Supplemental data to

Hepatic molecular signatures highlight the sexual dimorphism of Non-Alcoholic SteatoHepatitis (NASH).

by

Jimmy Vandel¹, Julie Dubois-Chevalier¹, Céline Gheeraert¹, Bruno Derudas¹, Violetta Raverdy², Dorothée Thuillier², Luc Van Gaal³, Sven Francque⁴, François Pattou², Bart Staels¹, Jérôme Eeckhoute¹ and Philippe Lefebvre¹,*

¹ Univ. Lille, Inserm, CHU Lille, Institut Pasteur de Lille, U1011-EGID, Lille, France
² Univ. Lille, Inserm, CHU Lille, U1190-EGID, Lille, France.
³ Department of Endocrinology, Diabetology and Metabolism, Antwerp University Hospital, Edegem (Antwerp), Belgium.
⁴ Department of Gastroenterology and Hepatology, Antwerp University Hospital, Edegem (Antwerp), Belgium.
⁵ Laboratory of Experimental Medicine and Paediatrics (LEMP), University of Antwerp, Wilrijk (Antwerp), Belgium.

* Correspondence: Dr P. Lefebvre, philippe-claude.lefebvre@inserm.fr
SUPPLEMENTARY METHODS

HUL cohort constitution.

Severely (BMI>35) or morbidly (BMI>40) obese patients were referred to the HUL bariatric surgery unit for evaluation for bariatric surgery (BS). All patients fulfilling the following criteria for bariatric surgery were prospectively included in the HUL cohort on the day of BS: 18 years or older at time of evaluation and meeting the criteria for BS according to French national guidelines: morbid or severe obesity with at least one comorbidity factor (i.e. arterial hypertension or diabetes mellitus) for at least 5 years and resistance to medical treatment; absence of medical or psychological contraindications for BS; social security insurance coverage; no current excessive drinking (average daily consumption of alcohol <20 g/d for women and <30 g/d for men), and no past excessive drinking for a period longer than 2 years at any time in the last 20 years; absence of long-term consumption of hepatotoxic drugs; negative screening for chronic liver disease. Informed written consent was obtained from all patients and the study was conducted in conformity with the Helsinki Declaration. The ethics committee approved the cohort and was supported by grants by the government and the French Ministry of Health (PHRC). After legal revision, a new approval was obtained in 2006 (n° CP06/49; NCT01129297).

Biopsy procedure: the indication for BS was confirmed after an extensive multidisciplinary preoperative evaluation, according to current French guidelines. Liver biopsies were systematically planned during the surgical procedure. A liver needle biopsy was performed during the first part of the surgical procedure after trocar insertion and abdominal exploration, within 10 minutes after pneumo-peritoneum installation. The MONOPTY needle biopsy system (16G, ref: 121620; C. R. Bard, Tempe, AZ) was used. Biopsies were routinely stained with H&E Saffron and Masson’s trichrome, Sirius Red, and Perl’s staining. Two pathologists were blinded to clinical and biological data and independently graded steatosis (0-3), lobular inflammation (0-3), and ballooning (0-2). Liver fibrosis was assessed using the Kleiner fibrosis score (F0, normal; F1 with mild or moderate pericellular fibrosis in zone 3 or portal fibrosis; F2, perivenular and pericellular fibrosis confined to zones 2 and 3, with or without portal or periportal fibrosis; F3, bridging or extensive fibrosis with architectural distortion and no clear-cut cirrhosis; and F4, cirrhosis).
Cohort stratification.

NASH was defined according to the following histological parameters: steatosis > 5%, lobular inflammation and ballooning > 0. Absence of NASH was defined by inflammation and ballooning scores =0, independently of the steatosis grade. Both NAFL (steatosis > 5%) and healthy liver (HL; steatosis ≤ 5%) patients were thus included in the "NoNASH" group. Patients from HUL and UZA cohorts were assigned to NASH and NoNASH groups using available histological data. The original stratification of the UKD cohort in HO, NAFL and NASH groups was left unchanged. NAFL and HO patients were thus included in the "NoNASH" group, whereas the NASH group was as described. The DU cohort, classified on the basis of the fibrosis grade, hindered a direct comparison with other cohorts.

Data pre-processing

Pre-processing of expression data was performed prior to differential analysis as detailed below. As a general approach, transcripts from collected datasets were first annotated, then corrected to remove experimental bias between cohorts. Finally, patients from each cohort were stratified as NASH and NoNASH individuals based on available histological parameters and criteria set for the HUL cohort.

Gene annotation: Gene annotations for the 4 datasets were directly imported from the corresponding ThermoFisher array web pages. The most recent annotation files were used (release 36). For each dataset, signals detected for multiple probes sharing the same gene annotation were averaged to generate mean gene expression values.

Experimental bias correction: Experimental bias between the 4 datasets was removed using COMBAT from the sva R package (v3.26.0), a process in which each dataset was considered as a single batch. Additionally, COMBAT models were adjusted with sex as a variable for HUL, UZA and UKD datasets to define sex-specific signatures. The HUL cohort was used as the reference dataset to correct separately experimental biases in the 3 other datasets.

Propensity matching.
Propensity matching was carried out using the Mahalanobis distance optimization method from the Matching R package (v4.9-3, maximum distance threshold = 1.5) to match highly similar patients.

**Signature definition**

As a general approach, signatures were built in 2 steps. First, DE gene sets were identified, then RF models were used to select a signature from each DE gene set. The procedure is detailed below:

**Differential analysis:** The differential analysis between NoNASH and NASH patients was performed using the Limma R package (v3.34.9) that computes moderated t-tests (1). The Limma model was designed using the NASH/NoNASH histological status, sex and insulin resistance factors in interaction. Additionally, to avoid potential confounding effects, pharmacological treatments (metformin and statins) and technical (batch effect) information were included in the Limma model as additive factors. Multiple testing correction was applied to resulting Limma statistics using a False Discovery Rate approach (FDR) (2). Three contrasts were evaluated by Limma, corresponding to a NASH vs NoNASH comparison as a function of sex or not.

To assess the stability of the differential analysis, a bootstrap procedure was applied (3). One hundred sub-populations were randomly sampled from the learning cohort using a 0.9 selection rate, then an independent differential analysis was performed on each sub-population. DE genes between NASH and NoNASH conditions with a FDR< 10% detected in at least 75% of the 100 sub-populations were considered as "reliable". Thus, 3 gene sets composed of reliable DE genes across sub-populations were obtained from the men contrast (termed "G\text{men}"), the women contrast (termed "G\text{women}") and the men+women contrast (termed "G\text{all}").

A gene signature was defined through RF modeling of G\text{men}, G\text{women} and G\text{all} with a Recursive Feature Elimination (RFE) strategy (4, 5). RF model is commonly used to classify samples based on observations of a feature set (gene set here), whereas RFE is a generic strategy to identify an optimal set of features (genes) to reach a better classification. In this study, these 2 approaches were used in tandem to identify the best subset of genes allowing to detect NoNASH vs NASH patients based on liver gene differential expression. Briefly, signatures were built in two steps for each G\text{x}. First, genes from G\text{x} were ranked based on their individual classification power computed on the learning cohort. Second, an optimal number of genes was defined in an incremental manner
following gene ranking and guided by the global classification power. By applying this procedure to the $G_{men}$, $G_{women}$ and $G_{all}$ gene sets, transcriptomic signatures referred below as “reference signatures” and termed $S_{men}$, $S_{women}$ and $S_{all}$ were identified. The main steps of the whole procedure are summarized in Figure 1.

Random signatures: For each identified reference signature ($S_{men}$, $S_{women}$, $S_{all}$), 200 guided random signatures were generated encompassing an identical number of genes randomly selected from the corresponding “reliable” gene set ($G_{men}$, $G_{women}$ and $G_{all}$ respectively). Additionally, 200 unguided random signatures were generated from the full list of annotated genes in the HUL dataset. These random signatures were evaluated using criteria similar to those used for reference signatures (see below).

Evaluation of signatures.

The reliability of reference signatures was evaluated using several methods:

AUC: Distinct RF models were learnt for each reference or random signatures to classify patients as NASH or NoNASH. Like any supervised approach, learning RF model requires a training set with known patient NASH status and a testing set to evaluate the predictive power of the model. Two training/testing set selection strategies were used to evaluate RF models. First, a cross-validation scheme was used to train and evaluate RF models using the learning cohort. The training set was randomly sampled to include 75% of the learning cohort and 200 iterative runs were performed to evaluate the classification power of the RF model against the testing set, made of the remaining 25% of the learning cohort. Second, RF models were trained using the entire learning cohort and evaluated against independent cohorts. Four different validating cohorts were considered: HUL patients not belonging to the learning cohort and the other 3 independent cohorts (UZA, UKD and DU). For both selection strategies, 3 RF models were trained from $S_{men}$, $S_{women}$ and $S_{all}$ genes expression in men, women and all patients of the training set respectively. Then each of these RF models, independently of the associated signature, was evaluated to classify successively men, women and all patients of the corresponding testing set. Thus a RF model learnt from $S_{women}$ gene expression patterns in women of a training set was evaluated to classify successively men, women and all patients of the associated testing set. Patient sex being unknown in the DU cohort, RF models were only evaluated to classify directly all patients
independently of sex. The classification power of RF models was measured using the Area Under the Receiver Operating Characteristic (ROC) Curve (AUC) metric.

Univariate classifier: RF models learnt from reference signatures were compared to "single gene" classifiers. This approach is based on the use of a single gene expression level as a metric to separate NASH from NoNASH patients. AUC was computed similarly to the RF model prediction, allowing the comparison of the predictive power of each individual gene vs reference signatures.

**Biological meaning**

Selected gene subsets were enriched for biological terms by scanning the Gene Ontology Biological Processes database (BP Direct GO) using DAVID (v 6.8) (6).
BIBLIOGRAPHY

1. Smyth GK: limma: Linear Models for Microarray Data. In: R. G, V.J. C, W. H, R.A. I, S. D, eds. Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Statistics for Biology and Health. New York, NY: Springer, 2005; 397-420.
2. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological) 1995;57:289-300.
3. Efron B, Tibshirani RJ. An Introduction to the Bootstrap: Chapman and Hall/CRC, 1994.
4. Breiman L. Random Forests. Machine Learning 2001;45:5-32.
5. Guyon I, Weston J, Barnhill S, Vapnik V. Gene Selection for Cancer Classification using Support Vector Machines. Machine Learning 2002;46:389-422.
6. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44-57.
Supplemental Figure 1: Flow chart for biopsy selection and HUL cohort stratification. The successive steps to constitute a validated cohort based on biopsy, transcriptomic and histological quality controls are described. N/A: not available; n: number of biopsies; Lob. Inflam.: lobular inflammation. Numbers between brackets separated by commas indicate biopsy numbers for each histological grade (steatosis, 0 to 3; ballooning, 0 to 2; lobular inflammation, 0 to 2).
Supplemental Figure 2: *Bootstrapped LIMMA for all patients*. The number of DEGs between NoNASH and NASH patients (FDR<10%) for all patients was assessed after 100 bootstrap executions. The mean value of DEGs is represented by a black dotted line.
Supplemental Figure 3: **Identification of reliable DE genes.** The absolute log₂ FC of DEGs was computed for the women (A) and all (B) learning cohort. Each significantly DEG (FDR<10%), is represented by a red dot. Gene reliability is represented through the number of bootstrap runs in which the gene remains significantly DE. Blue dots, represent the mean log₂(FC) for a given bootstrap run count. Dashed line: FC=1.5; dotted line: occurrence=75. The grey area represents reliable genes with occurrences ≥ 75.
Supplemental Figure 4: Variance of gene expression. The distribution of gene expression standard deviation in each sub-group of the HUL learning cohort (NASH/NoNASH; men or women) was calculated. Only genes belonging to $G_{\text{women}} \cup G_{\text{men}}$ were considered (n=1341). Kolmogorov-Smirnov two-sided test was applied to measure distribution difference between men and women successively in NASH and NoNASH groups. No significant difference (p-value<0.05) was observed neither in NASH (p-value=0.086) or NoNASH (p-value=0.949) groups.
Supplemental Figure 5: Overlap between differentially expressed genes in HUL sub-cohorts. We defined for each contrast (NASH vs NoNASH for women, men and all patients) DEGs with FDR < 10% and absolute log₂FC > log₂(1.5) in at least 75% of 100 bootstrapped Limma executions (selection rate=0.9). Men (blue), women (red) and all (yellow).
Supplemental Figure 6: Classification power (AUC) of RF models. RF were trained with a progressively reduced number of genes to identify an optimal subset of genes corresponding to the proposed signature for all patients, established by the second step of the RFE strategy. The red dotted line indicates the optimal number of genes yielding the highest AUC.
Supplemental Figure 7: Principal Component Analysis. PCA applied to men (left) and all patients (right) from learning cohort based on expression of all genes (A,C) and $s_{\text{men}}$ or $s_{\text{all}}$ genes (B,D). The percentage of global data variance explained by each component is indicated in axis labels (%var.). Each dot represents NoNASH (blue) or NASH (yellow) patient.
Supplemental Figure 8: Classification of the full learning cohort. (A) AUC distribution of RF models to predict all patients of the learning cohort in a cross-validation scheme. RF models learnt using $\mathcal{S}_{\text{all}}$ (red) were compared to RF models learnt using random signatures built from $\mathcal{G}_{\text{women}}$ (khaki), $\mathcal{G}_{\text{men}}$ (green), $\mathcal{G}_{\text{all}}$ (blue) and the full list of available genes (purple). Distribution means are represented as vertical dashed lines.
Supplemental Figure 9: AUC of single gene predictors. The ability of single genes composing $S_{all}$ to predict all patients of the learning cohort was evaluated. Mean AUC reached by RF model learnt from $S_{all}$ in a cross-validation scheme is represented by a red horizontal dashed line.
Supplemental Figure 10: *Single gene-based prediction using highest deregulated genes.*
The individual prediction power of highest deregulated genes in female, male and all patients was measured. The AUC of single gene predictors to predict women (A), men (B) and all patients (C) of the learning cohort was computed. Mean AUC reached by RF models learnt from the signature composed of these genes in a cross-validation scheme is represented through a red horizontal dashed line.
Supplemental Figure 11: Correlation networks. Pearson correlation ($\rho$) was computed for all pairs of genes belonging to $S_{\text{women}}$, $S_{\text{men}}$ and $S_{\text{all}}$. Resulting correlation networks for $S_{\text{women}}$ (A), $S_{\text{men}}$ (B) and $S_{\text{all}}$ (C) genes were drawn by keeping links when $\text{abs}(\rho) \geq 0.5$. Correlation intensity is indicated through a color gradient going from yellow ($\rho = \pm 0.5$) to red ($\rho = 1$) for positive correlations, and yellow to green ($\rho = -1$) for negative correlations.
Supplemental Figure 12: **Single gene-based prediction using highest deregulated genes for men cohorts prediction.** The individual prediction power of highest deregulated genes in the HUL men group was measured. AUC of corresponding single gene predictors to stratify men from the HUL (A), UZA (B), UKD (C) cohorts and all patients from the DU cohort (D) was computed. AUC reached by RF model learnt from the signature composed of these genes is represented by a red horizontal dashed line.
Supplemental Table 1. Characteristics of 420 NASH/NoNASH patients with quality-checked biopsies from the HUL cohort.

| Characteristics          | HL n=78 | NAFL n=274 | NASH n=68 |
|--------------------------|---------|------------|-----------|
|                           | Women (n; %) | 66; 85% | 200; 73% | 41; 60% |
| Biometric parameters     | Age (mean±sd) | 34.8±11 | 41.6±11 | 47.2±10 |
|                          | BMI (mean±sd) | 46.3±6 | 47.8±8 | 46.8±8 |
|                          | Body mass (mean±sd) | 129.4±23 | 134.6±25 | 133.5±28 |
| Liver histology          | Steatosis grade (n; %) | 0; 100% | 0; 0% | 0; 0% |
|                          | 1; 0% | 209; 76% | 14; 21% |
|                          | 2; 0% | 44; 16% | 28; 41% |
|                          | 3; 0% | 21; 8% | 26; 38% |
|                          | Lobular inflammation (n; %) | 78; 100% | 274; 100% | 0; 0% |
|                          | 1; 0% | 0; 0% | 47; 69% |
|                          | 2; 0% | 0; 0% | 21; 31% |
|                          | Ballooning (n; %) | 78; 100% | 274; 100% | 0; 0% |
|                          | 1; 0% | 0; 0% | 47; 69% |
|                          | 2; 0% | 0; 0% | 21; 31% |
|                          | Fibrosis (Kleiner) (n; %) | 69; 88% | 211; 77% | 8; 12% |
|                          | 1a; 2% | 7; 3% | 8; 12% |
|                          | 1b; 1% | 2; 1% | 9; 13% |
|                          | 1c; 4% | 26; 9% | 4; 6% |
|                          | 2; 0% | 10; 4% | 11; 16% |
|                          | 3q; 0% | 5; 2% | 12; 18% |
|                          | 3s; 0% | 4; 1% | 10; 15% |
|                          | 4; 0% | 0; 0% | 3; 4% |
| Liver functions          | AST (IU/L)(median; IQR) | 21; 8 | 22; 9 | 38; 21 |
|                          | ALT (IU/L)(median; IQR) | 20; 10 | 27; 16 | 47.5; 30 |
|                          | GGT (IU/L)(median; IQR) | 23.5; 21 | 31; 25 | 57; 40 |
| Metabolic parameters     | Diabetes (n; %) | 10; 13% | 87; 31% | 58; 85% |
|                          | Treated diabetes (n; %) | 10; 13% | 73; 27% | 52; 76% |
|                          | Fasting blood glucose (mM)(mean±sd) | 5.4±1.0 | 6.4±2.4 | 9.3±3.3 |
|                          | Fasting insulin (IU/mL)(median; IQR) | 12.5; 8.8 | 14.2; 11.2 | 23.2; 25.9 |
|                          | HbA1c (%)(median; IQR) | 5.4; 0.6 | 5.8; 0.8 | 7.8; 3.6 |
|                          | HOMA-IR (median; IQR) | 2.8; 2.2 | 3.6; 3.1 | 9.2; 11.9 |
|                          | Total cholesterol (mmol/L)(mean±sd) | 4.8±0.9 | 5.0±0.9 | 4.7±1.0 |
|                          | LDL cholesterol (mmol/L)(mean±sd) | 3.1±0.8 | 3.1±0.8 | 2.8±0.9 |
|                          | HDL cholesterol (mmol/L)(mean±sd) | 1.1±0.2 | 1.1±0.2 | 1.0±0.2 |
|                          | Triglycerides (mmol/L)(mean±sd) | 1.2±0.4 | 1.7±1.1 | 2.0±1.0 |
| Others                   | Diastolic blood pressure (mmHg)(mean±sd) | 73.0±13 | 77.4±14 | 77.0±12 |
|                          | Systolic blood pressure (mmHg)(mean±sd) | 130.3±16 | 137.3±20 | 139.9±19 |
**Supplemental Table 2.** Characteristics of 170 patients of the HUL learning cohort.

| Characteristics                          | HL n=16 | NAFL n=108 | NASH n=46 |
|------------------------------------------|---------|------------|-----------|
| **Biometric parameters**                 |         |            |           |
| Women (n; %)                             | 10; 62% | 52; 48%    | 23; 50%   |
| Age (mean±sd)                            | 42.6±13 | 43.6±11    | 47.9±10   |
| BMI (mean±sd)                            | 47.8±7  | 48.4±7     | 46.7±6    |
| (mean±sd) Body mass                      | 138.8±26| 140.6±25   | 135.3±24  |
| **Liver histology**                      |         |            |           |
| Steatosis grade (n; %)                   | 16; 100%| 0; 0%      | 0; 0%     |
| 1                                        | 0; 0%   | 77;71%     | 11; 24%   |
| 2                                        | 0; 0%   | 22;20%     | 19; 41%   |
| 3                                        | 0; 0%   | 9;8%       | 16; 35%   |
| Lobular inflammation (n; %)              | 16; 100%| 108; 100%  | 0; 0%     |
| 1                                        | 0; 0%   | 0; 0%      | 32; 70%   |
| 2                                        | 0; 0%   | 0; 0%      | 14; 30%   |
| Ballooning (n; %)                        | 16; 100%| 108; 100%  | 0; 0%     |
| 1                                        | 0; 0%   | 0; 0%      | 37; 80%   |
| 2                                        | 0; 0%   | 0; 0%      | 9; 20%    |
| Fibrosis (Kleiner) (n; %)                | 13; 81% | 79; 73%    | 8; 17%    |
| 1a                                       | 2; 13%  | 4; 4%      | 7; 15%    |
| 1b                                       | 0; 0%   | 0; 0%      | 8; 17%    |
| 1c                                       | 1; 6%   | 12; 11%    | 2; 4%     |
| 2                                        | 0; 0%   | 8; 7%      | 8; 17%    |
| 3q                                       | 0; 0%   | 4; 4%      | 9; 20%    |
| 3s                                       | 0; 0%   | 1; 1%      | 4; 9%     |
| 4                                        | 0; 0%   | 0; 0%      | 0; 0%     |
| **Liver functions**                      |         |            |           |
| AST (IU/L)(median; IQR)                  | 22; 8   | 23.5; 11   | 38; 21    |
| ALT (IU/L)(median; IQR)                  | 19.5; 5 | 28.5; 20   | 45.5; 29  |
| GGT (IU/L)(median; IQR)                  | 37.5; 20| 32; 25     | 54.5; 37  |
| **Metabolic parameters**                 |         |            |           |
| Diabetes (n; %)                          | 4; 25%  | 43; 40%    | 40; 87%   |
| Treated diabetes (n; %)                  | 5; 31%  | 37; 34%    | 35; 76%   |
| Fasting blood glucose (mM)(mean±sd)      | 6.1±1.6 | 6.6±2.1    | 9.2±3.3   |
| Fasting insulin (IU/mL)(median; IQR)     | 18.4; 6.7| 15.8; 10.2| 24.3; 21.7|
| HbA1c (%) (median; IQR)                  | 5.7; 0.6| 6.0; 0.9   | 7.7; 3.4  |
| HOMA-IR (median; IQR)                    | 4.6; 1.3| 4.5; 3.7   | 9.6; 10.9 |
| Total cholesterol (mmol/L)(mean±sd)      | 4.6±1.1 | 4.9±0.9    | 4.7±1.0   |
| LDL cholesterol                          | 2.8±0.9 | 3.0±0.8    | 2.8±0.9   |
| (mmol/L)(mean±sd)                        | 1.1±0.2 | 1.1±0.2    | 1.0±0.2   |
| HDL cholesterol                          | 1.4±0.5 | 1.9±1.4    | 2.0±0.8   |
| Triglycerides (mmol/L)(mean±sd)          |         |            |           |
| **Others**                               |         |            |           |
| Diastolic blood pressure (mmHg)(mean±sd) | 76.9±17 | 78.7±15    | 76.5±13   |
| Systolic blood pressure (mmHg)(mean±sd)  | 137.3±19| 140.0±21   | 139.2±17  |
### Supplemental Table 3. List of reliable genes with absolute log₂FC > \log(1.5) for men, women and all patients.

| Women         | Men          | All           |
|---------------|--------------|---------------|
| FABP5P7       | DEFB1        | DEFB1         |
| CXCL10        | KPNA2        | KPNA2         |
| CHST9         | WIPI1        | FAT1          |
| CYP2C19       | FAT1         | KRT18         |
| SPP1          | KRT18        | AJUBA         |
| FABP4         | MIR622       | MIR622        |
| AKR1B10       | FABP5P7      | FABP5P7       |
| FABP5P1       | HSPA4L       | HSPA4L        |
| HYDIN2        | CXCL10       | CXCL10        |
| HYDIN         | SLC25A33     | EFEMP1        |
| GPNMB         | ABCB4        | PLP2          |
| ANXA2P2       | EFEMP1       | CHST9         |
| ANXA2         | PLP2         | CYP2C19       |
| OLR1          | CHST9        | SPP1          |
| UBD           | HYOU1        | PRAMEF10      |
| HSPA5         | HMGCR        | FABP4         |
| CYP2C19       | DBH-AS1      | LRRC19        |
| SPP1          | IL32         | THY1          |
| PRAMEF10      | LAPTMM5      | ZMAT3         |
| LOC101928961  | VN1          | AKR1B10       |
| FABP4         | HSD17B7P2    | CCL2          |
| THY1          | IFI30        | FABP5P1       |
| ZMAT3         | SNX10        | CCL20         |
| AKR1B10       | CHI3L1       | HYDIN2        |
| ANXA5         | HSD17B7      | HYDIN         |
| FABP5P1       | ANXA2        | LYZ           |
| CCL20         | OLR1         | ME1           |
| SULF2         | MSMO1        | LUM           |
| LYZ           | SQLE         | LOC101926960  |
| LOC101928714  | BIRC3        | LINC00890     |
| PEPD          | GSTA7P       | GPNMB         |
| ME1           | FBP1         | DTNA          |
| LEAP2         | TNFAIP3      | CCND1         |
| LOC730101     | UBD          | EMP1          |
| PLIN2         | SULT1B1      | MT1F          |
| LUM           | MT1M         | HPS5          |
| PDE11A        | CYR61        | ITGBL1        |
| VCAN          | HMGCS1       | ID1           |
| SLC22A10      | CYFIP2       | MT1H          |
| GPNMB         | LDLR         | ANXA2P2       |
**Supplemental Table 4.** Top 10 gene ontology enrichments for reliable DEG sets with absolute log$_2$FC $>$ log$_2$(1.2) for men, women and all patients.

| GO terms                                           | Gene set               | Men (637 genes) |       | Women (41 genes) |       | All (454 genes) |       |
|---------------------------------------------------|------------------------|-----------------|-------|------------------|-------|------------------|-------|
|                                                   | rank                   | p-value         |       | rank             | p-value | rank             | p-value |
| Cholesterol biosynthetic process                  | 1                      | 2.5x10^{-4}***  |       | 2                | 6.5x10^{-9}*** |       |                   |       |
| Single organismal cell-cell adhesion             | 2                      | 4.4x10^{-6}***  |       | 3                | 3.9x10^{-4}  |       |                   |       |
| Cell-cell adhesion                               | 3                      | 2.5x10^{-5}     |       | 4                | 1.6x10^{-5}  |       |                   |       |
| Negative regulation of apoptotic process         | 4                      | 3.3x10^{-5}     |       | 5                | 5.1x10^{-4}  |       |                   |       |
| Hepatocyte apoptotic process                     | 5                      | 3.6x10^{-5}     |       | 6                | 2.6x10^{-3}  |       |                   |       |
| Leukocyte migration                              | 7                      | 3.9x10^{-5}     |       | 8                | 2.9x10^{-4}  |       |                   |       |
| Response to unfolded protein                     | 8                      | 5.0x10^{-5}     |       | 9                | 2.9x10^{-7}***|       |                   |       |
| ER to Golgi vesicle-mediated transport           | 9                      | 6.2x10^{-5}     |       | 10               | 6.3x10^{-5}** |       |                   |       |
| Retrograde vesicle-mediated transport, Golgi to ER | 10                     | 6.3x10^{-5}**   |       |                   |         |                   |       |
| Response to organonitrogen compound              | ∅                      | ∅               |       | 1                | 3.4x10^{-4}  | 139   | 3.7x10^{-2}       |       |
| Triglyceride catabolic process                   | ∅                      | ∅               |       | 2                | 1.6x10^{-3}  |       |                   |       |
| Cell adhesion                                    | 16                     | 2.1x10^{-4}**   |       | 3                | 3.9x10^{-9}***|       |                   |       |
| Response to drug                                 | 8                      | 5.4x10^{-5}     |       | 4                | 2.8x10^{-3}  |       |                   |       |
| Intestinal epithelial cell maturation            | 52                     | 6.0x10^{-3}     |       | 5                | 2.8x10^{-3}  |       |                   |       |
| G1/S transition of mitotic cell cycle            | ∅                      | ∅               | 6     | 1.7x10^{-2}      |       |                   |       |
| Positive regulation of protein phosphorylation   | 63                     | 8.1x10^{-3}     |       | 7                | 2.6x10^{-2}  |       |                   |       |
| Oxidation-reduction process                      | 49                     | 5.5x10^{-3}     |       | 8                | 2.8x10^{-2}  |       |                   |       |
| Response to corticosterone                       | ∅                      | ∅               | 9     | 3.5x10^{-2}      |       |                   |       |
| Response to vitamin D                            | ∅                      | ∅               | 10    | 3.5x10^{-2}      |       |                   |       |
| Inflammatory response                            | 37                     | 2.9x10^{-3}     |       |                   | 1       | 1.6x10^{-6}***   |       |
| Cholesterol biosynthetic process                 | 1                      | 2.5x10^{-6}**   |       |                   | 2       | 6.5x10^{-9}***   |       |
| Cell adhesion                                    | 16                     | 2.1x10^{-4}**   |       |                   | 3       | 9.5x10^{-9}***   |       |
| Cellular response to interleukin-1               | 116                    | 2.7x10^{-2}     |       |                   | 4       | 4.4x10^{-8}***   |       |
| Monocyte chemotaxis                              | 148                    | 4.6x10^{-2}     |       |                   | 5       | 2.6x10^{-7}***   |       |
| Response to drug                                 | 8                      | 5.4x10^{-5}     |       |                   | 6       | 2.9x10^{-7}***   |       |
| Extracellular matrix organization                | 11                     | 6.8x10^{-5}     |       |                   | 7       | 3.6x10^{-7}***   |       |
| Chemokine-mediated signaling pathway             | ∅                      | ∅               | 8     |                   | 8       | 4.0x10^{-7}***   |       |
| Extracellular matrix disassembly                 | 23                     | 7.4x10^{-4}**   |       |                   | 6       | 8.1x10^{-7}***   |       |
| Cellular response to tumor necrosis factor       | ∅                      | ∅               | 10    |                   | 10      | 9.0x10^{-7}***   |       |

P-values and Benjamini-Hochberg FDR were computed by DAVID using the Biological process Direct GO terms database, enrichments were ranked following p-values. Enrichments with corresponding FDR $<$ 10%, 1% and 0.1% are tagged with *, ** and *** respectively.
### Supplemental Table 5. List of signature genes.

| Women       | Men          | All           |
|-------------|--------------|---------------|
| TYMS        | SDCBP        | N4BP3         |
| EGFL8       | RPS6KA3      | UNC119        |
| LAMA3       | CFAP221      | CHST9         |
| STMN2       | HTRA1        | ZFP1          |
| CCL22       | TYMS         | FCAMR         |
| CXCL10      | RAB6A        | TESPA1        |
| MEAF6       | REXO2        | TNFRSF10A     |
| BCAT1       | IL32         | D58           |
| WDFY3-AS2   | KCNAB2       | CD58          |
| ANXA2P2     | RRM2B        | DHR59         |
| AKR1B10     | NIN          | LOC101926960  |
| FABP5P1     | TNFRSF10A    | LOC102723989  |
| DHR9        | LINC00375    | OLR1          |
| OLR1        | CCND1        | ASAP2         |
| UBD         | FAT1         | TREM2         |
| FABP5P7     |              | SERPINB8      |
| LPL         | PCOLCE2      |              |
| CYP2C19     | MMRN2        |              |
| TTC9        | RPS6KA1      |              |
| ITGAX       | LOC283922    |              |
|             | DMD          |              |
|             | MYL12B       |              |
|             | CES1         |              |
|             | SLC1A7       |              |
|             | ARM5C        |              |
|             | HYDIN2       |              |
|             | NUSAP1       |              |
|             | RAN          |              |
|             | SLITRK6      |              |
|             | CNKSR2       |              |
|             | CCL22        |              |
|             | OR56B1       |              |
|             | DDB2         |              |
|             | CENPU        |              |
|             | SALL4        |              |
|             | NNT          |              |
|             | BIRC3        |              |
|             | FGD5         |              |
|             | TMEM17       |              |
|             | NADSYN1      |              |
|             |              | ANXA2         |