are increased due to hyperphosphorylation⁷, and MBNL proteins functionally depleted due to sequestration by CUG^{exp} RNA⁶. We noticed that the *PCK1* 3' UTR contains a 36-nucleotide consecutive stretch of the CELF1 binding motifs TGTG/GTGT (Fig. 6a). To explore further the possibility that the transcript repression that we observed in DM1 urine exRNA may result from overactivity of CELF1 and/or functional loss of MBNL, we performed a binding site enrichment analysis of the 3'UTR of the ten downregulated genes using a chi-square test (Fig. 6b). Five transcripts show enrichment of CELF1 binding sites in the 3' UTR, including *PCK1*, *DPYS*, *GSTA1*, *ACY1*, and *BINI* (all $P < 0.05$), while a sixth, *PIGT*, fell just short of statistical significance ($P = 0.05$). Five transcripts show enrichment of MBNL binding sites, including *PCK1*, *HPD*, *GSTA1*, *HSD17B14*, and *BINI* (all $P < 0.05$). Two transcripts, *ETFB* and *GCDH*, showed no enrichment of either CELF1 or MBNL binding sites, although the 3' UTR of *ETFB* includes an 11-nucleotide stretch of overlapping CELF1 motifs, while *GCDH* contains a TGTG CELF1 motif overlapping with MBNL motifs upstream and downstream.

The first 12 nucleotides of exon 10 in the *CELF1* transcript are alternatively spliced, resulting in the formation of CELF1 protein either with or without a leucine-tyrosine-leucine-glutamine (LYLQ) amino acid sequence located in the linker region of the protein³⁴ (Fig. 6c). The LYLQ inclusion isoform is the predominant form in human striated muscle³⁵. To examine alternative splicing of *CELF1* in other human tissues, cells, and EVs, we developed a ddPCR assay to quantify the inclusion of the LYLQ-encoding nucleotides. As compared to skeletal muscle, full-length exon 10 inclusion was lower in UA kidney and bladder tissue, and higher in UA brain tissue (Fig. 6d). In urinary EVs, inclusion of full-length exon 10 was significantly lower in the DM1 group (mean 38.9%) than in either the DM2 (54.1%), DMD (59.6%), or UA (57.1%) groups ($P < 0.0001$). Inclusion of full-length exon 10 also was lower in mixed urine cells of the DM1 than the UA group (mean 54.6% vs. 67.9%; $P < 0.01$), while in serum EVs, the difference between the DM1 and UA groups was smaller (75.2 vs. 81.0%; $P < 0.05$) (Supplementary Fig. 12). To estimate the potential for CELF1 and/or MBNL proteins as regulators of *CELF1* alternative splicing, we performed a chi-square binding site enrichment analysis in the exon and intron regions. Multiple CELF1 and MBNL sites were enriched in the downstream intron while the MBNL motif CAGC (chi-square 11.5) was enriched in the target exon, including at and downstream of the alternative splice site (Fig. 6e).

Elevation of urinary metabolites in DM1

Inborn errors of metabolism (IEMs) are a heterogeneous group of disorders in which a single enzyme defect causes a block in a biochemical pathway, resulting in a clinically relevant deficiency of the product or accumulation of toxic intermediates upstream of the block³⁶. Clinical presentation of IEMs is highly variable and can occur from infancy through adulthood, depending on residual enzyme activity. Several of the genes that we found to be downregulated in DM1 urine exRNA are associated with IEMs and organic acidurias due to deficient activity of the enzymes encoded by these genes^{37–43} (Supplementary Table 2). To explore the possibility that the downregulated genes may induce metabolic disturbance in DM1, we performed organic acid analysis on random urine samples of DM1 individuals. Urine creatinine, the analyte used to normalize each acid,

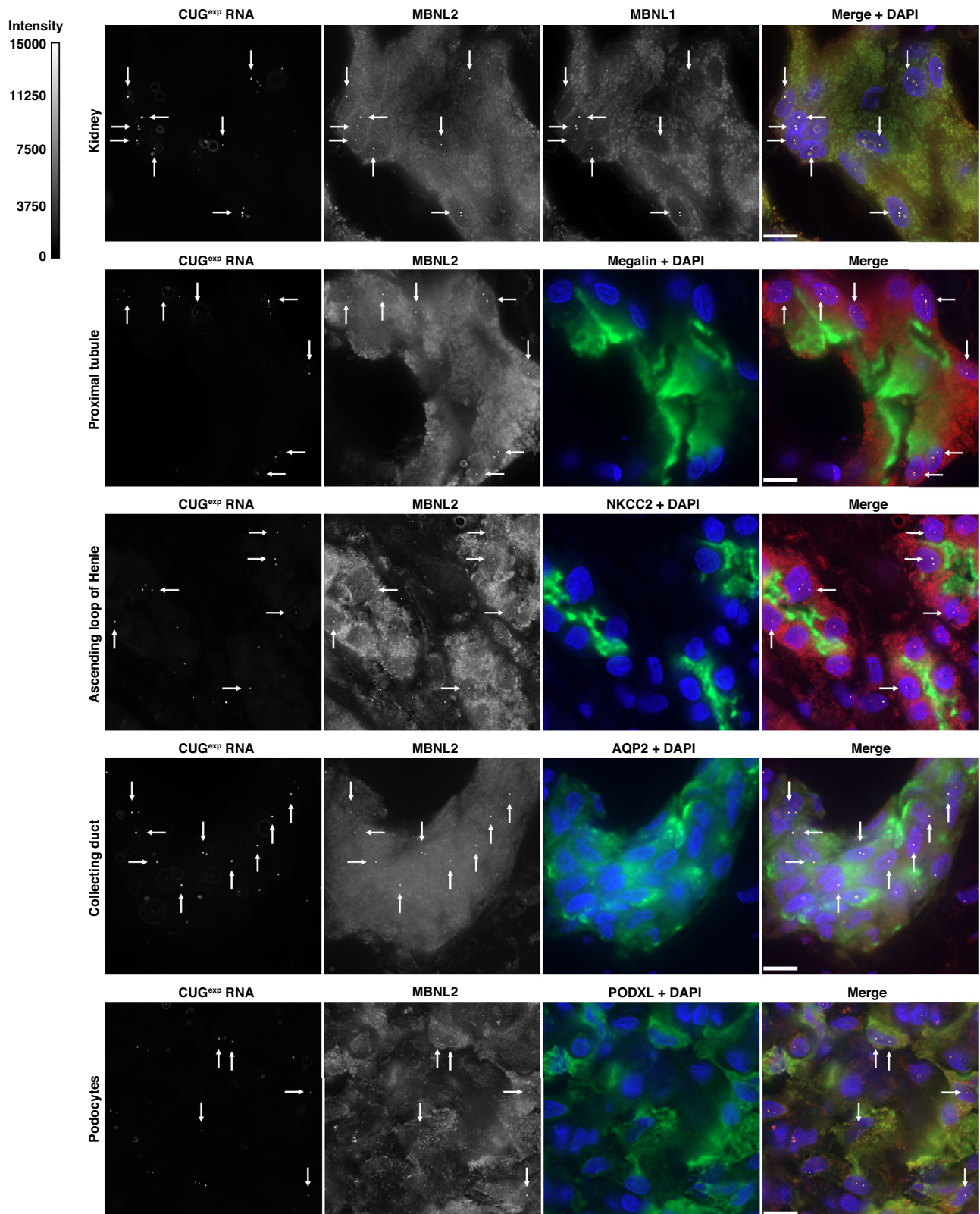


Fig. 4 | Localization of CUG^{exp} RNA within the nephron. We performed fluorescence in situ hybridization for CUG^{exp} RNA and immunofluorescence labeling for MBNL2 together with MBNL1, megalin (renal proximal tubule), sodium-potassium-chloride cotransporter 2 (NKCC2; loop of Henle), aquaporin 2 (AQP2; collecting duct), or podocalyxin (PODXL; podocytes) proteins in DM1 autopsy kidney

specimens ($N = 4$). The arrows indicate co-localization of CUG^{exp} RNA and MBNL2 protein with MBNL1 protein (upper row) and CUG^{exp} RNA and MBNL2 protein within the indicated regions of the nephron (lower four rows). In the merge images, yellow indicates CUG^{exp} RNA, red indicates MBNL2 protein, blue indicates nuclei, and green indicates MBNL1, megalin, NKCC2, AQP2, or PODXL. Bars = 10 μ m.

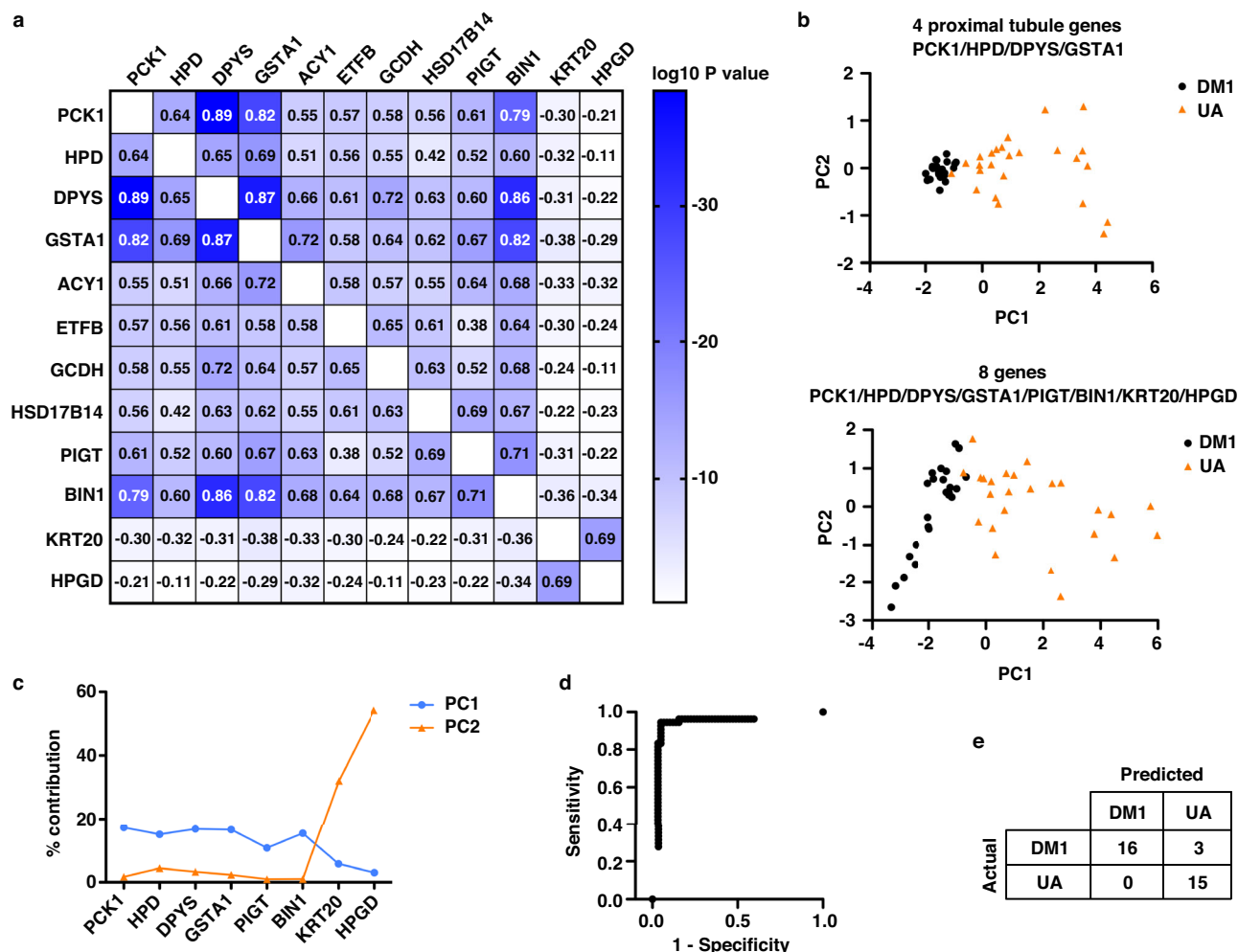


Fig. 5 | ddPCR correlations, principal component analysis (PCA), and predictive modeling. **a** ddPCR expression correlation matrix of 12 DEGs in urinary EVs (see Fig. 2). The Pearson correlation coefficient r for each comparison is shown within the matrix. The scale shows the \log_{10} of the P values for each correlation within the matrix (two-tailed t tests). **b** Principal component analysis of ddPCR gene expression in urinary EVs of DM1 ($N=24$) and UA ($N=25$) using four proximal tubule genes *PCK1*, *HPD*, *DPYS*, and *GSTA1* (upper) and eight genes comprised of the four proximal tubule genes along with *PIGT*, *BIN1*, *KRT20*, and *HPGD* (lower). The first two principal components PC1 and PC2 are shown. **c** Percent contribution of each of the eight genes to the separation of DM1 and UA along PC1 (blue)

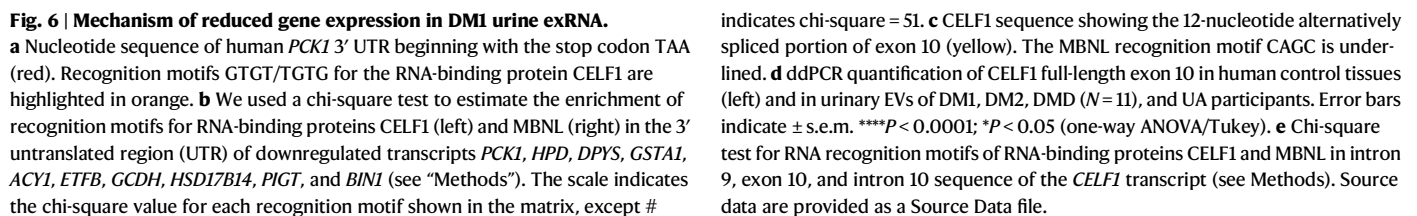
and PC2 (orange). **d** We combined 8-transcript PC data from the first 49 participants ($N=24$ DM1; 25 UA), then randomly assigned 76% ($N=37$), irrespective of genotype, to a training set, and the remaining 24% ($N=12$) to an independent validation set. Using the training set and a threshold of 0.5 (see “Methods”), we developed a predictive model that was 97.5% accurate in a fivefold cross-validation test. The receiver operating characteristic (ROC) curve is shown. **e** Confusion matrix for the held-out test set ($N=12$) and the next 22 individuals enrolled after model implementation ($N=34$ total). The model correctly predicted 16/19 DM1 and all 15 UA volunteers (31/34 = 91%). Source data are provided as a Source Data file.

was within the laboratory reference range for each specimen (Fig. 7a). Several samples contained elevated pyruvic acid, tricarboxylic acid (TCA) cycle intermediate 2-oxoglutaric acid, C6-C10 dicarboxylic acids, and/or tyrosine catabolism intermediates, consistent with deficiencies of *PCK1*, *ETFB*, and/or *HPD* enzymes (Fig. 7a and Supplementary Fig. 13). The most frequently elevated metabolite was 2-hydroxyglutaric acid, which can be elevated in disorders of fatty acid oxidation (e.g., *ETFB* deficiency) or produced by non-enzymatic reduction of 2-oxoglutaric acid⁴³. To screen further for metabolites associated with *ACY1* and *ETFB* deficiencies, we next examined an acylglycines panel in random urine samples of DM1 individuals. Several acylglycines and fatty acid oxidation intermediates were elevated, consistent with deficiencies of *ACY1* or *ETFB* (Fig. 7b and Supplementary Fig. 13). Elevated ethylmalonic acid, evident in more than half of the DM1 samples, can be seen in deficiencies of *PCK1* or *ETFB* enzymes. Glutaric acid, a metabolite that has been reported as elevated in *PCK1*, *ETFB*, or *GCDH* deficiencies, also was elevated in several DM1 samples, although 3-hydroxyglutaric acid, a metabolite specific

for *GCDH* deficiency, was absent in all samples. Of the 21 samples examined with the organic acids panel, 20 showed elevation of at least one metabolite, and 15 of the 16 samples examined with the acylglycines panel showed elevation of at least two metabolites (Supplementary Fig. 13).

Downregulated genes and urinary metabolites correlate with clinical measures

Next we examined the relationship between urinary EV gene expression and clinical parameters. The 6-min walk test (6MWT) is a measurement of total distance walked (meters) during a 6-min period as an indication of walking ability and endurance⁴⁴. To enhance the accurate assessment of 6MWT limitation in the DM1 population, we used existing predictive equations generated from healthy controls ages 18–50 years⁴⁵ and ages 51–80 years⁴⁶ (% predicted). Urinary EV expression of *PCK1*, *HPD*, *DPYS*, *GSTA1*, and *ACY1* correlated with 6MWT % predicted ($r=0.53$ – 0.71 ; $P=0.002$ to <0.0001) (Fig. 8a). In DM1, distal muscles in the ankles, hands, and wrists, are affected earlier



Progressive neuromuscular respiratory failure is a common cause of mortality in DM1⁴⁸. A decline in respiratory muscle strength may

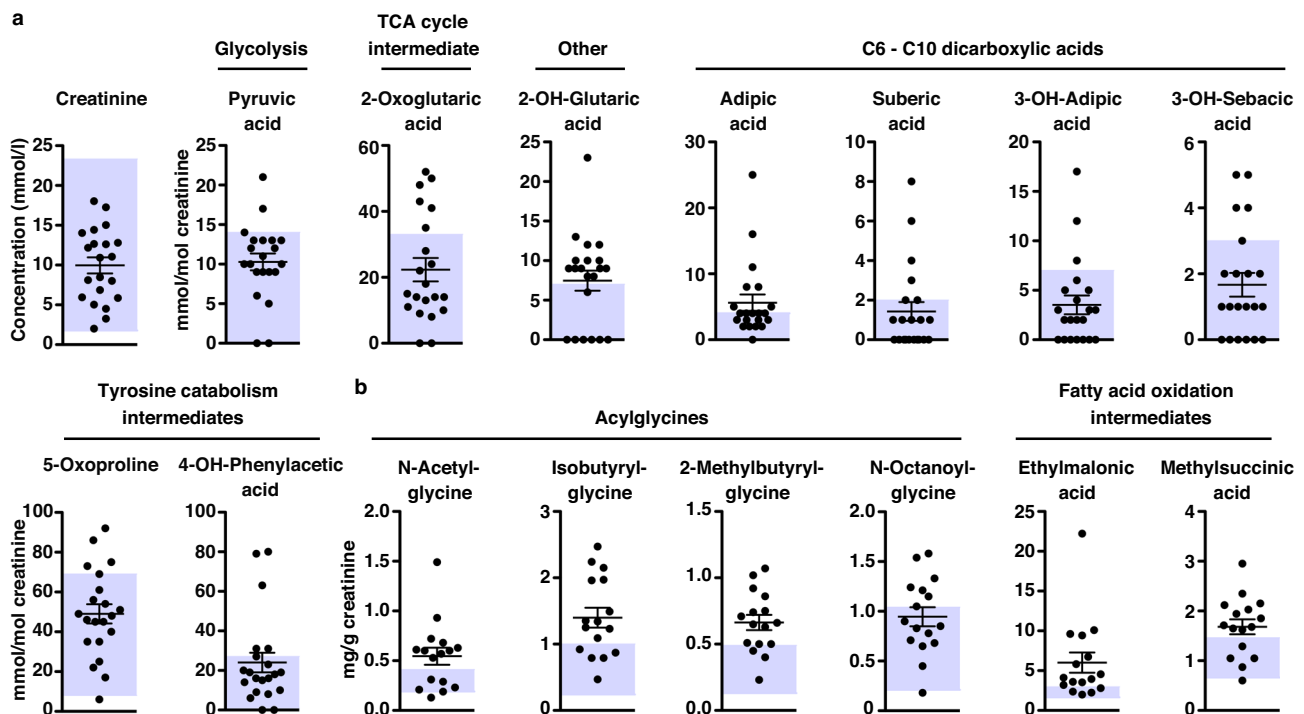


Fig. 7 | Elevated urine organic acids and acylglycines in DM1. **a** Elevated urine organic acids in DM1 ($N = 20$ individuals) include TCA cycle intermediates, C6-C10 dicarboxylic acids, and tyrosine catabolism intermediates (mmol metabolite/mol creatinine; Quest Diagnostics). **b** Elevated acylglycines and fatty acid oxidation

intermediates in DM1 ($N = 16$ individuals) (mg metabolite/g creatinine; Mayo Clinic Laboratories). The blue-shaded areas indicate the reference range for each metabolite. Error bars indicate \pm s.e.m. See Supplementary Fig. 13. Source data are provided as a Source Data file.

result from respiratory muscle sarcopenia. Peak expiratory flow rate (PEFR) is a pulmonary function measurement of the maximum speed that air is forcefully exhaled, and has been proposed as an indicator of respiratory muscle sarcopenia⁴⁹. Maximal inspiratory pressure (MIP) is a pulmonary function measurement of respiratory muscle strength. We found that urinary EV expression of *PCK1*, *HPD*, *DPYS*, and *ETFB* each correlated with PEFR % predicted ($r = 0.47$ – 0.59 ; $P = 0.006$ – 0.0002), and that *PCK1*, *HPD*, and *DPYS* each correlated with MIP % predicted in DM1 individuals ($r = 0.41$ – 0.50 ; $P = 0.013$ – 0.003) (Fig. 8d, e and Supplementary Fig. 14).

A urinary metabolite elevated in HPD deficiency, 5-oxoproline⁴³, correlated with several clinical measurements in DM1 participants, including ADF, handgrip, and pinch strength ($r = -0.59$ to -0.65 ; $P < 0.0001$), the Scale for the Assessment and Rating of Ataxia (SARA) score ($r = 0.52$; $P = 0.018$), electrocardiographic QRS duration (milliseconds) ($r = 0.58$; $P = 0.008$), and time to complete the 9-hole peg test (9HPT; seconds) ($r = 0.49$; $P = 0.001$) (Fig. 9a, b). A second metabolite associated with HPD deficiency, 4-hydroxyphenylacetic acid, also correlated with the 9HPT ($r = 0.65$; $P < 0.0001$). For these correlations, higher metabolite levels were associated with greater functional impairment.

Fatigue affects more than 90% of DM1 patients and is the symptom that has the greatest overall impact on their lives, even surpassing the limitations of mobility⁵⁰. A prior investigation found that serum levels of 5-oxoproline and 4-hydroxyphenylacetic acid were lower in individuals with exhaustion disorder as compared to healthy controls⁵¹. In our study, both urinary 5-oxoproline and 4-hydroxyphenylacetic acid inversely correlated with Fatigue and Daytime Sleepiness (FDSS) scores⁵² ($r = -0.58$ and -0.65 ; $P = 0.007$ and 0.002), with lower levels of these metabolites associated with greater fatigue in DM1 (Fig. 9c).

Loss of olfactory ability has been suggested as a possible indicator of central nervous system dysfunction in DM1⁵³. Using the

University of Pennsylvania Smell Identification Test (UPSIT)⁵⁴, DM1 participants scored lower (mean 27 of maximum score 40) than UA controls (mean 33) and with a wider range (10–38 in DM1, 24–38 in UA) ($P < 0.0001$) (Fig. 9d). Urinary metabolites N-acetylglycine (elevations associated with ACY1 deficiency), isobutyrylglycine (ETFB), 2-methylbutyrylglycine (ETFB), adipic acid (ETFB or PCK1), and aconitic acid (PCK1) all correlated with UPSIT scores in DM1 individuals ($r = -0.57$ to -0.70 ; $P = 0.034$ – 0.0008) (Fig. 9e).

Discussion

Renal involvement in DM1 has been mostly overlooked due to the absence of documented CKD. Here we use urinary EVs, ddPCR, DM1 autopsy kidney tissue, and predictive modeling to establish DM1 molecular pathogenesis in the kidney for the first time. The down-regulation of metabolism gene expression in DM1 urinary EVs and kidney autopsy tissue combined with the elevation of corresponding urinary metabolites that we observed is consistent with a functional deficit of the encoded enzymes. The enrichment of CELF1 and/or MBNL RNA recognition motifs in the 3' UTR of these genes argues in favor of transcript repression and/or instability as the mechanism of down-regulation. Our working model is that DM1 molecular pathogenesis within the nephron leads to (1) nuclear accumulation of *DMPK*-CUG^{exp} transcripts, (2) sequestration of MBNL proteins to form ribonuclear inclusions, (3) hyperphosphorylation of CELF proteins, (4) transcript repression from CELF gain-of-function (increased binding 3' UTR) and/or transcript instability from MBNL loss-of-function (reduced binding 3' UTR)⁸, (5) decreased transcript abundance, (6) functional deficiency of the encoded enzymes, and (7) elevation of corresponding metabolites (Fig. 9f). Together with the prior report of reduced renal function in DM1¹³, our results support longitudinal study of the relationship between GFR and urinary EV gene expression, urinary metabolites, and clinical severity in this population. Confirmation of CKD in DM1 may result in modification of clinical care

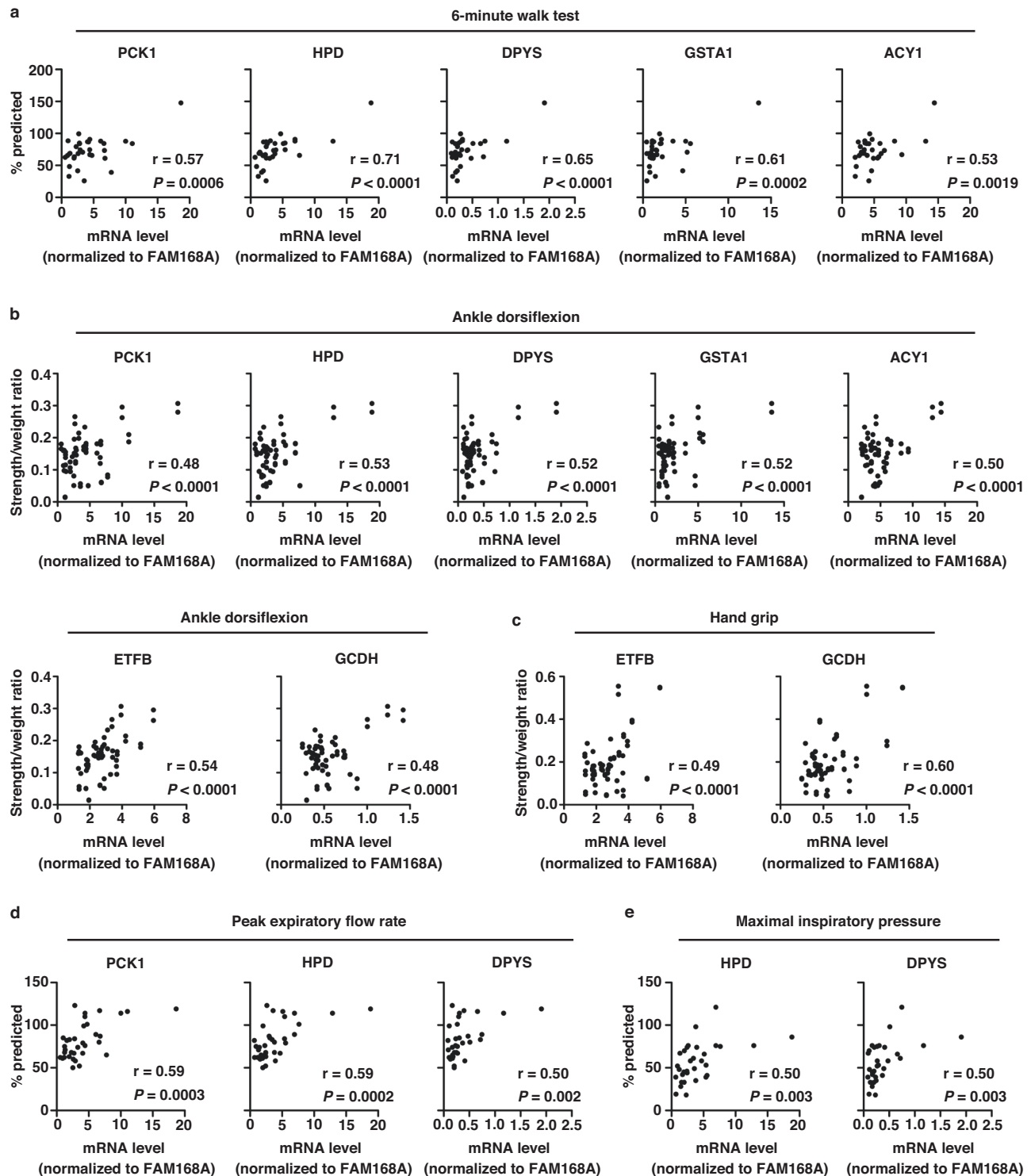


Fig. 8 | Correlation of urinary EV gene expression with functional measures. Correlation of ddPCR expression of the genes *PCK1*, *HPD*, *DPYS*, *GSTA1*, *ACY1*, *ETFB*, and *GCDH* in urine exRNA of DMI participants ($N = 33$) with clinical measurements of (a) 6-min walk test (6MWT) % predicted, (b) ankle dorsiflexion (ADF) strength-to-weight ratio ($N = 33$ left side + 33 right side = 66 total measures), and (c) handgrip (HG) strength-to-weight ratio ($N = 66$ total measures). The correlation coefficient and P value for each are shown (two-tailed t tests). (d) Correlation of ddPCR

expression of the genes *PCK1*, *HPD*, *DPYS* in urinary EVs with peak expiratory flow rate (PEFR) % predicted in DMI participants ($N = 33$). The correlation coefficient r and P values for each are shown (two-tailed t tests). (e) Correlation of ddPCR expression of *HPD* and *DPYS* in urinary EVs with maximal inspiratory pressure % predicted in DMI participants ($N = 33$). The correlation coefficient r and P values for each are shown (two-tailed t tests). See Supplementary Fig. 14. Source data are provided as a Source Data file.

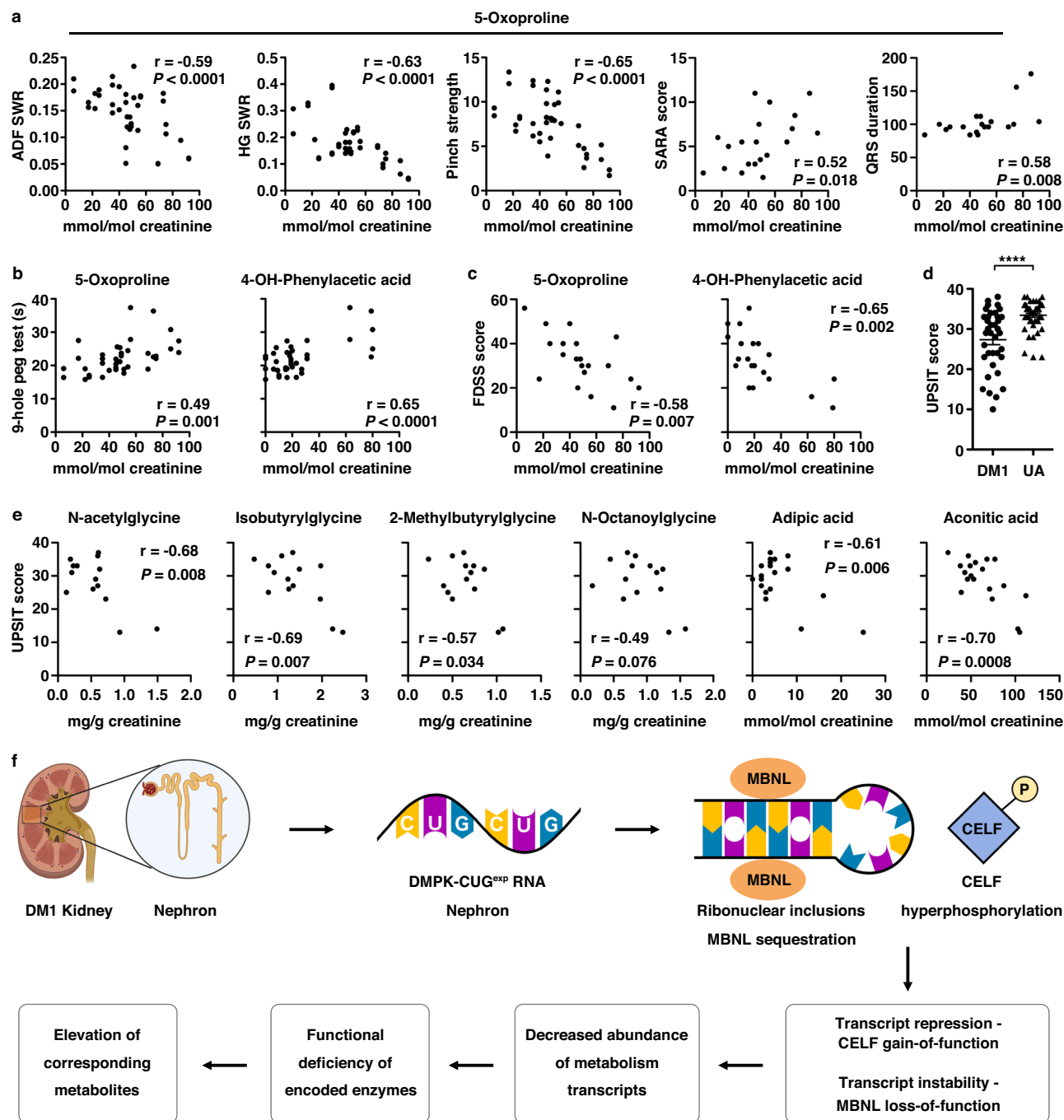


Fig. 9 | Correlation of urinary metabolites with functional measures.

a Correlation of urinary 5-Oxoproline levels with ankle dorsiflexion (ADF) strength-to-weight ratio (SWR), handgrip SWR, pinch (thumb and index finger) strength (pounds of force), scale for the assessment and rating of ataxia (SARA) score²⁶, and the electrocardiogram measurement QRS duration in DM1 individuals ($N = 20$; for strength measures, $N = 20$ left side and $N = 20$ right side for a total of 40 separate measures). **b** Correlation of urinary 5-oxoproline and 4-OH-phenylacetic acid levels with the time to complete the 9-hole peg test⁴⁴ (s) in DM1 individuals ($N = 20$ right hand and $N = 20$ left hand, for a total of 40 separate measures). **c** Correlation of urinary 5-oxoproline and 4-OH-phenylacetic acid levels with the Fatigue and Daytime Sleepiness (FDSS) score⁵² in DM1 ($N = 20$). The

correlation coefficient r and P values for each are shown (two-tailed t tests).

d University of Pennsylvania Smell Identification (UPSIT) scores⁵⁴ of DM1 ($N = 40$) and UA ($N = 43$) participants (maximum score = 40). **** $P < 0.0001$ (two-tailed t test). Error bars indicate \pm s.e.m. **e** Relationship of urine N-acetylglycine, isobutyrylglycine, 2-methylbutyrylglycine, N-octanoylglycine, adipic acid, and aconitic acid to UPSIT score in DM1 participants ($N = 14$ –19). The correlation coefficient r and P values for each are shown (two-tailed t tests). **f** Working model of DM1 molecular pathogenesis within the nephron leading to downregulation of metabolism genes and elevation of corresponding metabolites. The kidney and nephron diagrams were created in Biorender (Kumari; <https://BioRender.com/t33c158>)⁶¹. Source data are provided as a Source Data file.

recommendations to include monitoring of renal function and may impact eligibility for clinical trials.

The most common method to estimate GFR is the use of an equation based on serum creatinine, sex, and age¹¹. Creatinine is the

end product of muscle catabolism. A balance of its generation from muscle tissue and removal by the kidney is a major determinant of serum creatinine levels. In the setting of impaired kidney function, reduced renal clearance of creatinine causes an elevation of serum

creatinine, resulting in a lower estimated GFR. Due to reduced muscle mass in DM1, serum creatinine values can be low due to decreased production, resulting in an overestimation of renal clearance of creatinine and GFR. Cystatin C is an alternative means to estimate GFR that is less dependent on muscle mass than creatinine¹¹. The reliance on creatinine-based measure of GFR may have contributed to under-recognition of kidney involvement in DM1. Our findings support longitudinal study of the incidence, prevalence, and progression of impaired kidney function using cystatin C to estimate GFR in this population.

DM1 clinical features result from expression of *DMPK*-CUG^{exp} RNA in skeletal muscle, heart, cerebral cortex, smooth muscle (e.g., intestine, bladder), and kidney, among other tissues and cell types. To date, quantification of splice events or gene expression as measures of DM1 molecular pathogenesis has been limited to the examination of tissue biopsies, autopsy specimens, cultured cells, or model systems. Urinary EVs serve as liquid biopsies of kidney and bladder tissues, enabling the study of basic disease pathophysiology of these tissues in living individuals with DM1 for the first time. EV-based measurements are likely to be sensitive to changes in DM1 molecular disease activity that can indicate clinical progression or response to therapy. Multi-omics datasets combined with the use of advanced bioinformatics methods such as artificial intelligence (AI), machine learning (a form of AI), and predictive modeling already are being used to decipher molecular mechanisms of biology and disease^{55,56}. As AI becomes increasingly sophisticated, this practice figures to grow in the coming years, particularly for the study of human populations where mechanistic experimental studies are impractical, cost-prohibitive, and/or unethical. Given that vigorous interventional designs have confirmed the clinical consequences of only a single splicing change in DM1 over the past 20 years^{57,58}, the use of advanced bioinformatics methods is likely the only option to unravel the mechanisms by which hundreds of splice event and gene expression changes across multiple systems combine to produce the complex heterogeneous clinical features in this population. Our results support future EV multi-omics and advanced bioinformatics studies to expand our understanding of DM1 molecular pathogenesis, identify new therapeutic targets, and develop robust biomarkers.

CKD and DM1 each are associated with muscle wasting, insulin resistance, and impaired systemic metabolism of carbohydrates, proteins, and lipids^{2,59,60}. The degree of muscle wasting in CKD is related to the severity of renal dysfunction⁶⁰, suggesting the possibility that the reduced GFR observed in DM1¹³ may contribute to muscle wasting that these individuals experience. A recent study in mice found that (1) PCK1 is a key regulator of renal metabolism, (2) PCK1 expression is reduced in acute and chronic kidney disease, (3) metabolic alterations that arise from reduced PCK1 expression are important in the pathology of CKD progression, and (4) mitigation of PCK1 downregulation improved renal function, concluding that PCK1 is an attractive therapeutic target in kidney disease⁶¹. In separate studies, ASO-mediated reduction of PCK1 in rats induced insulin resistance⁶², while transgenic overexpression of PCK1 in mice improved muscle function, increased insulin sensitivity, and extended lifespan by nearly two years as compared to wild-type littermates^{63,64}. Our finding of *PCK1* downregulation in DM1 urinary EVs and kidney autopsy tissue supports further study of the relationship between *PCK1* expression, renal function, insulin resistance, and muscle wasting in DM1.

The development of biomarker tests for molecularly targeted therapies is considered the key to unlocking precision medicine⁶⁵. The FDA defines a monitoring biomarker as one that is measured repeatedly over time and can be used to assess, (1) disease progression, including the occurrence of new disease effects, worsening of previously existing abnormalities, or change in disease severity or specific abnormalities, or (2) response to a treatment, either favorable or unfavorable⁶⁶. The non-invasive nature of urinary EVs enables frequent

longitudinal sampling and the opportunity to detect a change in molecular disease status as an indicator of disease progression or drug response that may precede downstream changes in functional measures, an example of personalized medicine. In phase 2 clinical trials, EV biomarkers may facilitate patient stratification, improve understanding of how to use a drug, and perform dose finding, and in phase 3 trials guide the selection of dose or regimen.

An ideal property of a therapeutic drug for a multisystem disorder like DM1 is the efficient distribution to multiple affected tissues. To capture the full extent of drug response, DM1 may benefit from multiple biomarkers that report molecular pathogenesis in multiple systems. The main limitation of urinary EV biomarkers is that they indicate molecular pathogenesis restricted to the urinary tract. Muscle tissue biopsy biomarkers also have disadvantages, including the invasive nature that prevents frequent longitudinal monitoring and the lack of information for non-muscle tissues. However, based on the shared DM1 molecular pathogenesis across tissues, evidence of drug activity in one tissue type could be used as an indicator of drug activity in other tissues. In an earlier non-clinical study, we successfully demonstrated this concept. Systemic administration of a therapeutic antisense oligonucleotide (ASO) designed to induce RNase H cleavage of *Dmpk* RNA showed similar activity in mouse skeletal muscle, kidney, and bladder tissues¹⁹, suggesting that measurement of drug activity in either kidney or bladder tissue could predict drug activity in muscle tissue. The ASOs currently being tested in human DM1 clinical trials are expected to distribute widely to peripheral tissues after intravenous delivery, and none target skeletal muscle specifically. Study participants undergo multiple muscle tissue biopsies before and after treatment for quantification of transcriptomic biomarkers as indicators of drug response (clinicaltrials.gov IDs NCT06204809, NCT05481879, NCT06185764, NCT05027269). Since urinary EVs also carry splice event and gene expression biomarkers of DM1, they could be used as non-invasive liquid biopsies to report ASO drug activity in kidney and bladder tissues, and serve as predictors of therapeutic response in muscle tissue. The relationship between urinary EV and muscle tissue biomarkers could be determined by comparison of both biomarker sources from trial participants. If successful, urinary EVs could serve as companion diagnostics in DM1 therapeutic studies and post-marketing clinical applications to monitor response to treatment for the purpose of adjusting dose to achieve improved safety or effectiveness.

PCK1 is a cytosolic enzyme that catalyzes the rate-limiting step of gluconeogenesis and is a key contributor to glyceroneogenesis and TCA cycle cataplerosis⁶⁷. Electron transfer flavoprotein (ETF) is a dimeric enzyme in the mitochondrial matrix that mediates transfer of electrons from a series of flavoenzymes to the respiratory chain⁶⁸. Two separate genes, *ETFA* and *ETFB*, encode the ETF-alpha and ETF-beta subunits, respectively. Individuals with deficiency of PCK1 (OMIM 261680) or ETFB (OMIM 231680; multiple acyl CoA dehydrogenase deficiency [MADD]) may become symptomatic during times of increased metabolic demand, such as fasting, infections, febrile illnesses, prolonged exercise, medical procedures, or other physiologic stress, resulting in a characteristic urinary organic acid or acylglycine profile for each disorder^{37,42}. Elevation of urinary C6-C10 dicarboxylic acids, ethylmalonic acid, and glutaric acid can be seen with deficiency of either PCK1 or ETFB, and all of which we found elevated in individuals with DM1. The presence of elevated urinary TCA cycle intermediates together with elevated 2-hydroxyglutaric acid, dodecanedioic acid, and glycine conjugates of C4 and C5 acids suggests the co-existence of PCK1 and ETFB deficiency in DM1. The primary treatment strategy for PCK1 deficiency is to avoid fasting, while consumption of extra carbohydrate such as cornstarch or glucose polymer in the evening or after heavy exercise can be helpful³⁷. For ETFB deficiency, treatment consists of a high carbohydrate, low protein, and low fat diet, and the avoidance of fasting, with most individuals benefiting from supplementation with riboflavin and carnitine⁴². Our data

suggesting a functional deficiency of these enzymes in DM1 support further study on the effects of dietary modification as adjuvant therapy for disordered metabolism in these individuals.

PCK1 also plays a key role in serine biosynthesis⁶⁹, and plasma serine levels are correspondingly low in individuals with genetic PCK1 deficiency³⁷. HPD deficiency in Tyrosinemia type III causes elevated circulating tyrosine levels³⁹. The reduced expression of *PCK1* and *HPD* genes evident in DM1 urinary EVs suggests the possibility that functional deficiency of PCK1 and HPD enzymes may contribute to the unexplained low circulating levels of serine and elevated blood tyrosine levels previously observed in DM⁷⁰.

GSTA1 belongs to a family of glutathione S transferases (GSTs) that function to transfer glutathione to electrophilic regions of lipids as a means of neutralizing or detoxifying reactive oxygen species, drugs, anesthetic agents, intermediary metabolites, or environmental compounds. The reduced *GSTA1* expression evident in DM1 urinary EVs suggests the possibility that functional deficiency of the encoded *GSTA1* enzyme may contribute to the prolonged and heightened sensitivity to sedatives, opioids, and anesthetic agents that individuals with DM1 experience as compared to the general population³.

In our study, gene expression and *CELF1* alternative splicing in DM2 urinary EVs appeared more similar to DMD and UA than to DM1 samples, suggesting a less active involvement of DM2 molecular pathogenesis in urinary tract tissues. An important molecular difference between the two types of DM is that *CELF1* protein is hyperphosphorylated and accumulates in affected tissues of DM1⁷, but are non-elevated in DM2¹⁰. The shift in splicing of the *CELF1* transcript to reduce the LYQ isoform that we observed in DM1 urinary EVs may contribute to pathology by influencing the phosphorylation state and/or stability of the *CELF1* protein.

Both of the genes that we observed to be upregulated in DM1 urinary EVs, *KRT20* and *HPGD*, are enriched in bladder tissue, with *KRT20* limited to the umbrella cells of the bladder urothelium as a marker of differentiation⁷¹. The mechanism for the increased expression of these genes in DM1 urine exRNA is unknown, but may have to do with a response to the alteration of the metabolite profile in the urine.

Methods

Human research participants

The Mass General Brigham (MGB) IRB approved all studies involving human volunteers. The MGB IRB served as the primary IRB for Massachusetts General Hospital (MGH) and MGH Institute of Health Professions, and the reviewing IRB for Beth Israel Deaconess Medical Center (BIDMC), University of Texas Southwestern (UTSW), University of Pittsburgh (UP), Wake Forest University School of Medicine (WFSM), and Boston Children's Hospital (BCH) through the SMARTIRB online reliance system (<https://smartirb.org>). This cross-sectional discovery study is posted on clinicaltrials.gov website under identifiers NCT05016908, NCT05020002, and NCT05019625. Inclusion criteria for DM1 participants were age 5 years or older (14 years or older for muscle function testing), a diagnosis of DM1 based on genetic testing that identified a *DMPK*-CTG repeat expansion of ≥ 50 , or clinical diagnosis of DM1 and a first-degree relative with DM1 due to a *DMPK*-CTG repeat expansion of ≥ 50 . Only one DM1 participant has documented diabetes (type 2; treated with insulin) and no participants have documented chronic kidney disease. Inclusion criteria for DM2 were age 18 or older (rare or absent in children) and a *CNBP*-CCTG repeat expansion of > 75 . Inclusion criteria for participants with DMD were age 5 years or older and a sequence variant in the *DMD* gene consistent with clinical DMD. Inclusion criteria for unaffected (UA) healthy individuals were age 18 years or older and no known history of any muscular dystrophy. Prior to participation in the study, each participant provided informed consent or assent. Supplementary Table 1 outlines inclusion criteria and enrollment.

Biofluid collection and processing

Participants donated urine (approximate range from 40 to 120 milliliters) in a sterile specimen container (Covidien product number 2300SA). To maximize total exRNA recovery from each specimen, we processed the entire volume that was collected. The collection and processing of all samples was identical at each site. To remove urine cells, we centrifuged the entire volume at 2450×g for 10 min at room temperature, passed the supernatant through a 0.8-μm filter (Millipore product number SLAAR33SB) into sterile 50-ml tubes, and placed on wet ice or at 4 °C within 2 h of collection. Cell-free specimens were transferred to MGH unfrozen on cold packs either the same day by courier (BIDMC or BCH) or by overnight shipping (FedEx; UTSW, UP, or WFSM). After arrival at MGH, we proceeded with exRNA isolation from cell-free specimens either immediately or after storage at 4 °C overnight. To analyze total RNA from urine cells, we added TRIzol Reagent (Invitrogen product number 15596018) directly to the urine cell pellets after completion of the low-speed centrifugation step and proceeded according to the manufacturer's recommendations.

Retrieval of extracellular RNA and cDNA synthesis

We ultracentrifuged cell-free urine and serum samples at 100,000×g for 2 h at 4 °C, removed the supernatant, and extracted RNA from the translucent ribonucleoprotein pellet using 1 ml TRIzol Reagent according to the manufacturer's instructions. To enhance RNA pellet visibility, we added 10 μg linear acrylamide (Ambion product number AM9520) to each sample and mixed well prior to isopropanol precipitation. Pellets were re-suspended in molecular-grade water. We generated cDNA using Superscript III reverse transcriptase (Life Technologies product number 18080-044) and random primers (Invitrogen product number 48190-011) according to the manufacturer's instructions. To compare gene expression in EVs vs. that in non-EV-associated proteins and RNAs, we separated samples into equal volumes for parallel processing by (1) ultracentrifugation (UCF) or (2) centrifugal filtration (Centricon Plus-70; Millipore product number UFC703008) followed by size exclusion chromatography (SEC) (qEV Original/70 nm Gen2 columns; Izon Science product number ICO-70; qEV Concentration Kit; Izon Science product number RCT02) according to the manufacturer instructions. We extracted RNA in parallel from the UCF pellets, SEC EV fractions, and SEC protein fractions using TRIzol Reagent according to the manufacturer's instructions.

Library preparation, RNA sequencing, and differential expression analysis

We generated libraries from participants with DM1 and DMD ($N=4$ each group) using 10 ng urine exRNA from each sample and the Takara SMARTer Stranded Total RNA Sequencing kit V2–pico input mammalian (Takara product number 634412). Sequencing was performed in two separate batches using an Illumina NextSeq 500 platform at the George Washington University Genomics Core (<https://www.gwgenomics.org>) and a high-output 150-cycle kit to generate more than 83 million paired end reads per sample. Sequences were trimmed for adaptor sequence/low-quality sequence using *bbmap* (v. 35.85) with parameters *ktrim=r*, *k=21*, *mink=11*, *tbo*, *tpe*, *hdist=2*, *minlen=31*, *qtrim=r*, *trimq=10*, *maq=10*, *entropy=0.3*, *entropywindow=50*, and *entropyk=5*. The trimmed reads were aligned to the human GRCh38 genome and quantified using *Salmon* (v. 0.13.1)⁷². The specific mapping procedure involved aligning the reads using the *salmon* quant function with flags for sequence and GC bias corrections (*-seqBias -gcBias*), library type specification (*-l A*), and enhanced statistical modeling (*-numBootstraps=20*). Read count extraction and normalization also were performed using *Salmon* (v. 0.13.1) as part of the quantification process. The method integrates both alignment and quantification steps for accurate estimation of transcript abundance. Genes that have

the average expression of greater than 5 transcripts per million (TPM) in at least one group (DMI and/or DMD) were considered as expressed genes and included in the downstream analysis. One DMI and one DMD sample from batch 1 were excluded because they failed quality control and had too few reads. Differential expression analysis was performed using the edgeR (v.3.30.3) packages in R²⁰ with a *P* value < 0.05.

RT-PCR analysis of gene expression

We designed gene-specific primers using Primer3 software^{73,74} and amplified the target gene and reference gene *FAM168A* in the same reaction. Primer sequences are provided in Supplementary Table 3. We separated PCR products using agarose gels, stained with SYBR I green nucleic acid gel stain (Life Technologies product number S7567), and quantitated band intensities using a transilluminator, CCD camera, XcitaBlue™ conversion screen, and Image Lab image acquisition and analysis software (Image Lab v. 5.2.1; Bio-Rad), as previously described⁵⁸. Gene expression changes were calculated after normalization with *FAM168A* gene.

Droplet digital PCR (ddPCR) analysis of gene expression

To quantify gene expression in human control tissues, we purchased total RNA purified from UA human kidney tissue (Invitrogen product number AM7976), bladder tissue (Zyagen product number HR-902), skeletal muscle (Invitrogen product number AM7982), and brain (Invitrogen product number AM7962). We used Primer3 software^{73,74} to design primer-probe sets for target genes and the normalization control *FAM168A*. Custom primer sequences are provided in Supplementary Table 4. A commercially available primer-probe set for *GTF2B* (Applied Biosciences, FAM-MGB; assay ID Hs00976256_m1) was used as a second normalization control. We used ddPCR Supermix for probes (Bio-Rad product number 186-3010), an automated droplet generator and droplet reader (Bio-Rad QX200), and PCR cycling conditions according to the manufacturer's instructions as follows: enzyme activation 95 °C for 10 min (one cycle), denaturation 94 °C for 30 s followed by annealing/extension at 60 °C for 1 min (40 cycles), enzyme deactivation 98 °C for 10 min (one cycle), and hold at 4 °C, as previously described¹⁹. After PCR was complete, the plate was loaded into the droplet reader, processed/analyzed using QuantaSoft software (v. 1.7.4; Bio-Rad), and total events quantitated using the mean copy number per microliter of duplicate 20 µl assays from individual samples. We used the String database (v. 12.0) to search for known and predicted interactions between the protein products of differentially expressed genes²⁵.

Fluorescence in situ hybridization (FISH)

To identify CUG^{exp} ribonuclear inclusions, we used our modification of a previously published protocol^{75,76}. We fixed 8-µm muscle cryosections with 3% paraformaldehyde, pH 7.3 in 1× PBS, washed in 1× PBS, permeabilized nuclei using 0.5% Triton X-100 in 1× PBS for 5 min or 0.2% Triton X-100 in 1× PBS for 10 min at room temperature, and incubated in pre-hybridization solution (30% formamide/2× SSC) for 10 min at room temperature. Hybridization was for three hours at 37 °C in 33% formamide, 2× SSC, 0.2 mg/ml bovine serum albumin (NEB product number B9001S), 70 µg/ml yeast t-RNA (Invitrogen product number 15401-011), 2 mM ribonucleoside vanadyl complex (NEB product number S1402S), and 1 µg/ml 2'-O-methyl-modified RNA CAG repeat probe 5' labeled with Alexa647. Post-hybridization, we incubated sections in 30% formamide/2× SSC for 30 min at 42 °C followed by 1× SSC for 30 min at room temperature, and three washes in 1× PBS at room temperature before proceeding to labeling with primary antibody.

Hybridization probe. 5'-Alexa647-mCmAmGmCmAmGmCmAmGmCmAmGmCmAmGmCmAmGmCmA-3'

(HPLC-purified; “m” designates that RNA bases have 2'-O-methyl modifications; IDT).

Immunolabeling

To localize MBNL1 and MBNL2 proteins after FISH, we incubated tissue sections in anti-MBNL1 rabbit polyclonal antibody (2 µg/ml in PBS; Abcam product number ab45899) together with anti-MBNL2 (3B4) mouse monoclonal antibody (isotype IgG2b; 5 µg/ml PBS; Santa Cruz Biotechnology, Inc. product number sc-136167) overnight at 4 °C, as previously described⁵⁸. To localize ribonuclear inclusions within the nephron and podocytes, we performed FISH followed by co-labeling with anti-MBNL2 primary antibody, as above, and either anti-LRP2 (megalin; targeting the proximal tubule) rabbit polyclonal (1 µg/ml PBS; Abcam product number ab76969), anti-NKCC2 antibody (loop of Henle) rabbit polyclonal (5 µg/ml PBS; Invitrogen product number PA5-53263), anti-aquaporin 2 (collecting duct) rabbit monoclonal (0.3 µg/ml PBS; Abcam product number ab199975), or anti-podocalyxin (podocytes) goat polyclonal (5 µg/ml PBS; R&D Systems product number AF1658-SP) primary antibodies overnight at 4 °C. Secondary antibodies goat anti-mouse IgG2b Alexa 546 (Invitrogen product number A-21143) together with either goat anti-rabbit Alexa 488 (Invitrogen product number A-11034) or donkey anti-goat Alexa 488 polyclonal secondary antibody (Invitrogen product number A-32814) were applied at concentration 1 µg/ml PBS each for one hour at room temperature. To highlight nuclei, we added DAPI (33 ng/ml) together with the secondary antibody.

Quantitative fluorescence microscopy

To capture single images or z-series stacks, we used an AxioImager microscope (Zeiss), filters for DAPI (excitation/emission 365/445; Zeiss filter set 49), GFP (excitation/emission 470/525; Zeiss filter set 38), Cy3 (excitation/emission 550/605; Zeiss filter set 43 HE), and Cy5 (excitation/emission 640/690; Zeiss filter set 50), a Flash 4.0 LT sCMOS camera (Hamamatsu), and Volocity image acquisition software (version 6.3.1; Perkin Elmer). Objectives were 5× EC Plan-NEOFLUAR NA 0.16, 10× EC Plan-NEOFLUAR NA 0.3, 20× Plan-APOCHROMAT NA 0.8, 40× Plan-APOCHROMAT NA 1.4, and 63× Plan-APOCHROMAT NA 1.4. To quantitate fluorescence, we used Volocity quantitation and restoration software modules (v. 6.3.1; Perkin Elmer), as previously described⁷⁷.

Principle component analysis (PCA) and predictive modeling

To avoid co-linearity, we built a logistic regression model using the principal components from PCA as predictor variables. First, the samples were split randomly into the training set (76%) and validation set (24%), maintaining an equal proportion of each group in both sets. The ddPCR expression data was centered, scaled, and then used for PCA. A principal component analysis^{21,22} score for each urine exRNA sample was calculated using R statistical software³³ with a linear combination of four ddPCR gene expression values (*PCK1*, *HPD*, *DPYS*, and *GSTAI*). We used the same methods to combine ddPCR gene expression values of four additional transcripts (*PIGT*, *BINI*, *KRT20*, and *HPGD*) for an eight-transcript PCA. The PCA scores of the eight-transcript expression data were calculated using the *prcomp* function in the stats package (v. 3.6.2) in R. To train the logistic regression model, we used the caret package (v. 6.0-94) in R with the *glm* method and binomial family. Finally, the test group PCA scores were used for model evaluation. To test the performance of the model, we calculated the confusion matrix as the number of individuals correctly and incorrectly classified by group. The analyst testing the predictive model (ZL) was blinded to the disease status of the sample donors.

Analysis for CELF1 and MBNL1 RNA recognition motifs

We developed a chi-square test to estimate the enrichment of recognition motifs for RNA-binding proteins CELF1 and MBNL in the

3' untranslated region (UTR) of downregulated transcripts *PCK1*, *HPD*, *DPYS*, *GSTA1*, *ACY1*, *ETFB*, *GCDH*, *HSD17B14*, *PIGT*, and *BINI*, and in sequence adjacent to and including alternative exon 10 of the *CELF1* transcript. CELF1 motifs examined were TGTT, ATGT, TTGT, TGTC, GTGT, TGTA, GTTT, TGTG, GTCT, and TTTT (previously published⁸). MBNL motifs examined were GCTT, CGCT, TGCT, GCGC, CCGC, CTGC, GCTA, ACGC, CGCA, AGCT, TTGC, and CAGC (previously published⁸). To detect overrepresentation of a motif compared to what may be observed by chance, the chi-square test (R stats package `chisq.test` function, v. 3.6.2) was computed using the total number of ATGC motifs and N-mers that are not ATCG motifs in the 3' UTR (test region) vs. the same calculation in the non-test region of each gene.

Quantification of urine organic acids and acylglycines

For urine organic acid testing, we placed 6 ml aliquots of random urine samples at −20 °C without removal of cells or sediment and submitted to Quest Diagnostics for comprehensive organic acid quantification by gas chromatography/mass spectrometry (test code 35819). For urine acylglycine testing, we placed 8-ml aliquots of random urine samples at 4 °C overnight without removal of cells or sediment, and submitted to Mayo Clinic Laboratories for metabolite quantification by gas chromatography/mass spectrometry (test code AGU20).

Clinical assessments

A physical therapist or exercise physiologist evaluated 6-min walk test performance (total distance in meters and % predicted; a measure of walking endurance), muscle strength (bilateral ankle dorsiflexion, handgrip, and wrist extension) using quantitative hand-held dynamometers (microFET2 and microFET Handgrip; Hoggan Scientific, LLC), 9-hole peg test⁴⁴ (9HPT; a measure of fine motor dexterity), and the scale for the assessment and rating of ataxia (SARA; previously reported as useful in DM1)⁷⁸. To calculate the 6MWT% predicted, we used previously published reference equations for participants ages 18–50 years⁴⁵ or 51–80 years of age⁴⁶. The strength-to-weight ratio was calculated for each muscle group by dividing strength (pounds of force) by body weight (pounds)⁴⁷. Peak expiratory flow rate (PEFR; % predicted) and maximal inspiratory pressure (MIP; % predicted) were measured in a standard pulmonary function test laboratory at each performance site. Standard 12-lead electrocardiograms at each site measured QRS duration (milliseconds). Participants completed the Fatigue and Daytime Sleepiness Scale (FDSS), with scores determined by Rasch analysis³². To evaluate olfaction, we used the University of Pennsylvania Smell Identification Test (UPSIT; Sensonics International); individuals with a history of COVID-19 were excluded. We maintained study data using Research Electronic Data Capture (REDCap) software⁷⁹.

Statistics

For two-group and multigroup comparisons, we used unpaired two-tailed *t* tests or multigroup analysis of variance (ANOVA) (Prism v. 10.0 software, GraphPad, Inc.). We used the *F* test to compare variances between the two groups. For comparison of two groups that have a statistically significant difference in variance, we used *t* test with Welch's correction to determine the difference between groups. Group data are presented as mean ± s.e.m. To determine associations of gene expression with clinical parameters or urine metabolites, we used Pearson correlation coefficients. A *P* value < 0.05 was considered significant.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The RNA sequencing raw data generated in this study have been deposited in the Gene Expression Omnibus (GEO) database under the accession code GSE284252. The analyzed RNA sequencing data are available as source data with this paper. Source data are provided with this paper.

Code availability

All codes generated in this study are accessible at https://github.com/mindds/DM1_EV/.

References

- Turner, C. & Hilton-Jones, D. Myotonic dystrophy: diagnosis, management and new therapies. *Curr. Opin. Neurol.* **27**, 599–606 (2014).
- Krentz, A. J., Williams, A. C. & Nattrass, M. Insulin resistance in multiple aspects of intermediary metabolism in myotonic dystrophy. *Metabolism* **40**, 866–872 (1991).
- Ashizawa, T. et al. Consensus-based care recommendations for adults with myotonic dystrophy type 1. *Neurol. Clin. Pr.* **8**, 507–520 (2018).
- Brook, J. D. et al. Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **68**, 799–808 (1992).
- Overend, G. et al. Allele length of the DMPK CTG repeat is a predictor of progressive myotonic dystrophy type 1 phenotypes. *Hum. Mol. Genet.* **28**, 2245–2254 (2019).
- Miller, J. W. et al. Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J.* **19**, 4439–4448 (2000).
- Kuyumcu-Martinez, N. M., Wang, G. S. & Cooper, T. A. Increased steady-state levels of CUGBP1 in myotonic dystrophy 1 are due to PKC-mediated hyperphosphorylation. *Mol. Cell* **28**, 68–78 (2007).
- Wang, E. T. et al. Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. *Genome Res.* **25**, 858–871 (2015).
- Liquori, C. L. et al. Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* **293**, 864–867 (2001).
- Lin, X. et al. Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Hum. Mol. Genet.* **15**, 2087–2097 (2006).
- Chapter 1. Definition and classification of CKD. *Kidney Int. Suppl.* **3**, 19–62 (2013).
- Matsumura, T. et al. Renal dysfunction can be a common complication in patients with myotonic dystrophy 1. *J. Neurol. Sci.* **368**, 266–271 (2016).
- Aldenbratt, A., Lindberg, C. & Svensson, M. K. Reduced renal function in patients with myotonic dystrophy type 1 and the association to CTG expansion and other potential risk factors for chronic kidney disease. *Neuromuscul. Disord.* **27**, 1038–1042 (2017).
- Tkach, M. & Thery, C. Communication by extracellular vesicles: where we are and where we need to go. *Cell* **164**, 1226–1232 (2016).
- Erdbrugger, U. & Le, T. H. Extracellular vesicles in renal diseases: more than novel biomarkers? *J. Am. Soc. Nephrol.* **27**, 12–26 (2016).
- Kalluri, R. & McAndrews, K. M. The role of extracellular vesicles in cancer. *Cell* **186**, 1610–1626 (2023).
- Svenningsen, P., Sabaratnam, R. & Jensen, B. L. Urinary extracellular vesicles: origin, role as intercellular messengers and biomarkers; efficient sorting and potential treatment options. *Acta Physiol.* **228**, e13346 (2020).
- Pisitkun, T., Shen, R. F. & Knepper, M. A. Identification and proteomic profiling of exosomes in human urine. *Proc. Natl. Acad. Sci. USA* **101**, 13368–13373 (2004).

19. Antoury, L. et al. Analysis of extracellular mRNA in human urine reveals splice variant biomarkers of muscular dystrophies. *Nat. Commun.* **9**, 3906 (2018).
20. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
21. Raychaudhuri, S., Stuart, J. M. & Altman, R. B. Principal components analysis to summarize microarray experiments: application to sporulation time series. *Pac. Symp. Biocomput.* 455–466 https://doi.org/10.1142/9789814447331_0043 (2000).
22. Ringner, M. What is principal component analysis? *Nat. Biotechnol.* **26**, 303–304 (2008).
23. Mootha, V. K. et al. PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
24. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
25. Szklarczyk, D. et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* **51**, D638–D646 (2023).
26. Mathieu, J. et al. Assessment of a disease-specific muscular impairment rating scale in myotonic dystrophy. *Neurology* **56**, 336–340 (2001).
27. Lee, J. W., Chou, C. L. & Knepper, M. A. Deep sequencing in microdissected renal tubules identifies nephron segment-specific transcriptomes. *J. Am. Soc. Nephrol.* **26**, 2669–2677 (2015).
28. Limbutara, K., Chou, C. L. & Knepper, M. A. Quantitative proteomics of all 14 renal tubule segments in rat. *J. Am. Soc. Nephrol.* **31**, 1255–1266 (2020).
29. Dorrenhaus, A. et al. Cultures of exfoliated epithelial cells from different locations of the human urinary tract and the renal tubular system. *Arch. Toxicol.* **74**, 618–626 (2000).
30. Chhoy, P., Brown, C. W., Amante, J. J. & Mercurio, A. M. Protocol for the separation of extracellular vesicles by ultracentrifugation from in vitro cell culture models. *STAR Protoc.* **2**, 100303 (2021).
31. Jiang, H. et al. Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Hum. Mol. Genet.* **13**, 3079–3088 (2004).
32. Otero, B. A. et al. Transcriptome alterations in myotonic dystrophy frontal cortex. *Cell Rep.* **34**, 108634 (2021).
33. Mevik, B.-H. & Wehrens, R. The pls Package: principal component and partial least squares regression in R. *J. Stat. Softw.* **18**, 1–24 (2007).
34. Takahashi, N., Sasagawa, N., Suzuki, K. & Ishiura, S. The CUG-binding protein binds specifically to UG dinucleotide repeats in a yeast three-hybrid system. *Biochem. Biophys. Res. Commun.* **277**, 518–523 (2000).
35. Koshelev, M. et al. Heart-specific overexpression of CUGBP1 reproduces functional and molecular abnormalities of myotonic dystrophy type 1. *Hum. Mol. Genet.* **19**, 1066–1075 (2010).
36. Driesen, K. & Witters, P. Understanding inborn errors of metabolism through metabolomics. *Metabolites* **12**, 398 (2022).
37. Vieira, P. et al. Cytosolic phosphoenolpyruvate carboxykinase deficiency: expanding the clinical phenotype and novel laboratory findings. *J. Inher. Metab. Dis.* **45**, 223–234 (2022).
38. Goetz, M. et al. Genotypic and phenotypic spectrum of cytosolic phosphoenolpyruvate carboxykinase deficiency. *Mol. Genet. Metab.* **137**, 18–25 (2022).
39. Tomoeda, K. et al. Mutations in the 4-hydroxyphenylpyruvic acid dioxygenase gene are responsible for tyrosinemia type III and hawkinsinuria. *Mol. Genet. Metab.* **71**, 506–510 (2000).
40. van Kuilenburg, A. B. et al. Dihydropyrimidinase deficiency: phenotype, genotype and structural consequences in 17 patients. *Biochim. Biophys. Acta* **1802**, 639–648 (2010).
41. Sass, J. O. et al. Mutations in ACY1, the gene encoding aminoacylase 1, cause a novel inborn error of metabolism. *Am. J. Hum. Genet.* **78**, 401–409 (2006).
42. Grunert, S. C. Clinical and genetical heterogeneity of late-onset multiple acyl-coenzyme A dehydrogenase deficiency. *Orphanet J. Rare Dis.* **9**, 117 (2014).
43. Kumps, A., Duez, P. & Mardens, Y. Metabolic, nutritional, iatrogenic, and artifactual sources of urinary organic acids: a comprehensive table. *Clin. Chem.* **48**, 708–717 (2002).
44. Gagnon, C. et al. Report of the second outcome measures in myotonic dystrophy type 1 (OMMYD-2) international workshop San Sebastian, Spain, October 16, 2013. *Neuromuscul. Disord.* **25**, 603–616 (2015).
45. Halliday, S. J. et al. Six-minute walk distance in healthy young adults. *Respir. Med.* **165**, 105933 (2020).
46. Enright, P. L. & Sherrill, D. L. Reference equations for the six-minute walk in healthy adults. *Am. J. Respir. Crit. Care Med.* **158**, 1384–1387 (1998).
47. Hiraki, K. et al. Decreased physical function in pre-dialysis patients with chronic kidney disease. *Clin. Exp. Nephrol.* **17**, 225–231 (2013).
48. Mathieu, J. et al. A 10-year study of mortality in a cohort of patients with myotonic dystrophy. *Neurology* **52**, 1658–1662 (1999).
49. Kera, T. et al. Definition of respiratory sarcopenia with peak expiratory flow rate. *J. Am. Med. Dir. Assoc.* **20**, 1021–1025 (2019).
50. Heatwole, C. et al. Patient-reported impact of symptoms in myotonic dystrophy type 1 (PRISM-1). *Neurology* **79**, 348–357 (2012).
51. Hadrevi, J. et al. Plasma metabolomic patterns in patients with exhaustion disorder. *Stress* **22**, 17–26 (2019).
52. Hermans, M. C. et al. Fatigue and daytime sleepiness scale in myotonic dystrophy type 1. *Muscle Nerve* **47**, 89–95 (2013).
53. Masaoka, Y. et al. Impairment of odor recognition and odor-induced emotions in type 1 myotonic dystrophy. *Neurosci. Lett.* **503**, 163–166 (2011).
54. Doty, R. L., Shaman, P. & Dann, M. Development of the University of Pennsylvania Smell Identification Test: a standardized micro-encapsulated test of olfactory function. *Physiol. Behav.* **32**, 489–502 (1984).
55. Ostrowski, J. & Wyrwicz, L. S. Integrating genomics, proteomics and bioinformatics in translational studies of molecular medicine. *Expert Rev. Mol. Diagn.* **9**, 623–630 (2009).
56. Yue, R. & Dutta, A. Computational systems biology in disease modeling and control, review and perspectives. *NPJ Syst. Biol. Appl.* **8**, 37 (2022).
57. Wheeler, T. M. et al. Correction of CLC-1 splicing eliminates chloride channelopathy and myotonia in mouse models of myotonic dystrophy. *J. Clin. Invest.* **117**, 3952–3957 (2007).
58. Hu, N., Kim, E., Antoury, L. & Wheeler, T. M. Correction of Clcn1 alternative splicing reverses muscle fiber type transition in mice with myotonic dystrophy. *Nat. Commun.* **14**, 1956 (2023).
59. Wang, X. H., Mitch, W. E. & Price, S. R. Pathophysiological mechanisms leading to muscle loss in chronic kidney disease. *Nat. Rev. Nephrol.* **18**, 138–152 (2022).
60. Heitman, K., Alexander, M. S. & Faul, C. Skeletal muscle injury in chronic kidney disease—from histologic changes to molecular mechanisms and to novel therapies. *Int. J. Mol. Sci.* **25**, 5117 (2024).
61. Verissimo, T. et al. PCK1 is a key regulator of metabolic and mitochondrial functions in renal tubular cells. *Am. J. Physiol. Ren. Physiol.* **324**, F532–F543 (2023).
62. Beddow, S. A. et al. PEPCK1 antisense oligonucleotide prevents adiposity and impairs hepatic glycogen synthesis in high-fat male fed rats. *Endocrinology* **160**, 205–219 (2019).

63. Hakimi, P. et al. Overexpression of the cytosolic form of phosphoenolpyruvate carboxykinase (GTP) in skeletal muscle repatterns energy metabolism in the mouse. *J. Biol. Chem.* **282**, 32844–32855 (2007).
64. Hanson, R. W. & Hakimi, P. Born to run; the story of the PEPCK-Cmus mouse. *Biochimie* **90**, 838–842 (2008).
65. Lyman, G. H. & Moses, H. L. Biomarker tests for molecularly targeted therapies—the key to unlocking precision medicine. *New Engl. J. Med.* **375**, 4–6 (2016).
66. BEST (Biomarkers, EndpointS, and other Tools) Resource (FDA-NIH Biomarker Working Group, Silver Spring (MD), 2016). <https://www.ncbi.nlm.nih.gov/books/NBK326791/>
67. Hakimi, P. et al. Phosphoenolpyruvate carboxykinase and the critical role of cataplerosis in the control of hepatic metabolism. *Nutr. Metab.* **2**, 33 (2005).
68. Henriques, B. J., Katrine Jentoft Olsen, R., Gomes, C. M. & Bross, P. Electron transfer flavoprotein and its role in mitochondrial energy metabolism in health and disease. *Gene* **776**, 145407 (2021).
69. Kalhan, S. C. & Hanson, R. W. Resurgence of serine: an often neglected but indispensable amino acid. *J. Biol. Chem.* **287**, 19786–19791 (2012).
70. Moxley, R. T. et al. Insulin resistance and regulation of serum amino acid levels in myotonic dystrophy. *Clin. Sci.* **71**, 429–436 (1986).
71. Jafari, N. V. & Rohn, J. L. The urothelium: a multi-faceted barrier against a harsh environment. *Mucosal Immunol.* **15**, 1127–1142 (2022).
72. Patro, R. et al. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
73. Koressaar, T. & Remm, M. Enhancements and modifications of primer design program Primer3. *Bioinformatics* **23**, 1289–1291 (2007).
74. Untergasser, A. et al. Primer3-new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115 (2012).
75. Mankodi, A. et al. Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Hum. Mol. Genet.* **10**, 2165–2170 (2001).
76. Hu, N. et al. Antisense oligonucleotide and adjuvant exercise therapy reverse fatigue in old mice with myotonic dystrophy. *Mol. Ther. Nucleic Acids* **23**, 393–405 (2021).
77. Hu, N. et al. Non-invasive monitoring of alternative splicing outcomes to identify candidate therapies for myotonic dystrophy type 1. *Nat. Commun.* **9**, 5227 (2018).
78. DiPaolo, G. et al. Functional impairment in patients with myotonic dystrophy type 1 can be assessed by an ataxia rating scale (SARA). *J. Neurol.* **264**, 701–708 (2017).
79. Harris, P. A. et al. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inf.* **42**, 377–381 (2009).
80. Kumari, P. <https://BioRender.com/s97f089> (2025).
81. Kumari, P. <https://BioRender.com/t33c158> (2025).

Acknowledgements

The authors thank all of the volunteers who participated in this study. We also thank the Elaine and Richard Slye Fund (T.M.W.), the Department of Defense (award IDs W81XWH-20-1-0293 and W81XWH-19-1-0392; T.M.W.), National Institutes of Health (award ID R61NS117210; T.M.W.), the

Myotonic Dystrophy Foundation (Fellowship Award; P.K.), and Muscular Dystrophy Association (Development Grant ID 1072634; P.K.) for support, and the University of Rochester Wellstone Muscular Dystrophy Specialized Research Center Shared Scientific Resource Core (P50NS048843; Dr. Charles Thornton) for providing the DM1 kidney autopsy tissue.

Author contributions

P.K., L.M.S., Z.L., E.P.C., E.C., R.J., N.H., J.A.S., B.B.M., R.R.K., S.D., and T.M.W. performed experiments and analyzed data. P.G.P., C.L., K.O., H.G., and S.S. acquired data. E.L.T., B.T.D., S.B.R., S.T.I., P.R.C., and A.P. acquired data and supervised activities at their respective sites. P.K. and T.M.W. wrote the manuscript. All authors reviewed and approved the manuscript.

Competing interests

M.G.H. and T.M.W. have been awarded a patent (U.S. patent number 11,866,782 B2; International patent application number PCT/2017/043348) for the use of extracellular RNA to identify markers of muscular dystrophies. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-56479-5>.

Correspondence and requests for materials should be addressed to Sudeshna Das or Thurman M. Wheeler.

Peer review information *Nature Communications* thanks anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2025