Running title:

A methodological approach based on gold-nanoparticles followed by matrix assisted laser desorption ionization time of flight mass spectrometry-based profiling for the analysis of urine profiling of Knee Osteoarthritis

Statement of relevance

The American College of Rheumatology considers the following main criteria for diagnostic of knee osteoarthritis, KOA: (i) clinic criteria, (ii) clinic and laboratory criteria and (iii) clinic and radiologic criteria. The sensitivity and specificity for each one is, respectively, 95% and 69%; 92% and 75%; and 91 and 86%.

Urine gold-nanoparticle-based MALDI profiling can be used as a rapid tool to diagnostic people with KOA. This classification can be reached by profiling the samples or by searching for specific m/z peaks of classification using Urine gold-nanoparticle-based sample treatment. The method here proposed has sensitivity and specificity values of the same order of above-mentioned criteria of diagnostic, 97% and 69% respectively. When the proposed method is compared with those used in common practice it is found that sensitivity was the highest, thus with a low false negative rate for diagnostic KOA patients in the population studied. Specificity was lower but in the range accepted for diagnostic objectives. Further advantages of the method proposed are that urine is a sample that can be taken without invasive techniques, is easy of handling and that the proposed protocol is not expensive, no time-consuming.

A methodological approach based on gold-nanoparticles followed by matrix assisted laser desorption ionization time of flight mass spectrometry for the analysis of urine profiling of Knee Osteoarthritis

Rubén López-Cortés^a, Jacobo Formigo^b, Miguel Reboiro-Jato^c, Florentino Fdez-Riverola^c, Francisco J. Blanco^d, Carlos Lodeiro^{e,f}, Elisabete Oliveira^{e,f}, J. L. Capelo^{e,f*}, H. M. Santos^{e,f*}

^aRubén López-Cortés. Facultade de Bioloxía. Departamento de Bioquímica. Xenética e Inmunoloxía. Grupo de Biomarcadores Moleculares. Universidade de Vigo. Campus Lagoas-Marcosende E36310. Vigo. España.

^bDepartment of Physical Medicine and Rehabilitation. Complexo Hospitalario Universitario de Ourense. Spain.

^cSING Group. Informatics Department. Higher Technical School of Computer Engineering. University of Vigo. Ourense. Spain.

^dServicio de Reumatología, Unidad de Proteómica. ProteoRed/ISCIII. INIBIC-Hospital Universitario A Coruña. As Xubias 84. 15006. A Coruña. Spain.

^eBIOSCOPE group. REQUIMTE. Department of Chemistry. Faculty of Science and Technology. University NOVA of Lisbon. Caparica. Portugal.

^fProteoMass Scientific Society. Madan Parque. Rua dos Inventores. 2825-182. Caparica. Portugal.

*Corresponding Authors:

J. L. Capelo. Email: jlcm@fct.unl.pt URL: www.bioscopegroup.org

H. M. Santos. Email: hms@fct.unl.pt URL: www.bioscopegroup.org

BIOSCOPE group. REQUIMTE. Department of Chemistry. Faculty of Science and Technology. University NOVA of Lisbon. Caparica. Portugal and ProteoMass Scientific Society. Madan Parque. Rua dos Inventores. 2825-182. Caparica. Portugal.

Abstract

The aim of this work is to develop a nanoparticle-based methodology to find out biomarkers of diagnostic for knee osteoarthritis, KOA, through the use of matrix assisted laser desorption ionization time-of-flight-based mass spectrometry profiling. Urine samples used for this study were obtained from KOA patients (42 patients), patients with prosthesis (58 patients), and controls (36 individuals) with no history of joint disease. Gold-nano particle MALDI-based urine profiling was optimized and then applied over the 136 individuals. Jaccard index and 10 different classifiers over MALDI MS datasets were used to find out potential biomarkers. Then, the specificity and sensitivity of the method were evaluated. The presence of ten m/z signals as potential biomarkers in the healthy versus non-healthy approach suggests that patients (KOA and prosthesis) are differentiable from the healthy volunteers through profiling. The automatic diagnostic study confirmed these preliminary conclusions. The sensitivity and the specificity for the urine profiling criteria here reported, achieved by the C4.5 classifier, is 97% and 69% respectively. Thus, it is confirmed the utility of the method proposed in this work as an additional fast, non-expensive and robust test for KOA diagnostic. When the proposed method is compared with those used in common practice it is found that sensitivity is the highest, thus with a low false negative rate for diagnostic KOA patients in the population studied. Specificity is lower but in the range accepted for diagnostic objectives.

Keywords: urine, gold nanoparticle, knee osteoarthritis, MALDI, profiling, diagnostic

1. Introduction

Knee osteoarthritis (KOA) is an articular degenerative pathology defined as the loss of hyaline cartilage, subchondral sclerosis, bone degradation and synovial joint narrowing [1]. At a biochemical level, KOA is characterized by the degradation of several components of the extracellular matrix (ECM), as cartilage [2], as well as an increased local inflammation of the affected joint [3]. Several protein biomarkers such as monocyte chemotactic protein-1 (MCP-1) and fibronectin-aggrecan complex (FAC) correlate with KOA stage [4]. Furthermore, other molecules of cell signaling like IL-1 β , TNF, IL-6, IL-15, IL-17, IL-18, IL-21, leukemia inhibitory factor and IL-8 have been linked to the development of the disease.

OA is the most common articular pathology in developed countries [5]. Prevalence of symptomatic KOA increases from 10% in patients older than 20 to 14% in patients older than 65, reaching the maximum in the 70-79 age range, and thereafter decreasing [6,7]. KOA is more prevalent in females than in males and seems to be highly associated to elderly [8]. Usual clinical symptoms are pain and the rigidity of the affected joint [9]. Currently, OA is the first permanent impairment cause in the developed world [10]. This pathology is predicted to increase its incidence dramatically [11], as well as it is predicted that related health care costs will increase."

KOA is commonly diagnosed by radiologic and clinical analysis [12]. The clinical reference for diagnosis is by radiographing the affected articulation and staging by Kellgren-Lawrence scale from 0 to 4 [13]. However, clinicians usually report difficulties in making a correct diagnostic for patients with minimal radiological symptoms referring high pain and others with a clear joint wear that suffer neither pain nor functional alterations. Furthermore, actual diagnostic tests are not useful in early detection [14]. To help in early detection, some biomarkers of inflammation such as high sensitive C-reactive protein (hsCRP) [15] has been proposed as a complement to radiological tests albeit unspecific. Moreover, some other potential biomarkers are being proposed for validation in the last years as well as those that correlate with evolution of the pathology during recommended pharmacologic treatments [16,17,18]. Within the field of proteomics, methodologies widely used to study the protein content are the chromatographic techniques, mono and two-dimensional electrophoresis and

mass spectrometry [19]. Instead of identifying discrete expression biomarkers, a quick and inexpensive way to classify samples is using matrix assisted laser desorption-ionization, MALDI,-mass spectrometry-based fingerprint of peptides [20]. This methodology is applied increasingly by the scientific community due to its versatility and simplicity, even in the field of rheumatology [21]. Peptide fingerprinting only requires an easy analysis per sample and the results are obtained in some tens of seconds [22].

In previous studies we have described a high-throughput method for serum proteomic profiling in patients with multiple myeloma [23]. In the present work such method has been optimized and applied to the complexity of urine samples. The urine's protein content was concentrated by precipitation using citrate-coated gold nanoparticles (citrate@GNPS). Then, ultra fast ultrasound-assisted in-solution tryptic digestion of the proteins followed by direct peptide profiling by MALDI-MS was used to obtain the profiles of the patients. Then, profiles were used to classify the patients and a pattern of prediction, healthy versus diseased, was obtained.

2. Material and methods

2.1 Reagents

All reagents were HPLC-grade or higher. Sodium dodecyl sulfate (SDS) and formaldehyde were purchased from Panreac (Barcelona, Spain). β -mercaptoethanol was purchased from Merck (Hohenbrunn, Germany) and bromophenol-blue was purchased from Riedel-de Haen (Seelze, Germany). Trypsin, trifluoroacetic acid, bovine serum albumin, carbonic anhydrase DL-Dithiothreitol (DTT), Iodoacetamide (IAA), acrylamide/bis-acrylamide 30% solution (37.5:1), Tris-base, Coomassie Brilliant Blue R250 (CBB), sodium carbonate, sodium thiosulfate, Na_2 -EDTA, Sodium citrate tribasic, silver nitrate solution and the SigmaMarker wide range 6,500–200,000 Da were purchased from Sigma-Aldrich (St. Louis, MO, USA). Hydrogen tetrachloroaurate (III) hydrate (99,9%-Au) (49% Au) at 10% w/v was purchased from Strem Chemicals (Newburyport, MA, USA). α -Cyano-4-hydroxycinnamic acid, ammonium bicarbonate (AMBIC) and formic acid were purchased from Fluka (Steinheim, Germany). ZipTip® was purchased from Millipore. Creatinine urinary kit assay was purchased from Cayman chemicals (Ann Arbor, MI, USA).

2.2 Samples

Urine samples used for this study, first morning void, were obtained from knee osteoarthritic, KOA, patients (42 patients), patients with prosthesis (58 patients), and controls (36 individuals) with no history of joint disease. Samples were provided by the CHUO (Complejo Hospitalario Universitario de Ourense, Spain). Patients were of both genders and ages ranged from 47 to 86 years. Samples were taken from anonymous patients from the Hospital of Ourense, Spain. All of them signed an informed consent. The study was approved by the ethics committee of Galicia (Spain).

All patients that volunteered for this study were classified as normoproteinuric, being the average protein concentrations as follows: (i) samples from healthy volunteers: 69.3 ± 25.6 $\mu\text{g/ml}$, (ii) samples from KOA patients 50.8 ± 19.7 $\mu\text{g/ml}$ and (iii) samples from patients with knee joint prosthetic 56.5 ± 27.0 $\mu\text{g/ml}$. Table 1 summarizes the characteristics of the patients included in this study.

2.3 Pre-analytical treatment of urine samples and urine normalization

Prior to analysis, 20 ml of urine were centrifuged for 30 min at 5000g to remove cells and debris. Then, 3 ml of urine were diluted to 30 ml with ultrapure water, then urine was aliquoted in five subsamples and stored at -60 °C.

Total protein content was determined using a Bradford assay. Urine protein content was normalized to creatinine by the Jaffè assay from Cayman Chemicals (500701 kit, Ann Arbor, MI, USA). Colorimetric absorbance was measured at 492 nm in duplicate; results were expressed in mg protein/mmol creatinine.

2.4 Apparatus

Microplate-reader LT-4000 from Labtech (Uckfield, UK) was used for Bradford and Jaffè colorimetric assays. PowerPac Basic from Bio-Rad (CA, USA) was used as current supplier for SDS-PAGE protein separation. Image gels were obtained using a ProPicII (Digilab-Genomic Solutions, USA). Mass spectrometry analysis was performed using a MALDI Ultraflex from Bruker-Daltonics (Bremen, Germany),

equipped with a LIFT cell and a nitrogen laser of 50Hz. Centrifugation was done in a CM-50M centrifuge from ELMI (Riga, Latvia).

2.5 Synthesis of citrate-capped gold nanoparticles.

Gold nanoparticles, GNPs, were synthesized by the citrate reduction method in aqueous solution [23]. Briefly, 60 ml of sodium citrate tribasic solution (0.075% w/v) was heated to 100 °C, and then gold was added as 54 µL of 10% w/v of hydrogen tetrachloroaurate (III) hydrate solution. The reaction mixture was kept under refluxing until a deep red color was detected. Solution of nanoparticles is chilled at room temperature and stored at 4 °C for a maximum of one month.

2.6 Protein reduction and alkylation

Urine treated as described in 2.3 was 1/30 diluted with milliQ water, then the sample was made pH 8.8 with Tris 2.5 M. Disulphide bonds were reduced with DTT (final concentration 10 mM) and incubated at 37 °C for 1 h. Proteins were subsequently alkylated with IAA (final concentration 50 mM) at room temperature for 45 min.

2.7 “Protein solid-phase extraction with citrate-GNPs”

After protein reduction and alkylation, 1 ml of GNPs was added to the sample and pH was adjusted to 4.0 with 1 M of citric acid. To allow for the proteins to bind to gold-nanoparticles, samples were incubated for 2 h at 37°C. Finally, GNPs-protein complexes were harvested by centrifugation at 8600 g for 40 min and the resulting pellet was then washed twice with 1ml of citric acid/ sodium citrate buffer pH 4.0 to remove unbound protein.

2.8 SDS-PAGE

The GNPs-Protein pellet was resuspended in 5 µL of water plus 5 µL of sample buffer (65.8 mM Tris-HCl pH 6.8, 2.1% SDS (w/v), 26.3% (w/v) glycerol, 0.01% (w/v) of bromophenol blue and 10% (v/v) β-mercaptoethanol). Samples were heated to 95 °C for 5 min and then loaded into a 5% acrylamide/bis-acrylamide, stacking gel / 12.5% acrylamide/bis-acrylamide running gel, of 1mm thickness, and

separated at 200 V for 50 minutes, until the tracking dye front reaches the bottom of the gel. After gel electrophoresis, gels were rinsed with milli-Q water and then incubated for 2 hours in R-250 Coomassie staining solution at room temperature (previously prepared dissolving 1 g of R-250 in 200 ml of a mixture of 45% v/v methanol and 7.5% v/v glacial acetic acid). After staining, gels were rinsed 4×20 min with 100 mL of distilled water and further washed twice with 100 mL of 45% v/v methanol and 7.5% v/v glacial acetic acid until a clear background was observed. Gel imaging was carried out with a ProPicII-robot using 14 ms of exposure time and a resolution of 70 μm .

2.9 Protein digestion

GNPs-protein pellets were resuspended in 10 μL of ACN 50% (v/v) / Ambic 25 mM. Then, 5 μL were further diluted with Ambic 25 mM up to 50 μL and digested with trypsin. Digestion was performed using an ultrasonic cup-horn (Bandelin electronic, Berlin, Germany) for 5 min at 50% ultrasonic amplitude [24]. Digestion was stopped by the addition of 1 μL of 50% (v/v) formic acid. Samples were desalted with zip-tips and peptides were eluted with 90% (v/v) ACN/ 0.1% (v/v) TFA [25]. Eluted peptides were dried in a UNIVAPO 150 ECH from UniEquip (Martinsried, Germany) coupled with a refrigerated vacuum pump Unijet II (UniEquip). Prior to MS analysis, samples were resuspended in 10 μL of formic acid 0.3% (v/v) and 0.5 μL of sample was hand-spotted onto a ground steel MALDI plate (MTP348) and subsequently 1 μL of matrix solution containing 5 mg/ml of alpha-cyano-4-hydroxycinnamic acid in 50% (v/v) ACN / 0.1% (v/v) TFA was added and allowed to air dry.

2.10 MALDI MS analysis

MALDI-TOF/TOF mass spectrometer was operated in positive ion mode using a reflectron. Spectra were acquired in the m/z range of 600 to 3500. A total of 500 shots were acquired for each sample, five times (five different spots) at a laser frequency of 50 Hz. External calibration was performed with the $[\text{M}+\text{H}]^+$ monoisotopic peaks of bradykinin 1-7 (m/z 757.3992), angiotensin II (m/z 1046.5418), angiotensin I (m/z

1296.6848) substance P (m/z 1758.9326), ACTH clip 1-17 (m/z 2093.0862), ACTH 18-39 (m/z 2465.1983) and somatostatin 28 (m/z 3147.4710).

The 5 spectra acquired for each sample were pre-processed in two steps: (i) peaks with a signal-to-noise ratio lower than 6 were discarded, and (ii) peaks of the different spectra were matched and a consensus spectrum was created using the “Forward” algorithm included in the Mass-Up software [26] configured to use a tolerance of 300 ppm and a percentage of presence of 80%. After the intra-sample pre-processing, an inter-sample peak matching is performed with the Mass-Up software configured to use a tolerance of 300 ppm. This last step unifies the peak values among samples.

2.11 Jaccard Index (JI) and statistical analysis

In order to estimate the similarity of the profiles of the samples, the Jaccard Index (JI) was employed. JI is a statistic measure used for comparing the similarity of two sets. As Formula 1, F1, shows bellow, it is calculated as the ratio between the intersection and the union of the two sets under evaluation (i.e. A and B). Therefore, the JI value is in the range [0, 1], where 0 means that sets do not have common elements and 1 means that sets are equals.

$$F1: JI(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

For the evaluation of the statistical significance of the peaks only the presence or absence of the peaks in each sample was taken into account, discarding the use of the intensities. Following the recommendations of [27], Fisher’s exact test of independence was used for problems with two conditions, while randomization test of independence was used for the multiple conditions problems. For the randomization test, 100,000 simulations of the χ^2 test were done for each peak. In order to control the presence of false positive among the statistically significant peaks, p-values were corrected using the Benjamini-Hochberg’s false discovery rate procedure [28].

2.12 Automatic Diagnostic

The suitability of the approach presented in this work for automatic diagnostic of KOA was assessed by evaluating the performance of 10 different classifiers over MALDI MS datasets. Specifically, the classifiers evaluated were a K-Nearest Neighbours (KNN)

[29], a C4.5 decision tree [30], a Random Forest (RF) [31], a Bayes Network (BN), a Naïve Bayes Simple (NBS) [32], a Decision Table (DT) [33], a PART decision list [34], a Support Vector Machine (SVM) trained with the Sequential Minimum Optimization algorithm (SMO) [35], a SVM with a lineal kernel (SVM-LIN), and a SVM with a radial basis function kernel (SVM-RBF). Tests were run using the implementation of these algorithms provided by the Weka machine learning library [36] Default configuration was used for every algorithm, except for the IBK classifier, configured to optimize the number of neighbours to a maximum of 10, and for the SVM-LIN and SVM-RBF classifiers, optimized according to the suggestions of the authors of the SVM implementation included in Weka [37].

Each classifier was evaluated following a Leave-One-Out Cross Validation (LOOCV) strategy. Cross-validation techniques evaluate the performance of classifiers by repeatedly splitting an initial dataset into two subdatasets: (i) a training set, used to train the classifier, and (ii) a testing set, classified by the trained classifier. The performance of the classifier is estimated with the predictions for testing sets. When the LOOCV is used, a singleton testing set is created for each sample. The corresponding trainings sets will contain all the samples but the one contained in the testing set.

Training datasets were filtered with a custom feature selector. The tests of independence described in the previous section were applied to each training dataset and peaks with an corre p-value (q-values) above 0.05 were discarded. In order to avoid biasing the performance results, feature selector was applied independently to each of the training sets generated with the LOOCV.

Performance of the classifiers was assessed using the sensitivity, specificity, accuracy, and Mathews Correlation Coefficient (MCC) statistical measures. These are functions commonly used to evaluate the performance of binary classification tests, although accuracy can be used for multiclass classification tests. These functions are based on four measurements that, in the context of a case vs. control problem, can be described as follow:

- (i) true positive, TP, as the number of patients from the case group truly identified as belonging to the case group,
- (ii) false negative, FN, as the number of patients from the case group identified as belonging to the control group,
- (iii) false positive, FP, as the number of volunteers from the control group identified as belonging to the case group, and

(iv) true negative (TN) is the number of volunteers from the control group identified as belonging to the control group.

Sensitivity is defined as the $TP/(TP+FN)$ ratio, Specificity is defined as the $TN/(FP+TN)$ ratio. Accuracy is the proportion of correct predictions against the total number of predictions, and it can be calculated as the $(TP+TN)/(TP+TN+FP+FN)$ ratio. Finally, the MCC is a measure of the global quality of a binary classification related to the χ^2 statistic. It is a correlation coefficient between the observed and predicted binary classifications that takes values between -1, for total disagreement, and +1, for a perfect prediction. An MCC value of 0 means that prediction is no better than random. MCC can be calculated as given in F2.

$$F2: MCC = (TP \times TN - FP \times FN) / \sqrt{((TP+FP)(TP+FN)(TN+FP)(TN+FN))}$$

Additionally, the 95% confidence interval for the sensitivity, specificity and accuracy measurements was estimated by the Wilson score interval with continuity correction approach [38].

3.- Results and discussion

3.1.- Preservation of urine.

The first problem to be solved when dealing with urine is to preserve its integrity as much as possible [39]. To this end, the scientific community has developed a large number of protocols. By the year 2010, the Human Kidney and Urine Proteomics Project, HKUPP, addressed some protocols for preserving urine that were assayed in the first steps of this work [40,41]. Thus, the HKUPP's recommended protocol at that time was as follows: first, urine is centrifuged at 10 000 g to eliminate cell and debris (in our case 3 ml were centrifuged). Then urine was buffered with Tris buffer 1M (final pH 8.0, in our case 2 ml of Tris buffer 1M were added) and then the sample was immediately frozen to -60 °C. Next day samples were unfrozen but, unlikely, a precipitate was formed in all the samples prepared following the aforementioned protocol. To recover the precipitate, the samples were centrifuged at 6 000 rpm during 3 min. Then the supernatant was withdrawn and the precipitate was cleaned twice with 100 µL of acetone. Unlikely, it was not possible to dissolve this precipitate using 10 mM PBS and ultrasonication. As the formation of a precipitate after thawing urine samples usually causes the missing of unique proteins [42], the method recommended by the HKUPP by

the year 2010 was dismissed. Then, a number of protocols obtained also from the HKUPP and other literature were assayed [43-45]. In addition, and based on our own expertise in sample treatment of complex samples, it was decided to test sample dilution as a way to avoid the formation of a precipitate. The protocols that were tested are briefly described in Table 2. Of such methods, the one selected to avoid the formation of a precipitate after unfreezing the urine samples consisted first, in centrifuging the sample to eliminate cell debris and then, in diluting up to 10% v/v the sample with ultrapure water (3 ml to 30 ml). Following such protocol the formation of a precipitate after unfreezing the sample was not observed for any of the 136 samples assessed. One sample, treated as described above, was submitted for 1D-gel electrophoresis before freezing it at -60 °C and after 30 days of preservation. Results presented in Fig. 1 clearly show that the intensity and integrity of the bands remain the same, suggesting that the proteome was preserved. Therefore, this was the method of choice for preserving the urine samples, as it was not expected a storage period longer than 30 days.

3.2 Optimization of urine profiling using gold nanoparticles.

A large number of methods to concentrate the proteins present in urine have been reported in literature, being the most reported the one that uses precipitation and ultracentrifugation [46]; although filtration, dialysis and lyophilization have been also cited [47]. A comparison of the GNPs-based precipitation approach with the above-mentioned methods reveals that the GNPs-based method is less time-consuming and that it allows using larger urine volumes. It is worth to mention that urine protein concentration circa three times larger is required for protein precipitation methods that use chemical organic solvents if compared with the GNPs method. This factor acquires relevance for the case of urine samples from patients with renal disease [48].

Serum profiling using GNPs previously done in our laboratory has addressed a number of parameters that need to be optimized, as follow: (i) GNPs/protein ratio, (ii) pH range and (iii) incubation time. In addition, we have shown in previous work [23] that proteins with and without cysteine amino acids, a.a, can be separated one from each other using gold nanoparticles. Proteins without cysteine a.a. remain in the supernatant and can be further investigated, if necessary, as we have also recently shown [49].

In the following sections the optimization of such parameters is described.

3.2.1 Optimization of the GNPs/protein ratio

To assess the effect of the GNPs (ml)/protein ($\mu\text{g/ml}$) ratio in the protein separation method, three different volumes of the same urine, 1, 3 and 9 ml were diluted up to 30 ml. Each volume was prepared six times and the amount of protein for each one was estimated to be $1.37 \pm 0.04 \mu\text{g/ml}$, $4.1 \pm 0.5 \mu\text{g/ml}$ and $12 \pm 2 \mu\text{g/ml}$ (mean $\pm$ standard deviation), respectively. Each diluted sample (total of six replicates for each volume of 1, 3, and 9 ml), was first reduced with DTT and then alkylated with IAA. Afterwards, 1 ml of a solution containing GNPs was added, and the solutions were made pH 4 with citrate buffer 0.1 M. Then, solutions were incubated during 2h at 37°C. Afterwards, the samples were pelleted by centrifugation at 8 600 g during 40 min. Then, pellet was re-suspended in 10 μL of a 1:1 solution of ACN/Ambic (25 mM). Then, as described in the sample treatment section, 5 μL were used for 1D gel electrophoresis and 5 μL were used for MALDI profiling. For this case, samples were 1:1 mixed with α -CHCA. Each sample was spotted five times.

The results for 1D gel electrophoresis experiments are shown in Fig. 2. It can be noted that the sample with the largest urine volume (9 ml) produces a broad band in the zone corresponding to masses at about 250-150 KDa. Another important tip was obtained when the MALDI profiling was carried out. The spectra for each sample (a total of 5 MALDI replicates) were pre-processed using the protocol described in section 2.11. Then, Jaccard Index, JI, was used to compare the reproducibility of the spectra. Figure 1 supplementary material, Fig. 1SM, shows the MALDI spectra of all replicates. The average JI obtained were of 0.654 ± 0.105 , 0.567 ± 0.116 and 0.450 ± 0.115 , for the protein amounts of $1.37 \mu\text{g/ml}$, $4.11 \mu\text{g/ml}$ and $12.33 \mu\text{g/ml}$, respectively. An ANOVA test comparing the average JI of the three batches confirmed that the differences are significant with a p-value of 5.71×10^{-6} . In addition, a Tukey's HSD post-hoc test revealed that MALDI profiling reproducibility gets worst, as the amount of protein gets higher. Further information about this preliminary study is provided in Data 1 supplementary material D1SM. Because a compromise must be obtained between reproducibility and amount of total protein under assessment, the maximum amount of protein to work with was settled in $4.0 \mu\text{g/ml}$.

3.2.2 Optimization of the pH and time of incubation.

Two factors affecting protein separation using gold nanoparticles must be considered. On the one hand, proteins containing residues of cysteine must be present as is trough sulphur that the proteins and peptides attach to the gold [50]. On the other hand, peptides and proteins have positive charges depending on the pH of the solution, whilst gold nanoparticles present a negative zeta (ζ potential). For this reason pH must be carefully assessed. To test at which pH the best separation of proteins takes place, urine was incubated at different pHs ranging from 3.5 to 9.5. It is known that pH lower than 3 leads to GNPs precipitation [51] whilst if pH is higher than 8, the negative charge of the proteins hampers them of attaching the GNPs surface [52]. As shown in Fig. 3, the best separation was obtained at a pH range comprised between 3.5 and 4.5. At pH higher than 4.5 the amount of protein separated drops significantly, as it is shown by the lower intensity of the bands.

Regarding incubation time, in addition to the 2 h initially tested, the samples (with the nanoparticles) were also left in agitation overnight (12 h, in Fig. 3). As may be seen in Fig. 3, overnight agitation does not bring benefit. This trend was also confirmed through the MALDI analysis of the pellets, as for all samples left overnight the MALDI profiles were less intense and with less peaks than their counterpart of 2 h.

Therefore, the selected pH and incubation time for further experiments were set as 4.0 and 2 h, respectively.

3.3 Optimization of ultrasonic trypsination.

Once proteins are separated through GNPs, then the pool of peptides is obtained by digesting the proteins while they are still linked to GNPs. To make easier the study about protein digestion, once the sample was pelleted and then resuspended, 5 μ L were diluted up to 50 μ L by adding 45 μ L of AMBIC 25mM (5% v/v in ACN). Then, different ratios trypsin/GNPs and ultrasonication time were assayed to find out the best digestion protocol. The ultrasonic amplitude was settled at 25% or 50%. As may be seen in Fig. 2SM, a total of 8 different methods were assayed. The method providing the spectra with the highest number of m/z signals was as follows: 1/20 trypsin/protein ratio, 5 minutes (2x2.5 min) ultrasonication time and 25% ultrasonic amplitude.

4. Application to a large set of samples

Once the best pH, incubation time, ratio trypsin/GNPs and ultrasonic amplitude were determined, such conditions were used in a real case study as explained below.

A total of 136 individuals were recruited for this study. The characteristics of the groups are shown in Table 1. As stated in section 2.13, the suitability of the proposed approach for the automatic diagnostic of KOA was assessed by evaluating the performance of ten different classifiers over a dataset created with the results of the MALDI profiling as described in section 2.11.

The lists of m/z signals of KOA and prosthesis groups were compared but m/z signals with statistical significant difference were not found. This result suggests that these two groups of patients are closely related and that they cannot be differentiated using profiling (at least with the conditions used in this work). On basis on this finding, a different approach for diagnostic was proposed: Healthy individuals versus non-healthy individuals, where KOA and prosthesis patients were grouped together and compared with the healthy volunteers. The tests of independence described in Section 2.12 were used to identify peaks with statistical significant difference that could be used as biomarkers. Figure 4 lists the peaks with a corrected q -value < 0.05 for the healthy versus non-healthy approach. The presence of ten potential biomarker m/z signals in the healthy versus non-healthy approach suggests that patients (KOA and prosthesis) are differentiable from the healthy volunteers through profiling. The automatic diagnostic study confirmed these preliminary conclusions. Figure 5 shows one healthy spectrum as well as one patient spectrum with some characteristic m/z values highlighted. Tables 1SM and 2SM presents the results for the healthy versus non-healthy approach. As may be seen, the classifiers show a good performance, with accuracy above 85%, a MCC value above 0.7 and a sensitivity of 97% in almost every case. Specificity was slightly lower, with a peak value of 69% for the decision trees (C4.5 and RF) and the SVM-RBF.

5. Evaluation as criteria for clinic diagnostic

As a general role the sensitivity of a method of diagnostic, refers to its capacity to correctly identify a specific disease. On a similar way, the specificity of a method refers to its capacity to correctly identify patients without a specific disease. Although these

measures are often considered fixed properties of a diagnostic test, in reality they are subject to multiple sources of variation such as the population case mix and the severity of the disease under study. As a general role the American College of Rheumatology considers the following main criteria for diagnostic of KOA: (i) clinic criteria, (ii) clinic and laboratory criteria and (iii) clinic and radiologic criteria. The sensitivity and specificity for each one is, respectively, 95% and 69%; 92% and 75%; and 91 and 86% [53]. The sensitivity and the specificity for the urine profiling criteria here reported, achieved by the C4.5 classifier, is 97% and 69% respectively. Thus, it is confirmed the utility of the method proposed in this work as an additional fast, non-expensive and robust test for KOA diagnostic.

6. Conclusions

Urine gold-nanoparticle-based MALDI profiling of knee osteoarthritis can be used as a rapid tool to help in classifying people with such disease. This classification can be reached by profiling the samples or by searching for specific m/z peaks of classification. The method here proposed has sensitivity and specificity values of the same order of other classic criteria of diagnostic such as clinic and laboratory criteria or the clinic and radiologic criteria. When the proposed method is compared with those used in common practice it is found that sensitivity was the highest, thus with a low false negative rate for diagnostic KOA patients in the population studied. Specificity was lower but in the range accepted for diagnostic objectives.

Further advantages of the method proposed are that urine is taken without invasive techniques, is easy of handling, is not expensive and is not time-consuming.

The proposed method will require further validation by other research teams in large cohorts of patients.

Acknowledgements

E. Oliveira and H. M. Santos acknowledge the post-doctoral grants SFRH/BPD/72557/2010 and SRFH/BPD/73997/2010 respectively, provided by Fundação para a Ciência e a Tecnologia –Ministério da Educação e Ciência (FCT-MEC, Portugal). This work was supported in part by Xunta de Galicia (Spain) under projects 09CSA043383PR and 10CSA383009PR, by PROTEOMASS Scientific

Society (Portugal) and by Fondo de Investigacion Sanitaria-Spain (PI-12/0329).

Authors thank to REQUIMTE-FCT PEst-C/EQB/LA0006/2013. Maphre Foundation is also acknowledged for financial support.

The authors declare no conflict of interest

References

- [1] F. J. Blanco. *Reumatol. Clin.* 10 (2014) 4-5.
- [2] R. J. Boniface, P.R. Cain, C.H. Evans *Arthritis Rheum* 31 (1988) 258-266.
- [3] M. Kapoor, J. Martel-Pelletier, D. Lajeunesse, J. P. Pelletier, H. Fahmi. *Nat Rev Rheumatol.* 7 (2010) 33-42.
- [4] T. P. San Giovanni, S. R. Golish, A. Palanca, L. S. Hanna, G. J. Scuderi. *Foot and Ankle Int.* 33 (2012) 627-631.
- [5] Zhang Y., Jordan J.M. *Rheum Dis N Am* 34 (2008) 515-529.
- [6] Carmona L., Ballina F.J., Gabriel R., Laffon A., EPISER Study Group. *Ann Rheum Dis* 60 (2001) 1040-1045.
- [7] S. Issa, L. Sharma, *Curr. Opin Rheumatol* 8 (2006) 7-15.
- [8] Centers for Disease Control and Prevention. 2005. *MMWR Weekly* 58 (2009) 421-426.
- [9] I. Haq, E. Murphy, J. Dacre, *J. Postgrad. Med. J.* 79 (2003) 377-783.
- [10] E. Battel-Gualda, V. Jovani Casano, J. Ivorra Cortes, E. Pascual Gómez *Rev. Esp. Reumatol.* 25 (1988) 91-105.
- [11] J. M., Hootman, C. G. Helmick. *Arthritis Rheum.* 54 (2006) 226-229.
- [12] A. Alonso Ruíz in: *SER Sociedad Española de Reumatología (Ed.) Artrosis. Fisiopatología, diagnóstico y tratamiento.* Panamericana 2010. pp. 65-71.
- [13] J. H. Kellgren, J. S. Lawrence. *Ann. Rheum. Dis.* 16 (1957) 494-502.
- [14] R. P. Matveev, S. V. Bragina, *Human Ecol.* 9 (2012) 53-62.
- [15] T. Saxne, M. Lindell, B. Mansson, I.F. Petersson, H. D. Heinegard. *Rheumatology* 42 (2003) 903-904.
- [16] J. P. Pelletier, J. P. Raynald, J. Caron, F. Mineau. *Ann. Rheum. Dis.* 69 (2010) 2095-2101.
- [17] M. A. Karsdal, I. Byrjalsen I., A. C. Bay-Jensen, K. Henriksen. *BMC Musculoskelet Disord.* 11 (2010) 125-133.

- [18] G. J. Scuderi, S. R. Golish, F. F. Cook, J. M. Cuellar. Bone Joint J.- Series A 93 (2011) 336-340.
- [19] F. E. Ahmed. Expert Opin. Med. Diagn. 3 (2009) 429-444.
- [20] X. Zhang, S. Zhu, C. Deng, X. Zhang. Chem. Commun. 48 (2012) 2689-2691.
- [21] X. Zhang, Z. Yuan, B. Shen, M. Zhu. Int. J. Clin. Exp. Med. 12 (2012) 145-151.
- [22] R. Bakry. Anal. Chim. Acta 690 (2011) 26-34.
- [23] R. Lopez-Cortes, E. Oliveira, C. Nuñez, C. Lodeiro, María Páez de la Cadena, Florentino Fdez-Riverola, H. López-Fernández, Miguel Reboiro-Jato, Daniel Glez-Peña, José Luis Capelo, Hugo M. Santos Talanta 100 (2012) 239–245.
- [24] H. M. Santos, C. Mota, C. Lodeiro, I. Moura. Talanta 77 (2008) 870–875.
- [25] J. L. Capelo, R. J. Carreira, L. Fernandes, H. M. Santos. Talanta 80 (2010) 1476-1486
- [26] <http://sing.ei.uvigo.es/mass-up>, last time accessed 2014, July, the 20th
- [27] McDonald J.H. Handbook of Biological Statistics. Sparky House Publishing, Baltimore 2009. pp. 80-83.
- [28] Benjamini Y., Hochberg Y. J. Roy. Statist. Soc. Ser. B 57 (1995) 289–300.
- [29] D. Aha and D. Kibler Machine Lear. 6 (1991) 37-66.
- [30] Quinlan R. C4.5: Morgan Kaufmann Publishers. San Mateo, CA. 1993.
- [31] L. Breiman. Mach. Lear. 45 (2011) 5-32.
- [32] R. Duda and R. Hart. Pattern Classification and Scene Analysis. Wiley, New York 1973.
- [33] R. Kohavi. The Power of Decision Tables. 8th European Conference on Machine Learning 1995, 174-189.
- [34] E. Frank and I. H. Witten. Generating Accurate Rule Sets Without Global Optimization. Fifteenth International Conference on Machine Learning, 1998, 144-151.
- [35] T. Hastie and R. Tibshirani. Stanford Press. Advances in Neural Information Processing Systems. 1998.
- [36] M. Hall, E. Frank, G. Holmes, B. Pfahringer. SIGKDD Explorations 11 (2009) 10-18.
- [37] <http://www.csie.ntu.edu.tw/~cjlin/libsvm/> last time accessed 2014, July, the 20th
- [38] R. G. Newcombe. Statis. Med. 17 (1998) 873-890.
- [39] V. Thongboonkerd and P. Saetun. J. Proteome Res. 6 (2007) 4173-4181.
- [40] <http://www.hkupp.org/Urine%20collection%20Documents.htm>

- [41] T. Yamamoto. The 4th Human Kidney and Urine Proteome Project (HKUPP) Workshop 26 September 2009, Toronto, Canada. 10 (2010) 2069-2070.
- [42] D. Mataija-Botelho, P. Murphy, D. M. Pinto, D. L. MacLellan. *Proteomics Clin. Appl.* 3 (2009) 95-105.
- [43] V. J. Thongboonkerd. *J. Prot. Res.* 6 (2007) 3881-3890.
- [44] J. C. Sadler and W. A. Shanks. *Ann. Clin. Biochem.* 24 (1987) 111-112.
- [45] C. E. Thomas, W. Sexton, K. Benson, R. Sutphen, J. Koomen. *J. Cancer Epidemiol Biomarkers Prev.* 19 (2010) 953-959.
- [46] V. Thongboonkerd, K. R. McLeish, J. M. Arthur, J. B. Klein. *Kidney Int.* 62 (2002) 1461-1469.
- [47] V. Thongboonkerd, S. Chutipongtanate, R. J. Kanlaya, *Proteome Res.* 5 (2006) 183-191.
- [48] L. Santucci, G. Candiano, M. Bruschi, M. Bodria. *J. Nephrol.* 26 (2013) 610-616.
- [49] José E. Araújo, Carlos Lodeiro, José L. Capelo, Benito Rodríguez-González, Alcindo A. dos Santos, Hugo M. Santos, Javier Fernández-Lodeiro. *Nano Research* (2015) DOI 10.1007/s12274-014-0599-4
- [50] M. D. Li, T. L. Cheng, W. L. Tseng, *Electrophoresis* 30 (2009) 388-395.
- [51] A. Ros, M. Faupel, H. Mees H., J. van Oostrum. *Proteomics* 2 (2002) 151-156.
- [52] W. K. Lee, J. S. Ko, H. M. Kim. *J. Colloid Interface Sci.* 246 (2002) 70-77.
- [53] R. Altman, E. Asch, D. Bloch, G. Bole, D. Borenstein, K. Brandt. *Arthritis Rheum.* 29 (1986) 1039-1049.

Tables

Table 1: Data of the samples used in this study.

	Urine samples		
	Control	Arthritis	Prosthesis
Patients	36	42	58
% women	44%	81%	72%
Age	65 ± 6	67 ± 8	69 ± 8
Age Range	54-85	55-84	47-86
Total protein content ($\mu\text{g}/\text{ml}$)			
Concentration	69 ± 26	51 ± 20	57 ± 27
Range	19-138	19-89	11-143
Creatinine normalization ($\mu\text{g}/\mu\text{mol}$)			
Concentration	0.4 ± 0.2	0.3 ± 0.2	0.4 ± 0.2
Range	0.03-0.1	0.03-0.8	0.03-0.8

Table 2: Experimental conditions tested: 3 mL of Midstream urine were used, No preservative was added. Samples were centrifuged at 10 000 g and then they were stored at -60 °C. For all cases samples showed a precipitate when unfrozen unless samples diluted 1/10 or 1/15. Therefore the optimum condition used in this work for urine preservation was dilution with milli-q water, 3 urine ml to 30 ml.

	Urine storage protocols assessed									
	Buffer				Ratio water/urine (ml)					
	HKUPP	NH ₃ 0.1%	AMBIC 25 mM	Tris-HCl pH=10.6	1	2	3	4	9	14
Time (min)	10	10	10	10	30	30	30	30	30	30
pH	8.0	8.0	8.0	8.0	-	-	-	-	-	-
Frozen volume (ml)	5.0	4.2	11.3	4.2	6	9	12	15	30	45

Table 3: Distribution of peaks with q-value <0.05 (Benjamini-Hochberg FDR). Red spots indicate the presence of the peak in the sample whilst green ones denote its absence.

Bi-class classification			
Peak/ m/z	q-value	Control	Patients
855.0927	0.000001
861.0986	0.001327
877.0724	0.019962
914.4523	0.022626
927.5024	0.034843
954.5033	0.047381
1153.5605	0.023435
1175.5819	0.019522
2298.2195	0.016049
3301.0154	0.000830

Figures & Figure captions

Figure 1

SDS-PAGE (12%) Urine's proteome comparison before and after one frozen-thawed cycle. Day 0: Starting point, a part of the 10 v/v % diluted sample was handled before freezing. Day 30: thawing. Protein content was concentrated in a Vivaspin® 4 tube previously to gel loading. x1 refers to the stored sample without this pre-concentration step; x60 refers to 30 ml of sample pre-concentrated to 500 microl; x90 refers to 30 ml of sample pre-concentrated to 330 microl; x120 refers to 30 ml of sample pre-concentrated to 250 microl. Equal volumes of each sample (20 μ l) were loaded for assessing the integrity of bands.

Figure 2

Urine was diluted as indicated in the protocol described in the sample treatment section. The final protein concentrations were as follows: 1.37 $\mu\text{g/ml}$, 4.11 $\mu\text{g/ml}$, and 12.33 $\mu\text{g/ml}$. Each sample was treated by sextuplicate with 1 ml GNPs, 2 h incubation time, 37°C incubation temperature, pH 4.0. The harvested pellets (5 μL) were run in SDS-PAGE (12%).

SDS-PAGE gels (12%) of total protein content from urine. Urine was treated with 1mL GNPs at different pHs; 3.5, 4.5, 5.5, 6.5, 7.5, 8.5, 9.5. Incubation was performed either for 2 h, 37°C or for 2 h, 37°C followed by overnight agitation at room temperature. Pellet was harvested at 10 000 rpm, 4°C; resuspended in 10 µl ACN 50 v/v %/AMBIC 25 mM and directly loaded for electrophoresis.

Figure 4

Spectra of a control (red) and a knee osteoarthritis patient (blue). Some characteristic m/z peaks are highlighted.