Age and Sex Effects on Plasma Metabolite Association Networks in Healthy Subjects

Vignoli, A., Tenori, L., Luchinat, C., & Saccenti, E.

Age and sex effects on plasma metabolite association networks in healthy subjects

Alessia Vignoli[1], Leonardo Tenori[1,2], Claudio Luchinat[1,3,*], and Edoardo Saccenti[4,*]

[1]Magnetic Resonance Center (CERM), University of Florence, Sesto Fiorentino, Italy

[2]Department of Clinical and Experimental Medicine, University of Florence, Florence, Italy

[3]Department of Chemistry, University of Florence, Sesto Fiorentino, Italy

[4]Laboratory of Systems and Synthetic Biology, Wageningen University & Research, Wageningen, The Netherlands.

*Corresponding Authors:

Edoardo Saccenti, Email: esaccenti@gmail.com; Laboratory of Systems and Synthetic Biology, Wageningen University & Research, Stippeneng 4, 6708 WE, Wageningen, the Netherlands.; Tel: +31 (0)317 482018; Fax: +31 (0) 3174

Claudio Luchinat, Email: luchinat@cerm.unifi.it; Magnetic Resonance Center (CERM), University of Florence, Via Luigi Sacconi 6, 50019, Sesto Fiorentino (Florence), Italy; Tel: +39 055 457 4296.

Abstract

In the era of precision medicine, the analysis of simple information like sex and age can increase the potential to better diagnose and treat conditions that occur more frequently in one of the two sexes, present sex-specific symptoms and outcomes, or are characteristic of a specific age group. We present here a study of the association networks constructed from an array of 22 plasma metabolites measured on a cohort of 844 healthy blood donors. Through differential network analysis we show that specific association networks can be associated with sex and age: different connectivity patterns where observed suggesting sex-related variability in several metabolic pathways (branched-chain amino acids, ketone bodies and propanoate metabolism). Reduction in metabolite hubs connectivity was also found to be associated with age in both sex groups. Network analysis was complemented with standard univariate and multivariate statistical analysis that revealed age and sex specific metabolic signatures. Our results demonstrate that the characterization of metabolite–metabolite association networks is a promising and powerful a tool to investigate the human phenotype at a molecular level.

Keywords

NMR, Metabolomics, Differential network analysis, Metabolism, Network inference

INTRODUCTION

Simple type information like sex and age can prove valuable information within a precision medicine approach to investigate and to define disease risk and susceptibility in the human population. Although the molecular mechanisms are yet to be fully understood, it has been long known that differences between the two sexes affect manifestation, epidemiology and pathophysiology of many widespread diseases and, therefore, require different approaches to health care:[1] for instance, drug efficacy and toxicity profiles are affected by sexual dimorphisms[2] as well the levels and associations of insulin, cholesterol, very low density lipoprotein (VLDL), and certain triacylglycerols.[3] Consistently, also the human metabolome is affected by endogenous factors such as sex and age whose effects have been investigated through metabolomics-based studies.

Differences in the concentration levels of several urine and blood metabolites (creatinine, citrate, glycine and hippuric acid) or class of metabolites (breached-chain amino acids and lipids) have been found to be different between men and women.[4–7]

Age-related considerations are increasingly being taken into account in medical therapy, such as in the case of cancer treatment;[8] but while sexual dimorphisms is a dichotomic variable providing a clear discrimination of the groups of interest, age and ageing are elusive concepts and new insight into systemic metabolic patterns associated with this variable are needed. Indeed, not only the definition of young and old and the discriminant among them depends on socioeconomic and geographical considerations, but the metabolomics study of aging requires large-scale longitudinal studies with replication due to the high variability observed among subjects in term of metabolic phenotypes.[8–13] Metabolites such as creatinine, creatine, aromatic amino acids, glycine and glutamate have been found to be associated to age-related differences and the correlation between metabolome and metabolic modification with blood leukocyte telomere length (LTL) and epigenetic modifications have been also investigated,

identifying several classes of metabolites (glycerophosphoethanolamines, glycerophosphocholines, glycerolipids, bile acids, isoprenoids, fatty amides, and L-carnitine esters) significantly associated with LTL.[14–16] Attempts have also been made to characterize the real "metabolic" age via metabolomic approaches.[7,17 1-6]These findings provide ample evidences that the human metabolome is highly influenced by this factor whose effects sum with those induced by environment, stress conditions and other exogenous factors, contributing to the shaping of the human metabolic phenotype.[9–11,18]

However, despite the large efforts spent in characterizing age and sex-related characteristics of (part of) the human metabolome, a clear mechanistic interpretation and explanation of these findings is still missing. This reflects in part, our limited understanding and knowledge of secondary metabolism, and in part the fact that most metabolomics studies are exploratory in nature. In the attempt to move one step further towards mechanistic understanding, in this study, we take a systems biology approach to investigate sex and age-specific differences in the plasma metabolite profiles of healthy subjects by considering a group of young and middle aged healthy subjects of both sexes; participants were from a larger cohort of healthy blood donor volunteers who were analyzed for their plasma metabolite concentration profiles using Nuclear Magnetic Resonance (NMR) spectroscopy.[19]

In contrast with other age-related metabolomics studies, our data analysis strategy is not limited to standard univariate and multivariate analysis[20] to explore the patterns of variation of metabolites among the two age classes but deploy metabolite-metabolite association network analysis and inference.

Networks and network analysis are fundamental tools in systems biology and are being exploited more and more often to analyze, understand and interpret the complex patterns observed in metabolomics data.[21,22,19,13] The rationale underlying the use of network representation and analysis is that metabolite concentrations change in an orchestrated

fashion in such a way that the association among metabolite inferred from measured concentrations in biofluids can be considered, to some extent, related to the underlying structure of the biological networks. Indeed, different metabolite correlation patterns have been found to be associated with sexual dimorphism[23] or to change upon treatment or dietary intervention.[24,25]

The utility of network-based approaches to analyze metabolomics data relies on these two key concepts; in a metabolomics context networks are best exploited when compared across different conditions in a so-called differential-network analysis approach: different network characteristics and different patterns of association between metabolites can highlight possibly affected molecular mechanisms. When investigating age and sex-related effect we expect these endogenous factors to influence not only the metabolite levels but also the patterns of metabolite associations whose changes can be more significant than those of levels alone, as previously shown phenotyping studies[23,25]: through network analysis in a metabolomics context was possible, for instance, to identify possible mechanisms of action underlying latent cardiovascular risk status in healthy subjects[19] or modelling the variability observed in the human urinary metabolic phenotype.[13]

MATERIALS AND METHODS

Data description

The study population consists of 844 adult healthy volunteers (661 males, 183 females, median age 41±12 years, see Table 1 for an overview) recruited in collaboration with the Tuscany section of the Italian Association of Blood Donors (AVIS) in the Transfusion Service of the Pistoia Hospital (Ospedale del Ceppo, AUSL 3 - Pistoia, Italy). According to AVIS rules for blood donations, volunteers had to do not take (or have recently taken) drugs for a variable

period of time according to the active substance, the pharmacokinetics of the prescribed drug, and the disease being treated.

Plasma samples were obtained after overnight fasting and ethylenediaminetetraacetic acid (EDTA) was used as anticoagulant, but its presence does not affect significantly the quality of the samples for the NMR analysis.[26] Plasma samples were collected, and immediately stored at -80 °C according to the Standard Operating Procedures (SOPs) described in Bernini *et al.* [27] NMR spectra and associated clinical data were retrieved from the Open-Access Database Repository MetaboLights[28] with the accession number MTBLS147 (http://www.ebi.ac.uk/metabolights). Twenty samples from the original dataset were excluded from this analysis because sex or age information were missing. Age distributions for both sexes are reported as histograms in the Figure 1.

In the original paper,[26] multivariate and univariate analyses were applied to infer a pattern of metabolic alterations that correlates with cardiovascular risk. The same dataset was then re-analyzed in a following paper[19] to define metabolite probabilistic networks specific for low and high latent cardiovascular risk, using a new approach based on systems biology and metabolite-metabolite correlation networks. Starting from the data collected in the former, and using the computational methodology developed in the latter, the present paper shifts the focus from cardiovascular risk to age and sex characterization. Consistent with previous observations of sex specific differences at the metabolic level[22] we performed the analysis separately for males and females to minimize confounding effects.

NMR sample preparation

Frozen plasma samples were thawed at room temperature and shaken before use.[27] A total of 300 µL of a sodium phosphate buffer (10.05 g $Na_2HPO_4 \cdot 7H_2O$; 0.2 g NaN_3; 0.4 g sodium trimethylsilyl [2,2,3,3-$^2H_4$] propionate (TMSP) in 500 mL of H_2O with 20% (v/v) ^2H_2O; pH 7.4)

was added to 300 μL of each plasma sample, and the mixture was homogenized by vortexing for 30 s. A total of 450 μL of this mixture was transferred into a 4.25 mm NMR tube (Bruker BioSpin srl) for the analysis.

NMR analysis and processing

Monodimensional ^{1}H NMR spectra for all plasma samples were acquired using a Bruker 600 MHz spectrometer (Bruker BioSpin) operating at 600.13 MHz proton Larmor frequency and equipped with a 5 mm CPTCI ^{1}H-^{13}C-^{31}P and ^{2}H-decoupling cryoprobe including a z axis gradient coil, an automatic tuning-matching (ATM) and an automatic sample changer. A BTO 2000 thermocouple served for temperature stabilization within an uncertainty of approximately 0.1 K at the sample. Before measurement, samples were kept for at least 3 minutes inside the NMR probehead for temperature equilibration at 310 K.

Water suppressed Carr–Purcell–Meiboom–Gill (CPMG)[29] spin echo pulse sequence (RD-90°-(τ-180°-τ)n-acq) with a total spin echo (2nτ) of 300 ms was used in order to obtain monodimensional ^{1}H NMR spectra in which broad signals from high molecular weight metabolites (i.e. proteins, lipids, and lipoproteins) are attenuated. 64 FIDs were collected into 73728 data points over a spectral width of 12019 Hz, with a relaxation delay (RD) of 4 s and acquisition time of 3.1 s. Free induction decays were multiplied by an exponential function equivalent to a 1.0 Hz line-broadening factor before applying Fourier transformation. Transformed spectra were automatically corrected for phase and baseline distortions and calibrated (anomeric glucose doublet at 5.24 ppm) using TopSpin 3.2 (Bruker Biospin srl).

For the multivariate statistical analysis, each 1D spectrum in the range 0.2-10.00 ppm was segmented into 0.02 ppm chemical shift bins and the corresponding spectral areas were integrated using AMIX 3.8.4 software (Bruker BioSpin srl). The region between 4.2 and 6.0 ppm containing the residual water signal was removed, thus the dimension of the system was

reduced to 451 bins. The total spectral area was calculated on the remaining bins and total area normalization was carried out on the data prior to pattern recognition.

Metabolite identification and quantification

The spectral regions related to the metabolites were assigned in the CPMG spectra retrieved from Metabolights repository by using matching routines of AMIX 3.8.4 (Bruker BioSpin) in combination with the BBIOREFCODE (Bruker BioSpin), freely available dataset i.e. Human Metabolome DataBase (HMDB)[30,31] and published literature when available. The metabolites quantification was determined by an in-house developed software in MATLAB programming suite (Mathworks, MATLAB version R2014b). The algorithm is based upon the unconstrained non-linear minimization (fitting) of the metabolites NMR signals, employing a combination of Lorentzian-Gaussian functions. With this approach, each NMR region of interest is decomposed and deconvoluted into its component parts, and then integrated to obtain the metabolite concentrations in arbitrary units. A list of the quantified metabolites together with the corresponding spectral region of interest and signal multiplicity is given in Supplementary Table S1

Statistical analysis

Univariate analysis

Wilcoxon signed-rank test[32] was used to assess statistically the existence of differences in metabolite concentrations between the groups. Bonferroni correction method[33] was applied to reduce the risk of false positive due to multiple testing. An adjusted *P*-value < 0.05 was deemed significant. Cliff's *delta* (Δ)[34] was calculated using the R package "effsize" (https://cran.r-project.org/web/packages/effsize/index.html) to obtain a non-parametric estimation of the

effect size, used to complement the *P*-values obtained from Wilcoxon signed-rank test. Univariate analysis was applied exclusively on the quantified metabolite concentrations.

Multivariate analysis

Standard partial least square discriminant analysis[35] (PLS-DA) was applied to discriminate the groups of interests: PLS models were built to discriminate between Males and Females, between Young males and Old males and between Young females and Old females. Both the full NMR CPMG spectra and the set of 22 quantified metabolites were used for a total of 6 discriminant PLS models.

PLS-DA is a PLS regression where the continuous response Y vectors is substituted with a vector containing dummy variables (-1,1) indicating group belonging of the samples. PLS regression is a generalization of multiple regression where a set of new variables (component) is defined that explain as much as possible of the covariance between the predictor (X) and the response Y. The number of PLS components to fit the model to achieve an optimal prediction model is optimized using a double cross validation (2CV) cross-validation strategy as detailed in reference.[36] Briefly, the 2CV scheme consists of two nested loops CV1 and CV2: the CV1 is used to optimize the number of components to be used in the PLS-DA model while the CV2 is to assess final model performance. In the outer loop (CV2) the complete dataset is split into a test set and a rest set: the test set is set aside, and the rest set is used in the loop CV1. Within CV1 the rest set is split into a validation set and a training set. The training set is used to develop a series of PLS-DA models with different number of latent variables from which the samples in the validation set are predicted: the optimal number of components is chosen that maximize the AUROC (*i.e*. maximize the prediction power of the model). AUROC is the Area Under the Receiver Operating Curve and summarize both sensitivity and specificity of the discrimination

model and range between 0 and 1 with indicating perfect discrimination. The AUROC has been found to be most appropriate measure to optimize PLS-DA model.[36]

Finally, a PLS-DA model is built using the optimal number of components on the data in the test set to assess the quality of the PLS-DA model. Using a 2CV scheme calibration and quality assessment of the model are kept separated: in this way the samples used to assess the performance of the model are not used to calibrate the model and this guarantees unbiased model calibration and assessment. For an overview of the 2CV calibration procedure see Figure 1 in reference.[36] We used a 4:3 data splitting meaning that the full data was split in 4 parts: 1 part was retained as test set (CV2); the remaining 3 (the rest set) was split in 3 parts in the CV1 where a 3:2 split was used for model calibration. The overall 2CV procedure was repeated 100 times to account for possible differences due to the random data splitting. AUROC and Accuracy for a given discrimination model are given as average values over the 100 repetitions. Significance of the classification results was assessed by means of a permutation test using 10^3 permutations.[36]

In the case of the comparison of Male *vs* Females the group size is unbalanced. To compensate for this 150 male samples and 150 female samples are randomly chosen from the full data set and subjected to PLS-DA modelling as described above. The resampling procedure is then performed 100 times to account for variability in the sampling procedure. The AUROC values reported are average over the 100 resamplings.

Network Analysis

Network reconstruction

Three different algorithms were used to infer metabolite-metabolite association networks together with the standard correlation approach. A brief description of the methods is presented here: we refer to the original publications for more details.[19,37,38] The following

synthetic description is based, with minimal adaption, on the one provided in reference.[38] Networks were built using only the quantified metabolites.

Method based on correlations (CORR)

The association between any pair of metabolites is measured through the absolute value of Pearson's correlation.

The CLR algorithm

The CLR (Context Likelihood of Relatedness) algorithm[39] uses mutual information as a measure of the similarity between the profiles of any two metabolites in the data which is then compared against the local context for each possible interaction: in this way possible spurious (indirect) associations are removed. These results in a weighted adjacency matrix that can be transformed into a binary adjacency matrix imposing a threshold of 0 (default value) on its entries.

The ARACNE algorithm

As CLR, ARACNE (Algorithm for the Reconstruction of Accurate Cellular Networks)[40] uses mutual information (MI) as a measure of the similarity between the profiles of any two metabolites. The properties of MI are used to prune the network of spurious interactions. Specifically, edges for which mutual independence cannot be ruled out at a given level are removed from the network. The default 0 threshold has been used.

The PCLRC algorithm

The PCLRC[19] (Probabilistic Context Likelihood of Relatedness) algorithm is a modification of the CLR algorithm (using correlation instead of MI to measure similarity between metabolite

profiles) and on iteratively sampling the dataset resulting in a weighted adjacency matrix containing an estimate of the association likeliness between any two metabolites. The default values of 25-75% data split and 90% confidence level have been used. An R implementation of this algorithm is available at semantics.systemsbiology.nl.

Construction of plasma metabolite networks

The plasma metabolite-metabolite association networker taking a wisdom of crowd approach as detailed in reference.[13] For each set of samples (Males, Females, Old M, Young M, Old F, Young F) adjacency matrices $(a_{ij})_m$ (with m=1,2 ,... 4) were obtained using the above described methods. The entries of such matrices matrix are real numbers in the range [-1, 1] for correlation matrices, in the [0, +∞] range for mutual information matrices, or [0, 1] for probabilistic networks, indicating the strength or the likelihood of the metabolite-metabolite associations. These matrices are binarized to 0 and 1 imposing a threshold on the values[9–11,13]:

$$\{a_{ij}\} \rightarrow \begin{cases} 1 & \text{if } a_{ij} > \tau_m \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The values of τ_m tm depend on the method considered: 0 for ARACNE and CLR, 0.95 for PCLRC, and 0.6 for the correlation map, as previously detailed.[37] The four networks were then superimposed

$$\{a_{ij}\} \rightarrow \sum_{m=1}^{4} \{a_{ij}\}_m \quad (2)$$

The final adjacency matrix, representing the metabolite network was defined by retaining only those links inferred by three or more methods, as suggested in reference [37]

$$\{a_{ij}\} \rightarrow \begin{cases} 1 & \text{if } a_{ij} \geq 3 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

In total, six networks were defined: Males, Females, Old M, Young M, Old F, and Young F.

Network vertexes are colored according to their modularity calculated using the R package "igraph".[41] The modularity of each graph with respect to vertex type measures how separated are the different vertex types from each other. It is defined as

$$Q = \frac{1}{2m}\sum_{i,j}^{m}\left(A_{ij} - \frac{k_i \cdot k_j}{2m}\right)\cdot \delta(c_i c_j) \quad (4)$$

where m is the number of edges, A_{ij} is the element of the adjacency matrix **A** in row i and column j, k_i is the degree of i, k_j is the degree of j, c_i is the type (or component) of i, c_j that of j, The sum goes over all i and j pairs of vertices, and $\delta(x, y)$ is 1 if x=y and 0 otherwise.

Ethical Issues

The original data were collected in accordance with the 1964 Helsinki declaration and its later amendments.

RESULTS AND DISCUSSION

Sex related effects on blood metabolite profiles

Discrimination analysis among the NMR profiles of males and females was performed using PLS-DA (Figure 2A). The data set was unbalanced, with far more males (n = 661) than females (n = 163), reflecting the sex bias observed among blood donor volunteers in Italy.[42] To avoid possible bias resulting from different sample size of the two groups, the analysis was performed by resampling n = 150 subjects from both the male and female groups and taking the average over 10^2 repetitions.

We obtained excellent discrimination between the plasma profiles (binned NMR spectra) of males (n =150) and females (n =150) (see Table 2), a result in line with previous observations.[22] Discrimination accuracy was also high when only a reduced set of 22 quantified metabolites was used, indicating that sex-specific biological information is thoroughly

represented by a limited number of metabolites, consistently to what observed in the case of metabolite and metabolic profiling in urines.[13]

Univariate analysis was applied only on the quantified metabolites and several were found to be statistically different between males (using the full cohort n = 661) and female (n = 163) (*P*-value<0.05 after Bonferroni correction): creatine showed higher levels in females; instead, phenylalanine, glutamine, proline, histidine, glutamate, tyrosine, valine, propylene glycol, leucine, isoleucine, creatinine, and acetone were higher in males (see Supplementary Table S2 for a summary). Higher levels of creatinine in males has been reported since long time and has been found to relate to muscle mass. Overall we observed systematic higher plasma concentrations of amino acids, especially BCAA, in males which may be linked to differences in muscle mass metabolism,[43] to the larger muscle mass of men[44] or to the higher protein intake of men with respect to women.[45]

Our results are in line with those of Krumsiek and coworkers[22] who addressed sex-specific differences in the metabolism of healthy subjects, although two different analytical approaches were applied (Mass Spectrometry *vs* NMR spectroscopy) and the study size was different. Interestingly, we found creatinine and tyrosine trends to be consistent with their findings but in contrast with what reported by Dunn *et al*. in their UK population study[46] where lower levels of creatinine and tyrosine in males were observed. Moreover, we did not observe sex-related differences in glucose and lactate plasma levels.

Sex related effects on plasma metabolite association networks

Plasma metabolite-metabolite association networks were estimated taking a so-called "wisdom of crowd" approach, *i.e.* combing the results of four different methods to avoid bias induced by the choice of a particular method, following an established practice.[13,37,47] Sex specific networks are shown in Figure 3A (males, constructed using n = 661 samples) and 3B

(females, constructed using n = 163s). Networks were arranged and colored according to metabolite modularity (see Equation 4). In both sex-specific networks two aminoacidic clusters are visible; moreover, in males also metabolites related to glucose/energetic metabolism formed a cluster. These evidences are in line with those observed by Krumsiek and coworkers.[22]

We observed that, in general, the female-specific network is less densely connected that the males-specific network (see Figure 3 for network representation and Figure 4 for an overview of metabolite connectivity) and this different topology is likely to reflect underlying metabolic differences, but it could be also affected by the differences in the sample size. Highly connected metabolites, the so called "hubs" play a special role in biological network and network analysis since in many cases, for example in gene co-expression and regulatory networks,[48,49] metabolic networks, protein−protein interaction networks,[50−52] and cell−cell interaction networks,[53] there is evidence of few highly connected nodes that are considered to play crucial biological roles; for instance it has been shown that, in yeast, proteins that are highly connected are essential for survival.[48,50]

Consistently with previous approaches,[54] at a first stage, we considered as hubs those metabolites with degree larger than 5 and clustering coefficient <0.03. In male-specific networks only leucine, glucose, lactate, valine, acetoacetate, and creatinine satisfied these selection criteria, while in the female-specific network only valine was found to be a hub according to this classification. Interestingly, valine is a hub metabolite in both networks, yet it resulted more connected in males than in females (9 *vs* 5 connections); 4 connections are in common (isoleucine, leucine, phenylalanine, and tyrosine) while citrate is a connection found only in females, and acetoacetate, creatinine, creatine, acetone, and histidine are connected with valine only in males. It is worth of noting that citrate excretion levels differ significantly in males and females, and in particular the sex differences has been found to increase with age.[55]

We used high connectivity and low clustering coefficient to exclude metabolites that could be highly interconnected because participating to same molecular machine, such as amino acids deriving from protein metabolism and catabolism, while focussing on metabolites that could be pivotal in discriminating between males and females network topology. Indeed, according to this criterion such as isoleucine and leucine, whose levels are also dependent of protein intake, are excluded.

The hub metabolites in the male-specific network can be related to the propionate metabolism and the valine, leucine and isoleucine degradation. Branched amino acids cannot be synthesized de novo, so their homeostasis is maintained by degradation and dietary intake only,[56] which we speculate may be differentially regulated in males and females given the different connectivity observed in the sex-specific networks. Whether protein metabolism and catabolism are different in males and females is an open question, and there are conflicting results given also the different methodologies used in such studies. However, the present results seem to suggest the existence of sex-specific differences in protein metabolism: whether these differences are due to different body mass composition or different protein intake or being attributable to sex hormones as suggested[57,58] has to be ascertained.

Propionate and BCAA metabolism overlap at the gene level while propanoate is involved in BCAA metabolism but is also involved in the short-chain fatty acid metabolism, and this result can be reconciled with the hypothesis of sexual dimorphism in human lipid metabolism[59] for which evidence has been provided also in metabolomics studies.[23,60]

Similarly to previous studies[54] we also analyzed the networks using a hub definition of connectivity greater than 5 without considering the clustering coefficient, and we focused solely on those metabolites for which the difference Δ between a metabolite connectivity in the male and female specific network was larger than 2. In addition to the valine, leucine and isoleucine catabolism and propanoate metabolism we found three other pathways possibly

associated with sex-specific metabolite connectivity differences, namely synthesis and degradation of ketone bodies, and pyruvate metabolism. That ketone bodies catabolism emerges as a discriminant between males and females is interesting; blood samples have been collected after overnight fasting, and it is known that during short term fasting the decrease of plasma glucose and the increase of ketone body levels are greater in females than in males[61–65] and evidences have been given for the existence of sexual dimorphism for what concerns lipid metabolism in response to short-term fasting.[61] However, as for the case of protein metabolism, the causes are not known and a role of sexual hormones has been proposed.[61] Pyruvate is produced during glycolysis and it has been suggested that females have a significantly lower overall capacity for aerobic oxidation and for anaerobic glycolysis than males.[66,67]

Age effects on blood metabolite networks

While sexual dimorphism provides a clear discriminant for a comparative analysis, dividing the population under investigation in age groups is less straightforward. According to the World Health Organization,[68] the chronological age of 65 years is used for the definition of 'elderly' or older person. However, in this study the cohort is made of blood donor volunteers, and in Italy the maximum age for blood donation is set to 65. Here, the average age is 41±11 years for males and 42±12 years for females; we are thus in the presence of a relatively young, healthy, and homogenous population. To set boundaries for discriminant analysis we take a pragmatic approach, by taking the lower (*L*) and upper (*U*) tertiles of the age distribution and labelling as Old those individual with age > U and as Young those with age < L. For males we had $L = 35$ and $U = 45$ years, and for females $L = 37$ and $U = 48$ years (see Table 1 for the size of the resulting sub-groups).

Discriminant analysis was performed using PLS-DA on these age-defined groups (Figure 2B-C: Young males ($n = 234$) *vs* old males ($n = 213$); and Young females ($n = 63$) *vs* Old females

(n = 60)). As shown in Table 2 we found good discrimination accuracy between the two age classes as previously reported,[7,69,70] with discrimination higher in males than in females, when analyzing both binned spectra and the array of 22 metabolites. This can be due to the smaller sample size in the case of females (447 males vs 123 females) which can affect the discrimination power of the statistical model, or reflecting the fact that males age faster than females[71] resulting in larger age related differences in males in respect with females. Furthermore, in women oral contraceptive use and menopausal state have been found to altered significantly the metabolome, proving an additional source of variability and possible confounding factors.[72,73] However, given the relatively young age of the study volunteers the hypothesis of menopausal status can be probably ruled out.

Univariate analysis of the metabolite concentrations was also performed for age-comparisons in both sexes: acetate, and histidine exhibited statistically higher concentrations in Young males; conversely, alanine and creatine were elevated in Old males (Supplementary Table S3). For the female cohort, glucose, glutamine, glycine, tyrosine, and creatine presented statistically higher concentrations in Old females (Supplementary Table S4). Interestingly, only in women we can observe an increasing of glucose with the age, this evidence could imply an impairment of the glucose metabolism and therefore, a latent risk of developing Type II diabetes.

The analysis of the metabolite connectivity (Δ degree) was also performed for the comparison of Young males (234 subjects) *vs* Old males (213 subjects), and of Young females (63 subjects) *vs* Old females (60 subjects). For males 3-hydroxybutyrate, proline, citrate, and creatinine showed increased connectivity in the young cohort ($\Delta \geq 2$); instead, acetone, leucine, and propylene glycol decreased their connectivity with age (Figure 4B). Isoleucine, phenylalanine, glycine, glutamine, glutamate, formate, creatine, and acetoacetate did not display any connectivity variations in the male cohort ($\Delta = 0$).

For females, isoleucine, alanine, tyrosine, 3-hydroxybutyrate, proline, leucine and glucose were found to be more connected in the young cohort ($\Delta \geq 2$); instead, glutamine and creatine did not exhibit any changes in connectivity (Figure 4C).

As a general remark, we can observe a shrinkage of the number of connections in the older individuals, both for males (Figure 3C and E) and females (Figure 3D and F). Older males showed two hubs (valine and acetone) that are not present in younger ones; conversely, no hubs were found in females independent of age. Interestingly, Soltow *et al*.[74] in their study on metabolism and ageing in common marmosets (*Callithrix jacchus*, a premiere primate model for studies of aging) reported that metabolite connectivity decreases with age and these results is in line with what we are reporting in this paper. These evidences could indicate that further efforts in this direction and using these types of analyses (global metabolic profiling combined with network analysis) in well-designed models may reveal biomarkers associated with age-related phenotype and disease.

Valine was found to be a hub in older males but not in younger ones; although valine is not a hub in the networks of both young and old women a decrease in connectivity is also present (see Figure 3 and 4) . Studies in yeast[75] have revealed that addition of serine, threonine, and valine in growth media promotes cellular sensitization and aging by activating different interconnected pathways; furthermore, the restriction of the intake of proteins or certain amino acids, including valine, has been associated with extended longevity and reduced incidence and/or progression of multiple age-related diseases.[76] Consistently, our data could imply a pivotal role of valine in the molecular mechanism of ageing, with an increasing of its "metabolic importance" with the age.

CONCLUSIONS

We have presented network reconstruction and analysis of experimentally identified relationships between metabolites, and applied a differential network approach to analyze sex and age differences in a cohort of healthy subjects. Our results show significant differences between males and females for what concerns both metabolite concentrations and connections implying variations in the regulation of metabolic activity involving branches amino acids, lipids and ketone bodies. As in the case of previous studies,[6,7,14–17] providing a rationale for the explanation of the functional role of the metabolites for which concentrations changes due to age or sexual dimorphism is complicate but we can hypothesize that hormonal and body composition differences between men and women are the underlying causes for many of our observations as already observed for the proteome.[3]

The network approach seems to provide more insights than the standard approach showing a decrease of the connectivity with age in both sex groups, demonstrating that this evidence is peculiar of ageing even if the biological meaning of our result needs to be further investigated.

Although the age range is limited, and sex is unequally biased towards males, this study provides important information on how common variables influence expression of the metabolic phenotype. Both age and sex are recognized confounders, understanding these differences will be a critical component for development of metabolomics as a population screening and precision medicine platform.

SUPPORTING INFORMATION

The following files are available free of charge at ACS website http://pubs.acs.org:

- Supplementary Table S1. List of metabolites identified on 1D ^{1}H NMR spectra of AVIS healthy blood donors.
- Supplementary Table S2. Pairwise comparison of plasma metabolite concentration levels in males and females.

- Supplementary Table S3. Pairwise comparison of plasma metabolite concentration in Young and Old male subjects.
- Supplementary Table S4. Pairwise comparison of plasma metabolite concentration in Young and Old female subjects'.

Acknowledgements

This work was partly supported by the FP7 European Commission-funded INFECT project (Contract No. 305340), the H2020 EC-funded project PhenoMeNal (Grant 654241) and the H2020 EC-funded project PROPAG-AGEING (Grant 634821). Jacques Vervoort is acknowledged for fruitful discussion of the manuscript. We acknowledge AVIS Toscana (in the people of the Luciano Franchi and Donata Marangio), AVIS Pistoia (in the person of Alessandro Pratesi), and the technical staff of Transfusion Service of the Pistoia Hospital for volunteer recruitment and sample collection.

Author Contributions

Claudio Luchinat, Edoardo Saccenti, Leonardo Tenori designed the study. Edoardo Saccenti, and Alessia Vignoli performed statistical data analyses. Claudio Luchinat, Edoardo Saccenti, Leonardo Tenori, and Alessia Vignoli interpreted the data and results, prepared the manuscript, and were responsible for its final content. All authors read and approved the final version of the manuscript.

Conflicts of Interest

The authors declare no competing financial, and non-financial interests.

REFERENCES

(1) Regitz-Zagrosek, V. Sex and gender differences in health. *EMBO Rep.* **2012**, *13* (7), 596–603.

(2) Gandhi, M.; Aweeka, F.; Greenblatt, R. M.; Blaschke, T. F. Sex differences in pharmacokinetics and pharmacodynamics. *Annu. Rev. Pharmacol. Toxicol.* **2004**, *44*, 499–523.

(3) Curran, A. M.; Fogarty Draper, C.; Scott-Boyer, M.-P.; Valsesia, A.; Roche, H. M.; Ryan, M. F.; Gibney, M. J.; Kutmon, M.; Evelo, C. T.; Coort, S. L.; et al. Sexual dimorphism, age and fat mass are key phenotypic drivers of proteomic signatures. *J. Proteome Res.* **2017**.

(4) Kochhar, S.; Jacobs, D. M.; Ramadan, Z.; Berruex, F.; Fuerholz, A.; Fay, L. B. Probing gender-specific metabolism differences in humans by nuclear magnetic resonance-based metabonomics. *Anal. Biochem.* **2006**, *352* (2), 274–281.

(5) Bertram, H. C.; Duus, J. Ø.; Petersen, B. O.; Hoppe, C.; Larnkjaer, A.; Schack-Nielsen, L.; Mølgaard, C.; Michaelsen, K. F. Nuclear magnetic resonance-based metabonomics reveals strong sex effect on plasma metabolism in 17-year-old Scandinavians and correlation to retrospective infant plasma parameters. *Metabolism.* **2009**, *58* (7), 1039–1045.

(6) Mittelstrass, K.; Ried, J. S.; Yu, Z.; Krumsiek, J.; Gieger, C.; Prehn, C.; Roemisch-Margl, W.; Polonikov, A.; Peters, A.; Theis, F. J.; et al. Discovery of Sexual Dimorphisms in Metabolic and Genetic Biomarkers. *PLOS Genet.* **2011**, *7* (8), e1002215.

(7) Rist, M. J.; Roth, A.; Frommherz, L.; Weinert, C. H.; Krüger, R.; Merz, B.; Bunzel, D.; Mack, C.; Egert, B.; Bub, A.; et al. Metabolite patterns predicting sex and age in participants of the Karlsruhe Metabolomics and Nutrition (KarMeN) study. *PloS One* **2017**, *12* (8), e0183228.

(8) Repetto, L.; Venturino, A.; Fratino, L.; Serraino, D.; Troisi, G.; Gianni, W.; Pietropaolo, M. Geriatric oncology: a clinical approach to the older patient with cancer. *Eur. J. Cancer Oxf. Engl. 1990* **2003**, *39* (7), 870–880.

(9) Assfalg, M.; Bertini, I.; Colangiuli, D.; Luchinat, C.; Schafer, H.; Schutz, B.; Spraul, M. Evidence of different metabolic phenotypes in humans. *Proc.Natl.Acad.Sci.U.S.A* **2008**, *105* (1091–6490 (Electronic)), 1420–1424.

(10) Bernini, P.; Bertini, I.; Luchinat, C.; Nepi, S.; Saccenti, E.; Schafer, H.; Schutz, B.; Spraul, M.; Tenori, L. Individual human phenotypes in metabolic space and time. *J ProteomeRes* **2009**, *8* (1535–3893 (Print)), 4264–4271.

(11) Ghini, V.; Saccenti, E.; Tenori, L.; Assfalg, M.; Luchinat, C. Allostasis and Resilience of the Human Individual Metabolic Phenotype. *J. Proteome Res.* **2015**, *14* (7), 2951–2962.

(12) Mäkinen, V.-P.; Ala-Korpela, M. Metabolomics of aging requires large-scale longitudinal studies with replication. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (25), E3470.

(13) Saccenti, E.; Menichetti, G.; Ghini, V.; Remondini, D.; Tenori, L.; Luchinat, C. Entropy-Based Network Representation of the Individual Metabolic Phenotype. *J. Proteome Res.* **2016**, *15* (9), 3298–3307.

(14) Menni, C.; Kastenmüller, G.; Petersen, A. K.; Bell, J. T.; Psatha, M.; Tsai, P.-C.; Gieger, C.; Schulz, H.; Erte, I.; John, S.; et al. Metabolomic markers reveal novel pathways of ageing and early development in human populations. *Int. J. Epidemiol.* **2013**, *42* (4), 1111–1119.

(15) Jové, M.; Maté, I.; Naudí, A.; Mota-Martorell, N.; Portero-Otín, M.; De la Fuente, M.; Pamplona, R. Human Aging Is a Metabolome-related Matter of Gender. *J. Gerontol. A. Biol. Sci. Med. Sci.* **2016**, *71* (5), 578–585.

(16) Zhao, J.; Zhu, Y.; Uppal, K.; Tran, V. T.; Yu, T.; Lin, J.; Matsuguchi, T.; Blackburn, E.; Jones, D.; Lee, E. T.; et al. Metabolic profiles of biological aging in American Indians: the Strong Heart Family Study. *Aging* **2014**, *6* (3), 176–186.

(17) Hertel, J.; Friedrich, N.; Wittfeld, K.; Pietzner, M.; Budde, K.; Van der Auwera, S.; Lohmann, T.; Teumer, A.; Völzke, H.; Nauck, M.; et al. Measuring Biological Age via Metabonomics: The Metabolic Age Score. *J. Proteome Res.* **2016**, *15* (2), 400–410.
(18) Saccenti, E.; Tenori, L.; Verbruggen, P.; Timmerman, M. E.; Bouwman, J.; Greef, J. van der; Luchinat, C.; Smilde, A. K. Of Monkeys and Men: A Metabolomic Analysis of Static and Dynamic Urinary Metabolic Phenotypes in Two Species. *PLOS ONE* **2014**, *9* (9), e106077.
(19) Saccenti, E.; Suarez-Diez, M.; Luchinat, C.; Santucci, C.; Tenori, L. Probabilistic networks of blood metabolites in healthy subjects as indicators of latent cardiovascular risk. *J. Proteome Res.* **2015**, *14* (2), 1101–1111.
(20) Saccenti, E.; van Duynhoven, J.; Jacobs, D. M.; Smilde, A. K.; Hoefsloot, H. C. J. Strategies for individual phenotyping of linoleic and arachidonic Acid metabolism using an oral glucose tolerance test. *PloS One* **2015**, *10* (3), e0119856.
(21) Krumsiek, J.; Suhre, K.; Illig, T.; Adamski, J.; Theis, F. J. Gaussian graphical modeling reconstructs pathway reactions from high-throughput metabolomics data. *BMC Syst. Biol.* **2011**, *5*, 21.
(22) Krumsiek, J.; Mittelstrass, K.; Do, K. T.; Stückler, F.; Ried, J.; Adamski, J.; Peters, A.; Illig, T.; Kronenberg, F.; Friedrich, N.; et al. Gender-specific pathway differences in the human serum metabolome. *Metabolomics Off. J. Metabolomic Soc.* **2015**, *11* (6), 1815–1833.
(23) Szymańska, E.; Bouwman, J.; Strassburg, K.; Vervoort, J.; Kangas, A. J.; Soininen, P.; Ala-Korpela, M.; Westerhuis, J.; van Duynhoven, J. P. M.; Mela, D. J.; et al. Gender-dependent associations of metabolite profiles and body fat distribution in a healthy population with central obesity: towards metabolomics diagnostics. *Omics J. Integr. Biol.* **2012**, *16* (12), 652–667.
(24) Jansen, J. J.; Szymańska, E.; Hoefsloot, H. C. J.; Jacobs, D. M.; Strassburg, K.; Smilde, A. K. Between Metabolite Relationships: an essential aspect of metabolic change. *Metabolomics* **2012**, *8* (3), 422–432.
(25) Jansen, J. J.; Szymańska, E.; Hoefsloot, H. C. J.; Smilde, A. K. Individual differences in metabolomics: individualised responses and between-metabolite relationships. *Metabolomics* **2012**, *8* (1), 94–104.
(26) Bernini, P.; Bertini, I.; Luchinat, C.; Tenori, L.; Tognaccini, A. The cardiovascular risk of healthy individuals studied by NMR metabonomics of plasma samples. *J. Proteome Res.* **2011**, *10* (11), 4983–4992.
(27) Bernini, P.; Bertini, I.; Luchinat, C.; Nincheri, P.; Staderini, S.; Turano, P. Standard operating procedures for pre-analytical handling of blood and urine for metabolomic studies and biobanks. *J. Biomol. NMR* **2011**, *49* (3–4), 231–243.
(28) Kale, N. S.; Haug, K.; Conesa, P.; Jayseelan, K.; Moreno, P.; Rocca-Serra, P.; Nainala, V. C.; Spicer, R. A.; Williams, M.; Li, X.; et al. MetaboLights: An Open-Access Database Repository for Metabolomics Data. *Curr. Protoc. Bioinforma.* **2016**, *53*, 14.13.1-18.
(29) Meiboom, S.; Gill, D. Modified Spin-Echo Method for Measuring Nuclear Relaxation Times. *Rev. Sci. Instrum.* **1958**, *29* (8), 688–691.
(30) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; et al. HMDB: the Human Metabolome Database. *Nucleic Acids Res.* **2007**, *35* (Database issue), D521-526.
(31) Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E.; et al. HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic Acids Res.* **2013**, *41* (D1), D801–D807.
(32) Wilcoxon, F. Individual Comparisons by Ranking Methods. *Biom. Bull.* **1945**, *1* (6), 80.
(33) Bonferroni, C. E. Il calcolo delle assicurazioni su gruppi di teste. In *Studi in Onore del Professore Salvatore Ortu Carboni*; Rome, 1935; pp 13–60.
(34) Cliff, N. *Ordinal Methods for Behavioral Data Analysis*; Psychology Press: Mahwah, N.J, 1996.

(35) Barker, M.; Rayens, W. Partial least squares for discrimination. *J. Chemom.* **2003**, *17* (3), 166–173.
(36) Szymańska, E.; Saccenti, E.; Smilde, A. K.; Westerhuis, J. A. Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics* **2012**, *8* (1), 3–16.
(37) Suarez-Diez, M.; Saccenti, E. Effects of Sample Size and Dimensionality on the Performance of Four Algorithms for Inference of Association Networks in Metabonomics. *J. Proteome Res.* **2015**, *14* (12), 5119–5130.
(38) Suarez-Diez, M.; Adam, J.; Adamski, J.; Chasapi, S. A.; Luchinat, C.; Peters, A.; Prehn, C.; Santucci, C.; Spyridonidis, A.; Spyroulias, G. A.; et al. Plasma and Serum Metabolite Association Networks: Comparability within and between Studies Using NMR and MS Profiling. *J. Proteome Res.* **2017**, *16* (7), 2547–2559.
(39) Faith, J. J.; Hayete, B.; Thaden, J. T.; Mogno, I.; Wierzbowski, J.; Cottarel, G.; Kasif, S.; Collins, J. J.; Gardner, T. S. Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles. *PLoS Biol.* **2007**, *5* (1), e8.
(40) Margolin, A. A.; Nemenman, I.; Basso, K.; Wiggins, C.; Stolovitzky, G.; Dalla Favera, R.; Califano, A. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* **2006**, *7 Suppl 1*, S7.
(41) Csardi, G.; Nepusz, T. The igraph software package for complex network research. *InterJournal* **2006**, *Complex Systems*, 1695.
(42) Bani, M.; Giussani, B. Gender differences in giving blood: a review of the literature. *Blood Transfus.* **2010**, *8* (4), 278–287.
(43) Honda, T.; Kobayashi, Y.; Togashi, K.; Hasegawa, H.; Iwasa, M.; Taguchi, O.; Takei, Y.; Sumida, Y. Associations among circulating branched-chain amino acids and tyrosine with muscle volume and glucose metabolism in individuals without diabetes. *Nutr. Burbank Los Angel. Cty. Calif* **2016**, *32* (5), 531–538.
(44) Margolis, L. M.; Pasiakos, S. M.; Karl, J. P.; Rood, J. C.; Cable, S. J.; Williams, K. W.; Young, A. J.; McClung, J. P. Differential effects of military training on fat-free mass and plasma amino acid adaptations in men and women. *Nutrients* **2012**, *4* (12), 2035–2046.
(45) Schmidt, J. A.; Rinaldi, S.; Scalbert, A.; Ferrari, P.; Achaintre, D.; Gunter, M. J.; Appleby, P. N.; Key, T. J.; Travis, R. C. Plasma concentrations and intakes of amino acids in male meat-eaters, fish-eaters, vegetarians and vegans: a cross-sectional analysis in the EPIC-Oxford cohort. *Eur. J. Clin. Nutr.* **2016**, *70* (3), 306–312.
(46) Dunn, W. B.; Lin, W.; Broadhurst, D.; Begley, P.; Brown, M.; Zelena, E.; Vaughan, A. A.; Halsall, A.; Harding, N.; Knowles, J. D.; et al. Molecular phenotyping of a UK population: defining the human serum metabolome. *Metabolomics Off. J. Metabolomic Soc.* **2015**, *11*, 9–26.
(47) Marbach, D.; Costello, J. C.; Küffner, R.; Vega, N. M.; Prill, R. J.; Camacho, D. M.; Allison, K. R.; DREAM5 Consortium; Kellis, M.; Collins, J. J.; et al. Wisdom of crowds for robust gene network inference. *Nat. Methods* **2012**, *9* (8), 796–804.
(48) Carter, S. L.; Brechbühler, C. M.; Griffin, M.; Bond, A. T. Gene co-expression network topology provides a framework for molecular characterization of cellular state. *Bioinforma. Oxf. Engl.* **2004**, *20* (14), 2242–2250.
(49) Stuart, J. M.; Segal, E.; Koller, D.; Kim, S. K. A gene-coexpression network for global discovery of conserved genetic modules. *Science* **2003**, *302* (5643), 249–255.
(50) Jeong, H.; Mason, S. P.; Barabási, A. L.; Oltvai, Z. N. Lethality and centrality in protein networks. *Nature* **2001**, *411* (6833), 41–42.
(51) Rzhetsky, A.; Gomez, S. M. Birth of scale-free molecular networks and the number of distinct DNA and protein domains per genome. *Bioinforma. Oxf. Engl.* **2001**, *17* (10), 988–996.

(52) Yook, S.-H.; Oltvai, Z. N.; Barabási, A.-L. Functional and topological characterization of protein interaction networks. *Proteomics* **2004**, *4* (4), 928–942.
(53) Hartwell, L. H.; Hopfield, J. J.; Leibler, S.; Murray, A. W. From molecular to modular cell biology. *Nature* **1999**, *402* (6761 Suppl), C47-52.
(54) Lu, X.; Jain, V. V.; Finn, P. W.; Perkins, D. L. Hubs in biological interaction networks exhibit low changes in expression in experimental asthma. *Mol. Syst. Biol.* **2007**, *3* (1), n/a-n/a.
(55) Caudarella, R.; Vescini, F.; Buffa, A.; Stefoni, S. Citrate and mineral metabolism: kidney stones and bone disease. *Front. Biosci. J. Virtual Libr.* **2003**, *8*, s1084-1106.
(56) Wanders, R. J. A.; Duran, M.; Loupatty, F. J. Enzymology of the branched-chain amino acid oxidation disorders: the valine pathway. *J. Inherit. Metab. Dis.* **2012**, *35* (1), 5–12.
(57) Phillips, S. M.; Atkinson, S. A.; Tarnopolsky, M. A.; MacDougall, J. D. Gender differences in leucine kinetics and nitrogen balance in endurance athletes. *J. Appl. Physiol. Bethesda Md 1985* **1993**, *75* (5), 2134–2141.
(58) McKenzie, S.; Phillips, S. M.; Carter, S. L.; Lowther, S.; Gibala, M. J.; Tarnopolsky, M. A. Endurance exercise training attenuates leucine oxidation and BCOAD activation during exercise in humans. *Am. J. Physiol. Endocrinol. Metab.* **2000**, *278* (4), E580-587.
(59) Mittendorfer, B. Sexual dimorphism in human lipid metabolism. *J. Nutr.* **2005**, *135* (4), 681–686.
(60) Stroeve, J. H. M.; Saccenti, E.; Bouwman, J.; Dane, A.; Strassburg, K.; Vervoort, J.; Hankemeier, T.; Astrup, A.; Smilde, A. K.; van Ommen, B.; et al. Weight loss predictability by plasma metabolic signatures in adults with obesity and morbid obesity of the DiOGenes study. *Obes. Silver Spring Md* **2016**, *24* (2), 379–388.
(61) Mittendorfer, B.; Horowitz, J. F.; Klein, S. Gender differences in lipid and glucose kinetics during short-term fasting. *Am. J. Physiol. Endocrinol. Metab.* **2001**, *281* (6), E1333-1339.
(62) Paula, F. J.; Pimenta, W. P.; Saad, M. J.; Paccola, G. M.; Piccinato, C. E.; Foss, M. C. Sex-related differences in peripheral glucose metabolism in normal subjects. *Diabete Metab.* **1990**, *16* (3), 234–239.
(63) Yki-Järvinen, H. Sex and insulin sensitivity. *Metabolism.* **1984**, *33* (11), 1011–1015.
(64) Rizza, R. A.; Mandarino, L. J.; Gerich, J. E. Dose-response characteristics for effects of insulin on production and utilization of glucose in man. *Am. J. Physiol. - Endocrinol. Metab.* **1981**, *240* (6), E630–E639.
(65) Zderic, T. W.; Coggan, A. R.; Ruby, B. C. Glucose kinetics and substrate oxidation during exercise in the follicular and luteal phases. *J. Appl. Physiol.* **2001**, *90* (2), 447–453.
(66) Green, H. J.; Fraser, I. G.; Ranney, D. A. Male and female differences in enzyme activities of energy metabolism in vastus lateralis muscle. *J. Neurol. Sci.* **1984**, *65* (3), 323–331.
(67) Russ, D. W.; Lanza, I. R.; Rothman, D.; Kent-Braun, J. A. Sex differences in glycolysis during brief, intense isometric contractions. *Muscle Nerve* **2005**, *32* (5), 647–655.
(68) WHO | Proposed working definition of an older person in Africa for the MDS Project http://www.who.int/healthinfo/survey/ageingdefnolder/en/ (accessed Feb 8, 2017).
(69) Chaleckis, R.; Murakami, I.; Takada, J.; Kondoh, H.; Yanagida, M. Individual variability in human blood metabolites identifies age-related differences. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (16), 4252–4259.
(70) Saito, K.; Maekawa, K.; Pappan, K. L.; Urata, M.; Ishikawa, M.; Kumagai, Y.; Saito, Y. Differences in metabolite profiles between blood matrices, ages, and sexes among Caucasian individuals and their inter-individual variations. *Metabolomics* **2014**, *10* (3), 402–413.
(71) Blagosklonny, M. V. Why men age faster but reproduce longer than women: mTOR and evolutionary perspectives. *Aging* **2010**, *2* (5), 265–273.
(72) Ruoppolo, M.; Campesi, I.; Scolamiero, E.; Pecce, R.; Caterino, M.; Cherchi, S.; Mercuro, G.; Tonolo, G.; Franconi, F. Serum metabolomic profiles suggest influence of sex and oral contraceptive use. *Am. J. Transl. Res.* **2014**, *6* (5), 614–624.

(73) Auro, K.; Joensuu, A.; Fischer, K.; Kettunen, J.; Salo, P.; Mattsson, H.; Niironen, M.; Kaprio, J.; Eriksson, J. G.; Lehtimäki, T.; et al. A metabolic view on menopause and ageing. *Nat. Commun.* **2014**, *5*, ncomms5708.
(74) Soltow, Q. A.; Jones, D. P.; Promislow, D. E. L. A Network Perspective on Metabolism and Aging. *Integr. Comp. Biol.* **2010**, *50* (5), 844–854.
(75) Longo, V. D.; Shadel, G. S.; Kaeberlein, M.; Kennedy, B. Replicative and chronological aging in Saccharomyces cerevisiae. *Cell Metab.* **2012**, *16* (1), 18–31.
(76) Mirzaei, H.; Suarez, J. A.; Longo, V. D. Protein and Amino Acid Restriction, Aging and Disease: from yeast to humans. *Trends Endocrinol. Metab. TEM* **2014**, *25* (11), 558–566.

Figures Captions

Figure 1. Age distribution for: A) Female; B) Male. Median age, mean age and standard deviation are also reported.

Figure 2. Scatter plot of the first two latent variables (LV) of the PLS-DA models on NMR plasma metabolite concentration profiles of: A) Male (blue dots, n = 661) and Female (pink dots, n = 183); B) Young (light blue, n = 234) and Old (dark blue, n = 213) male subjects; C) Young (light pink, n = 63) and Old (dark pink, n = 60) female subjects. The corresponding discrimination accuracy obtained using *k*NN for each different model are given in Table 2.

Figure 3. Metabolite-metabolite association networks reconstructed from the plasma metabolite concentration profiles of: A) males; B) females; C) Young males; D) Young females; E) Old males and F) Old females. Vertexes are colored according to metabolite modularity (see Equation (4)). Abbreviations: AAC (Acetoacetate), ACE (Acetate), ACN (Acetone), ALA (Alanine), BHB (3-hydroxybutyrate), CIT (Citrate), CR (creatine), CRN (creatinine), FOR (formate), GLN (glutamine), GLS (glucose), GLU (glutamate), GLY (glycine), HIS (Histidine), ILE (Isoleucine), LAC (Lactate), LEU (Leucine), PGL (Propylene glycol), PHE (Phenylalanine), PRO (Proline), TYR (Tyrosine), VAL (Valine).

Figure 4. Differences in metabolite connectivity (node degree) between A) Males and Females; B) Young and old males and C) Young and old females

Tables

Table 1. Demographic characteristics of the healthy blood donor cohort examined in this study.

	Subjects	**Median age (yrs)**
Female (total)	183	43
Young Female (yrs < 37)	63	28
Old Female (yrs > 48)	60	55
Male (total)	661	40
Young Male (yrs < 35)	234	30
Old Male (yrs > 45)	213	52

Table 2. The average AUROC (Area Under the Receiver Operative Curve) and Accuracy for the PLS-DA model; values are given together with the standard error calculated over 10^2 repetitions of the double cross validation procedure (2CV); in the case of the PLS-DA model of Males vs Females a 100x resampling procedure has been utilized to compensate for group unbalancedness (see Material and Methods for more details). The significance (*P*-value) of the discriminant model has been assessed by mean of permutation test (=10^3 permutations).

PLS-DA	**AUROC**	**Accuracy**	***P*-value**
Full NMR CPMG spectra			
Males *vs* Females	0.97±0.01	0.95±0.02	0.001
Young *vs* Old Males	0.97±0.01	0.92±0.02	0.001
Young *vs* Old Females	0.98±0.02	0.93±0.04	0.001
Quantified metabolites			
Males *vs* Females	0.95±0.01	0.85±0.03	0.001
Young *vs* Old Males	0.78±0.02	0.71±0.02	0.001
Young *vs* Old Females	0.85±0.05	0.76±0.05	0.001