Proteomic-based research strategy identified laminin subunit alpha 2 as a potential urinary specific biomarker for the medullary sponge kidney disease

Antonia Fabris[1,6], Maurizio Bruschi[2,6], Laura Santucci[2], Giovanni Candiano[2], Simona Granata[1], Alessandra Dalla Gassa[1], Nadia Antonucci[1], Andrea Petretto[3], Gian Marco Ghiggeri[4], Giovanni Gambaro[5], Antonio Lupo[1] and Gianluigi Zaza[1]

[1]Renal Unit, Department of Medicine, University Hospital of Verona, Verona, Italy; [2]Laboratory on Physiopathology of Uremia, Istituto Giannina Gaslini, Genoa, Italy; [3]Laboratory of Mass Spectrometry, Core Facility, Istituto Giannina Gaslini, Genova, Italy; [4]Division of Nephrology, Dialysis and Transplantation, Istituto Giannina Gaslini, Genoa, Italy; and [5]Division of Nephrology and Dialysis, Columbus-Gemelli University Hospital Catholic University, School of Medicine, Rome, Italy

Medullary sponge kidney (MSK) disease, a rare kidney malformation featuring recurrent renal stones and nephrocalcinosis, continues to be diagnosed using expensive and time-consuming clinical/instrumental tests (mainly urography). Currently, no molecular diagnostic biomarkers are available. To identify such we employed a proteomic-based research strategy utilizing urine from 22 patients with MSK and 22 patients affected by idiopathic calcium nephrolithiasis (ICN) as controls. Notably, two patients with ICN presented cysts. In the discovery phase, the urine of 11 MSK and 10 controls, were randomly selected, processed, and analyzed by mass spectrometry. Subsequently, several statistical algorithms were undertaken to select the most discriminative proteins between the two study groups. ELISA, performed on the entire patients' cohort, was used to validate the proteomic results. After an initial statistical analysis, 249 and 396 proteins were identified exclusive for ICN and MSK, respectively. A Volcano plot and ROC analysis, performed to restrict the number of MSK-associated proteins, indicated that 328 and 44 proteins, respectively, were specific for MSK. Interestingly, 119 proteins were found to differentiate patients with cysts (all patients with MSK and the two ICN with renal cysts) from ICN without cysts. Eventually, 16 proteins were found to be common to three statistical methods with laminin subunit alpha 2 (LAMA-2) reaching the higher rank by a Support Vector Machine, a binary classification/prediction scheme. ELISA for LAMA-2 validated proteomic results. Thus, using high-throughput technology, our study identified a candidate MSK biomarker possibly employable in future for the early diagnosis of this disease.

Correspondence: *Gianluigi Zaza, Renal Unit, Department of Medicine, University Hospital of Verona, Piazzale A. Stefani 1, 37126, Verona (VR), Italy. E-mail: gianluigi.zaza@univr.it*

[6]These authors contributed equally to the study.

Received 20 July 2016; revised 5 September 2016; accepted 15 September 2016

Kidney International (2016) ■, ■–■; http://dx.doi.org/10.1016/j.kint.2016.09.035

KEYWORDS: biomarker; laminin subunit alpha 2; MSK; proteomic

In the past 10 years, worldwide important research programs have been undertaken to better identify the pathologic network and the genetic bases involved in the onset and development of medullary sponge kidney (MSK) disease, a rare congenital condition (prevalence of ∼5 cases per 10,000–100,000 in the general population) typically associated with nephrocalcinosis and nephrolithiasis, urinary acidification and concentration defects, and cystic anomalies in the precalyceal ducts.[1]

MSK disease, although rare in the general population, is relatively frequent in renal stone formers, reaching a high degree of association in some published medical records (as high as 20%), and it has been also linked to developmental disorders (e.g., congenital hemihypertrophy and Beckwith-Wiedemann syndrome) and kidney developmental anomalies (e.g., horseshoe kidney, unilateral renal aplasia, contralateral congenital small kidney).[2–5]

MSK disease occurrence in childhood and its relationship with systemic congenital malformations support the hypothesis of an inherited condition.[6] As recently reported by our group, familial clustering of MSK disease is identified in several cases with autosomal dominant inheritance with reduced penetrance and variable expressivity.[7] Additionally, recent evidence showed that mutations or polymorphisms of the glial cell line–derived neurotrophic factor and receptor tyrosine kinase genes, disrupting the "ureteric bud–metanephric mesenchyme" interface, could be responsible for the disease pathogenesis.[8,9]

Despite the aforementioned associations and the latest progress in understanding the biological mechanisms associated with MSK disease, the pathogenesis of this disorder is

only partially defined, and further studies are necessary to achieve better understanding of the mechanisms.

Furthermore, at the moment, no diagnostic biomarkers are available for clinical purposes. In fact, because of its silent manifestation, currently MSK disease can only be clinically and laboratory suspected. Demonstration of nephrocalcinosis and multiple small calcium concretions seen on the papillary medulla on a radiography of the kidney or on a computed tomography scan without contrast or hyperechoic papillae on ultrasound is diagnostic.[10] The essential characteristic of the diagnostic image should be the identification of ectatic pre-calyceal papillary collecting ducts that represent the anatomic landmark of MSK disease.[11]

Because of its high spatial resolution, i.v. urography represented the gold standard for diagnosing MSK disease,[12,13] but unfortunately this test has been largely abandoned, revealing the necessity to identify MSK disease molecular biomarkers that could help clinicians to easily identify patients with this disease and to start preventive therapies to minimize renal and systemic complications.

Therefore, in our research project, a proteomic high-throughput methodology was used for the first time to identify new biological elements involved in the pathophysiology of MSK disease and to select specific MSK disease–associated proteins that, in future, whether validated in a large cohort, could represent valuable new diagnostic disease biomarkers for use in daily clinical practice. This will represent a significant step forward beyond current state of the art.

RESULTS

Hierarchical clustering analysis such as multidimensional scaling and Spearman's correlogram are able to discriminate the MSK from ICN based on their urinary proteomic profile

Multidimensional scaling (MDS) performed using the whole patients' urinary proteomic profile obtained by mass spectrometer analysis was able to clearly discriminate MSK disease from ICN (Figure 1). Interestingly, ICN with renal cysts are located on the edge of the two cohorts revealing an intermediate phenotype of these subjects.

Concordantly to MDS, correlogram (Figure 2a) was able to differentiate the 2 study groups. The average Spearman coefficient (R^2) was higher in the 2 study groups (0.45 ± 0.85). The mean of the coefficient of variation for the 21 biological replicates was 0.43, and no relationship was found between individual proteomes.

A total of 1529 proteins were identified by mass spectrometry (MS) analysis, and 884 (58%) overlapped the 2 study groups. Instead, 249 proteins (16%) and 396 (26%) were exclusive for ICN and MSK disease, respectively (Figure 2b) (Supplementary Tables S1 and S2).

To better describe differences between the 2 cohorts of patients and to restrict the number of MSK disease–associated proteins, we used a volcano plot and receiver-operating characteristic (ROC) curve analysis. The 2 methods revealed that 328 (Figure 3a) and 44 (Supplementary Table S3)

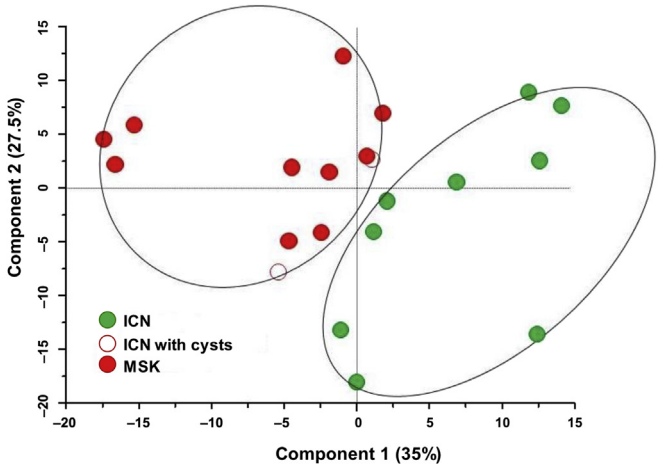

Figure 1 | Multidimensional scaling (MDS) analysis of biochemical clinical data of medullary sponge kidney (MSK) disease and idiopathic calcium nephrolithiasis (ICN) patients. Two-dimensional scatterplot of MDS analysis of MSK disease (red circles) and ICN (green circles) patients show 2 separate clusters corresponding to the 2 cohorts of patients. ICN with renal cysts are located on the edge of the 2 cohorts revealing an intermediate phenotype of these subjects.

proteins, respectively, were specific for MSK disease. Interestingly, then, 119 proteins (Supplementary Table S4) differentiated patients with cysts (a group including all MSK disease + 2 ICN patients with renal cysts) from ICN patients without cysts (Figure 3b). In particular, a total of 22 and 15 proteins were simultaneously up-regulated or down-regulated in the comparisons between MSK disease versus ICN and cysts versus no cysts.

Sixteen proteins were common in the 3 methods, as shown in the Venn diagram (Figure 4a) and listed in Table 1. The proteome profile (Figure 4b) summarized results and reported the name of the 16 highlighted proteins.

To have another classification method to establish the priority and relevance of proteins and to further reduce the choice of the highlighted potential biomarkers identified by means of statistical analysis, we also used a Support Vector Machine. A Support Vector Machine is a nonprobabilistic machine-learning method of binary classification/prediction proposed by Vapnik.[14] It constructs a hyperplane or a set of hyperplanes in a high- or infinite-dimensional space by different kernel functions to achieve a high accuracy of classification. Here, we used the nonlinear kernel function to establish a ranked proteins list.

Using Support Vector Machine, the majority of 44 proteins resulting from the ROC curve analysis and all of the 16 proteins obtained from the combination of the 3 statistic methods were lower than the hundredth position of the ranked proteins list. This result confirms the high ability of these proteins to distinguish MSK disease from ICN (Figure 4).

Five proteins have been consistently selected by all 4 methods: glypican-1, plexin domain–containing protein 1, beta-hexosaminidase, epididymis-specific alpha-mannosidase, and laminin subunit alpha 2 (LAMA-2). This last protein was the most significant biomarker identified by the 4 methods