TITLE PAGE

Manuscript Title

Identification of novel plasma proteomic biomarkers of Dupuytren Disease

Authors

Blake Hummer [1]

Paola Sebastiani [2,3] ORCID 0000-0001-6419-1545

Anastasia Leshchyk [4] ORCID 0000-0002-2830-6701

Anastasia Gurinovich [2,3] ORCID 0000-0001-9808-9397

Cecilie Bager [5] ORCID 0000-0003-2430-9777

Morten Karsdal [5] ORCID 0000-0001-5026-8740

Signe Holm Nielsen [5] ORCID 0000-0002-4612-1014

Charles Eaton* [1] ORCID 0000-0002-8186-3611

Affiliations

1. Dupuytren Research Group, West Palm Beach, FL
2. Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA
3. Department of Medicine, Tufts University, Boston, MA
4. Department of Medicine, Computational Biomedicine Section, Chobanian & Avedisian School of Medicine, Boston University, Boston, MA
5. Nordic Bioscience, Herlev, Denmark

* Corresponding Author Charles Eaton MD, Dupuytren Research Group 1850 Forest Hill Blvd STE 201, West Palm Beach, FL 33406 1.561.429.8279 c.eaton.md@dupuytrens.org

Conflict of Interest

The authors have declared that no conflict of interest exists.

Identification of novel plasma proteomic biomarkers of Dupuytren Disease

ABSTRACT

Dupuytren Disease (DD) is a chronic progressive disease that can result in disabling hand deformities. The most common treatments have high rates of complications and early recurrence. Dupuytren lacks a staging biomarker profile to develop preventive therapeutics to improve long-term outcomes. This multi-omic study aimed to create a DD blood proteomic biomarker profile by comparing DD plasma to a healthy control group. We measured circulating collagen metabolism peptides and found normal Collagen I synthesis but impaired Collagen I degradation in DD. We measured 6995 serum proteins and identified 68 with statistically significant differences from the control group. We developed two Diagnostic Proteomic Risk Scores (DPRS) based on hypothesis-free and hypothesis-based analyses. In independent data, our hypothesis-free and the hypothesis-based DPRS distinguished Dupuytren from control subjects with 76.5% and 70.6% accuracy, respectively. Our hypothesis-based DPRS also distinguished DD subjects with different disease progression rates based on subject age at the time of their first corrective procedure (p=0.0018). This pilot study is the first to provide evidence that Collagen I accumulation in DD is due to impaired degradation rather than increased collagen synthesis. It also describes novel DPRS that have potential as diagnostic and staging biomarker panels for Dupuytren disease.

Identification of novel plasma proteomic biomarkers of Dupuytren Disease

MAIN TEXT

Introduction

Dupuytren disease (DD) is the most common of the heritable fibrotic disorders, affecting an estimated 10 million Americans (1). DD is a clinical diagnosis made by examination. Dupuytren Contracture (DC) refers to progressive DD-related deformities that prevent the fingers from straightening. DC has social, functional, and health impacts. DC is associated with loss of productivity, higher healthcare costs, and a greater prevalence of benign diseases (2).

The greatest need in DD care is biological treatments to prevent contracture progression and recurrence. The authors have not found any published reports of nonsurgical treatments with adequately powered statistically significant evidence of preventing DC recurrence. As with other fibrotic diseases, developing effective treatments requires a repeatable biomarker profile to stage disease activity and measure therapeutic response. The authors have not found any reports of such a panel for Dupuytren disease. This pilot study aimed to develop preliminary data on a diagnostic blood biomarker profile of DD.

DC is a common outcome of DD. Of those diagnosed with DD at age 65, one in four men and one in eight women will require at least one corrective procedure over their lifetime (3). Current treatments are reparative, not preventive. Corrective procedures often result in partial, temporary improvement (4). The risk of complications from DD surgery is greater than other commonly performed elective hand operations (5-8). Recurrent DC is the most common diagnosis leading to elective finger amputation (9). There is a 3% risk of finger amputation with primary fasciectomy (the most common surgery) (10, 11) and an 8 to 12% amputation rate as a result of repeat fasciectomy for recurrence (8, 12).

DD is associated with European (13) and Neanderthal ancestry (14). Inheritance follows a pattern of autosomal dominance with wide phenotypic variation. Each child of a parent with DD

has a 50% chance of developing the clinical disease (15). Genetic factors account for about 80% of the risk of developing DD (16). Family history, age, and gender are the strongest predictors of developing disease. Local trauma, diabetes, smoking, alcoholism, heavy manual labor, and BMI each have minor effects on incidence. A genetic risk score (GRS) based on 11 SNPs correlates with DD diagnosis (17), and a weighted GRS based on 26 SNPs correlates with post-procedure recurrence (18).

DD affects the palm skin, subcutaneous tissues, and fascia (19). In early disease, fibroblasts, myofibroblasts, and macrophages accumulate, with perivascular clusters of myofibroblasts, pericytes, and mononuclear immune cells (20). Involved tissues show increased expression of immune (20), hematopoietic, and mesenchymal cell markers (19), changes in cytokines and signaling factors (21), and extracellular matrix components (22, 23). If early DD progresses to contracture, the local histology progresses to nearly acellular densely packed collagen.

Physical examination is the basis of DD staging and treatment. To date, there are no noninvasive tests to stage DD biological activity. Imaging studies are unhelpful for staging. A diagnostic biopsy can trigger disease activity and progression, DD-specific tissue findings do not always correlate with blood measurements (24-26), and local DD-specific gene expression does not reliably correlate with local protein levels (27). Physical examination does not reflect biological activity (28) nor predict clinical progression (29). Clinical outcome benchmarks are of limited value because Dupuytren lacks an unambiguous definition of recurrence. Therefore, blood remains the most promising source for a disease activity profile for drug development.

We conducted a discovery study to compare levels of circulating proteins and peptides in subjects with DD versus controls using ELISA, aptamer, and mass spectrometry methods. We found statistically significant differences in disease versus control cohorts in independent analyses of peptides and proteins. We also found statistically significant differences in protein aptamers in DD subjects with different rates of clinical progression of untreated disease.

RESULTS

Assessment of Collagen Metabolism in DD.

We evaluated DD collagen synthesis and degradation markers using the competitive ELISA method (Nordic Biosciences; Herlev, Denmark). We measured plasma collagen metabolism markers and calculated these ratios (30). Collagen synthesis and degradation processes release protein fragments specific to the process, collagen type, and enzyme involved. The ratio of synthesis to degradation-related fragments correlates with net accumulation or loss of that collagen type.

We measured markers of types I, II, III, IV, V, VI, and IX collagen metabolism in blood from a cohort of healthy controls (n = 14) and subjects with DD (n = 13). We found significantly low levels of markers of Collagen I degradation C1M (Figure 1A, $p = 0.004$), Collagen V synthesis Pro-C5 (Figure 1B, $p = 0.046$), and Collagen VII synthesis Pro-C7 (Figure 1C, $p = 0.048$) in the DD cohort compared to the control group. The remaining collagen biomarkers we measured did not show significant differences between the disease and control groups (Table 1).

We used collagen synthesis to degradation ratios to calculate collagen turnover rates in DD subjects and healthy controls. We calculated these ratios for collagen types I, II, III, IV, and V. The Collagen I synthesis to degradation ratio Pro-C1/C1M was significantly greater in the DD cohort compared to the control group (Figure 1D, $p = 0.024$). The remaining ratios either were not significant (Col II, III, IV, VI) or could not be calculated because the degradation marker was either unavailable (Col V, IX) or was not detectible in both disease and control cohorts (Col VII).

Assessment of Differentially Expressed Blood Proteins in DD.

Identification of novel plasma proteomic biomarkers of Dupuytren Disease

We performed two orthogonal unbiased proteomic methods, the aptamer-based SomaScan (SomaLogic; Boulder, CO) and liquid chromatography-tandem mass spectrometry (LC-MS/MS) (MacCoss Lab, University of Washington, WA, USA). The SomaScan is a highly multiplexed proteomic assay that leverages SOMAmers (slow off-rate modified aptamers). SOMAmers are short single-stranded DNA sequences that selectively bind proteins in their native conformation, with affinities comparable with antibodies. LC-MS/MS is a highly sensitive technique that combines liquid chromatography to separate molecules, followed by tandem mass spectrometry, which can detect and quantify clinically relevant analytes.

We used sixty-nine samples for SomaScan testing after quality control (36 DD and 33 controls). These represent 42 subjects, 27 of whom had second samples obtained six to fourteen months after the first (Figure 2). The SomaLogic panel measures 6995 protein binding sites applying to 5444 proteins.

Hypothesis-free approach

The 6995 aptamer measurements identified 54 proteins with statistically significant differences between DD and control groups, with 24 overexpressed and 30 underexpressed after adjusting for multiple comparisons (Suppl Table 1). Figure 3 is a heatmap of these 54 protein expressions.

After identifying the 54 differentially expressed (DE) plasma proteins, we performed pathway analysis to identify network relationships using STRING (31). We found significant protein-protein interactions (PPI) in 29, with multiple enriched categories (Figure 4C, p-value = 4.55E-04). Suppl Table 2 summarizes categories of enriched gene groups, and Figure 5 provides a visual cross-reference of functional categories and individual genes. The dominant categories identified involve coagulation and complement cascades, the extracellular matrix, enzyme inhibition, cell membrane structures, phosphorylation, and signaling pathways.

Hypothesis-free Diagnostic Proteomic Risk Score (DPRS)

Using the SomaScan results, we aimed to develop a hypothesis-free indicator of DD diagnosis with a calculated Diagnostic Proteomic Risk Score (DPRS) similar to the widely used polygenic risk score model (32). The DPRS is a single score used to estimate an individual's disease risk based on plasma protein abundance.

We developed a hypothesis-free DPRS estimation based on two of the 54 proteins. These proteins, OSCAR (overexpressed) and SPART (underexpressed), had a low relative correlation (Pearson's correlation coefficient = -0.19) (Table 2), suggesting minimal interaction. Using this hypothesis-free DPRS as a diagnostic measure for DD achieved an accuracy of 79% in the training set (AUC 0.8) and 76.5% in the test set (Figure 4A), indicating a predictive value for discriminating between disease and control. Figure 4B shows a boxplot of control and DD DPRS. The DPRS calculated for controls and DD significantly differed by the Wilcoxon test (p-value = 2.5E-06).

Hypothesis-based approach

Because of the statistical limitations in analyzing 6995 candidate biomarkers, we also conducted a hypothesis-based approach to analyze SomaScan data with selected candidate biomarkers to decrease the sample size. We created a list of likely DD-related proteins and their parent genes from a search of 2547 full-text DD-related publications published between 1980 and 2023. This literature search identified 328 proteins available on the SomaLogic panel of 6995 protein binding sites (Suppl Table 3). On evaluating this 328 protein subset, 23 had statistically significant differences between DD and control, with nine overexpressed and fourteen underexpressed after adjusting for multiple comparisons (Suppl Table 4).

After identifying 23 DE plasma proteins, we performed PPI networks and enrichment analysis using STRING. We found significantly enriched PPI in 21 proteins with multiple enriched

categories (Figure 6C, p-value = 1.35e-07). We found more enrichment categories, more category overlap, and greater significance in the hypothesis-based than the hypothesis-free analysis. These findings may reflect a higher-yield group of candidate markers, selection bias, or both. Suppl Table 5 lists categories of enriched gene groups and Figure 7 provides a visual cross-reference of functional categories and individual genes.

Hypothesis-based Diagnostic Proteomic Risk Score

Of the 23 DE proteins identified, we applied the DPRS computation procedures, identifying 11 proteins with statistically significant differences from controls. We used these proteins for DPRS estimation (Table 3). Of these, the genes of six were underexpressed (*SMAD1, SH3BP2, LCN2, USP8, CSNK1G2, SERPINH1*) and five overexpressed (*CSMD1, POSTN, ACAN, AOC3, DDR2*). Six proteins in the hypothesis-based DPRS were also in the group of 54 proteins with significant differential expression in the hypothesis-free analysis (Table 3). Employing the hypothesis-driven DPRS generation, our predictive model for DD achieved a training accuracy of 78.85% (AUC 0.9) and a test accuracy of 70.59% (Figure 6A). Boxplots in Figure 6B show the DPRS distributions of Control and DD groups differ significantly, confirmed by the Wilcoxon test (p-value = 2.2E-09).

Evaluation of DRPS for Prognosis

We explored the potential of each DPRS to profile DD biological severity and prognosis. We segregated our DD samples by whether they were younger than 50 (11 specimens) or 50 and older (16 specimens) at the time of their first DD. To avoid selection bias in this analysis, we excluded subjects younger than 50 from this analysis. A one-way ANOVA analysis unveiled a statistically significant variance in the means of hypothesis-driven generated DPRS across groups with different rates of pretreatment contracture progression (Figure 8, p-value = 0.0018).

Conversely, we found no associations in the DPRS generated through the hypothesis-free approach.

Mass Spectrometry

The LC-MS dataset contained 48 samples: 27 DD and 21 controls. This assay identified 561 peptides. None had statistically significant differences in abundance across groups (FDR > 0.25).

Mass-Spectrometry-SomaScan Concordance

To further validate our findings from the SomaScan, we evaluated the MS results for concordance. Of the 54 significant proteins from the SomaScan data analysis, six were present in the Mass Spectrometry findings: AOC3, POSTN, SERPINC1, TF, PCSK9, and KNG1. We ran the same analyses on these six proteins using the Mass Spectrometry data. AOC3, POSTN, and SERPINC1 had adjusted p-values (FDR) <0.25 in both Mass Spec and SomaLogic analyses (Table 4).

DISCUSSION

The key objective of this study is to discover candidate biomarkers that correlate with the diagnosis of DD as the first step toward a minimally invasive diagnostic DD blood test.

Dynamic biomarker panels are essential to developing and testing preventive therapeutics for chronic progressive diseases like DD. The relevance of this discovery study is that DD currently lacks such biomarkers, and DD biomarker sources are limited. Blood proteomic analysis is the most promising and unexplored source of biomarkers suitable for biological trend analysis.

Identification of novel plasma proteomic biomarkers of Dupuytren Disease

Existing knowledge of DD biology provides limited proteomic guidance. There are no broad surveys of circulating proteins in DD. Proteomic findings may not parallel genomic findings because of genetic paralogues, epigenetic modulation, transcriptomic artifacts of cell culture, and other factors. Blood findings may not parallel tissue findings because of rapid degradation, intracellular or membrane-bound locations, post-translational variants, protein binding, and other mechanisms. Because of this, new and diverse discovery research is necessary to profile DD-associated blood protein biomarkers.

There are two common biomarker discovery strategies: hypothesis-free and hypothesis-based. The hypothesis-free approach (unbiased or non-hypothesis-based) screens many markers and is the standard in systems biology and discovery research. The hypothesis-based strategy evaluates selected markers based on published information and is the most common model of DD research. We used both methods in this study.

Collagen accumulation is the hallmark of DD. DD-affected tissues have unique proportions of collagen types I (33), III (34), IV (23), and V (33). The most significant changes from preclinical to end-stage disease are an increase in total Collagen I and an increase in the proportion of Collagen III (33).

Our hypothesis-based approach used ELISA to compare collagen metabolism markers in DD subjects with those in the control group. Collagen turnover releases circulating peptides unique to the synthesis and degradation of each collagen type. In the absence of disease, synthesis and degradation are tightly coupled. Plasma synthesis to degradation marker ratios can indicate net collagen gain, loss, or homeostasis (30). We measured plasma levels of sixteen collagen metabolism markers involving eight collagen types.

Our findings are summarized in Table 1 and discussed below.

The DD cohort had normal Collagen I synthesis marker Pro-C1 levels but significantly decreased Collagen I degradation marker C1M compared to controls (Figure 1A), resulting in an elevated synthesis-to-degradation marker ratio of DD compared to controls (Figure 1D). These findings suggest collagen accumulation in DD results from impaired Collagen I degradation rather than excess collagen synthesis.

DD-related Collagen I fibrils have two differences from normal collagen, increasing their resistance to enzymatic degradation. The first difference is homotrimer Collagen I. Typical Collagen I fibrils are heterotrimers of two collagen I α1 (*COL1A1*) strands and one collagen I α2 (*COL1A2*) strand. DD-affected tissues have an increased percentage of α1 (I) homotrimer Collagen I (35). The second is that DD collagen has a higher percentage of crosslinks stabilized by hydroxylation and glycosylation (36). Both homotrimer Collagen I (37) and glycosylation-stabilized crosslinks (38) increase the resistance of DD collagen to enzymatic cleavage.

Collagen-degrading enzymes have reduced activity in DD. TIMP1 (Tissue Inhibitor of Metalloproteinase 1) inhibits multiple MMPs, and both TIMP1 and TIMP1/MMP ratios are elevated in the circulation of DD patients (24). MMP14 partial loss-of-function variants are associated with DD (39). MMP14 is necessary for MMP2 activation, and both enzymes facilitate myofibroblast contraction (40). Clinical support for a relationship between impaired collagen degradation and DD is the rapid onset of DD coincident with metalloproteinase inhibitor treatment (41). Enzymatic cleavage of collagen fibrils is necessary for matrix shortening (42), and impaired cleavage would slow this process. Slow matrix remodeling could explain why DD-related tissue shortening progresses more slowly than wound healing despite similar biology and why DD surgery has an increased risk of delayed wound healing (43).

The hallmark of fibrotic diseases is collagen accumulation due to greater collagen synthesis than degradation. Increased collagen synthesis contributes to collagen accumulation in

pulmonary fibrosis (44), liver fibrosis (45), primary biliary cirrhosis (46), renal fibrosis (47), and cardiovascular disease (48). Our finding of normal levels of circulating Collagen I synthesis biomarkers in DD suggests that the mechanism of collagen accumulation in DD is not due to the increased collagen synthesis associated with these diseases. Although often used interchangeably, neither collagen accumulation itself nor upregulated collagen gene expression is de facto evidence of increased collagen synthesis. Cell culture studies show increased *COL1A1* and *COL1A2* mRNA expression in DD-derived fibroblasts (27, 49), but these RNA findings are not associated with increased collagen I protein production (27). Our finding of normal collagen I synthesis and impaired degradation impacts future research because DD drug development has focused on interventions to reduce collagen synthesis rather than increase collagen degradation (50).

Collagen II markers were normal. We evaluated Collagen II markers because of the correlation between DD and psoriasis (51) and the finding of abnormal Collagen II in psoriasis (52). We found no significant differences between groups in Collagen II synthesis marker PRO-C2_HP and degradation marker C2M_HP or their ratios. The CALC2 synthesis marker was undetectable in both cohorts, preventing the calculation of CALC2 synthesis-to-degradation ratios.

Collagen III synthesis markers were normal, consistent with a prior report of normal circulating Collagen III synthesis marker amino-terminal propeptide of type III procollagen (PIIINP) in DD subjects (53). Given the accumulation of Collagen III in DD, we expected to find elevated Collagen III synthesis to degradation marker ratios. Instead, we found normal Collagen III synthesis to degradation ratios. This finding may reflect that other MMPs and proteases also degrade Collagen III, and we did not measure all potential collagen degradation fragments. After the unexpected finding of no significant differentials of Col III synthesis (PRO-C3_roHP) and degradation (C3M_HP) markers compared to controls, we assayed a different Col III

degradation marker CTX-III_HP on the remaining aliquots. This assay was not initially planned, and only 17 (10 DD, 7 Controls) of the initial 27 (13 DD, 14 Controls) leftover aliquots had sufficient residual volumes for this assay. Although this small subgroup's CTX-III violin plot appearance suggested reduced degradation in DD compared to controls (Suppl Figure 1), this difference was not significant.

We found significantly lower Collagen V synthesis marker Pro-C5 levels in the DD cohort than controls (Figure 1B). This finding is unexplained. An increased percentage of Collagen V has been reported in DD tissues (33), suggesting a high Collagen V synthesis to degradation ratio. However, we could not calculate this ratio because a Collagen V degradation marker was unavailable.

The DD cohort also had significantly lower Collagen VII synthesis marker Pro-C7 than controls (Figure 1C). We could not calculate Collagen VII synthesis to degradation ratios because Collagen VII degradation markers were undetectable in both cohorts. Although there are no reports of Collagen VII in DD, its role in matrix mechanobiology and elevated tissue levels in systemic sclerosis (54) makes it a potential marker for collagen metabolism abnormalities in DD.

We found no significant differences in the measurements of the remaining collagen turnover markers. We measured Collagen IV markers because Collagen IV is increased in DD tissues (23). We assayed Collagen VI markers and ratios because prior pathway analysis suggested a role for Collagen VI in DD and other fibrotic diseases (55). We selected Collagen IX markers based on unpublished data suggesting a role in DD. None of these markers or ratios showed significant differences between disease and control cohorts.

Separate from the Nordic collagen fragment analysis, we analyzed whole proteins with the SomaLogic aptamer-based platform. The SomaLogic panel measures 6995 protein binding sites

associated with 5444 proteins. Suppl Table 6 lists the genes mentioned in this manuscript and their alias names.

SomaLogic hypothesis-free analysis identified 54 distinct proteins with significantly different plasma levels than controls. Twenty-three proteins had increased levels, and 29 had lower DD levels than controls (Suppl Table 1). Our findings are consistent with existing literature. Below, we discuss DD literature and pathway analysis of these 54 proteins.

The body of DD research publications mentions only seven of these 54 differentially expressed proteins: Aggrecan (*ACAN*), Amine Oxidase Copper Containing 3 (*AOC3*), Caspase-3 (*CASP3*), Periostin (*POSTN*), SMAD Family Member 1 (*SMAD1*), Spartin (*SPART*), and Ubiquitin carboxyl-terminal hydrolase 8 (*USP8*). Our differential expression findings concur with reports of *ACAN* (56), *AOC3* (20), and *POSTN* (57) overexpression in tissue and *SMAD1* (58), *SPART* (59), and *USP8* (59) underexpression in cell culture. CASP3 levels were low in our findings but upregulated in a report of transcriptomic profiling (56). This finding is unexplained. *ACAN* is also near the variant rs6496519 associated with DD (60). Suppl Table 1 summarizes the biological and clinical relevance of these 54 genes to DD.

We constructed PPI networks and enrichment analysis of these 54 plasma proteins, revealing significantly enriched protein-protein interactions in 32 with multiple enriched protein groups and categories (Figure 5, Suppl Table 2). Abnormalities in the KEGG category has04610 (complement and coagulation cascades) are consistent with DD pathway meta-analysis findings (61). Complement fragments induce adhesion molecules that act on endothelial cells, consistent with DD-related perivascular clusters of immune cells and small vessel occlusion (20). Complement fragments also activate mast cells that release TNF, a known DD pathway component (62), and histamine, a driver of itching, which is common in early-stage DD (1). Cell membrane and extracellular matrix signaling receptor binding of growth factors, cytokines, and other signaling factors plays a pivotal role in DD pathobiology (21). DD plasma has elevated

MMP inhibitors (63), a driver of DD-related collagen accumulation. Protein phosphorylation processes are consistent with increased phosphorylated STAT1, STAT3, SMAD2, SMAD3, and ERK1/2 in DD tissues (58, 64, 65). The enriched Reactome HSA-162582 (Signal Transduction) pathway is noteworthy, as it includes signaling by TGF-Beta, WNT, and NOTCH, all prominent in DD biology (66).

This hypothesis-free analysis provided data to generate a Diagnostic Proteomic Risk Score (DPRS) calculated on two proteins: high levels of Osteoclast-associated immunoglobulin-like receptor (*OSCAR*) and low levels of Spartin (*SPART*) (Table 2). Figure 4A shows the AUC-ROC of this panel (AUC = 0.8), and Figure 4B shows boxplots of the DPRS based on this panel to distinguish DD from control cohorts.

Our finding of elevated OSCAR levels is interesting because Collagen I is an OSCAR ligand. OSCAR also stimulates T-cells to release TNF (67), a component of DD biology (62). Our finding of decreased *SPART* expression is consistent with a report of low *SPART* expression in DD fibroblasts (59). SPART is also associated with Epidermal Growth Factor Receptor (EGFR) degradation and transport to the cell membrane (68), consistent with an increased ratio of intracellular to cell membrane EGFR levels found in DD (69).

Separately, using the SomaLogic platform, we also conducted a hypothesis-based analysis of 328 candidate DD-related proteins identified in a literature review (Suppl Table 3). From these candidate proteins, we identified 23 with significantly different plasma levels compared to controls. Nine proteins had higher DD levels, and fourteen had lower DD levels than controls. A group of nine proteins overlapped hypothesis-free and hypothesis-based groups. Our findings are consistent with existing literature. Suppl Table 4 summarizes the potential relationships of these genes to DD.

Similar to our hypothesis-free analysis, we performed a STRING PPI analysis of the hypothesis-based findings to identify network relationships of the 23 proteins. We found significantly enriched protein-protein interactions in 21, with multiple enriched categories (Suppl Table 5, p-value = 1.35e-07). This analysis implicates additional enrichment categories to those from the hypothesis-free analysis. Myofibroblast differentiation, proliferation, and motility, WNT signaling, and interleukin pathways are unsurprising, given their known roles in DD biology (17, 26). Cancer pathway associations are also unsurprising, given the correlations of DC with carcinomas and sarcomas (70, 71). Infection-related pathways were unexpected and exclusively involved viral and parasitic infection categories. This finding might be due to a hidden variable of tissue repair biology, which overlaps interleukin-driven parasitic immune defense biology and interleukin-related DD biology. The possible relationships between viral infection pathways and DD are unclear.

This hypothesis-based analysis provided data to generate a second Diagnostic Proteomic Risk Score (DPRS) calculated on eleven proteins (Table 3). Figure 6A shows the AUC-ROC of this panel (AUC = 0.9), and Figure 6B shows boxplots of the DPRS based on this panel to distinguish DD from control cohorts.

We evaluated the potential of each DPRS as a staging instrument. Recurrence after a procedure is the clinical benchmark of Dupuytren's biological severity. Traditionally, surgeons use clinical factors such as age of onset, affected family members, and disease locations other than the palms to predict post-procedure recurrence. We could not use these to correlate DD subsets with our DPRS because each criterion was mandatory for initial inclusion into the study. Recurrence is more common in patients younger at their first corrective procedure (72, 73). We evaluated the ability of each DPRS to distinguish Dupuytren subjects who had their first corrective procedure before age 50 from those having their first correction at age 50 or older. Unlike the hypothesis-free DPRS, the hypothesis-based DPRS distinguished DD subjects

requiring their first corrective procedure before age 50 from those undergoing their first correction after age 50 (p=0.0018), suggesting potential use as a staging and predictive tool.

Although Mass Spectrometry did not reveal statistically significant differences in abundance across cohorts, there was statistically significant concordance between Mass Spec and SomaScan findings. The Mass Spec data contained 6 of the 54 proteins with statistically significant differences between disease and control cohorts using the hypothesis-free SomaScan data analysis: Amine Oxidase Copper Containing 3 *AOC3*, Periostin *POSTN*, Serpin Family C Member 1 *SERPINC1*, Transferrin *TF*, Proprotein Convertase Subtilisin/Kexin Type 9 *PCSK9*, and Kininogen 1 *KNG1*). . Table 4 compares the results of the SomaLogic and Mass Spectrometry data analysis. A pathway analysis of these six proteins identified two highly significant Reactome categories (Suppl Table 7). In addition to protein phosphorylation noted above, Reactome HSA-381426 involving Insulin-like Growth Factor *IGF* and Insulin-like Growth Factor Binding Proteins *IGFBP*s is relevant because of the role of IGF-II and IGFBP-6 in regulating Dupuytren fibroblast proliferation and contractility (74).

In summary, we used ELISA, aptamer, and mass spectrometry assays to identify differential plasma proteomic signatures between DD and control cohorts. We found significant DD-related proteomic abnormalities in both collagen metabolism peptide fragments and whole proteins in circulation. This study is the first to have evidence suggesting that net collagen accumulation in DD derives from normal collagen synthesis and impaired collagen degradation, challenging the model of excess collagen synthesis and distinguishing DD from other organ fibroses. We separately performed mass spectrometry on these cohorts. Mass spec identified 561 plasma entities but no significant differentials, and we believe this represents the technical limitations of our mass spec techniques on plasma. We independently analyzed a set of 6995 aptamers applying to 5444 proteins from the same cohorts. We first used a hypothesis-free analysis and found 54 statistically significant differences between DD and control cohorts. We then analyzed

a literature-based subset of 328 proteins and identified 23 proteins with statistically significant differences. We developed two Dupuytren Diagnostic Proteomic Risk Scores separately based on hypothesis-free and hypothesis-driven analyses. We used pathway analyses to identify DD protein-protein network enrichments, revealing previously unreported DD pathway categories. Finally, we showed our hypothesis-based DPRS correlated with the subject's age at the time of the first DD procedure, an index of biological severity. The hypothesis-based DPRS had a greater AUC-ROC value and statistical significance in distinguishing Dupuytren vs. control than the hypothesis-free DPRS. It also distinguished Dupuytren subjects by age at the time of their first Dupuytren procedure. These findings suggest that a validated hypothesis-based DPRS could guide management and develop therapeutic drugs.

This study has limitations. The small number of samples in all assays and the large number of aptamers in the SomaScan assay were statistical disadvantages. Although our findings align with current research on Dupuytren proteomics, there are differences. Elevated serum TIMP1 (24) and TNF (26) have been reported in DD, but did not have statistically significant differential expression between disease and control groups in our analysis. This may reflect that Somalogic assays target specific protein conformations, whereas ELISA measures all conformations.

We make the following recommendations based on our novel findings of circulating protein and peptide interactions in DD. We recommend validating Collagen I, III, V, and VII metabolite studies. Both degradation-resistant collagen and impaired enzyme activity likely contribute to impaired collagen degradation. The relative contribution of each needs further investigation, and normalizing impaired collagen degradation should be a lead target for therapeutic drug development. We also recommend a validating study of 68 unique proteins identified in the protein aptamer analysis (54 in the hypothesis-free analysis and 23 in the hypothesis-based analysis, with 9 in both groups). Testing only the subset of the 68 most pertinent proteins would increase the power to determine the most promising biomarkers but would involve the following

logistical challenges. Mass spectrography discovery was unrevealing; small custom Somalogic (https://somalogic.com/) and Olink (https://olink.com) assays are unavailable, and ELISA assays of all would be prohibitively expensive. An alternative option would be to create a custom multiplex panel for the 13 proteins in our two DPRS (21). Because of the limited number of collagen metabolite samples, we could not develop a combined protein/peptide risk score. We recommend research to develop a protein/peptide risk score. We recommend a larger study to validate the correlation of our hypothesis-based DPRS on age at the first DD procedure. Our findings strongly support the need to fund and conduct larger validating studies of circulating collagen metabolites and proteins in DD. This biomarker research is fundamental to developing DD diagnostic tests to diagnose, stage, and inform research to develop therapeutic drugs for DD.

METHODS

Sex as a biological variable

Our study involved male and female subjects. Sex was not considered a biological variable.

Subjects

We selected disease cohort subjects from the International Dupuytren Database (IDDB), an online patient-reported registry developed by the Dupuytren Research Group (DRG) in collaboration with the National Databank of Rheumatic Diseases (NDB). Control subjects were enrolled by word of mouth and from the NDB database (https://dupuytrens.org/enroll-in-the-iddb-2/). All subjects in the IDDB met the entry criteria and signed an Informed Consent Form.

Entry Criteria

Identification of novel plasma proteomic biomarkers of Dupuytren Disease

The DD cohort met criteria based on clinical evidence of DD with strong traditional predictors of recurrent deformity after a corrective procedure: age of disease onset younger than 50; two or more first-degree relatives with DD; current DD nodules; three or more separate locations of DD or two palm locations and Ledderhose disease (a related disorder). Control cohort subjects were healthy individuals with no clinical evidence of DD, no personal history of DD or Ledderhose disease, and no family history of DD.

We confirmed study eligibility by requiring candidates to upload recent images of the palmar and dorsal sides of each hand. The Principal Investigator/ senior author confirmed entry criteria and reviewed images to confirm clinical evidence of active DD in the disease cohort and the lack of these findings in the control cohort.

Blood Sample Collection

The study used a mobile phlebotomy service to collect blood specimens (Phlebotek Solutions Corporation, Fort Lauderdale, FL). Once confirmed that the subject met all study entry criteria, they were contacted by DRG and later by Phlebotek to schedule their blood draw. The phlebotomist visited the subjects in their homes and collected two tubes of blood. The last 16 blood draws included a third 2.5ml PAXgene tube collection for additional genomic testing.

Processing of the EDTA tubes for the proteomic testing was standardized. Immediately after collection, blood samples were centrifuged for 15 minutes at 2000 g to generate plasma. The plasma was harvested from the top of the container and pooled before aliquoting into storage cryotubes. The last 500 uL of plasma was discarded to avoid contamination with platelets. Samples were immediately packaged for shipping on dry ice.

Number of subjects

One hundred subjects were planned (50 DD/50 Control). Blood collection launched just as the

COVID pandemic began (i.e., February 2020). Due to COVID-related issues, we amended the protocol to enroll 50 subjects. Forty-seven subjects were enrolled;46 subjects had samples drawn. One sample was disqualified due to failed shipping refrigeration. Samples from the remaining 45 subjects (25 DD, 20 Control) had mass-spectrometry (Bioproximity, Manassas, VA). Twenty-seven subjects (13 DD, 14 control) had second blood samples drawn six to 14 months after their first draw for mass-spectrometry testing by a second lab (MacCoss Lab, Seattle, WA). Additional proteomic testing was conducted on remaining aliquots; 1) n= 27 redrawn samples had ELISA tests for collagen metabolites (Nordic Bioscience, Herlev, DK) 2) n= 72 from both initial and redraw specimens had Aptamer-based testing (SomaLogic, Boulder, CO). Figure 2 outlines the distribution of samples and laboratories.

Blinding

All subjects were assigned and identified as participant numbers. Investigators were blinded to the sample group allocation. All samples remained blinded throughout the processing and data collection in the laboratories.

Assays

Nordic Bioscience Collagen Biomarker Assay

Circulating collagen synthesis and degradation markers were measured using proprietary assays (30). Synthesis biomarkers included collagens I (PRO-C1), II (PRO-C2), III (PRO-C3), IV (PRO-C4), V (PRO-C5), VI (PRO-C6), VII (PRO-C7), and IX (PRO-C9). Circulating degradation markers were measured for collagens I (C1M), II (CALC2, C2M), III (C3M, CTX-III), IV (C4M), VI (C6M), and VII (C7M). All biomarkers were assessed by competitive ELISAs and validated to measure in human plasma samples. The inter- and intra-assay coefficients of variation are >15% and 10%, respectively, for all assays.

Mass Spectrometry

A novel untargeted proteomics assay using a large-scale liquid chromatography-tandem mass spectrometry (LC-MS/MS) approach called data-independent acquisition (DIA) was used to collect proteomic data. Each sample (10 µl plasma), following tryptic digestion, was diluted 50:50 by the addition of 4% sodium dodecyl sulfate (SDS) buffer to a final concentration of ~2% SDS. Approximately 200 µg of protein was added to 96 well-plates, and the remainder of the sample preparation was performed using the single-pot, solid-phase enhanced sample preparation for proteomics methodology (SP3) (75, 76). The entire process of protein capture, reduction and alkylation, protein cleanup, digestion, and peptide cleanup was performed using a KingFisher magnetic bead robot (76). Peptides were reconstituted in HPLC starting aqueous buffer (5% acetonitrile and 0.1% formic acid) with analysis performed on the high-resolution quadrupole-orbitrap hybrid mass spectrometer (Exploris 480, Orbitrap Fusion Lumos, or Orbitrap Fusion Eclipse).

Data was acquired in a single, 2-hour LC-MS/MS run (90 min HPLC gradient) on a Thermo Quadrupole-Orbitrap mass spectrometer. Crucially, the DIA method comprehensively samples all peptides between 400 and 1000 m/z using a repeated cycle of 75 MS/MS scans with an 8 m/z-wide isolation width. Using a multiplexing scheme based on staggering the isolation window in alternating cycles of MS/MS scans, the precursor selectivity of the resulting data is 4 m/z after computational demultiplexing (77). From the resulting data, fragment ion chromatograms were extracted for any peptide precursor between 400 and 1000 m/z. To determine which peptides are confidently detected in each acquisition, the software tool EncyclopeDIA (78) was used to query the data using a reference chromatogram library generated from a prior analysis of the tissue. Chromatogram-based relative-quantitation was then performed for each peptide detected. All measurements are made relative to a common reference. In the case of plasma, a common reference is a pooled standard from 400 age matched samples from subjects who are

both healthy and span the phenotypes being studied processed in each sample batch. Within each 96-well plate, one copy of the common reference standard is placed in each row for a total of 8 copies per plate (79). Before quantification, it is confirmed that peptides are stable during the time scale of the data collection, have a suitable linear response, and have suitable intra and inter-data precision.

SomaScan Assay

SomaLogic, Inc. (Boulder, CO, USA) performed proteomic analysis of the plasma samples using a SomaScan platform that quantifies the presence of 6995 human proteins (secreted proteins, extracellular domains, and intracellular proteins) belonging to broad biological groups, including receptors, kinases, cytokines, proteases, growth factors, protease inhibitors, hormones, and structural proteins (80). Most of these proteins involve signal transduction pathways, stress response, immune processes, phosphorylation, proteolysis, cell adhesion, cell differentiation, and intracellular transport (81).

Proteins were measured using a Slow Off-Rate Modified Aptamer (SOMAmer)-based capture array that uses chemically modified single-stranded DNA sequences capable of uniquely recognizing the individual proteins, transforming a protein signal to a nucleotide signal that can be quantified using relative fluorescence on microarrays. Each SOMAmer array was validated for its specificity, upper and lower limits of detection, and intra- and interassay variability. Plasma dilutions (0.005%, 0.05%, and 20%) were applied to capture low-, medium- and high-abundance proteins. Positive and negative controls were also positioned on the array to confirm if the experimental procedure was performed correctly.

After the hybridization step, the microarrays were washed and scanned using a specific scanner containing a laser that excites the fluorescence of the fluorochrome used in the labeling step.

The amount of the emitted signal is directly proportional to the dye on the microarray spot. The scanner can measure this quantity and process a digital image that reconstructs the position of each signal on the microarray based on the spot of its origin. The images were then analyzed by image analysis software that generates a text file containing a summary of the information on the intensity of the pixels contained in the spot and in the background. This information was processed, and a final value was generated that summarized the level of each detectable protein on the microarray.

Quality control was performed at the sample and SOMAmer levels and involved using control SOMAmer arrays on the microarray and calibration samples. Hybridization controls measured the sample-by-sample variation in hybridization, while the median signal over all SOMAmer arrays measured the technical variability. The SomaScan measures were finally reported as relative fluorescence units (RFU) in a summary ADAT file.

Protein-protein interaction network and pathway analysis

STRING (12.0) is a web tool that provides and illustrates protein-protein interaction and networks (31). STRING integrates different types of evidence from public databases, including genomic context, high-throughput, and biochemical interaction experiments, known protein complexes, computational predictions, and text-mining from scientific literature. Given a list of gene names, STRING will construct protein-protein interaction networks where each gene is considered as a node, with edges representing the predicted functional associations.

Networks were generated using the whole genome as the statistical background, and interaction sources included text mining, experiments, databases, co-expression, neighborhood, gene fusions, and co-occurrence. Interaction scores were computed by combining the probabilities from all evidence channels and corrected for the probability of randomly observing an

interaction. All nodes and interactions were computed with a required medium confidence interaction score (0.400) and an FDR<=0.05.

Analysis and statistics

Nordic collagen data turnover rates were calculated by performing ratios of the propeptide over the mature peptide (e.g., PRO-C1/ C1M). This ratio is an estimate of the formation/degradation of each collagen type and can be interpreted as a net gain of collagen/loss of collagen remodeling (30). All comparisons were done using Student's T-test.

A series of quality control and data cleaning steps were performed on the SomaLogic proteomics data. All non-human proteins were excluded from the analyses. Samples that did not pass the default normalization acceptance criteria for all row scale factors were removed. All aptamers that did not pass the default quality control criteria across all plates/sets were filtered out. The data had no missing values, and the protein values were log-transformed. No further normalization was performed.

We identified differentially expressed aptamers using linear regression of each log-transformed protein as the outcome and disease status as the main predictor. We adjusted each analysis by age and sex. We analyzed all repeated measurements of each individual and used generalized estimating equations to account for technical and biological variability using the gee package in R. Due to the high number of aptamers tested in the SomaScan assay (6995), we were at a statistical testing disadvantage regarding multiple corrections. Therefore, we calculated adjusted p-values using Benjamini Hochberg methods. Our SomaLogic analyses only report proteins with statistically significant differences from control at an FDR-adjusted p-value < 0.25.

For the MS protein data, all zeroes were replaced with NAs across samples, and protein values were log-transformed. The MS dataset was subset based on proteins that overlapped with the signature proteins associated with DD previously identified using the SomaScan proteomics

data. We used a generalized estimation equation model to account for the repeated blood draws, adjusted by age and gender.

Proteomic Risk Score

To prepare the data for the Diagnostic Proteomic Risk Score (DPRS) calculation, we first removed all highly correlated proteins (Pearson correlation > 0.8). 3369 out of 6995 aptamers were highly correlated and were excluded from the DPRS calculation. The DPRS calculation contained the following two steps. First, we fit a logistic regression for each protein with disease status (i.e., case or control) as the outcome variable adjusted by age and gender. We used a generalized estimation equation model for repeated measurements of two independent blood draws. We included the same subject's repeated measurements and used the generalized estimation equations model with the gee R package (82) to estimate the regression coefficients. Next, the DPRS was computed using summary statistics from each regression model. Only the proteins with statistically significant differences from control at an FDR-adjusted p-value < 0.25 were included in the score calculation. The DPRS was calculated by summing the beta estimated effect for all significant proteins, each multiplied by normalized protein abundance and divided by its standard error. Finally, we fit a logistic regression to examine the association of the computed DPRS with DD onset. The dataset was split into training (75%) and testing utilizing the caret R package for a leave-one-out cross-validation approach for model fitting and assessment (25%) sets (83).

All analyses were done using R version 4.2.1.

Data Availability

All the data and analytic scripts that support the findings of this study are openly available in the Github repository Dup_Blood_Proteomic_Data at https://github.com/Dupuytren/Dup_Blood_Proteomic_Data.

Study Approval

The Institutional Review Board Ascension Via Christi Hospitals Wichita, Inc. approved this study and the associated informed consent form. The written informed consent was signed and received before subject participation. This written informed consent included a description of the use of the photographs/ images to be reviewed for study entry. The records of these signed informed consents have been retained. All samples remained blinded throughout the processing and data collection in the laboratories.

References

1. Eaton C. In: Wolfe S, Pederson W, Kozin S, and Cohen M eds. *Green's Operative Hand Surgery*. Philadelphia: Elsevier; 2022:148-81.
2. Macaulay D, Ivanova J, Birnbaum H, Sorg R, and Skodny P. Direct and indirect costs associated with Dupuytren's contracture. *J Med Econ.* 2012;15(4):664-71.
3. Broekstra DC, Kuo RYL, Burn E, Prieto-Alhambra D, and Furniss D. Dupuytren Disease: Prevalence, Incidence, and Lifetime Risk of Surgical Intervention. A Population-Based Cohort Analysis. *Plast Reconstr Surg.* 2023;151(3):581-91.
4. Mendelaar NHA, Poelstra R, van Nieuwenhoven CA, Slijper HP, Feitz R, Hovius SER, et al. Outcome of Recurrent Surgery in Dupuytren's Disease: Comparison with Initial Treatment. *Plast Reconstr Surg.* 2019;144(5):828e-35e.
5. Goodman AD, Gil JA, Starr AM, Akelman E, and Weiss A-PC. Incidence of Serious Complications in Hand Surgery: A 10-Year Review. *The Journal of Hand Surgery.* 2016;41(9):S27-S8.

6. Schindelar L, McEntee R, D'Amore T, Beredjiklian P, and Lutsky K. Unplanned Return to the Operating Room in Upper-Extremity Surgery: Incidence and Reason for Return. *J Hand Surg Am.* 2021;46(8):715 e1- e12.

7. Denkler K. Surgical complications associated with fasciectomy for dupuytren's disease: a 20-year review of the English literature. *Eplasty.* 2010;Jan 27(10):e15.

8. Alser O, Craig RS, Lane JCE, Prats-Uribe A, Robinson DE, Rees JL, et al. Serious complications and risk of re-operation after Dupuytren's disease surgery: a population-based cohort study of 121,488 patients in England. *Sci Rep.* 2020;10(1):16520.

9. Degreef I, and De Smet L. Dupuytren's disease: a predominant reason for elective finger amputation in adults. *Acta Chir Belg.* 2009;109(4):494-7.

10. Jensen CM, Haugegaard M, and Rasmussen SW. Amputations in the treatment of Dupuytren's disease. *J Hand Surg Br.* 1993;18(6):781-2.

11. Norotte G, Apoil A, and Travers V. Résultats à plus de dix ans de la maladie de Dupuytren. *Annales de Chirurgie de la Main.* 1988;7(4):277-81.

12. Honner R, Lamb DW, and James JI. Dupuytren's contracture. Long term results after fasciectomy. *J Bone Joint Surg Br.* 1971;53(2):240-6.

13. Kurbel S, and Samarzija Z. Similar Distributions of Dupuytren's Contracture and Y-Chromosome Haplogroup I Among Modern Europeans Suggest Simultaneous Spreading of These Traits Some 40 to 10 KYA. *Coll Antropol.* 2016;40(1):63-4.

14. Ågren R, Patil S, Zhou X, Palotie A, Daly M, Riley-Gills B, et al. Major Genetic Risk Factors for Dupuytren's Disease Are Inherited From Neandertals. *Molecular Biology and Evolution.* 2023;40(6).

15. Dolmans GHCG, Wijmenga C, Ophoff R, and Werker PMN. *Dupuytren's Disease and Related Hyperproliferative Disorders*. 2012:115-22.

16. Larsen S, Krogsgaard DG, Aagaard Larsen L, Iachina M, Skytthe A, and Frederiksen H. Genetic and environmental influences in Dupuytren's disease: a study of 30,330 Danish twin pairs. *J Hand Surg Eur Vol.* 2015;40(2):171-6.

17. Dolmans GH, Werker PM, Hennies HC, Furniss D, Festen EA, Franke L, et al. Wnt signaling and Dupuytren's disease. *N Engl J Med.* 2011;365(4):307-17.

18. Riesmeijer SA, Manley OWG, Ng M, Nolte IM, Broekstra DC, Werker PMN, et al. A Weighted Genetic Risk Score Predicts Surgical Recurrence Independent of High-Risk Clinical Features in Dupuytren's Disease. *Plast Reconstr Surg.* 2019;143(2):512-8.

19. Iqbal SA, Hindocha S, Farhatullah S, Paus R, and Bayat A. *Dupuytren's Disease and Related Hyperproliferative Disorders*. 2012:167-74.

20. Mayerl C, Del Frari B, Parson W, Boeck G, Piza-Katzer H, Wick G, et al. Characterisation of the inflammatory response in Dupuytren's disease. *J Plast Surg Hand Surg.* 2016;50(3):171-9.

21. Gonga-Cave BC, Pena Diaz AM, and O'Gorman DB. Biomimetic analyses of interactions between macrophages and palmar fascia myofibroblasts derived from Dupuytren's disease reveal distinct inflammatory cytokine responses. *Wound Repair Regen.* 2021;29(4):627-36.

22. Halliday NL, Rayan GM, Zardi L, and Tomasek JJ. Distribution of ED-A and ED-B containing fibronectin isoforms in Dupuytren's disease. *J Hand Surg Am.* 1994;19(3):428-34.

23. Magro G, Fraggetta F, Travali S, and Lanzafame S. Immunohistochemical expression and distribution of alpha2beta1, alpha6beta1, alpha5beta1 integrins and their extracellular ligands, type IV collagen, laminin and fibronectin in palmar fibromatosis. *Gen Diagn Pathol.* 1997;143(4):203-8.

24. Ulrich D, Hrynyschyn K, and Pallua N. Matrix metalloproteinases and tissue inhibitors of metalloproteinases in sera and tissue of patients with Dupuytren's disease. *Plast Reconstr Surg.* 2003;112(5):1279-86.

25. Jupp O, Pullinger M, Marjoram T, Lott M, Chojnowski AJ, and Clark IM. *Dupuytren Disease and Related Diseases - The Cutting Edge*. 2017:55-61.

26. Wang ML, Rajpar I, Ruggiero NA, Fertala J, Steplewski A, Beredjiklian PK, et al. Circulating inflammatory cytokines alter transcriptional activity within fibrotic tissue of Dupuytren's disease patients. *J Orthop Res.* 2021.

27. van Beuge MM, Ten Dam EJ, Werker PM, and Bank RA. Matrix and cell phenotype differences in Dupuytren's disease. *Fibrogenesis Tissue Repair.* 2016;9:9.

28. Tripkovic I, Ogorevc M, Vukovic D, Saraga-Babic M, and Mardesic S. Fibrosis-Associated Signaling Molecules Are Differentially Expressed in Palmar Connective Tissues of Patients with Carpal Tunnel Syndrome and Dupuytren's Disease. *Biomedicines.* 2022;10(12).

29. Broekstra DC, Lanting R, Werker PMN, and van den Heuvel ER. Disease Course of Primary Dupuytren Disease: 5-Year Results of a Prospective Cohort Study. *Plast Reconstr Surg.* 2022;149(6):1371-8.

30. Karsdal MA, Leeming DJ, Henriksen K, Bay-Jensen AC, Nielsen SH, and CL B. *Biochemistry of Collagens, Laminins and Elastin.* Boston: Elsevier; 2019.

31. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 2023;51(D1):D638-D46.

32. Choi SW, Mak TS, and O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc.* 2020;15(9):2759-72.

33. Bailey AJ. In: McFarlane R, McGrouther D, and Flint M eds. *Dupuytren's Disease Biology and Treatment*. Edinburg: Churchill Livingstone; 1990.

34. Bailey AJ, Sims TJ, Gabbiani G, Bazin S, and LeLous M. Collagen of Dupuytren's disease. *Clin Sci Mol Med.* 1977;53(5):499-502.

35. Williamson K, Cooper G, Lee KJ, Beamish EL, Simpson D, Carter A, et al. Active synthesis of type I collagen homotrimer in Dupuytren's fibrosis is unaffected by anti-TNF-α treatment. *bioRxiv preprint* 2023.

36. Bazin S, Le Lous M, Duance VC, Sims TJ, Bailey AJ, Gabbiani G, et al. Biochemistry and histology of the connective tissue of Dupuytren's disease lesions. *Eur J Clin Invest.* 1980;10(1):9-16.

37. Han S, Makareeva E, Kuznetsova NV, DeRidder AM, Sutter MB, Losert W, et al. Molecular mechanism of type I collagen homotrimer resistance to mammalian collagenases. *J Biol Chem.* 2010;285(29):22276-81.

38. Snedeker JG, and Gautieri A. The role of collagen crosslinks in ageing and diabetes - the good, the bad, and the ugly. *Muscles Ligaments Tendons J.* 2014;4(3):303-8.

39. Itoh Y, Ng M, Wiberg A, Inoue K, Hirata N, Paiva KBS, et al. A common SNP risk variant MT1-MMP causative for Dupuytren's disease has a specific defect in collagenolytic activity. *Matrix Biol.* 2021;97:20-39.

40. Wilkinson JM, Davidson RK, Swingler TE, Jones ER, Corps AN, Johnston P, et al. MMP-14 and MMP-2 are key metalloproteases in Dupuytren's disease fibroblast-mediated contraction. *Biochim Biophys Acta.* 2012;1822(6):897-905.

41. Krzeski P, Buckland-Wright C, Balint G, Cline GA, Stoner K, Lyon R, et al. Development of musculoskeletal toxicity without clear benefit after administration of PG-116800, a matrix metalloproteinase inhibitor, to patients with knee osteoarthritis: a randomized, 12-month, double-blind, placebo-controlled study. *Arthritis Res Ther.* 2007;9(5):R109.

42. Castella LF, Buscemi L, Godbout C, Meister JJ, and Hinz B. A new lock-step mechanism of matrix remodelling based on subcellular contractile events. *J Cell Sci.* 2010;123(Pt 10):1751-60.

43. Bismil Q, Bismil M, Bismil A, Neathey J, Gadd J, Roberts S, et al. The development of one-stop wide-awake dupuytren's fasciectomy service: a retrospective review. *JRSM Short Rep.* 2012;3(7):48.

44. Organ LA, Duggan AR, Oballa E, Taggart SC, Simpson JK, Kang'ombe AR, et al. Biomarkers of collagen synthesis predict progression in the PROFILE idiopathic pulmonary fibrosis cohort. *Respir Res.* 2019;20(1):148.

45. Liu T, Wang X, Karsdal MA, Leeming DJ, and Genovese F. Molecular serum markers of liver fibrosis. *Biomark Insights.* 2012;7:105-17.

46. Vesterhus M, Nielsen MJ, Hov JR, Saffioti F, Manon-Jensen T, Leeming DJ, et al. Comprehensive assessment of ECM turnover using serum biomarkers establishes PBC as a high-turnover autoimmune liver disease. *JHEP Rep.* 2021;3(1):100178.

47. Han DC, Isono M, Hoffman BB, and Ziyadeh FN. High glucose stimulates proliferation and collagen type I synthesis in renal cortical fibroblasts: mediation by autocrine activation of TGF-beta. *J Am Soc Nephrol.* 1999;10(9):1891-9.

48. Genovese F, Goncalves I, Holm Nielsen S, Karsdal MA, Edsfeldt A, Nilsson J, et al. Plasma levels of PRO-C3, a type III collagen synthesis marker, are associated with arterial stiffness and increased risk of cardiovascular death. *Atherosclerosis.* 2024;388:117420.

49. Satish L, Gallo PH, Baratz ME, Johnson S, and Kathju S. Reversal of TGF-beta1 stimulation of alpha-smooth muscle actin and extracellular matrix components by cyclic AMP in Dupuytren's-derived fibroblasts. *BMC Musculoskelet Disord.* 2011;12:113.

50. Lambi AG, Popoff SN, Benhaim P, and Barbe MF. Pharmacotherapies in Dupuytren Disease: Current and Novel Strategies. *J Hand Surg Am.* 2023.

51. Patel M, Freeman NR, Dhaliwal S, Wright N, Daoud Y, Ryan C, et al. The prevalence of Dupuytren contractures in patients with psoriasis. *Clin Exp Dermatol.* 2014;39(8):894-9.

52. Holm Nielsen S, Magee C, Groen SS, Sinkeviciute D, Bay-Jensen AC, Karsdal MA, et al. Differentiating patients with psoriasis from psoriatic arthritis using collagen biomarkers. *Clin Exp Rheumatol.* 2022.

53. Gassier N, Brenner P, Berger A, and Delbruck A. In: Berger A, Delbruck A, Brenner P, and Hinzmann B eds. *Dupuytren's Disease Pathobiochemistry and Clinical Management*. Berlin: Springer-Verlag; 1994:94-8.

54. Sand JMB, Lamy P, Juhl P, Siebuhr AS, Iversen LV, Nawrocki A, et al. Development of a Neo-Epitope Specific Assay for Serological Assessment of Type VII Collagen Turnover and Its Relevance in Fibroproliferative Disorders. *Assay Drug Dev Technol.* 2018;16(2):123-31.

55. Williams LM, McCann FE, Cabrita MA, Layton T, Cribbs A, Knezevic B, et al. Identifying collagen VI as a target of fibrotic diseases regulated by CREBBP/EP300. *Proc Natl Acad Sci U S A.* 2020;117(34):20753-63.

56. Rehman S, Salway F, Stanley JK, Ollier WE, Day P, and Bayat A. Molecular phenotypic descriptors of Dupuytren's disease defined using informatics analysis of the transcriptome. *J Hand Surg Am.* 2008;33(3):359-72.

57. Vi L, Feng L, Zhu RD, Wu Y, Satish L, Gan BS, et al. Periostin differentially induces proliferation, contraction and apoptosis of primary Dupuytren's disease and adjacent palmar fascia cells. *Exp Cell Res.* 2009;315(20):3574-86.

58. Krause C, Kloen P, and Ten Dijke P. Elevated transforming growth factor beta and mitogen-activated protein kinase pathways mediate fibrotic traits of Dupuytren's disease fibroblasts. *Fibrogenesis Tissue Repair.* 2011;4(1):14.

59. Forrester HB, Temple-Smith P, Ham S, de Kretser D, Southwick G, and Sprung CN. Genome-wide analysis using exon arrays demonstrates an important role for expression

of extra-cellular matrix, fibrotic control and tissue remodelling genes in Dupuytren's disease. *PLoS One.* 2013;8(3):e59056.

60. Ng M, Thakkar D, Southam L, Werker P, Ophoff R, Becker K, et al. A Genome-wide Association Study of Dupuytren Disease Reveals 17 Additional Variants Implicated in Fibrosis. *Am J Hum Genet.* 2017;101(3):417-27.
61. Park TH, Kim D, Lee YS, and Kim SY. A meta-analysis to identify novel diagnostic and therapeutic targets for Dupuytren's disease. *Wound Repair Regen.* 2020;28(2):202-10.
62. Verjee LS, Verhoekx JS, Chan JK, Krausgruber T, Nicolaidou V, Izadi D, et al. Unraveling the signaling pathways promoting fibrosis in Dupuytren's disease reveals TNF as a therapeutic target. *Proc Natl Acad Sci U S A.* 2013;110(10):E928-37.
63. Johnston P, Chojnowski AJ, Davidson RK, Riley GP, Donell ST, and Clark IM. A complete expression profile of matrix-degrading metalloproteinases in Dupuytren's disease. *J Hand Surg Am.* 2007;32(3):343-51.
64. Akbar M, Garcia-Melchor E, Chilaka S, Little KJ, Sood S, Reilly JH, et al. Attenuation of Dupuytren's fibrosis via targeting of the STAT1 modulated IL-13Ralpha1 response. *Sci Adv.* 2020;6(28):eaaz8272.
65. Kida H, Jiang JJ, Matsui Y, Takahashi I, Hasebe R, Kawamura D, et al. Dupuytren's contracture-associated SNPs increase SFRP4 expression in nonimmune cells including fibroblasts to enhance inflammation development. *Int Immunol.* 2023.
66. Riesmeijer SA, Kamali Z, Ng M, Drichel D, Piersma B, Becker K, et al. A genome-wide association meta-analysis implicates Hedgehog and Notch signaling in Dupuytren's disease. *Nat Commun.* 2024;15(1):199.
67. Schultz HS, Nitze LM, Zeuthen LH, Keller P, Gruhler A, Pass J, et al. Collagen induces maturation of human monocyte-derived dendritic cells by signaling through osteoclast-associated receptor. *J Immunol.* 2015;194(7):3169-79.

68. Bakowska JC, Jupille H, Fatheddin P, Puertollano R, and Blackstone C. Troyer syndrome protein spartin is mono-ubiquitinated and functions in EGF receptor trafficking. *Mol Biol Cell.* 2007;18(5):1683-92.

69. Augoff K, Tabola R, Kula J, Gosk J, and Rutowski R. Epidermal growth factor receptor (EGF-R) in Dupuytren's disease. *J Hand Surg Br.* 2005;30(6):570-3.

70. Wilbrand S. *Medicine*. Uppsala: Acta Universitatis Upsaliensis; 2002:53.

71. Kuo RYL, Ng M, Prieto-Alhambra D, and Furniss D. Dupuytren's Disease Predicts Increased All-Cause and Cancer-Specific Mortality: Analysis of a Large Cohort from the U.K. Clinical Practice Research Datalink. *Plast Reconstr Surg.* 2020;145(3):574e-82e.

72. van Rijssen AL, ter Linden H, and Werker PM. Five-year results of a randomized clinical trial on treatment in Dupuytren's disease: percutaneous needle fasciotomy versus limited fasciectomy. *Plast Reconstr Surg.* 2012;129(2):469-77.

73. Wilbrand S, Ekbom A, and Gerdin B. The sex ratio and rate of reoperation for Dupuytren's contracture in men and women. *J Hand Surg Br.* 1999;24(4):456-9.

74. Raykha C, Crawford J, Gan BS, Fu P, Bach LA, and O'Gorman DB. IGF-II and IGFBP-6 regulate cellular contractility and proliferation in Dupuytren's disease. *Biochim Biophys Acta.* 2013;1832(10):1511-9.

75. Hughes CS, Moggridge S, Muller T, Sorensen PH, Morin GB, and Krijgsveld J. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc.* 2019;14(1):68-85.

76. Leutert M, Rodriguez-Mias RA, Fukuda NK, and Villen J. R2-P2 rapid-robotic phosphoproteomics enables multidimensional cell signaling studies. *Mol Syst Biol.* 2019;15(12):e9021.

77. Amodei D, Egertson J, MacLean BX, Johnson R, Merrihew GE, Keller A, et al. Improving Precursor Selectivity in Data-Independent Acquisition Using Overlapping Windows. *J Am Soc Mass Spectrom.* 2019;30(4):669-84.

78. Searle BC, Pino LK, Egertson JD, Ting YS, Lawrence RT, MacLean BX, et al. Chromatogram libraries improve peptide detection and quantification by data independent acquisition mass spectrometry. *Nat Commun.* 2018;9(1):5128.

79. Hoofnagle AN, Becker JO, Oda MN, Cavigiolio G, Mayer P, and Vaisar T. Multiple-reaction monitoring-mass spectrometric assays can accurately measure the relative protein abundance in complex mixtures. *Clin Chem.* 2012;58(4):777-81.

80. Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN, et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One.* 2010;5(12):e15004.

81. Berrone E, Chiorino G, Guana F, Benedetti V, Palmitessa C, Gallo M, et al. SOMAscan Proteomics Identifies Novel Plasma Proteins in Amyotrophic Lateral Sclerosis Patients. *Int J Mol Sci.* 2023;24(3).

82. Network TCRA. gee: Generalized Estimation Equation Solver. https://cran.r-project.org/package=gee. Accessed 01/06/2023.

83. Kuhn M. *Journal of Statistical Software*. American Statistical Association; 2008.

Author Contributions

BH: Analyzing data, writing the manuscript

PS: Analyzing data, writing the manuscript

AL: Analyzing data

AG: Analyzing data

CB: Analyzing data

SHN: Analyzing data

MK: Analyzing data

CE: Study design, Analyzing data, writing the manuscript

Acknowledgments

This work was funded entirely by the Dupuytren Research Group, a 501 (c) (3) nonprofit organization.

Gail Solomon assisted with manuscript writing.

PS, AL, AG: grant UM1TR004398 (the Tufts NIH CTSA award).

Databank and study support from FORWARD, the National Databank for Rheumatic Diseases, is gratefully acknowledged.

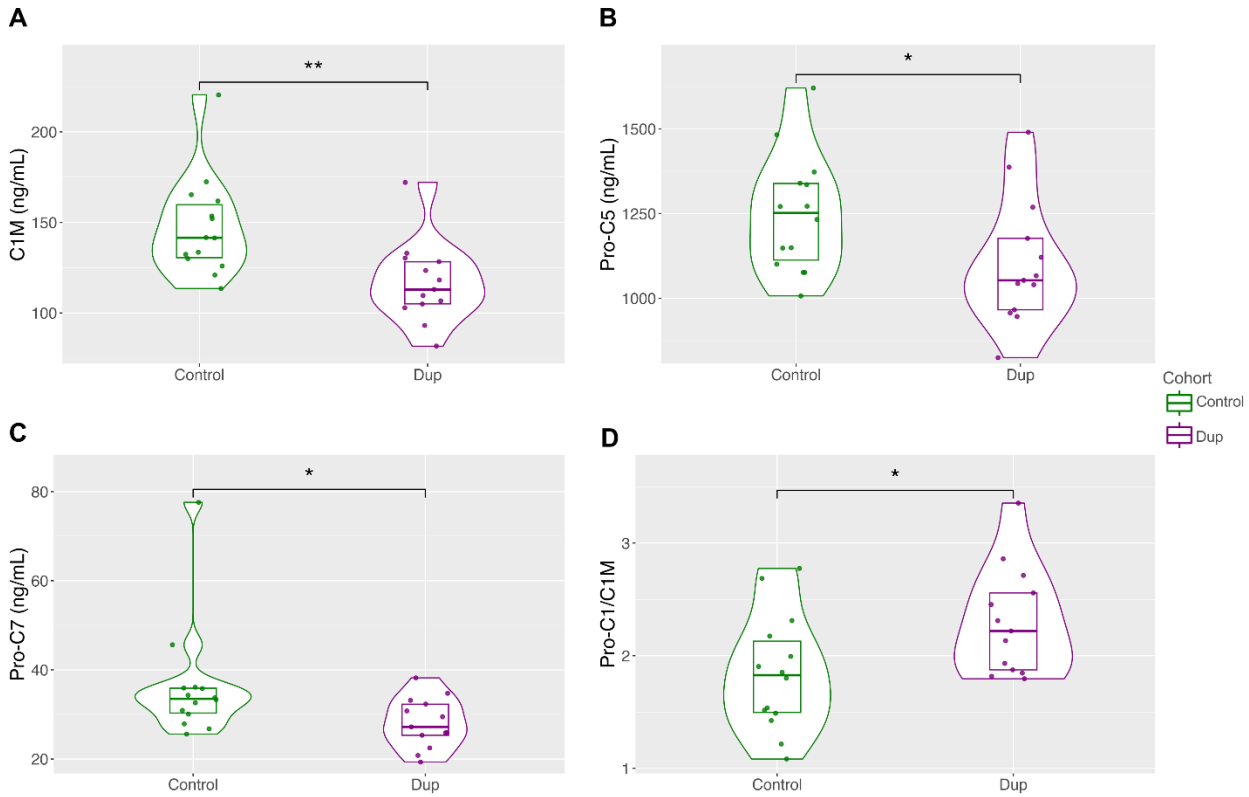

Figure 1. Differential production of Dupuytren vs. control collagen metabolism markers. A) Collagen I degradation marker C1M. B) Collagen I synthesis to degradation marker ratio Pro-C1/C1M. C) Collagen V synthesis marker Pro-C5. D) Collagen VII synthesis marker Pro-C7. * p-value <0.05, **p-value <0.005.

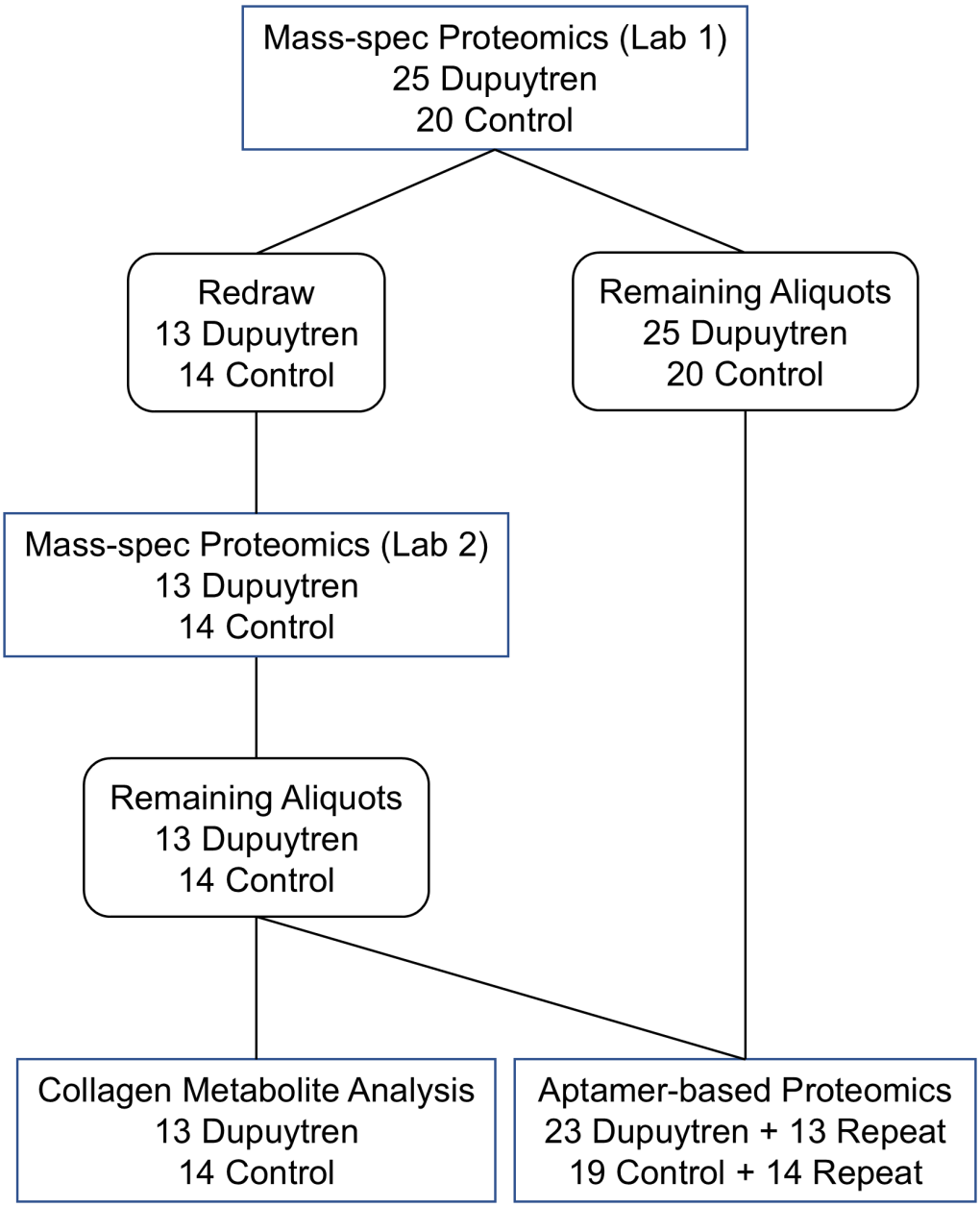

Figure 2. Distribution of samples to laboratories. Samples were initially collected from 45 subjects. Seventeen of these subjects had second samples drawn a minimum of six months after the first draw. Mass spec lab 1 performed analysis on the initial 45 samples. Mass spec lab 2 and collagen metabolite analyses were performed on redraw specimens. Aptamer-based analysis was performed on both initial and redraw specimens. Three of the original 45 samples for Aptamer analysis failed quality control, leaving 23 DD an 19 Controls available for analysis.

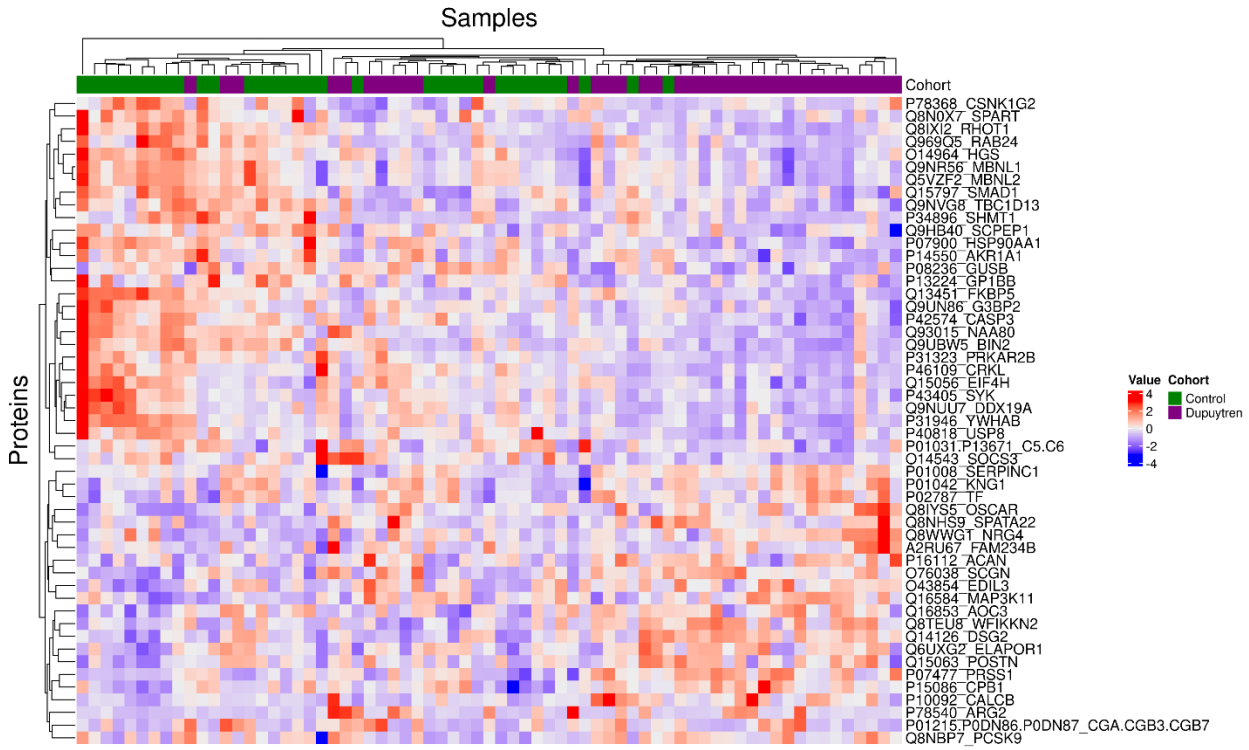

Figure 3. Heatmap of log-transformed standardized values of the 54 significant proteins in the hypothesis-free analysis.

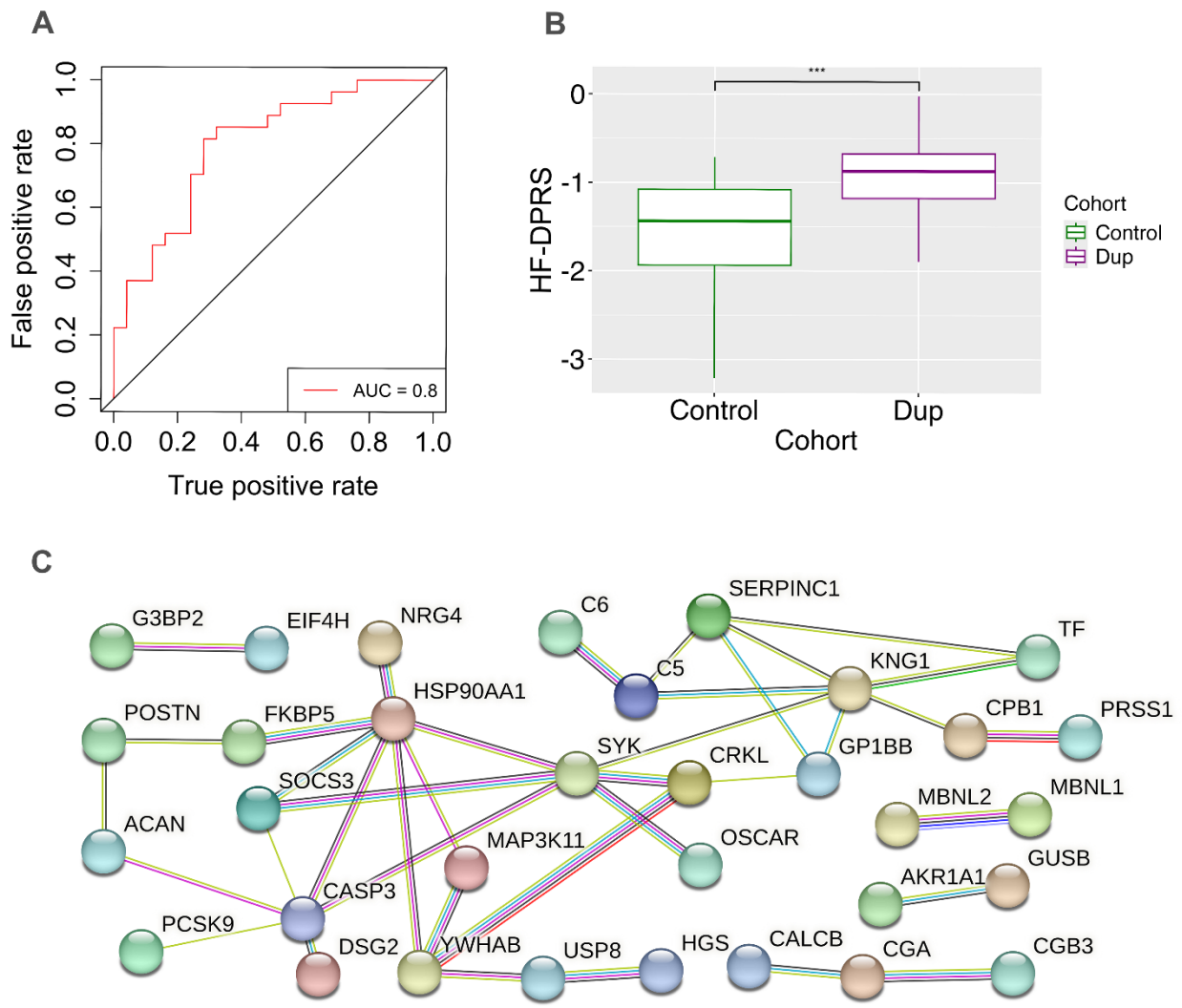

Figure 4. Hypothesis-Free DPRS and Protein-protein interactions. A) The Hypothesis-free Diagnostic Proteomic Risk Score has an AUC-ROC value of 0.8. B) Hypothesis-free Diagnostic Proteomic Risk Scores (HF-DPRS) of the control vs Dupuytren cohorts. p-value = 5.28E-06. C) Network expression of 54 differentially expressed proteins. Of the 54 proteins, 30 had strong pathway connections in the network: ACAN, AKR1A1, C5, C6, CALCB, CASP3, CGA, CGB3, CPB1, CRKL, DSG2, EIF4H, FKBP5, G3BP2, GP1BB, GUSB, HGS, HSP90AA1, KNG1, MAP3K11, MBNL1, MBNL2, NRG4, OSCAR, PCSK9, POSTN, PRSS1, SERPINC1, SOCS3, SYK, TF, USP8, and YWHAB.

	C&C	ECM	ENZ	MBR	PPL	SGN
ACAN		■				
AKR1A1		■				
BIN2		■				
C5	■	■	■			■
C6	■	■				
CALCB		■				■
CASP3			■			■
CGA		■				■
CGB3		■				
CGB7		■				
CPB1		■				
CRKL				■	■	■
CSNK1G2						■
DSG2		■				■
EDIL3		■				
FKBP5		■				■
GUSB		■				
HGS		■				■
HSP90AA1						■
KNG1	■		■			■
MAP3K11						■
NRG4						■
PCSK9				■		
PRKAR2B			■			■
RHOT1						■
SERPINC1	■		■			
SOCS3			■		■	
SYK				■	■	
TF				■		
USP8				■		
WFIKKN2			■			
YWHAB			■		■	

Figure 5. Protein-protein network enrichment categories in the hypothesis-free analysis. 32 of the 54 differentially expressed proteins had significant network enrichment (p.adj <0.05) in the following biological categories. C&C: coagulation and complement cascades, ECM: extracellular matrix, ENZ: Enzyme inhibition, MBR: cell membrane structures, PPL: phosphorylation, SGN: other signaling pathways.

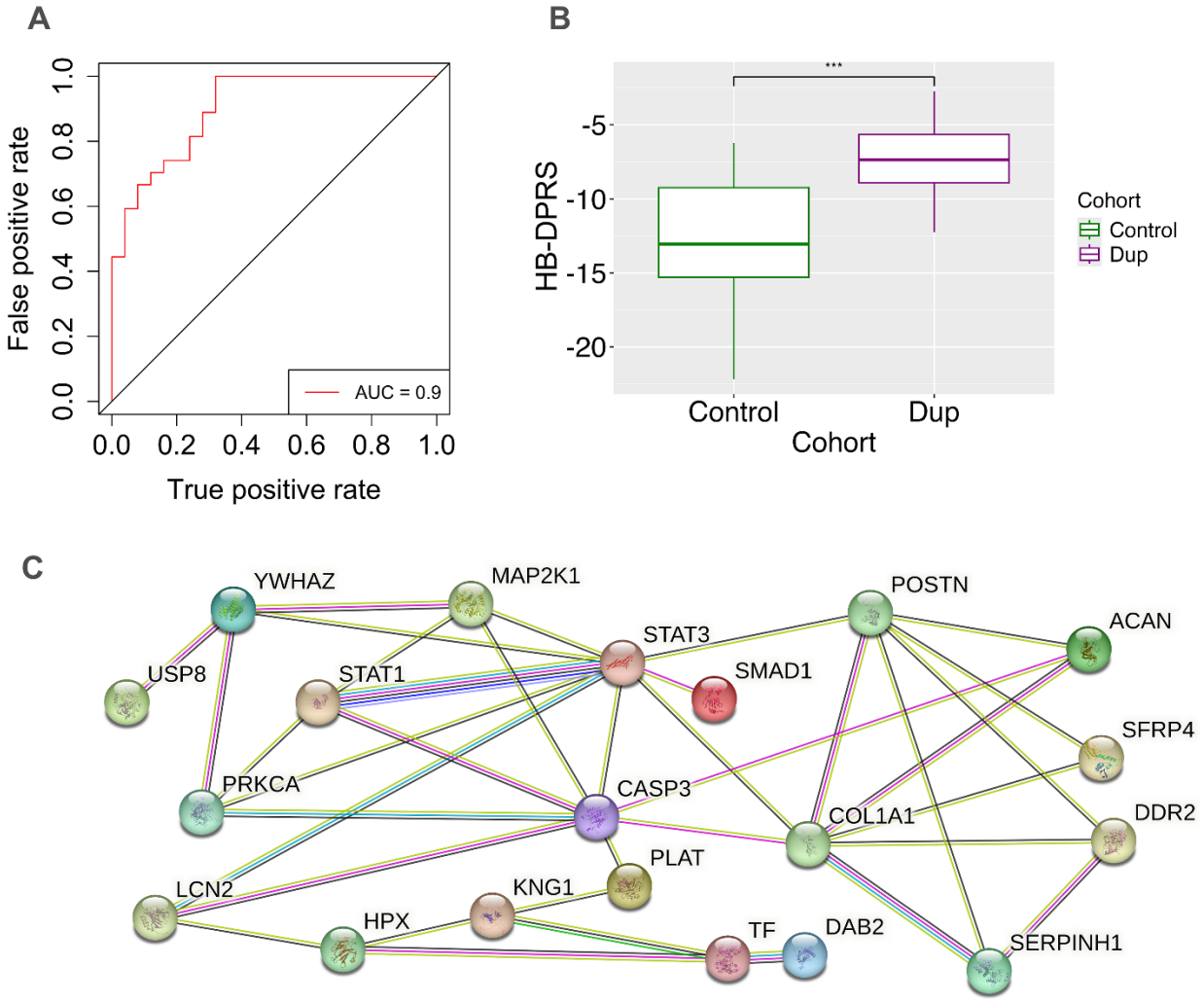

Figure 6. Hypothesis-based DPRS and Protein-protein interactions. A) The Hypothesis-based Diagnostic Proteomic Risk Score has an AUC-ROC value of 0.9. AUC-ROC (Area Under the Curve - Receiver Operating Characteristics) is an index of predictive accuracy. B) Hypothesis-free Diagnostic Proteomic Risk Scores (HF-DPRS) of the control vs Dupuytren cohorts. (p-value = 7.79E-09). C) Network expression of 23 differentially expressed proteins. Of the 23 proteins, 20 had strong pathway connections: DAB2, SMAD1, USP8, PLAT, SFRP4, ACAN, DDR2, HPX, KNG1, LCN2, SERPINH1, TF, MAP2K1, PRKCA, STAT1, YWHAZ, POSTN, COL1A1, CASP3, and STAT3.

	CAN	C&C	ECM	INL	INF	MBR	MFB	PPL	SGN	WNT
ACAN			■							
CASP3	■		■	■	■		■	■	■	
COL1A1	■	■	■		■	■	■	■	■	■
CSNK1G2								■		■
DAB2							■	■		■
DDR2			■			■	■	■		
HPX			■					■	■	
KNG1	■	■	■		■		■			
LCN2				■						
MAP2K1	■		■	■	■	■	■	■	■	
PLAT	■	■					■	■	■	
POSTN			■							
PRKCA	■	■	■		■	■	■	■	■	■
SERPINH1			■							
SFRP4							■			■
SMAD1							■	■		
STAT1	■		■	■	■	■	■	■	■	
STAT3	■		■	■	■	■	■	■	■	
TF	■	■					■	■		
USP8							■	■		■
YWHAZ	■	■		■	■		■	■		■

Figure 7. Protein-protein network enrichment categories in the hypothesis-based analysis. 21 of the 23 differentially expressed proteins had significant network enrichment (p.adj <0.05) in the following biological categories. CAN: Cancer, C&C: coagulation and complement cascades, ECM: extracellular matrix, INL: interleukin interactions, INF: infection, MBR: cell membrane structures, MFB: myofibroblast differentiation, proliferation, and motility: PPL: phosphorylation, SGN: other signaling pathways, WNT: WNT pathways. Hypothesis-based selection bias from a preselected list of candidate proteins may influence the enrichment results.

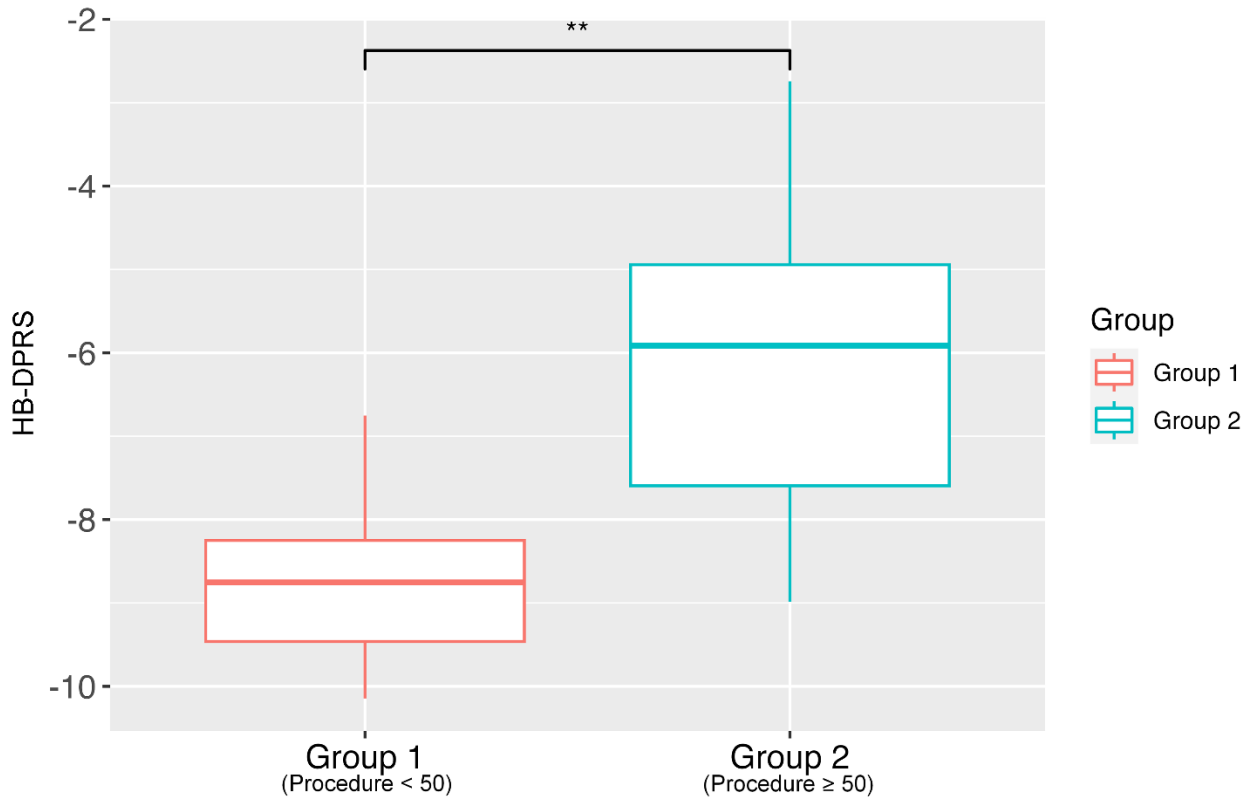

Figure 8. Hypothesis-based DPRS correlates with age at the time of the first Dupuytren procedure. This 11-protein Hypothesis-Based Dupuytren Proteomic Risk Score (HB-DPRS) distinguished DD subjects with different disease progression rates based on whether the subject was younger than 50 vs. 50 or older during their first corrective procedure (p=0.0018).

Table 1. Collagen epitope markers and findings

Assay	D/C	Description	Represents	p-value	SE	FC (D/C)
PRO_C1_HP	ns	Internal epitope in the N-terminal propeptide of type I collagen	Col I Synthesis	0.9437	C 10.99 D 11.89	0.93
C1M_HP	Down**	MMP generated fragment of the a1 chain of type I collagen	Col I Degradation	0.0036	C 5.25 D 4.30	0.73
PRO-C2_HP	ns	N-terminal fragment of type IIB procollagen	Col II Synthesis	0.7808	C 1.45 D 3.11	1.01
CALC2	ND	C-terminal propeptide of type II collagen	Col II Synthesis	NA	NA	NA
C2M_HP	ns	MMP generated fragment of type II collagen	Col II Degradation	0.1607	C 1.52 D 1.16	1.29
PRO-C3_roHP	ns	N-terminal propeptide of type III collagen	Col III Synthesis	0.9951	C 1.04 D 1.24	0.93
C3M_HP	ns	MMP-9 generated fragments of type III collagen	Col III Degradation	0.1615	C 16.96 D 13.61	0.86
CTX-III_HP	ns	MMP generated fragment of crosslinked type III collagen	Col III Degradation	0.4887	C 40.88 D 53.26	0.91
PRO-C4_HP	ns	Internal epitope in the 7S domain of type IV collagen	Col IV Synthesis	0.0739	C 195.98 D 183.12	0.84
C4M_HP	ns	MMP generated fragment of the a1 chain of type IV collagen	Col IV Degradation	0.1014	C 8.89 D 7.42	0.84
PRO-C5_HP	Down*	C-terminal propeptide of type V collagen	Col V Synthesis	0.0458	C 33.17 D 35.88	0.82
PRO-C6_roHP	ns	C-terminal propeptide of type VI collagen alpha-3 chain	Col VI Synthesis	0.7222	C 0.27 D 0.27	0.92
C6M_HP	ns	MMP2 generated fragment of type VI collagen	Col VI Degradation	0.1097	C 22.31 D 23.41	0.83
PRO_C7 HP	Down*	BMP-I mediated neo-epitope in the NC-2 propeptide of type VII collagen	Col VII Synthesis	0.0480	C 2.49 D 1.08	0.72
C7M	ND	Neo-epitope of MMP13 degradation of type VII collagen	Col VII Degradation	NA	NA	NA
PRO-C9	ns	C-terminus of NC-I domain of the a1 chain of type IX collagen	Col IX Synthesis	0.9658	C 0.80 D 0.65	0.92

Synthesis and degradation markers of Collagen I, II, III, IV, V, VII, and IX. Compared to controls, there were significant reductions in markers of Collagen I degradation, Collagen V synthesis, and Collagen VII synthesis. D/C: ratio of Dupuytren to control values. FC: fold change. **: <0.005, *: <0.05. ns: not significant. ND: undetectable in both disease and control cohorts.

Table 2. Hypothesis-free Diagnostic Proteomic Risk Score

Protein	Eff	SE	p-value	FDR
Spartin	-11.62125	2.97479	0.000094	0.18
Oscar	13.99037	3.596067	0.0001	0.18

Proteins of the Hypothesis-free DPRS. Eff = log (odds ratio comparing Dup to control) for a change of 1 in the log (protein data). SE: Standard error. FDR: False Discovery Rate (Adjusted p-value).

Table 3. Hypothesis-based Diagnostic Proteomic Risk Score

Protein	Eff	SE	p-value	FDR
SMAD1*	-6.3649	1.877804	0.0007	0.12
CSMD1	5.397535	1.907423	0.00466	0.18
SH3BP2	-4.59151	1.582341	0.00371	0.18
POSTN*	4.305245	1.502627	0.00417	0.18
ACAN*	5.589134	1.75746	0.00147	0.12
LCN2	-4.73895	1.734879	0.0063	0.19
AOC3*	7.050175	2.62144	0.00716	0.2
DDR2	1.63418	0.595212	0.00604	0.19
USP8*	-3.30095	1.041269	0.00152	0.12
CSNK1G2*	-10.3331	3.593633	0.00404	0.18
SERPINH1	-2.6813	0.799538	0.0008	0.12

Proteins of the Hypothesis-based DPRS. Eff = log (odds ratio comparing Dup to control) for a change of 1 in the log (protein data). SE: Standard error. FDR: False Discovery Rate (Adjusted p-value). * Five also had significant differential expression in the 54 proteins identified in the hypothesis-free analysis.

Table 4. Mass Spectrometry (MS) – SomaScan (SS) concordance

Protein	Est MS	p-val MS	FDR MS	Est SS	p-val SS	FDR SS
AOC3	0.214854	0.076413	0.232677	0.116624	0.000845	0.217972
POSTN	0.410212	0.098842	0.232677	0.182892	0.000312	0.211479
SERPINC1	0.071331	0.116338	0.232677	0.091474	0.001500	0.225251
TF	0.035200	0.567393	0.829196	0.076715	0.000025	0.172072
PCSK9	0.165811	0.755318	0.829196	0.185675	0.001371	0.222267
KNG1	0.010065	0.829196	0.829196	0.065561	0.000490	0.217972

Six proteins identified by MS were also in the 54 proteins found in the hypothesis-free SS analysis. AOC3, POSTN, SERPINC1, TF, PCSK9, and KNG1. Of these, AOC3, POSTN, and SERPINC1 had False Discovery Rate (FDR, Adjusted p-value) <0.25 in both Mass Spec and SomaLogic analyses.