

1 ***Novel protein markers of androgen activity in humans: proteomic***
2 ***study of plasma from young chemically castrated men.***

3 ***Aleksander Giwercman^{1*}, K Barbara Sahlin^{2,3*}, Indira Pla^{2,3}, Krzysztof Pawlowski^{2,4,6}, Carl***
4 ***Fehniger^{2,3}, Yvonne Lundberg Giwercman⁵, Irene Leijonhufvud¹, Roger Appelqvist^{2,3}, György***
5 ***Marko-Varga^{3,7}, Aniel Sanchez^{2,3†}, and Johan Malm^{2,3†}***

6 ¹ Molecular Reproductive Medicine, Department of Translational Medicine, Lund University,
7 Malmö, Sweden

8 ² Section for Clinical Chemistry, Department of Translational Medicine, Lund University, Skåne
9 University Hospital Malmö, 214 28 Malmö, Sweden

10 ³ Clinical Protein Science & Imaging, Biomedical Centre, Department of Biomedical Engineering,
11 Lund University, BMC D13, 221 84 Lund, Sweden

12 ⁴ Department of Experimental Design and Bioinformatics, Faculty of Agriculture and Biology,
13 Warsaw University of Life Sciences SGGW, Warszawa, Poland

14 ⁵ Molecular Genetic Reproductive Medicine, Department of Translational Medicine, Lund
15 University, Malmö, Sweden

16 ⁶ Department of Molecular Biology, University of Texas Southwestern Medical Center, Dallas,
17 TX, USA

18 ⁷ First Department of Surgery, Tokyo Medical University, 6-7-1 Nishishinjiku Shinjiku-ku, Japan

19 * † The authors share the first and last positions.

20 ***Short title:*** *Novel protein markers of androgen activity*

21 ***Key words:*** testosterone, hypogonadism, proteomics, physiology, biomarker

22 Corresponding author and to whom reprint requests should be addressed: Aleksander Giwercman. Dept.
23 of Translational Medicine. Lund University. CRC 91-10-058; Jan Waldenströms gata 35; SE 214 28
24 Malmö, Sweden. E-mail: aleksander.giwercman@med.lu.se

25

26 Abstract

27 **Background** Reliable biomarkers of androgen activity in humans are lacking. The aim of this
28 study was, therefore, to identify new protein markers of biological androgen activity and test their
29 predictive value in relation to low vs. normal testosterone values and some androgen deficiency
30 linked pathologies.

31 **Methods** Blood samples from 30 healthy GnRH-antagonist treated males were collected at three
32 time points: a) before GnRH antagonist administration; b) 3 weeks later, just before testosterone
33 undecanoate injection, and c) after additional 2 weeks. Subsequently they were analysed by mass
34 spectrometry to identify potential protein biomarkers of testosterone activity. Levels of proteins
35 most significantly associated with testosterone fluctuations were further tested in a cohort of 75
36 hypo- and eugonadal males suffering from infertility. Associations between levels of those
37 markers and cardio-metabolic parameters, bone mineral density as well as androgen receptor CAG
38 repeat lengths, were explored.

39 **Results** Using ROC analysis, 4-hydroxyphenylpyruvate dioxygenase (4HPPD), insulin-like
40 growth factor-binding protein 6 (IGFBP6) and fructose-bisphosphate aldolase (ALDOB), as well
41 as a Multi Marker Algorithm, based on levels of 4HPPD and IGFBP6, were shown to be best
42 predictors of low (< 8 nmol/L) vs. normal (> 12 nmol/L) testosterone. They were also more
43 strongly associated with metabolic syndrome and diabetes than testosterone levels. Levels of
44 ALDOB and 4HPPD levels also showed association with AR CAG-repeat lengths.

45 **Conclusions** We identified potential new protein biomarkers of testosterone action. Further
46 investigations to elucidate their clinical potential are warranted.

47 **Funding:** The work was supported by ReproUnion 2.0 (grant no 20201846), which is funded by
48 the Interreg V EU program.

49 **Introduction**

50 The male sex hormone, testosterone (T), plays an important physiological role in regulating
51 function of both reproductive and non-reproductive organs in males as well as in females. In males,
52 the diagnosis of T deficiency (i.e. male hypogonadism) is based on the presence of low serum T
53 levels combined with clinical symptoms, which, however, are not pathognomonic for this
54 condition (1).

55 The most common way of assessing T activity is by measuring the total concentration of this
56 hormone in a fasting morning blood sample. However, total T does not accurately reflect
57 biological androgenic activity (BAA), which might be considered a more useful biological and
58 clinical marker.

59 The association between T levels and BAA is affected by several biological mechanisms such as
60 the concentration of binding proteins, body mass index, certain diseases (e.g. diabetes), androgen
61 receptor (AR) sensitivity (2), and different co-factors (3). So far, no reliable algorithms for
62 translating T levels into BAA are available, but could be useful for example in the diagnosis of
63 male hypogonadism.

64 A correct hypogonadism diagnosis is important for proper identification of men for whom
65 androgen replacement therapy is warranted. However, the treatment of men with hypogonadism
66 represents a clinical challenge, because the symptoms associated with the condition are highly
67 unspecific (4). Furthermore, there are limitations in using the level of T in defining testosterone
68 deficiency. Generally, in many clinical guidelines, total T concentration below 8 nmol/L indicates
69 an insufficient hormone concentration, whereas levels above 12 nmol/L are considered normal (5).
70 Apart from the fact that men with testosterone levels between 8 and 12 nmol/L cannot be assigned

71 to any of these distinct categories, those presenting with a lower or higher hormone concentration
72 may also be misclassified due to an abnormal concentration of sex hormone binding protein
73 (SHBG). Low SHBG levels, as often seen in obese men, may imply low total T despite unaffected
74 BAA. On the other hand, some degree of reduced androgen sensitivity may be associated with
75 decreased BAA despite normal or high testosterone levels (6).

76 Hypogonadism has been identified as a predictor of several non-communicable chronic diseases
77 as well as premature mortality (7). Understanding the biology of androgen action may therefore
78 contribute to clarifying the pathogenetic mechanisms linking androgen deficiency to comorbid
79 conditions. Thus, the approach based on measuring the protein levels downstream of androgen
80 action is a feasible and logical concept for identifying clinical and biological useful markers of
81 BAA.

82 Proteomics is a technique aimed to study biological systems based on qualitative and quantitative
83 measuring of proteins and, thereby, integrate the cellular output related to transcription as well as
84 translation. Mapping the quantitative protein response downstream of androgen action may
85 provide new clinically valuable markers of BAA. In order to identify such markers, we compared
86 the protein profile of healthy individuals before and after T deprivation. Subsequently, we assessed
87 the predictive value of the newly identified protein markers in relation to hypogonadism and risk
88 of pathologies related to T deficiency.

89

90

91 **Subjects and Methods**

92 *Study outline*

93 The study was set up to I) identify new protein markers of BAA in healthy subjects; II) test the
94 markers' predictive values in relation to biochemically diagnosed hypogonadism, metabolic
95 syndrome (MetS), cardiovascular risk lipid profile (CVRLP), diabetes mellitus II (DM) and low
96 bone density (LBD) in infertile men; III) analyse androgen dependence of the identified proteins
97 by assessing how their levels associate with androgen receptor gene CAG-repeat length.

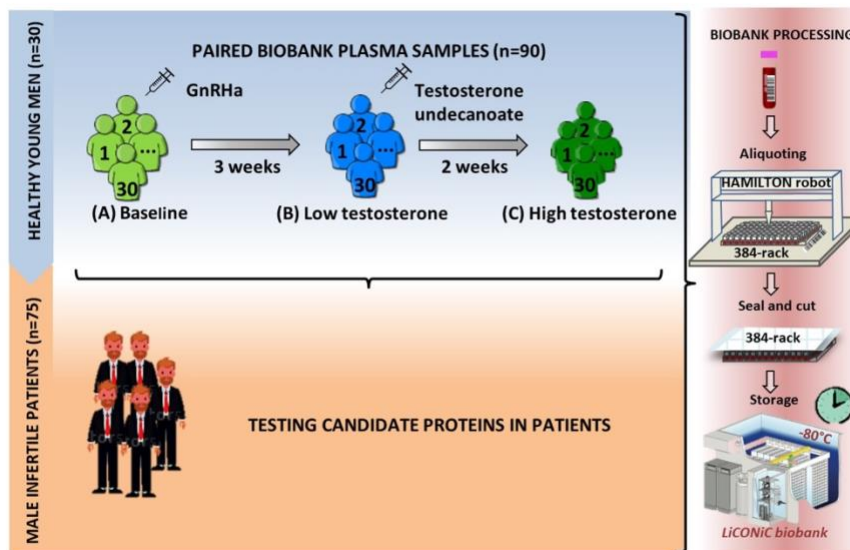
98 *Subjects*

99 All subjects were enrolled with informed written consent. The two studies from which they were
100 recruited were approved by the Swedish Ethical Review Authority (Approval number: DNR
101 2014/311, date of approval 8 May 2014; DNR 2011/1, date of approval 11 January 2011).

102 The first part of the study includes plasma samples obtained from 30 healthy men (biological
103 replicates) aged 19-32 years, BMI 19.1-26.9 kg/m². They, underwent chemical castration by
104 subcutaneous administration of 240 mg GnRH-antagonist (Degeralix®, Ferring Pharmaceuticals,
105 Saint-Prex, Switzerland) followed by remediation of testosterone levels by intramuscular injection
106 of 1000 mg testosterone undecanoate (Nebido®, Bayer AG, Leverkusen, Germany) after the
107 duration of three weeks (8,9). Blood samples were collected at baseline (A); three weeks later (B),
108 and, at the end of the study, after two additional weeks (C) (**Figure 1**).

109 To test the clinical predictive value of the proteins identified in the castrated men, we used a cohort
110 of 75 serum samples from 75 men (biological replication, subject age 32-43 years) previously
111 recruited for a study on hypogonadism among men from infertile couples (10). Eighty-five patients
112 were randomly selected from 213 infertile men and 223 age-matched controls. The selected

113 patients for the present study had the span of subnormal to upper normal range of T. Out of the 85
114 patients, ten patients were excluded; seven due to Klinefelter syndrome, one due to missing value
115 of T and the last two were statistical outliers, which were removed after considering the possible
116 causes. One patient had a high level of T (41.6 nmol/L) without androgen replacement therapy
117 (ART) and the other because he was the only one diagnosed with obstructive azoospermia.
118 Background characteristics of these patients can be found in **Table 1a** and **Table 1b**.



119
120 **Figure 1.** Study design. First, a model of 30 young healthy males was evaluated by proteomics at
121 three time points (A, B and C), where testosterone changes were induced. Identified proteins
122 proposed as candidate biomarkers were then evaluated in a cohort of infertile males. In both steps
123 of the study, the quality of the blood samples was ensured by following an automated workflow for
124 sample aliquoting and storage (-80°C).

Table 1a. Background characteristics of infertile patients.

	Mean (SD)	N
Age at inclusion (years)	37.8 (5.5)	75
BMI	27.2 (4.3)	72
Total testosterone (nmol/L)	12.8 (6.8)	75
FSH (IU/L)	15.8 (14.3)	75
LH (IU/L)	7.5 (5.7)	75
Estradiol (pmol/L)	96 (36-321)*	75
ApoB/ApoA1	0.7 (0.2)	68
HOMA-IR	1.6 (0.4-13.9)*	75
DEXA score	-0.5 (1.3)	74
CAG (repeated length)	22 (14-31)*	74

*Characteristics values are expressed as mean (SD), except for those that did not follow a normal distribution (non-Gaussian) and which are shown as median (min-max) *.*

Table 1b. Characteristics of the cohort of infertile patients.

	n (%)
Smoker	9 (12.0)
Current diseases	36 (48.0)
Insulin medication	1 (1.3)
Current ART	8 (10.7)
CVRLP	20 (27.0)
Insulin resistance	20 (27.0)
Diabetes Mellitus 2	4 (5.3)
Metabolic Syndrome (MetS)	14 (20.9)
Low bone density	23 (30.6)

Variables are expressed as frequency (n) and percentage of occurrence (%) - according to the total number of patients (N=75).

ART: Androgen Replacement Therapy;

CVRLP: cardiovascular risk lipid profile

125 The following comorbidities in the cohort of infertile patients were defined MetS, IR, CVRLP,
126 DM and LBD. MetS was determined according to the criteria defined at the National Cholesterol
127 Education Program Adult Treatment Panel III 2002 (**Table S1**). Homeostatic Model Assessment
128 of Insulin Resistance (HOMA-IR) was calculated as (glucose x insulin)/22.5 and IR was defined
129 as HOMA-IR > 2.5 (11). CVRLP was defined as the ratio apolipoprotein B/apolipoprotein A1 >

130 0.9 (12). DM was set at fasting blood glucose > 7 mmol/L (13). LBD was determined based on the
131 DEXA lumbar Z score with the cut off at < -1 (14). The methods for laboratory tests, CAG-repeat
132 length and proteomics are described in the supplementary file 1.

133 *Statistical analysis*

134 We briefly describe the statistical analyses performed. A full description of the statistical analysis
135 is available in supplementary file 2. Proteomics data pre-processing was done using Perseus
136 v1.6.7.0 (18) software and unless other software is specified, the statistical analyzes were
137 performed using R software (19,20).

138 *Healthy human model*

139 Protein intensities were Log2 transformed and standardized by Subtract Median normalization.
140 Differentially expressed proteins were determined by one-way repeated measures ANOVA
141 followed by a pairwise t-test (two-tails and paired). Adjusted *p*-values < 0.05 were considered
142 significant. The power of the candidate biomarkers to discriminate between normal and low T was
143 evaluated by doing receiver operating characteristic (ROC) analysis. Significant proteins (between
144 A-B with significant recovery in B-C) with a) area under the curve (AUC) > 0.80 or b) AUC
145 between 0.75 – 0.80 and highly enriched in liver tissues (Human Proteome Map (24) and Kampf
146 C. et. al. (25)) were considered candidate biomarker. These candidates were included as predictors
147 in a stepwise regression (method: backward) to select the best combination of markers that predict
148 the odds of being low T. Bootstrap resampling with replacement method was applied to assess
149 consistency. A new variable called Multi Marker Algorithm (MMA) was derived from the
150 predicted log-odds (of being low T) obtained from a binomial logistic regression analysis (see

151 supplementary file 2) and it was evaluated together with marker candidates proteins.

152 *Infertile cohort of patients*

153 The normal distribution of the variables that describe background characteristics of the infertile
154 cohort of patients (**Table 1a**) was evaluated by Kolmogorov-Smirnov test. The intensities of the
155 candidate biomarkers were Log2 transformed to achieve normal distributions. In this cohort, MMA
156 variable was created to predict the odds of suffering low T or other medical conditions associated
157 with low T levels.

158 Overall changes between more than two groups were evaluated by one-way ANOVA followed by
159 a pairwise FDR correction (29) and adjusted p -values < 0.05 were considered significant. In order
160 to know if the changes in the candidate markers occur with the change in T as observed in the
161 healthy human model, three groups of patients were created based on total T concentration (group
162 1: low T (LT) ≤ 8 nmol/L (n= 22); group 2: borderline low T (BL_T) between 8 and 12 nmol/L
163 (n=17); group 3: normal T (NT) > 12 nmol/L (n= 36)). The power of the candidate biomarkers to
164 discriminate between NT and LT or BL_T, or to distinguish patients with medical conditions
165 associated with low T levels (MetS, IR, CVRLP, DM or LBD) was evaluated by a ROC analysis.
166 The DeLong's test (paired) was used to compare the AUCs.

167 To investigate the androgenic dependence of the candidate biomarkers, we looked for associations
168 between their expression and the androgen receptor (AR) CAG repeat length (33–36): reference
169 group 1: patients with CAG repeat length 21 and 22 (n= 18); group 2: patients with CAG repeat
170 length < 21 (n= 26) and group 3: patients with CAG repeat length >22).

171

172 **Results**

173 *Proteins differentially expressed in chemically castrated men*

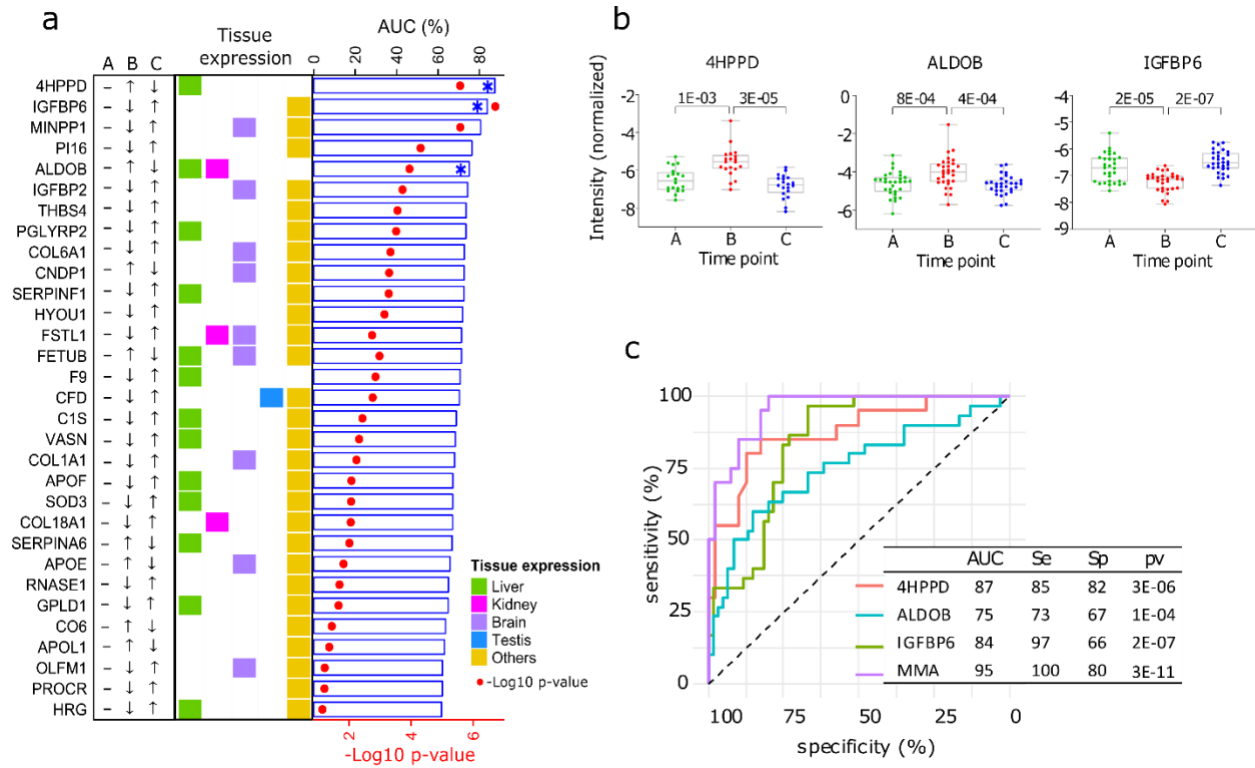
174 In total, in the healthy men, the expression level of 31 out of 676 proteins was statistically
175 significantly associated with T concentration (**Tables S2**). The levels of 23 proteins changed in
176 the same direction as T, whereas, the remaining eight markers changed in an opposite way. LH
177 and FSH changed significantly in A-B but not in B-C. The protein changes visualised as boxplots
178 can be found in the Supplementary **Figure S1**.

179 *Proteins capable to distinguish between low and normal testosterone*

180 Based on p -values for AUC in the ROC analysis, among healthy young men, 90% of the 31
181 proteins distinguished the low T time point (B) from the normal ones (A and C) with statistical
182 significance (Table S3, **Figure 2a**). ROC-AUC values greater than 0.80 were obtained for the
183 proteins 4-hydroxyphenylpyruvate dioxygenase (4HPPD) and insulin-like growth factor-binding
184 protein 6 (IGFBP6). Additionally, fructose-bisphosphate aldolase (ALDOB) was the only protein
185 enriched in liver tissue with ROC-AUC between 0.75 and 0.80.

186 The stepwise regression method selected 4HPPD and IGFBP6 as the best markers to be combined
187 to predict the odds of being low T, and thus, they were the basis for the new variable MMA (see
188 Material and Methods). MMA together with 4HPPD, ALDOB and IGFBP6 were selected as
189 potential candidate markers for the diagnosis of BAA (**Figures 2b and 2c**). The expression of the
190 4HPPD and ALDOB proteins were significantly increased at low T ($p<0.001$; $p<0.001$) and
191 remediated in response to the T treatment, whereas IGFBP6 expression was significantly decreased
192 ($p<0.001$) by castration.

193

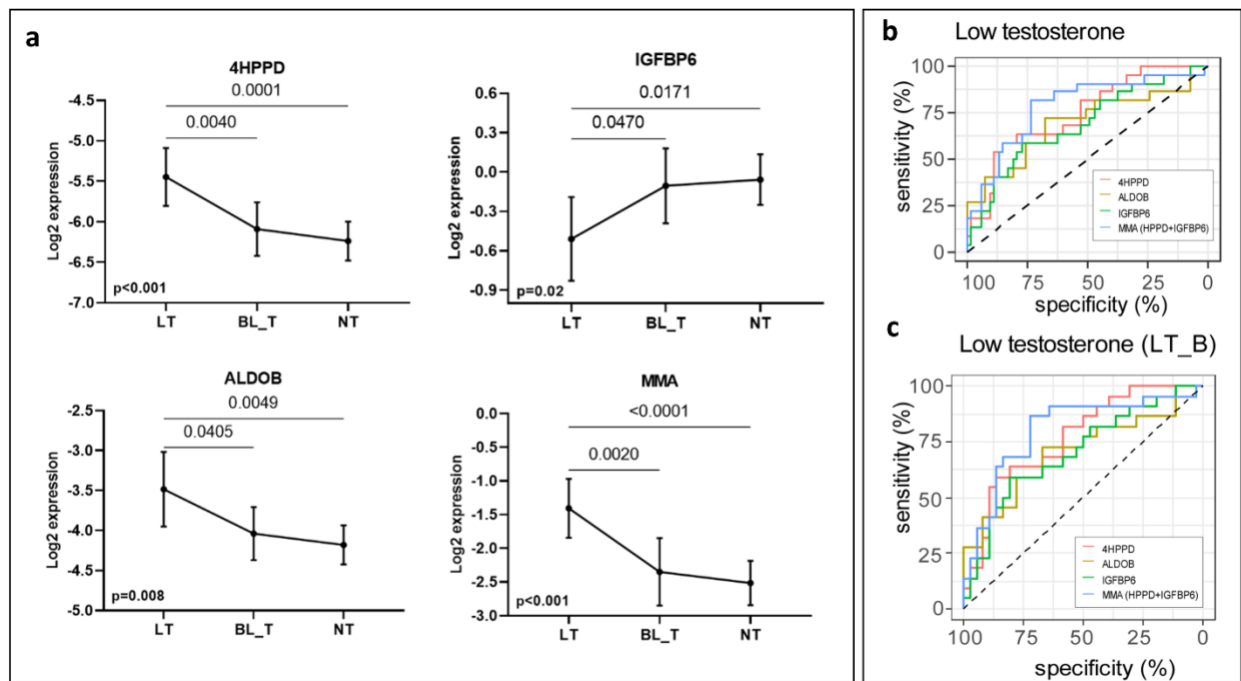


194

195 **Figure 2.** Proteins influenced by testosterone in the model of young healthy males. a) Top 25
 196 significant proteins selected in the healthy human model (ROC- $p < 0.01$, Table S3). The arrows
 197 indicate the direction of change in protein expression in the different conditions. The tissue with
 198 highest expression of each protein is indicated in colours. Also, results from the ROC analysis are
 199 shown as bar chart (AUC) and heat-map (p -values). b) Boxplot (mean (min; max)) of the top three
 200 significant proteins proposed as biomarker candidates, able to discriminate between low and
 201 normal testosterone (Table S4). The adjusted p -values are specified on top of the comparative
 202 horizontal lines. c) ROC of the analytes proposed as biomarker candidates, including Multi
 203 Marker Algorithm (MMA).

204 *Testing of the candidate biomarkers in infertile men*

205 The three proteins and MMA showed statistically significant differences (4HPPD: $p < 0.001$,
206 ALDOB: $p = 0.003$, IGFBP6: $p = 0.016$, MMA: $p < 0.001$, **Figure 3a**) between the three groups
207 defined according to the total T levels (**Table S2**). 4HPPD, ALDOB, and MMA showed a negative
208 association with T changes, while IGFBP6 displayed a positive association. The three proteins and
209 MMA significantly distinguished the patients with LT from BL_T/NT (4HPPD: AUC=0.75, $p =$
210 0.001; ALDOB: AUC= 0.70, $p = 0.008$; IGFBP6: AUC= 0.69, $p = 0.01$; MMA: AUC= 0.79, $p <$
211 0.001) (**Figure 3b**). Additionally, the power to discern low T values improved for all the
212 biomarkers tested when the patients with BL_T were excluded (**Table 2**).



213
214 **Figure 3.** New markers to discern states of different testosterone levels in men investigated for
215 infertility ($n=75$). **a**) Patients grouped by three levels of total testosterone: low testosterone (LT)
216 ≤ 8 nmol/L ($n=22$), border line testosterone (BL_T) between 8 and 12 nmol/L ($n=17$) and normal
217 testosterone (HT) > 12 nmol/L ($n=36$). Each group is represented by the mean and its 95% CI.

218 *Horizontal lines indicate significant differences between groups and the adjusted p-values are*
219 *specified on top of these lines (Table S5). **b**) ROC analysis to discriminate patients with LT in the*
220 *entire cohort and **c**) in a cohort that excluded patients with borderline testosterone levels (LT_B).*
221 *MMA is based on is the combination of levels of the proteins 4HPPD and IGFBP6.*

222 **Table 2. Comparison of ROC-Areas Under the Curve for testosterone and the candidate biomarkers in relation to the prediction of**
 223 **hypogonadism and its sequels in patients.**

Analyte	Low T (T ≤ 8 nmol/L)		Low T* (T ≤ 8 nmol/L)		IR (HOMA-ir > 2.5)		DM -		LBD (z-score < -1)		CVRLP (ApoB/ApoA1 ≥ 0.9)		MetS	
	AUC (Sp,Se)	p	AUC (Sp,Se)	p	AUC (Sp,Se)	p	AUC (Sp,Se)	p	AUC (Sp,Se)	p	AUC (Sp,Se)	p	AUC (Sp,Se)	p
4HPPD	0.75 (85,59)	<u>8.38E-04</u>	0.77 (86,59)	<u>2.66E-04</u>	0.79 (84,70)	<u>1.20E-04</u>	0.89 (93,75)	<u>9.00E-03</u>	0.64 (76,61)	<u>2.24E-02</u>	0.74 (46,90)	<u>5.79E-03</u>	0.74 (95,50)	<u>5.31E-03</u>
ALDOB	0.69 (68,73)	<u>8.25E-03</u>	0.70 (67,73)	<u>5.39E-03</u>	0.73 (71,75)	<u>2.85E-03</u>	0.85 (63,100)	<u>1.80E-02</u>	0.57 (98,27)	1.93E-01	0.71 (83,55)	<u>4.64E-02</u>	0.74 (82,64)	<u>6.02E-03</u>
IGFBP6	0.69 (77,59)	<u>1.05E-02</u>	0.70 (81,59)	<u>4.89E-03</u>	0.57 (42,80)	3.50E-01	0.59 (39,100)	5.40E-01	0.63 (43,78)	2.75E-01	0.59 (48,80)	1.87E-01	0.65 (45,93)	8.45E-02
Testosterone	-	-	-	-	0.71 (76,70)	<u>4.96E-03</u>	0.55 (32,100)	7.24E-01	0.75 (74,74)	<u>5.12E-04</u>	0.66 (65,85)	<u>6.55E-03</u>	0.56 (72,57)	5.08E-01
MMA	0.79 (74,82)	<u>9.23E-05</u>	0.80 (72,86)	<u>3.86E-05</u>	0.79 (82,70)	<u>1.46E-04</u>	0.92 (84,100)	<u>5.00E-03</u>	0.78 (82,65)	<u>1.40E-02</u>	0.75 (73,75)	<u>3.65E-03</u>	0.78 (63,86)	<u>1.57E-03</u>

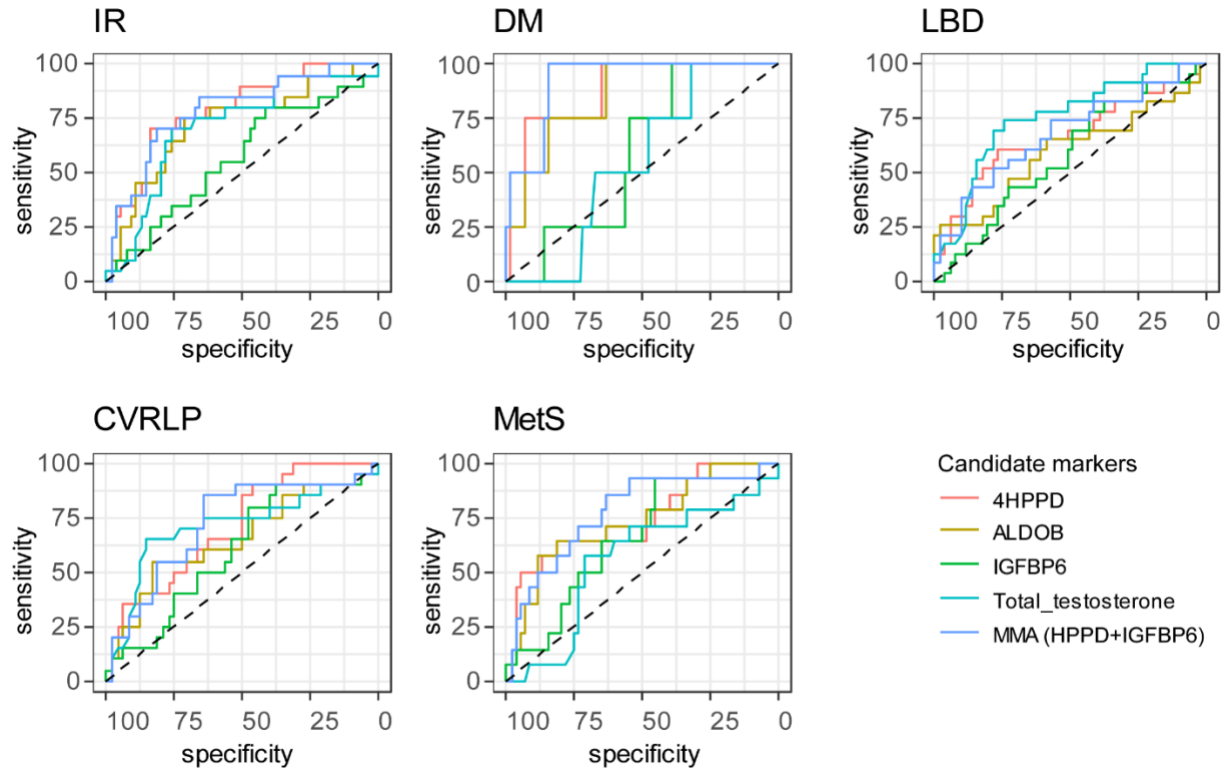
Significant p-values are highlighted in bold and underlined. * Excluding patients with testosterone values from the borderline low testosterone ($8 < BL_T \leq 12$).

IR: insulin resistance; DM: diabetes mellitus type 2; LBD: low bone density; CVRLP: cardiovascular risk lipid profile ; MetS: metabolic syndrome

AUC: area under the curve; Sp: Specificity in %; Se: Sensitivity in %

225 *Ability to distinguish men with abnormal metabolic comorbidities or reduced bone mineral*
226 *density*

227 The AUCs for 4HPPD, ALDOB and MMA in relation to risk of DM and MetS were statistically
228 significant whereas for T the p-value for AUC was 0.72. The AUCs for 4HPPD, ALDOB and
229 MMA were also statistically significantly larger than this for T (DM: 4HPPD ($p= 0.005$), ALDOB
230 ($p= 0.009$), MMA ($p= 0.002$); MetS: 4HPPD ($p= 0.032$), ALDOB ($p= 0.030$), MMA ($p= 0.002$)),
231 (**Table 2**). Additionally, the AUC values in relation to CVRLP and IR were numerically higher
232 for 4HPPD, ALDOB and MMA than for T, however, the differences between the AUC values
233 were not statically significant (CVRLP marker vs T: 4HPPD ($p= 0.97$), ALDOB ($p= 0.61$), MMA
234 ($p= 0.87$); IR marker vs T: 4HPPD vs T ($p= 0.30$), ALDOB vs T ($p= 0.88$), MMA ($p= 0.31$)).
235 4HPPD and MMA statistically significantly distinguished between patients with normal bone
236 density and LBD. The same was true for T, the differences between the AUC for T and those for
237 4HPPD and MMA not being statistically significant (LBD vs T: 4HPPD ($p= 0.28$), MMA ($p=$
238 0.30)). No statistical significance, in relation to prediction of LBD was found for IGFBP6 (**Figure**
239 **4**)



240

241 **Figure 4** Results from ROC analysis to determine whether the analytes discriminate between the
242 presence of comorbidities or not. Analytes included in the analysis are 4HPPD, IGFBP6, ALDOB
243 and MMA (combination of 4HPPD and IGFBP6). AUC, p-values can be found in Table 2. IR:
244 insulin resistance; DM: type 2 diabetes mellitus; LBD: low bone density; CVRLP: cardiovascular
245 risk lipid profile; MetS: metabolic syndrome.

246

247 **Association of the candidate biomarkers with androgen receptor CAG-repeat length**

248 Statistically significant inter-CAG-group overall differences were observed for 4HPPD ($p=0.012$)
249 and ALDOB ($p=0.008$) (**Figure 5**). Additionally, the protein expressions were significantly higher
250 in the groups with <21 and >22 CAG repeat length as compared with the reference (**Table 3** and
251 **Figure 5**). However, we did not observe any statistically significant association between CAG

252 number and expression of IGFBP6.

253 **Table 3. Ratio between mean concentrations of 4HPPD and ALDOB in men with CAG repeat**

254 **length < 21 or >22 as compared to the reference group**

<i>Proteins</i>	<i>Overall</i>	<i><21 vs reference</i>		<i>>22 vs reference</i>	
	<i>p-value</i>	<i>ratio (95% CI)</i>	<i>p value*</i>	<i>ratio (95% CI)</i>	<i>p value*</i>
4HPPD	0.012	1.34 (1.02 to 1.76)	0.032	1.62 (1.23 to 2.13)	0.001
ALDOB	0.008	1.35 (1.03 to 1.78)	0.029	1.72 (1.29 to 2.32)	<0.001

Ratio: ratio between mean concentration in <21 or >22 groups divided by the reference group (21 and 22).

** p-value of the post hoc constrain between groups*

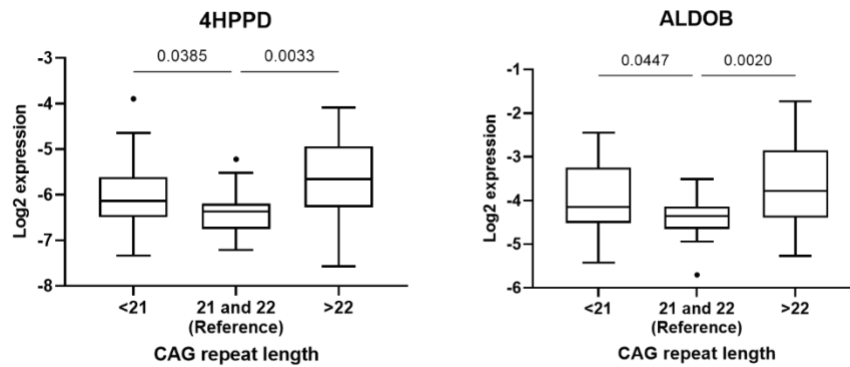


Figure 5. Association between androgen receptor CAG <21 (n=26), and CAG<22 (n=30) and 4HPPD and ALDOB, respectively, with CAG=21 and 22 (n=18) set as reference.

255

256 ***Discussion***

257 We have identified three plasma proteins, which are potential markers of BAA. In young healthy
258 men the three markers ALDOB, 4HPPD and IGFB6 were strongly associated with T levels. In a
259 slightly older cohort of infertile men, these markers were indicative of T deficiency (5).
260 Furthermore, levels of two of the markers, ALDOB and 4HPPD, were more strongly associated
261 with risk of metabolic disturbances than total T. The association seemed to become stronger by
262 creating a combined marker MMA, based on both 4HPPD and IGFBP6 levels. Finally, the
263 androgen dependence of ALDOB and 4HPPD was confirmed by the association between the
264 concentration of those proteins and the length of androgen receptor CAG repeats.

265 The ALDOB is a glycolytic enzyme, predominantly expressed in liver and kidney, that catalyzes
266 the reversible cleavage of fructose-1,6-bisphosphate into glyceraldehyde 3-phosphate and
267 dihydroxyacetone phosphate. The B isoform of aldolase, e.g. ALDOB, in the liver is under dietary
268 control (37). Ingestion of fructose induces ALDOB mRNA expression in the liver, which is
269 otherwise low in fasting conditions. In humans, absence of functional ALDOB enzyme due to
270 mutations in the ALDOB gene cause hereditary fructose intolerance, characterized by metabolic
271 disturbances that include hypoglycemia, lactic acidosis, and hypophosphatemia (38). An
272 upregulation of ALDOB in human pancreatic β -cells occurs upon the development of
273 hyperglycemia and may contribute to the impairment of insulin secretion in humans (39). In a
274 study on goats, the ALDOB gene was found to be downregulated at the time of post-natal initiation
275 of spermatogenesis (40). This finding is in accordance with our data showing that rising
276 testosterone is inhibiting ALDOB.

277 Similar to ALDOB, we found that 4HPPD was negatively associated with T levels. This enzyme

278 is involved in the catabolic pathway of tyrosine and catalyzes the conversion of 4-
279 hydroxyphenylpyruvate to homogentisic acid in the tyrosine catabolism pathway (41). The
280 expression of the gene is regulated by hepatocyte-specific and liver-enriched transcription factors,
281 as well as by hormones (24). Tyrosine has previously been reported to be upregulated in
282 hypogonadal men and both tyrosine and phenylalanine levels were suggested as predictors of the
283 risk of developing diabetes many years before manifest disease (43–45). In male tyrosine
284 hydroxylase knock-out mice normal body weight, puberty onset, and basal gonadotropin levels in
285 adulthood were evident, although T was significantly elevated in adult mice (46).

286 The last marker, IGFBP6 is expressed in most tissues (24) and is one of the binding proteins for
287 Insulin-Like Growth Factor (IGF). The principal function of IGFBP-6 is inhibiting IGF-II actions,
288 whereby IGF-II induced cell proliferation, differentiation, migration and survival is reduced.
289 Serum levels of IGFBP-6 increase gradually with age and are higher in men than in women, but
290 there are conflicting studies of the direct effects of sex steroids on IGFBP-6 expression in different
291 tissues (47). A positive association between the levels of T and IGFBP-6 have previously been
292 found (48) (49). The latter study has a somewhat similar set up as the present study, based on
293 chemical castration with a GnRH agonist, and has also focused on identifying novel markers of
294 BAA, but with candidate markers previously identified as being associated with changes in fat-
295 free mass. The study showed that early increases in IGFBP6 levels in men receiving testosterone
296 were associated with increases in fat-free mass and muscle strength.

297 Altogether, our findings may not only be clinically valuable in developing new methods of
298 assessing BAA but also add to our understanding of the biological role of T in human metabolism,
299 regulation of testicular function (ALDOB), as well as muscle strength and body composition

300 (IGFBP6). The combined parameter MMA may in this context be an important tool in the detection
301 of long-term morbidity, such as bone mineral density and cardio-metabolic risk, even before
302 clinical diagnosis.

303 Our study has some strengths and limitations. Using a chemical castration model in young healthy
304 men, we were able to identify proteins influenced by androgens and select those that were most
305 strongly associated with T levels. By utilizing proteomics, we had an explorative approach to
306 identify new markers of BAA without being restricted by previously published findings.

307 Another strength is the depth of the analysis due to the depletion abundant proteins from plasma.
308 We were able to identify more than 450 proteins, which were identified in the same concentration
309 range as 87 % of FDA-approved biomarkers (50). If depletion is not performed the detection level
310 is dampened by the components from the digested abundant proteins as the proteins removed are
311 of highest abundance in plasma and plasma proteome and exceed some lower abundance proteins
312 by 10 orders of magnitude (51).

313 Although more clinical testing is needed, we have provided preliminary results showing that these
314 protein markers may also be clinically useful. We have previously shown that median length CAG
315 number is associated with most active AR (52). Thus, the fact that we find the lowest ALDOB
316 and 4HPPD levels in those subjects having AR CAG repeat length close to median, confirms that
317 the candidate markers identified in the present study are androgen dependent.

318 A limitation of our study is that the lack of reliable criteria for clinical hypogonadism, which made
319 it impossible to test the power of the new markers in men in whom androgen replacement is
320 needed. Furthermore, the clinical part of the study was limited, because we do not have sufficient

321 information about potential factors influencing the inter- and intra-individual variation in the levels
322 of these proteins and, thereby, their suitability as clinical markers.

323 Furthermore, we are not reporting absolute values of quantifications for the potential markers, but
324 the relative quantifications for comparing the protein expression between groups. This kind of
325 comparative proteomics is favorable in research studies, in which preliminary results of protein
326 changes between groups are obtained. Also, the sample processing is complicated putting high
327 demands on the lab. In proteomics, there can be a high variation between labs in reporting absolute
328 concentration proteins in plasma, especially when small sample sizes are reported (53). In order to
329 obtain trustworthy absolute concentration ranges or determine the activity level of the enzymes, it
330 is necessary to analyze the potential markers in large cohorts including both healthy subjects and
331 patients.

332 In this study, we have applied an immunoassay for measuring T levels, although some consider
333 liquid chromatography-tandem mass spectrometry (LC-MS/MS) as gold standard in assessment of
334 sex hormone levels. However, worldwide the former is most commonly used for T measurements.
335 Additionally, in prediction of cardiometabolic risk, assessment of T by LC-MS/MS was not shown
336 to be superior to that performed by standard methodology (54).

337 In conclusion, we have identified three new potential biomarkers of BAA. Those proteins – alone
338 or in combination – are promising as useful parameters in the clinical diagnosis of male
339 hypogonadism and in the prediction of its long-term sequelae, as well as in studying the biology
340 of androgen action. More extensive testing is vital to elucidate their BAA potential, not only in
341 men but also in women and in prepubertal boys.

342 **Data availability**

343 The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium
344 via the PRIDE (55) partner repository with the dataset identifier PXD024448.

345 Supplementary tables (datasets):

346 <https://doi.org/10.6084/m9.figshare.14875431>

347 Source data of the Figures can be found on:

348 <https://doi.org/10.6084/m9.figshare.14875431>

349 Supplementary Figure S1:

350 <https://doi.org/10.6084/m9.figshare.14876562>

351 R code:

352 https://github.com/indirapla/TP1_proteins_marker_of_androgen_activity

353

354 **Declaration of conflict of interest:** The authors declare that there is no conflict of interest that
355 could be perceived as prejudicing the impartiality of the research reported.

356

357

358 **References**

- 359 1. Traish AM, Miner MM, Morgentaler A, Zitzmann M. Testosterone deficiency. *Am J Med.*
360 2011;124(7):578–87.
- 361 2. Zitzmann M. The Role of the CAG Repeat Androgen Receptor Polymorphism in
362 Andrology. In: *Advances in the Management of Testosterone Deficiency*. Basel: KARGER;
363 2008. p. 52–61.
- 364 3. Furuya K, Yamamoto N, Ohyabu Y, Morikyu T, Ishige H, Albers M, et al. Mechanism of
365 the tissue-specific action of the selective androgen receptor modulator s-101479. *Biol*
366 *Pharm Bull.* 2013 Mar;36(3):442–51.
- 367 4. Miner M, Barkin J, Rosenberg MT. Testosterone deficiency: myth, facts, and controversy.
368 *Can J Urol.* 2014 Jun;21 Suppl 2(3):39–54.
- 369 5. Arver S, Lehtihet M. Current guidelines for the diagnosis of testosterone deficiency. Vol.
370 37, *Frontiers of Hormone Research*. *Front Horm Res*; 2009. p. 5–20.
- 371 6. Diaz-Arjonilla M, Schwarcz M, Swerdloff RS, Wang C. Obesity, low testosterone levels
372 and erectile dysfunction. Vol. 21, *International Journal of Impotence Research*. *Int J Impot*
373 *Res*; 2009. p. 89–98.
- 374 7. Muraleedharan V, Jones TH. Testosterone and mortality. *Clin Endocrinol (Oxf).*
375 2014;81(4):477–87.
- 376 8. Sahlin KB, Pla I, Sanchez A, Pawłowski K, Leijonhufvud I, Appelqvist R, et al. Short-term

- 377 effect of pharmacologically induced alterations in testosterone levels on common blood
378 biomarkers in a controlled healthy human model. *Scand J Clin Lab Invest.* 2019 Nov
379 18;80(1):1–7.
- 380 9. Pla I, Sahlin KB, Pawłowski K, Appelqvist R, Marko-Varga G, Sanchez A, et al. A pilot
381 proteomic study reveals different protein profiles related to testosterone and gonadotropin
382 changes in a short-term controlled healthy human cohort. *J Proteomics.* 2020 May
383 30;220:1–3.
- 384 10. Bobjer J, Bogefors K, Isaksson S, Leijonhufvud I, Åkesson K, Giwercman YL, et al. High
385 prevalence of hypogonadism and associated impaired metabolic and bone mineral status in
386 subfertile men. *Clin Endocrinol (Oxf).* 2016 Aug;85(2):189–95.
- 387 11. Wickramasinghe VP, Arambepola C, Bandara P, Abeysekera M, Kuruppu S, Dilshan P, et
388 al. Insulin resistance in a cohort of 5-15 year old children in urban Sri Lanka. *BMC Res*
389 *Notes.* 2017 Jul;10(1).
- 390 12. Walldius G, Jungner I. Apolipoprotein B and apolipoprotein A-I: Risk indicators of
391 coronary heart disease and targets for lipid-modifying therapy. Vol. 255, *Journal of Internal*
392 *Medicine.* 2004. p. 188–205.
- 393 13. *Standards of Medical Care in Diabetes-2010.* 2010;
- 394 14. Isaksson S, Bogefors K, Åkesson K, Egund L, Bobjer J, Leijonhufvud I, et al. Risk of low
395 bone mineral density in testicular germ cell cancer survivors: association with
396 hypogonadism and treatment modality. *Andrology.* 2017;5(5):898–904.

- 397 15. Lundin KB, Giwercman A, Richthoff J, Abrahamsson PA, Giwercman YL. No association
398 between mutations in the human androgen receptor GGN repeat and inter-sex conditions.
399 Mol Hum Reprod. 2003;9(7–8):375–9.
- 400 16. Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, et al.
401 Measurement of protein using bicinchoninic acid. Anal Biochem. 1985;150(1):76–85.
- 402 17. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline:
403 an open source document editor for creating and analyzing targeted proteomics
404 experiments. Bioinformatics. 2010 Apr 1;26(7):966–8.
- 405 18. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus
406 computational platform for comprehensive analysis of (prote)omics data. Nat Methods.
407 2016 Sep 27;13(9):731–40.
- 408 19. RStudio Team. RStudio: Integrated Development for R. RStudio. Boston; 2016.
- 409 20. Team RC. R: A language and environment for statistical computing. R Foundation for
410 Statistical Computing. 2016.
- 411 21. Marshall KW, Mohr S, Khettabi F El, Nossova N, Chao S, Bao W, et al. A blood-based
412 biomarker panel for stratifying current risk for colorectal cancer. Int J Cancer. 2009
413 Mar;126(5):1177–86.
- 414 22. Bowers AJ, Zhou X. Receiver Operating Characteristic (ROC) Area Under the Curve
415 (AUC): A Diagnostic Measure for Evaluating the Accuracy of Predictors of Education

- 416 Outcomes. *J Educ Students Placed Risk*. 2019 Jan;24(1):20–46.
- 417 23. Šimundić A-MM. Measures of Diagnostic Accuracy: Basic Definitions. *Ejifcc*. 2009
418 Jan;19(4):203–11.
- 419 24. Kim M-SS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, et al. A draft map
420 of the human proteome. *Nature*. 2014 May 29;509(7502):575–81.
- 421 25. Kampf C, Mardinoglu A, Fagerberg L, Hallström BM, Edlund K, Lundberg E, et al. The
422 human liver-specific proteome defined by transcriptomics and antibody-based profiling.
423 *FASEB J • Res Commun*.
- 424 26. Kholodenko I V, Yarygin KN. Cellular Mechanisms of Liver Regeneration and Cell-Based
425 Therapies of Liver Diseases. *Biomed Res Int*. 2017;v.2017:17 pages.
- 426 27. Schmucker DL, Sanchez H. Liver Regeneration and Aging: A Current Perspective. *Curr*
427 *Gerontol Geriatr Res*. 2011;2011.
- 428 28. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950 Jan;3(1):32–5.
- 429 29. Benjamini Y, Krieger AM, Yekutieli D. Adaptive linear step-up procedures that control the
430 false discovery rate. *Biometrika*. 2006 Sep 1;93(3):491–507.
- 431 30. Chan I, Tang Fui MN, Zajac JD, Grossmann M. Assessment and management of male
432 androgen disorders: An update. *Aust Fam Physician*. 2014;43(5):277–82.
- 433 31. Lunenfeld B, Arver S, Moncada I, Rees DA, Schulte HM. How to help the aging male?

- 434 Current approaches to hypogonadism in primary care. Vol. 15, *Aging Male*. *Aging Male*;
435 2012. p. 187–97.
- 436 32. Zitzmann M, Faber S, Nieschlag E. Association of specific symptoms and metabolic risks
437 with serum testosterone in older men. *J Clin Endocrinol Metab*. 2006;91(11):4335–43.
- 438 33. Casella R, Maduro MR, Lipshultz LI, Lamb DJ. Significance of the polyglutamine tract
439 polymorphism in the androgen receptor. *Urology*. 2001;58(5):651–6.
- 440 34. Stanworth RD, Kapoor D, Channer KS, Jones TH. Androgen receptor CAG repeat
441 polymorphism is associated with serum testosterone levels, obesity and serum leptin in men
442 with type 2 diabetes. *Eur J Endocrinol*. 2008;159(6):739–46.
- 443 35. Kim JW, Bae YD, Ahn ST, Kim JW, Kim JJ, Moon DG. Androgen Receptor CAG Repeat
444 Length as a Risk Factor of Late-Onset Hypogonadism in a Korean Male Population. *Sex*
445 *Med*. 2018 Sep 1;6(3):203–9.
- 446 36. Ferlin A, Bartoloni L, Rizzo G, Roverato A, Garolla A, Foresta C. Androgen receptor gene
447 CAG and GGC repeat lengths in idiopathic male infertility. *Mol Hum Reprod*.
448 2004;10(6):417–21.
- 449 37. Munnich A, Besmond C, Darquy S, Reach G, Vaulont S, Dreyfus JC, et al. Dietary and
450 hormonal regulation of aldolase B gene expression. *J Clin Invest*. 1985;75(3):1045–52.
- 451 38. Hannou SA, Haslam DE, McKeown NM, Herman MA. Fructose metabolism and metabolic
452 disease. Vol. 128, *Journal of Clinical Investigation*. American Society for Clinical

- 453 Investigation; 2018. p. 545–55.
- 454 39. Gerst F, Jaghutriz BA, Staiger H, Schulte AM, Lorza-Gil E, Kaiser G, et al. The expression
455 of aldolase B in islets is negatively associated with insulin secretion in humans. *J Clin*
456 *Endocrinol Metab.* 2018;103(12):4373–83.
- 457 40. Bo D, Jiang X, Liu G, Xu F, Hu R, Wassie T, et al. Multipathway synergy promotes
458 testicular transition from growth to spermatogenesis in early-puberty goats. *BMC*
459 *Genomics.* 2020 May 25;21(1).
- 460 41. Hager SE, Gregerman RI, Kox WE. p-Hydroxyphenylpyruvate oxidase of liver. *J Biol*
461 *Chem.* 1957;225(2):935–47.
- 462 42. Barroso F, Correia J, Bandeira A, Carmona C, Vilarinho L, Almeida M, et al. Tyrosinemia
463 type III: A case report of siblings and literature review. *Revista Paulista de Pediatria.* 2020.
- 464 43. Wang TJ, Larson MG, Vasan RS, Cheng S, Rhee EP, McCabe E, et al. Metabolite profiles
465 and the risk of developing diabetes. *Nat Med.* 2011 Apr;17(4):448–53.
- 466 44. Fanelli G, Gevi F, Belardo A, Zolla L. Metabolic patterns in insulin-sensitive male
467 hypogonadism article /101/58. *Cell Death Dis.* 2018 Jun 1;9(6).
- 468 45. Guasch-Ferré M, Hruby A, Toledo E, Clish CB, Martínez-González MA, Salas-Salvadó J,
469 et al. Metabolomics in prediabetes and diabetes: A systematic review and meta-analysis.
470 *Diabetes Care.* 2016.
- 471 46. Stephens SBZ, Rouse ML, Tolson KP, Liaw RB, Parra RA, Chahal N, et al. Effects of

- 472 selective deletion of tyrosine hydroxylase from kisspeptin cells on puberty and reproduction
473 in male and female mice. *eNeuro*. 2017;
- 474 47. Bach LA, Fu P, Yang Z. Insulin-like growth factor-binding protein-6 and cancer. *Clinical*
475 *Science*. 2013.
- 476 48. Rooman RPA, Op De Beeck L, Martin M, van Doorn J, Mohan S, Du Caju MVL.
477 Ethinylestradiol and testosterone have divergent effects on circulating IGF system
478 components in adolescents with constitutional tall stature. *Eur J Endocrinol*. 2005
479 *Apr*;152(4):597–604.
- 480 49. Huang G, Rocha G V., Pencina KM, Cox K, Krishnan V, Henriksen K, et al. Circulating
481 Biomarkers of Testosterone’s Anabolic Effects on Fat-Free Mass. *J Clin Endocrinol Metab*.
482 2019 *Apr 11*;104(9):3768–78.
- 483 50. Anderson NL. The Clinical Plasma Proteome: A Survey of Clinical Assays for Proteins in
484 Plasma and Serum. *Clin Chem*. 2010 *Feb*;56(2):177–85.
- 485 51. Tu C, Rudnick PA, Martinez MY, Cheek KL, Stein SE, Slebos RJC, et al. Depletion of
486 abundant plasma proteins and limitations of plasma proteomics. *J Proteome Res*. 2010
487 *Oct*;9(10):4982–91.
- 488 52. Nenonen H, Björk C, Skjaerpe PA, Giwercman A, Rylander L, Svartberg J, et al. CAG
489 repeat number is not inversely associated with androgen receptor activity in vitro. *Mol Hum*
490 *Reprod*. 2009 *Nov*;16(3):153–7.

- 491 53. Nanjappa V, Thomas JK, Marimuthu A, Muthusamy B, Radhakrishnan A, Sharma R, et al.
492 Plasma Proteome Database as a resource for proteomics research: 2014 update. *Nucleic*
493 *Acids Res.* 2014 Jan;42(D1):D959–65.
- 494 54. Haring R, Baumeister SE, Nauck M, Volzke H, Keevil BG, Brabant G, et al. Testosterone
495 and cardiometabolic risk in the general population - the impact of measurement method on
496 risk associations: A comparative study between immunoassay and mass spectrometry. *Eur*
497 *J Endocrinol.* 2013 Oct;169(4):463–70.
- 498 55. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al.
499 The PRIDE database and related tools and resources in 2019: improving support for
500 quantification data. *Nucleic Acids Res.* 2019 Jan 8;47(D1):D442–50.

501

502 SUPPLEMENTARY MATERIAL

503 *Supplementary method description*

504 Laboratory tests and dual X-ray absorptiometry

505 The analysis of the clinical chemistry parameters as well as bone mineral density were conducted
506 as previously described (1) in the study on infertile men. Markers measured in fasting morning
507 plasma samples included: total T, luteinizing hormone (LH), follicle stimulating hormone (FSH),
508 estradiol (E2), glucose, insulin, apolipoprotein A1, and apolipoprotein B. The parameters were
509 analyzed at the Department of clinical chemistry, Skåne University Hospital, Malmö, Sweden.
510 Total testosterone, LH and FSH were measured by electrochemiluminescence immunoassay
511 (ECLIA) and an immunofluorometric method (DELFI A Estradiol, Wallac OY, Finland) was
512 utilized to measure E2. Glucose was measured by an automated hexokinase method and insulin
513 was measured by immunometric sandwich assay. Bone mineral density (BMD, g/cm²) at total hip
514 (TH) and lumbar spine L1–L4 (LS) was measured using dual X-ray absorptiometry.

515 CAG-repeat length

516 Androgen receptor CAG-repeat length was determined in all the infertile men. For the analysis,
517 DNA was extracted from peripheral leukocytes. The androgen receptor CAG stretch was amplified
518 by polymerase chain reaction (PCR), purified and directly sequenced in a Beckman Coulter CEQ
519 2000 XL (Beckman Coulter, Bromma, Sweden) as previously described (2).

520 Proteomics experiments

521 All chemical reagents for the proteomic analysis not specified in the text were purchased from

522 Sigma Aldrich (St. Louis, MO). Modified porcine trypsin was from Promega (Madison, WI), water
523 from Milli-Q Ultrapure Water system (Millipore, Billerica, MA). Water and organic solvents for
524 LC-MS were of LC-MS quality (Merck, Darmstadt, Germany).

525 The plasma and serum samples from respective cohort (healthy men and infertile patients) were
526 randomized regarding the order before starting sample preparation. Quantitation of total protein
527 content in the samples was performed using the bicinchoninic acid (BCA) assay (3). Following
528 the instructions provided by the manufacturer, the top seven most-abundant proteins (albumin,
529 IgG, IgA, transferrin, haptoglobin, antitrypsin and fibrinogen) were depleted from approximately
530 10 μ L of each sample using a MARS7 column (Agilent, Santa Clara, CA). The buffer was
531 exchanged to 1.6% SDC, 50 mM NH_4HCO_3 with Amicon ultra-centrifugal filters (0.5 mL, 10 kDa
532 cut-off, Millipore, Tullagreen, Ireland). The disulphide bonds were reduced by adding DTT to a
533 final concentration of 10 mM and incubated 1 h at 37 °C. The free thiol groups were alkylated by
534 adding iodoacetamide to a final concentration of 25 mM, and the reaction proceeded for an
535 additional 30 min at room temperature in darkness (reduction and alkylation occurred in the
536 Amicon filter). The buffer was exchanged to 50 mM NH_4HCO_3 and the samples were resuspended
537 in 100 μ L 50mM NH_4HCO_3 (30 μ g protein after BCA quantitation) and digested with trypsin at
538 an enzyme-to-substrate mass ratio of 1:30 (w/w) for 16 h at 37°C. The remaining SDC was
539 precipitated by adding 20% formic acid prior to filtering the samples through a polypropylene filter
540 plate with a hydrophilic PVDF membrane (mean pore size 0.45 μ m, Porvair Filtration Group,
541 Fareham, UK).

542 A Q-Exactive Plus mass spectrometer connected to an Easy-nLC 1000 liquid chromatography
543 system (Thermo Scientific, San José, CA, USA) was used to analyse the samples for automated

544 Data Dependent Acquisition (DDA) methods. Peptides (2 μ L, 1 μ g on the column) were loaded
545 onto an Acclaim PepMap 100 precolumn (75 μ m \times 2 cm, Thermo Scientific, San José, CA, USA)
546 and separated on an easy-Spray column (25 cm \times 75 μ m i.d., PepMap C18 2 μ m, 100 Å) at a flow
547 rate of 300 nL/min and a column temperature of 35°C. Solvent A (0.1% formic acid) and solvent
548 B (0.1% formic acid in acetonitrile) were used to create a nonlinear gradient to elute the peptides
549 (90 min).

550 The MS1 spectra of the peptides were acquired in the orbitrap mass analyser from m/z 400–1600
551 and a resolution of 70,000 (at m/z 200). The target automated gain control (AGC) value and
552 maximum injection time (IT) were 1e6 and 100 ms, respectively. The ten most intense peaks with
553 charge state ≥ 2 were fragmented by higher-energy collisional dissociation (HCD) with 26%
554 normalised collision energy. Tandem mass spectra were acquired in the orbitrap mass analyser at
555 a resolution of 35,000 (at m/z 200), a target AGC value of 5e4 and a maximum IT of 100 ms. The
556 underfill ratio was 10% and dynamic exclusion was 45 s. The normalised collision energy was
557 26%.

558 Selected proteins from the healthy individuals were monitored also in patients by Multiple
559 Reaction Monitoring (MRM) method. A TSQ Quantiva mass spectrometer equipped with an
560 EASY-Spray NG ion source and connected to an EASY n-LC 1000 liquid chromatography system
561 (Thermo Scientific, San José, CA, USA) was used for the MRM method. Peptides (5 μ L, 1 μ g on
562 the column) were loaded onto an Acclaim PepMap 100 precolumn (100 μ m \times 2 cm, Thermo
563 Scientific, San José, CA, USA) and separated on an easy-Spray column (15 cm \times 75 μ m i.d.,
564 PepMap C18 3 μ m, 100 Å) at a flow rate of 300 nL/min and a column temperature of 35°C. Solvent
565 A (0.1% formic acid) and solvent B (0.1% formic acid in acetonitrile) were used to create a

566 nonlinear gradient to elute the peptides (60 min). Internal references, isotopically labelled synthetic
567 peptides (UVic-Genome BC Proteomics Centre, BC, Canada) were added in order to determine
568 relative quantifications of proteins. Data was acquired in scheduled MRM mode with 5 min
569 detection windows. SRM transitions were acquired in Q1 and Q3 operated at unit resolution (0.7
570 full width at half maximum, FWHM), the collision gas pressure in Q2 was 1.5 mTorr. The cycle
571 time was 2 s, and calibrated radio frequency (RF) and S-lens values were used. At least three
572 transition per precursor were monitored and in most cases the ratio is reported between the relative
573 total intensities vs the references.

574 For peptide and protein identification the Proteome Discoverer v2.2 (PD) software (Thermo
575 Scientific, San José, CA) was used. Peptides were identified using SEQUEST HT against UniProt
576 (<http://www.uniprot.org>) human database (Human 9606, Reviewed, 20 165) integrated into PD.
577 The search was performed with the following parameters: cysteine carbamidomethylation as a
578 static modification, oxidation of methionine as a dynamic modification, 10 ppm precursor
579 tolerance and 0.02 Da fragment tolerance. Up to one missed cleavage for tryptic peptides was
580 allowed. According to the PD software, the filters applied were high- (FDR<0.01) and medium
581 confidence (FDR<0.05) at peptide- and protein levels, respectively. The peptide and protein
582 quantifications were based on the quantification of MS peptide signals (label-free quantification).
583 Label-free quantification used the Minora Feature Detector node in the processing workflow, and
584 the Precursor Ions Quantifier node and the Feature Mapper in the consensus workflow.

585 To analyse the MRM data, all MS files were imported into Skyline v3.5(4) (MacCoss Lab
586 Software, Seattle, WA). The peak integration was performed automatically by the software, using
587 Savitzky-Golay smoothing and to confirm the correct peak detection it was in addition, manually

588 inspected. The relative total intensities vs the references (ratios), were reported as expression
589 values. Only the values that passed the inspected intensity selection criteria according to PD or in
590 Skyline were reported.

591 References (Supplementary methods)

592

593 1. Bobjer J, Bogefors K, Isaksson S, Leijonhufvud I, Åkesson K, Giwercman YL, et al. High
594 prevalence of hypogonadism and associated impaired metabolic and bone mineral status in
595 subfertile men. Clin Endocrinol (Oxf) [Internet]. 2016 Aug [cited 2018 Nov 26];85(2):189–
596 95. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26857217>

597 2. Lundin KB, Giwercman A, Richthoff J, Abrahamsson PA, Giwercman YL. No association
598 between mutations in the human androgen receptor GGN repeat and inter-sex conditions.
599 Mol Hum Reprod [Internet]. 2003 [cited 2021 Jan 14];9(7–8):375–9. Available from:
600 <https://pubmed.ncbi.nlm.nih.gov/12802043/>

601 3. Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, et al.
602 Measurement of protein using bicinchoninic acid. Anal Biochem. 1985;150(1):76–85.

603 4. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline:
604 an open source document editor for creating and analyzing targeted proteomics
605 experiments. Bioinformatics. 2010 Apr;26(7):966–8.

606

608 *Supplementary Statistical analyses*

609 *Identification of candidate marker proteins from chemically castrated men*

610 Proteins identified by mass spectrometry techniques in samples from healthy subjects were pre-
611 processed in Perseus v1.6.7.0 (1) software. Intensities were Log₂ transformed to follow a normal
612 distribution and standardized by subtracting the median of its respective sample from each value.
613 The statistical analyses were performed using R software (2,3) and SPSS Statistics 21.0 (IBM,
614 Somers, IL, USA). Differentially expressed proteins were determined by doing one-way repeated
615 measures ANOVA (R function: ezANOVA{ez}) to reveal overall differences between conditions
616 (normal (A), low (B) and restored (C) T) and differences between individual conditions were
617 detected by performing a post-hoc test based on pairwise t-test (two-tails and paired) (R function:
618 pairwise.t.test{stats}). The 'pairwise.t.test' function utilized the proteins with significant overall
619 changes (ANOVA *p*-value < 0.05) to perform pairwise comparisons between conditions while
620 corrected for multiple pairwise testing. Proteins with adjusted *p*-values ('*fdr*' method) < 0.05
621 following the pairwise t-test were considered significant.

622 Because gonadotropins (FSH, LH) did not change their levels between conditions B and C, while
623 T did recover its value (4), proteins that changed with statistical significance between A-B and
624 recovered their values significantly between B and C were considered mainly influenced by T
625 changes and, hence, were selected for further analyses.

626 Based on the outcome of the first selection step of the analysis, we performed receiver operating
627 characteristic (ROC) analysis to select proteins capable to discriminate between normal and low
628 T. Two categories of samples were created: 'low T' that included samples evaluated at time point

629 B and ‘normal T’ that included samples evaluated at time point A and C. Although there is no
630 consensus about the threshold for sufficient diagnostic accuracy, the area under the curve (AUC)
631 can be interpreted as the following rough classifying system: > 0.90 = excellent; $0.80 - 0.90$ = very
632 good; $> 0.70 - 0.80$ = good; $0.60 - 0.70$ = sufficient and $0.50 - 0.60$ = poor (5–7). Consequently,
633 based on the ROC-AUC values, we created two cut-off values to select the final protein
634 biomarkers: a) AUC > 0.80 (i.e. very good) and b) AUC $0.75 - 0.80$ (i.e. good) and highly enriched
635 in liver tissues (according to the Human Proteome Map (8) and Kampf C. et. al. (9)), which, unlike
636 other organs, appears to be relatively protected from age-related changes (10,11). The Human
637 Proteome Map tool (8) was used to determine the tissue-specific localization of the proteins.

638 Because the combination of different markers may improve the discriminative power to diagnose
639 hypogonadism and predict its long term sequelae, proteins selected from the ROC-AUC analysis
640 were included as predictors in a stepwise regression (method: backward) to select the best
641 combination of markers that predict the odds of being low T. Bootstrap resampling with
642 replacement method was applied to assess consistency of predictors selected with the stepwise
643 regression. A new variable called Multi Marker Algorithm (MMA) was derived from the predicted
644 log-odds (of being low T) obtained from a binomial logistic regression analysis, in which the
645 dichotomized variable ‘testosterone level’ (0: normal T (time points A, C) and 1: low T (time point
646 B)) was used as the dependent variable and the intensities of the selected proteins from the stepwise
647 regression were the predictors' variables. Finally, MMA variable was evaluated together with
648 marker candidates proteins. Subsequently, an ROC analysis was performed based on the predicted
649 values and the Youden index method (12) was used to select the optimal cut-off point for the
650 sensitivity and specificity of each marker. All analyses described in this part were performed in R

651 software and plotted using `ggroc{pROC}` function.

652 *Background characteristics of the infertile cohort of patients*

653 To describe background characteristics of the infertile cohort of patients (**Table 1a**), the mean and
654 standard deviation (SD) was calculated for variables that followed a normal distribution or the
655 median and minimum-maximum (min-max) for variables that were not normally distributed (non-
656 Gaussian). The normal distribution was checked by applying a Kolmogorov-Smirnov test.

657 *Testing of the candidate biomarkers in infertile men*

658 To further perform statistical analyses, the quantified intensities of the candidate biomarkers in
659 infertile men were Log2 transformed to achieve normal distributions. Markers selected from the
660 stepwise regression in the healthy human model were also considered in this cohort to create the
661 MMA variable. In this case, MMA was created to predict the odds of suffering low T or other
662 medical conditions associated with low T levels. Unless other software is specified, the statistical
663 analyzes described in the next steps were performed using R software (2,3).

664 *Distinguishing infertile men with different testosterone levels*

665 In order to know if the change in the proteins occur with the change in T as observed in the healthy
666 human model, we created three groups of patients based on total T concentration values commonly
667 described in clinical guidelines (13) to define T deficiency: Group 1 contains patients commonly
668 defined with low T (LT): ≤ 8 nmol/L (n = 22), Group 2 contains patients defined with borderline
669 low T (BL_T): between 8 and 12 nmol/L (n=17), and Group 3 contains patients with normal T
670 (NT): > 12 nmol/L (n = 36). Differences among the three groups of patients were detected by

671 doing a one-way ANOVA per protein (outcome variable: protein intensity) followed by a pairwise
672 analysis that was corrected to control the FDR provoked by multiple pairwise comparisons (Two-
673 stage linear step-up procedure of Benjamini, Krieger and Yekutieli (14), performed in GraphPad
674 Prism software, version 9.00 for Windows, CA, USA). Adjusted p-values < 0.05 were considered
675 significant. An ROC analysis was performed to discriminate the patients with LT from those with
676 BL_T or NT. The same analysis was repeated after excluding the BL_T patients as this group
677 includes males in the gray zone for hypogonadism diagnosis (15–17).

678 *Distinguishing between men with abnormal cardio-metabolic parameters or reduced bone mineral*
679 *density*

680 In order to evaluate the diagnostic value of our new protein markers in relation to conditions
681 associated with low T levels, we performed a ROC analysis to discriminate patients with MetS,
682 IR, CVRLP, DM or LBD within the patient cohort using the expression level of the candidate
683 biomarkers and T. Based on the ROC curves, the DeLong's test (paired) was applied in R (roc.test
684 {pROC} function) to compare the AUCs of the candidate biomarkers vs the AUC of T levels to
685 discriminate the above mentioned pathological conditions.

686 *Association with Androgen receptor CAG-repeat length*

687 To investigate the androgenic dependence of the levels of the newly defined protein markers, we
688 looked for associations between CAG repeat length and concentration of those markers. The
689 infertile cohort was separated into three groups based on CAG repeat length of the androgen
690 receptor (AR); a group of patients with CAG repeat length 21 and 22 (n=18) was defined as
691 reference group (18–21). Patients with CAG repeat length less than the reference group (< 21)

692 were included in a group (n=26) and patients with CAG repeat length longer than the reference
693 group (>22) made up the third group (n=30). Using the software GraphPad Prism version 9.0.0 for
694 Windows (CA, USA), a one-way ANOVA followed by a pairwise test (two-stage linear step-up
695 procedure of Benjamini, Krieger and Yekutieli (14)) which controls FDR for multiple pairwise
696 comparison was conducted and adjusted p-values < 0.05 were considered significant.

697 *References (Supplementary statistics)*

- 698 1. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus
699 computational platform for comprehensive analysis of (prote)omics data [Internet]. Vol. 13,
700 Nature Methods. Nature Publishing Group; 2016 [cited 2020 Nov 20]. p. 731–40. Available
701 from: <https://pubmed.ncbi.nlm.nih.gov/27348712/>
- 702 2. RStudio Team. RStudio: Integrated Development for R. RStudio. Boston; 2016.
- 703 3. Team RC. R: A language and environment for statistical computing. R Foundation for
704 Statistical Computing. 2016.
- 705 4. Sahlin KB, Pla I, Sanchez A, Pawłowski K, Leijonhufvud I, Appelqvist R, et al. Short-term
706 effect of pharmacologically induced alterations in testosterone levels on common blood
707 biomarkers in a controlled healthy human model. Scand J Clin Lab Invest. 2019 Nov;1–7.
- 708 5. Marshall KW, Mohr S, Khettabi F El, Nossova N, Chao S, Bao W, et al. A blood-based
709 biomarker panel for stratifying current risk for colorectal cancer. Int J Cancer [Internet].
710 2009 Mar 1 [cited 2020 Nov 20];126(5):NA-NA. Available from:
711 <http://doi.wiley.com/10.1002/ijc.24910>

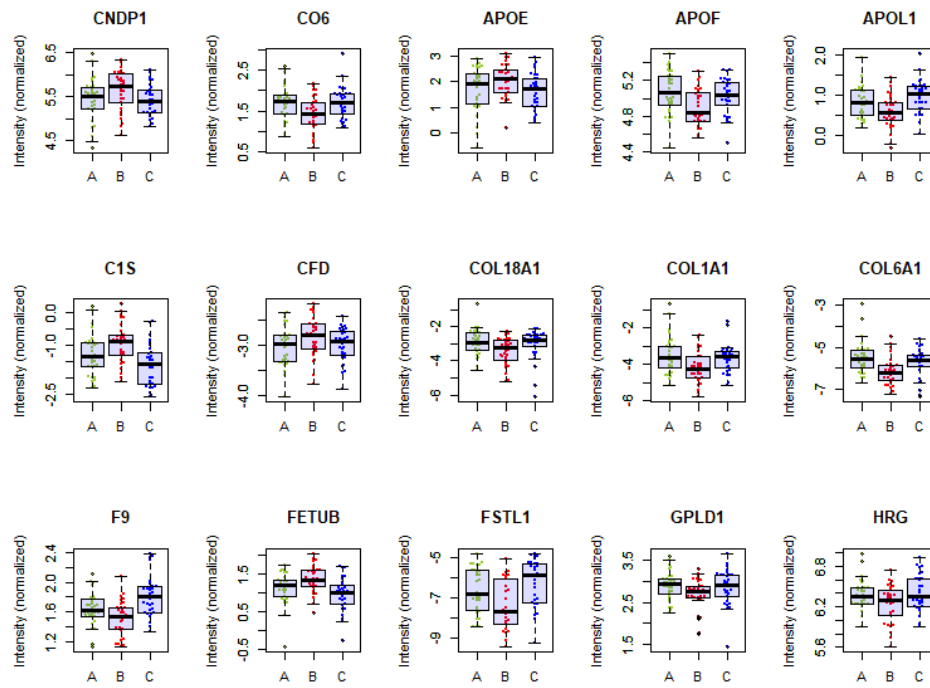
- 712 6. Bowers AJ, Zhou X. Receiver Operating Characteristic (ROC) Area Under the Curve
713 (AUC): A Diagnostic Measure for Evaluating the Accuracy of Predictors of Education
714 Outcomes. *J Educ Students Placed Risk* [Internet]. 2019 Jan 2 [cited 2020 Nov
715 20];24(1):20–46. Available from:
716 <https://www.tandfonline.com/doi/abs/10.1080/10824669.2018.1523734>
- 717 7. Šimundić A-M. Measures of diagnostic accuracy: basic definitions.
- 718 8. Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, et al. A draft map of
719 the human proteome. *Nature* [Internet]. 2014 [cited 2020 Dec 4];509(7502):575–81.
720 Available from: <https://pubmed.ncbi.nlm.nih.gov/24870542/>
- 721 9. Kampf C, Mardinoglu A, Fagerberg L, Hallström BM, Edlund K, Lundberg E, et al. The
722 human liver-specific proteome defined by transcriptomics and antibody-based profiling.
723 *FASEB J* [Internet]. 2014 Jul 1 [cited 2021 Apr 29];28(7):2901–14. Available from:
724 <https://pubmed.ncbi.nlm.nih.gov/24648543/>
- 725 10. Kholodenko I V, Yarygin KN. Cellular Mechanisms of Liver Regeneration and Cell-Based
726 Therapies of Liver Diseases. 2017;
- 727 11. Schmucker DL, Sanchez H. Liver Regeneration and Aging: A Current Perspective. *Curr*
728 *Gerontol Geriatr Res*. 2011;2011.
- 729 12. Youden WJ. Index for rating diagnostic tests. *Cancer* [Internet]. 1950 Jan 1 [cited 2020 Nov
730 20];3(1):32–5. Available from:
731 <https://acsjournals.onlinelibrary.wiley.com/doi/full/10.1002/1097->

- 732 0142%281950%293%3A1%3C32%3A%3AAID-CNCR2820030106%3E3.0.CO%3B2-3
- 733 13. Arver S, Lehtihet M. Current guidelines for the diagnosis of testosterone deficiency
734 [Internet]. Vol. 37, Frontiers of Hormone Research. Front Horm Res; 2009 [cited 2020 Nov
735 19]. p. 5–20. Available from: <https://pubmed.ncbi.nlm.nih.gov/19011285/>
- 736 14. Benjamini Y, Krieger AM, Yekutieli D. Adaptive linear step-up procedures that control the
737 false discovery rate. *Biometrika*. 2006 Sep;93(3):491–507.
- 738 15. Chan I, Fui MN, Zajac J, Grossmann M. Assessment and management of male androgen
739 disorders: an update - PubMed [Internet]. *Aust Fam Physician*. [cited 2020 Nov 20].
740 Available from: <https://pubmed.ncbi.nlm.nih.gov/24791767/>
- 741 16. Lunenfeld B, Arver S, Moncada I, Rees DA, Schulte HM. How to help the aging male?
742 Current approaches to hypogonadism in primary care [Internet]. Vol. 15, *Aging Male*.
743 *Aging Male*; 2012 [cited 2020 Nov 20]. p. 187–97. Available from:
744 <https://pubmed.ncbi.nlm.nih.gov/23067307/>
- 745 17. Zitzmann M, Faber S, Nieschlag E. Association of specific symptoms and metabolic risks
746 with serum testosterone in older men. *J Clin Endocrinol Metab* [Internet]. 2006 [cited 2020
747 Nov 20];91(11):4335–43. Available from: <https://pubmed.ncbi.nlm.nih.gov/16926258/>
- 748 18. Casella R, Maduro MR, Lipshultz LI, Lamb DJ. Significance of the polyglutamine tract
749 polymorphism in the androgen receptor. *Urology*. 2001;58(5):651–6.
- 750 19. Stanworth RD, Kapoor D, Channer KS, Jones TH. Androgen receptor CAG repeat

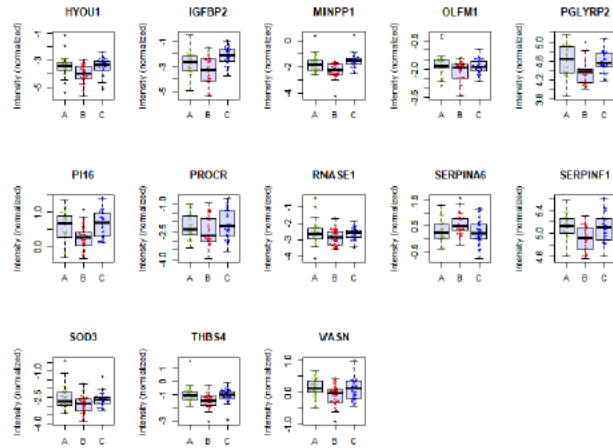
- 751 polymorphism is associated with serum testosterone levels, obesity and serum leptin in men
752 with type 2 diabetes. *Eur J Endocrinol.* 2008;159(6):739–46.
- 753 20. Kim JW, Bae YD, Ahn ST, Kim JW, Kim JJ, Moon DG. Androgen Receptor CAG Repeat
754 Length as a Risk Factor of Late-Onset Hypogonadism in a Korean Male Population. *Sex*
755 *Med* [Internet]. 2018;6(3):203–9. Available from:
756 <https://doi.org/10.1016/j.esxm.2018.04.002>
- 757 21. Ferlin A, Bartoloni L, Rizzo G, Roverato A, Garolla A, Foresta C. Androgen receptor gene
758 CAG and GGC repeat lengths in idiopathic male infertility. *Mol Hum Reprod.*
759 2004;10(6):417–21.
- 760

Supplementary Figure

Fig. S1. Boxplots based on the Log₂ intensities of the differentially expressed proteins. The healthy human model (i.e. castrated men).



It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).



762

763 **Supplementary tables**

Table S1. Metabolic syndrome criteria (NCEP, 2002)	
	NCEP-ATP III
Hypertension	Systolic BP \geq 130 mmHg and diastolic BP \geq 85 mmHg or antihypertensive treatment
Obesity	Waist circumference \geq 102 cm
Insulin resistance	Fasting P-Glucose \geq 6.1 mmol/L or treatment for DM
Dyslipidemia	P-HDL-cholesterol $<$ 1.0 mmol/L or statin usage
Hypertriglyceridemia	Fasting P-Triglycerides \geq 1.7 mmol/L
Metabolic syndrome	\geq 3 criteria
<i>NCEP-ATP III, National Cholesterol Education Program Adult Treatment Panel III 2002; BP, blood pressure; DM, diabetes mellitus; HDL, high-density lipoprotein.</i>	
Supplemental references	
NCEP. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. <i>Circulation</i> 2002;	

764

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/) .

Table S2. Significant proteins selected in the healthy human model from Step 1 and 2.
* bootstrap Coefficient Interval (CI)

#	Analyt	ROC		Log2.FC (B-A)	Log2.FC (C-B)	Log2.FC (C-A)	ANOVA p.val	post hoc paired t-test			significant changes	Sensitivity	Specificity
		AUC (%) [95% CI]*	p-value					adj.p (B - A)	adj.p (B - C)	adj.p (A - C)			
1	4HPPD	87 [77;96]	0.00000265	0.933	-1.253	-0.319	0.0000008	0.0010467	0.0000316	0.0225833	A-B; B-C; A-C;	85	82
2	IGFBP6	84 [75;91]	0.00000020	-0.511	0.748	0.237	0.0000000	0.0000188	0.0000002	0.0214264	B-C; A-C;A-B;	97	66
3	ALDOB	75 [63;86]	0.00011329	0.603	-0.658	-0.055	0.0000142	0.0007867	0.0003664	0.9998779	A-B; B-C;	73	67
4	MINPP1	80	0.00000261	-0.513	0.770	0.257	0.0000044	0.0014091	0.0000121	0.0948021	A-B; B-C;		
5	PII6	76	0.00004969	-0.347	0.446	0.099	0.0000002	0.0005094	0.0000089	0.0603732	A-B; B-C;		
6	IGFBP2	74	0.00019011	-0.640	1.185	0.545	0.0000000	0.0019273	0.0000000	0.0012160	A-B; B-C; A-C;		
7	THBS4	74	0.00027500	-0.478	0.381	-0.097	0.0026760	0.0062151	0.0062151	0.5141803	A-B; B-C;		
8	PGLYRP2	73	0.00030386	-0.247	0.226	-0.020	0.0001658	0.0004491	0.0004491	0.7698965	A-B; B-C;		
9	COL6A1	73	0.00046401	-0.655	0.403	-0.252	0.0005029	0.0004095	0.0283727	0.1351927	A-B; B-C;		
10	CNDP1	73	0.00051078	0.336	-0.660	-0.324	0.0000172	0.0223163	0.0000401	0.0090046	A-B; B-C; A-C;		
11	SERPINF1	73	0.00052732	-0.206	0.194	-0.012	0.0000030	0.0000539	0.0001855	0.7517160	A-B; B-C;		
12	HYOU1	72	0.00072252	-0.552	0.508	-0.043	0.0040436	0.0043378	0.0174434	0.8124464	A-B; B-C;		
13	FSTL1	71	0.00181638	-0.672	1.130	0.458	0.0042720	0.0313565	0.0237510	0.1408092	A-B; B-C;		
14	FETUB	71	0.00104480	0.235	-0.393	-0.158	0.0005134	0.0472545	0.0003591	0.0964466	A-B; B-C;		
15	F9	71	0.00141005	-0.112	0.287	0.176	0.0000028	0.0355388	0.0000087	0.0036811	A-B; B-C; A-C;		
16	CFD	70	0.00173224	-0.238	0.375	0.137	0.0001221	0.0142391	0.0000846	0.1215239	A-B; B-C;		
17	C1S	69	0.00371295	-0.170	0.123	-0.047	0.0006948	0.0002338	0.0137485	0.3332313	A-B; B-C;		
18	VASN	68	0.00473487	-0.229	0.226	-0.003	0.0037176	0.0065467	0.0065467	0.9715587	A-B; B-C;		
19	COL1A1	68	0.00578005	-0.711	0.540	-0.172	0.0011780	0.0073810	0.0074089	0.3411692	A-B; B-C;		
20	APOF	67	0.00838313	-0.273	0.269	-0.004	0.0078117	0.0094437	0.0094437	0.9747496	A-B; B-C;		
21	SOD3	67	0.00859700	-0.356	0.321	-0.034	0.0074186	0.0185355	0.0101178	0.7844501	A-B; B-C;		
22	COL18A1	67	0.00881574	-0.521	0.381	-0.140	0.0105982	0.0204482	0.0299038	0.4457830	A-B; B-C;		
23	SERPINA6	67	0.00974125	0.241	-0.277	-0.036	0.0155489	0.0216347	0.0216347	0.7296250	A-B; B-C;		
24	APOE	66	0.01506488	0.216	-0.217	-0.002	0.0079678	0.0216084	0.0216084	0.9821841	A-B; B-C;		
25	RNASEI	65	0.02036546	-0.286	0.258	-0.028	0.0399674	0.0382668	0.0382055	0.8438576	A-B; B-C;		
26	GPLD1	65	0.02179807	-0.214	0.178	-0.036	0.0067828	0.0190366	0.0232320	0.5912440	A-B; B-C;		
27	CO6	64	0.03599301	0.196	-0.163	0.033	0.0254637	0.0260492	0.0471678	0.5009590	A-B; B-C;		
28	APOL1	63	0.04338613	0.30	-0.40	-0.10	0.0074690	0.0330538	0.0188492	0.4331313	A-B; B-C;		
29	OLFM1	62	0.06086609	-0.344	0.260	-0.084	0.0053622	0.0187015	0.0276303	0.3909882	A-B; B-C;		
30	PROCR	62	0.06205427	-0.282	0.375	0.093	0.0001748	0.0064107	0.0029897	0.1488471	A-B; B-C;		
31	HRG	62	0.07226674	-0.138	0.159	0.021	0.0072758	0.0292888	0.0215359	0.6555564	A-B; B-C;		

765

766