Supplementary information for:

Label-free analysis of gingival crevicular fluid (GCF) by surface enhanced Raman scattering (SERS)

Stefano Fornasaro,*[a] Federico Berton, [b] Claudio Stacchi, [b] Federica Farina, [b] Alessandro Esposito, [a] Valter Sergo, [a,c] Roberto Di Lenarda, [b] and Alois Bonifacio. [a]

[a.] *Raman Spectroscopy Lab, Department of Engineering and Architecture, University of Trieste, 34100 Trieste, Italy.*
[b.] *Department of Medical, Surgical and Health Sciences, University of Trieste, 34100 Trieste, Italy.*
[c.] *Faculty of Health Sciences, University of Macau, SAR Macau, China.*

* corresponding author sfornasaro@units.it

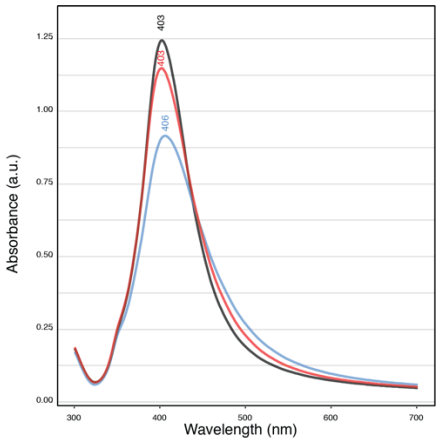

Figure S1 Representative UV-Vis absorption spectra of three different batches of Ag nanoparticles respectively synthesized by a modified Lee-Meisel (Lee and Meisel, J. Phys. Chem. 1982, 86 (17), 3391-3395) method. Spectra were chosen to represent the spectral variation between batches, which in some cases was indicative of the presence of some nanoparticle aggregates. The extinction maxima were 403-406 nm, which, according to literature (see for instance Steinigeweg and Schluecker, Chem.Comm. 2012, 48, 8682-8684; Wan et al. J. Colloid Interf. Sci. 2013, 394, 263-268, but there are many others which characterized the Lee-Meisel colloids), corresponds to nanoparticles having a variety of shapes and a broad size distribution (40-100 nm).

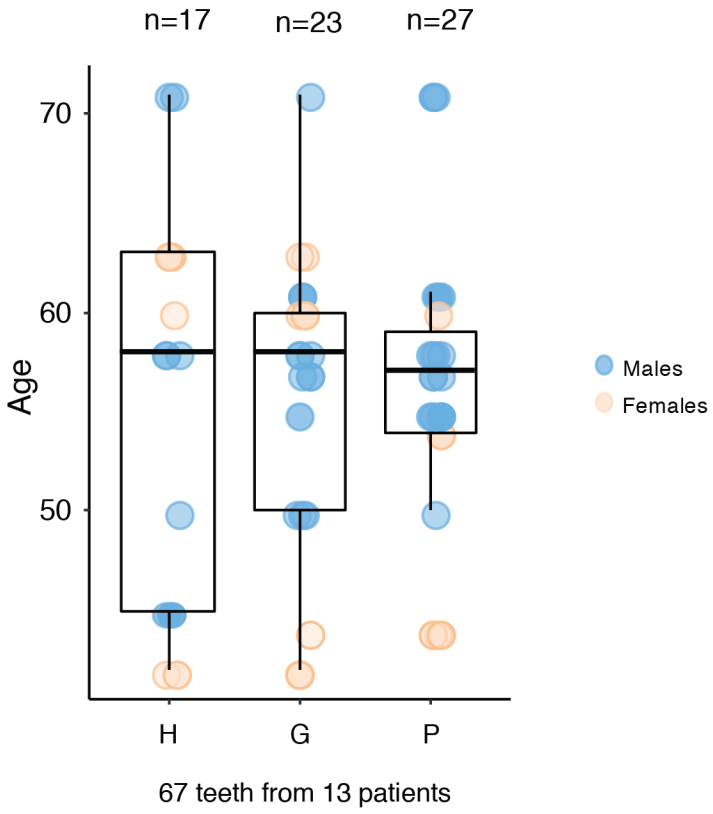

Figure S2 Age distribution

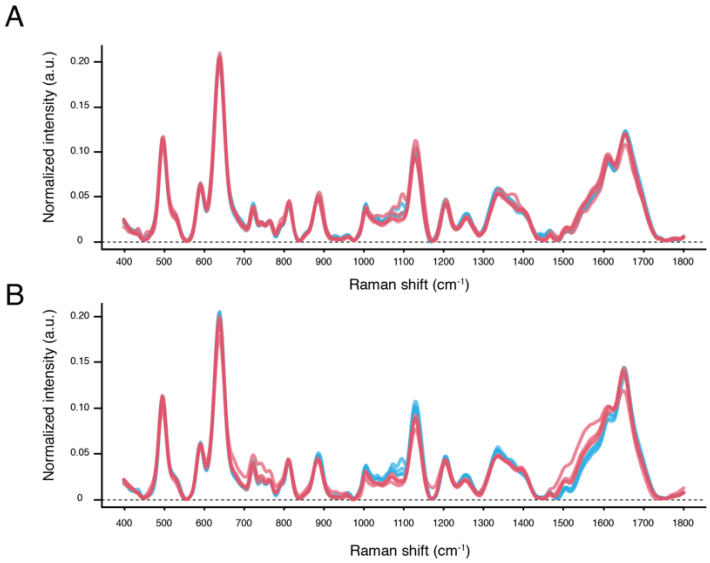

Figure S3 Stability of the SERS signal of human serum (ERM® certified Reference Material) on our substrate after one hour (A), and 48 hours (B). Four technical replicates from two different paper strips are shown overlaid. All spectra were collected with a 785 nm excitation.

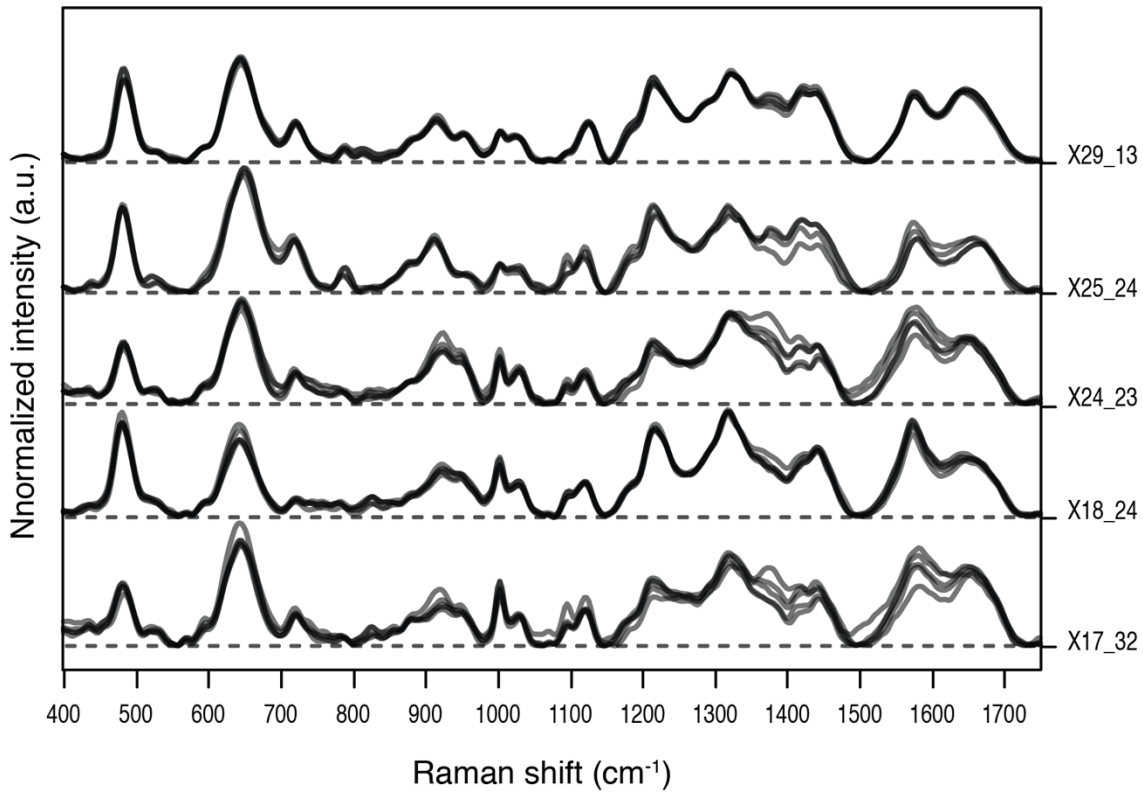

Figure S4 Intra-substrate variability. Five spectra (taken from different positions randomly selected on the substrates surface) from five different samples are shown overlaid. All spectra were collected with a 785 nm excitation. Each sample is labelled with a "patient_element" code.

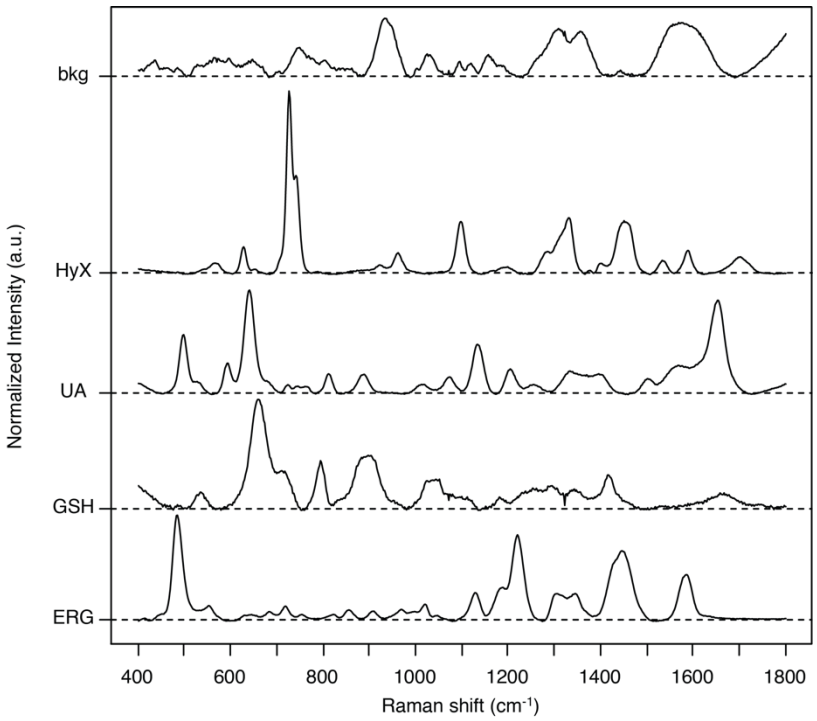

Figure S5 Normalized Surface-enhanced Raman scattering (SERS) spectra of reference compounds. ERG, ergothioneine; GSH, reduced glutathione; UA, uric acid; HyX, hypoxanthine; bkg, naked Periopaper strip covered with Ag NPs.

Supplementary methods

To evaluate the performance of the spectral fitting, the lack of fit (Lof) explained variance (R2) and relative fitting error (RFE) were calculated as follows:

$$LOF = \sqrt{\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(g_{ij} - \hat{g}_{ij})^2}{\sum_{i=1}^{I}\sum_{j=1}^{J} g_{ij}^2}} \times 100$$

$$R^2 = \left(1 - \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(g_{ij} - \hat{g}_{ij})^2}{\sum_{i=1}^{I}\sum_{j=1}^{J} g_{ij}^2}\right) \times 100$$

Where g_{ij} is the original elements of the matrix, with I rows and J columns, and $\hat{g}_{ij}$ is the elements of the matrix predicted by the fitting method. Although LOF and R^2 have the same interpretation, LOF is more sensitive to fitting differences.

Another common approach to quantify the reliability of a spectral fitting routine is to compare the norm of the fit to the norm of the input signal, a metric that has been termed the relative fitting error, RFE:

$$RFE = \frac{\|S - R\|}{\|S\|}$$

where S is the measured spectral data and R are the residuals (minimized by least-squares algorithm). || || indicates the norm of the vector. An RFE of 100% indicates a perfect fit and approaches zero as the fit degrades (i.e. the optimal fit lacks the ability to represent the input signal as a linearly weighted sum of the references supplied).

Table T1 Quality metrics to evaluate the performance of the spectral fitting

Quality metrics	value
LoF (%)	4.67 (1.75 – 8.73)
R^2 (%)	99.07 (98.52 – 99.60)
RFE (%)	97.10 (96.94 – 99.15)

LoF, Lack of Fit; R^2, explained variance; RFE, relative fitting error

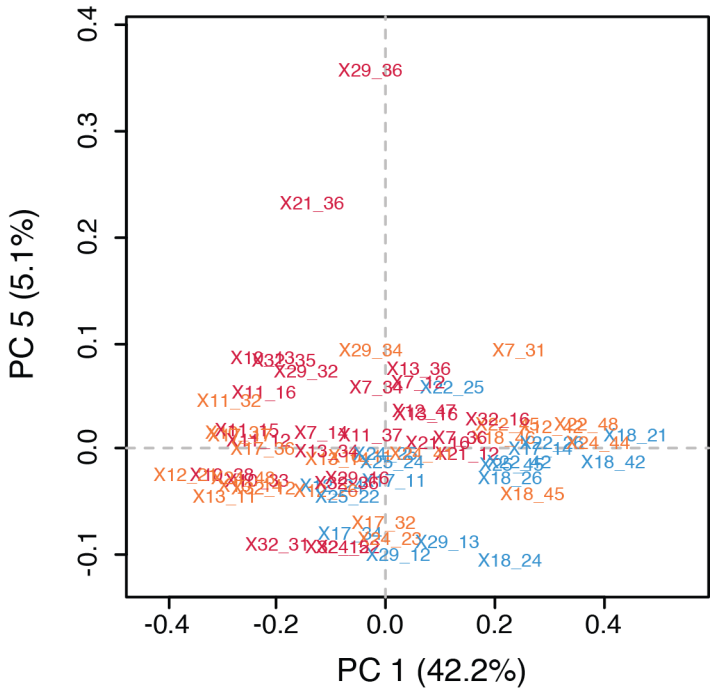

Figure S6 Score plot of the 1st principal component (PC1) against the 5th principal component (PC5). Samples (corresponding to single dental element) are depicted with different colors according to the diagnosis (healthy dental elements, H, blue; periodontitis, P, red; gingivitis, G, orange). The data variance (%) explained by each PC is shown between parentheses for both PC1 and PC5. Each sample is labelled with a "patient_element" code.

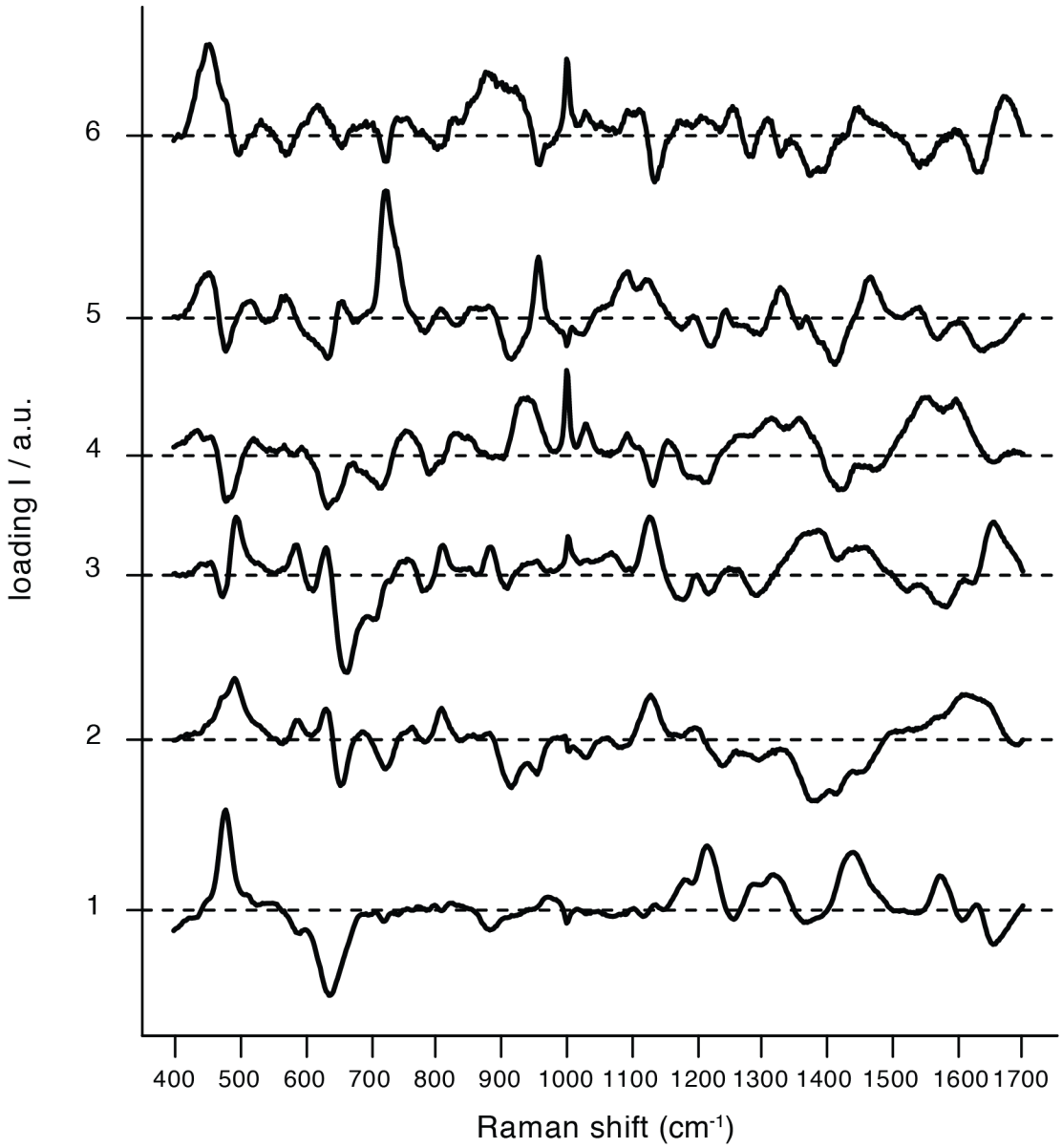

Figure S7 Loadings for first six principal components.

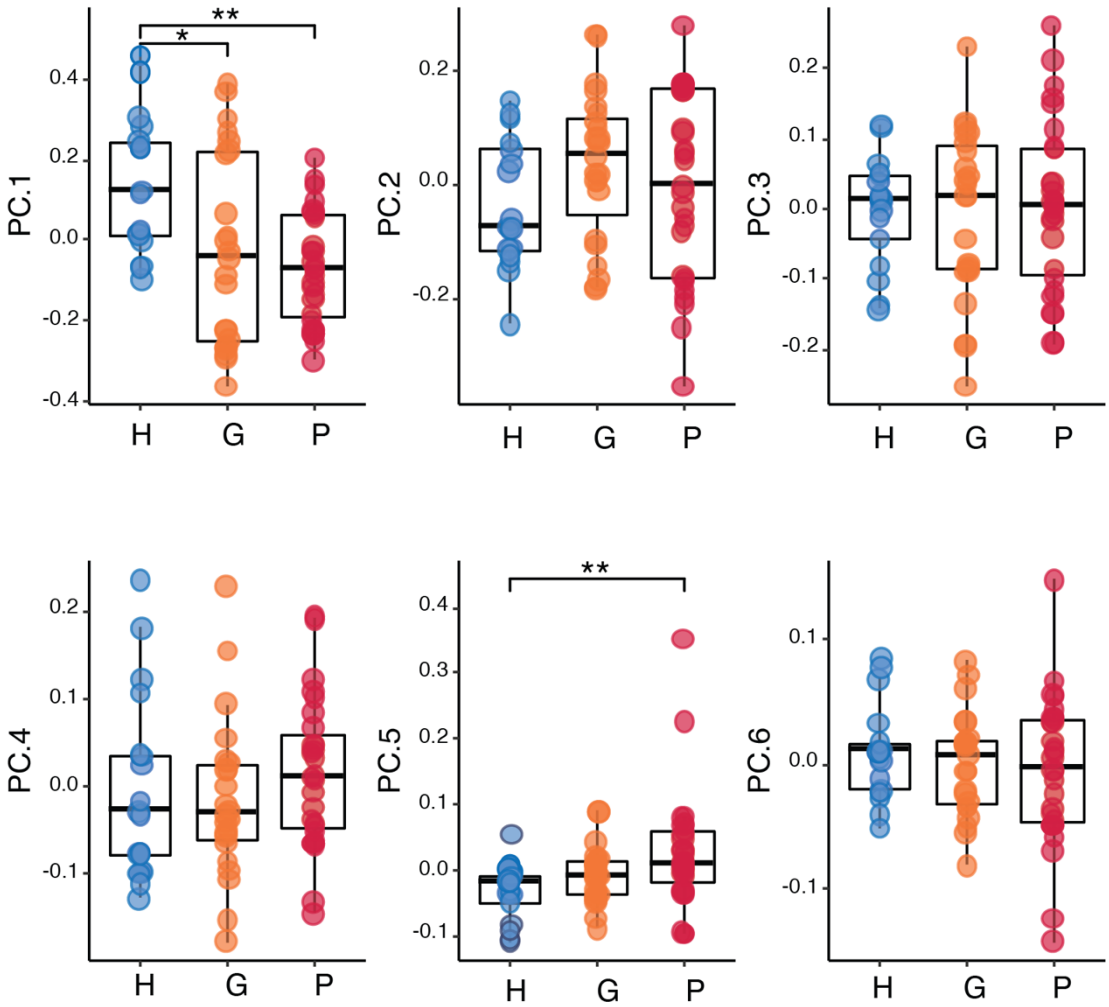

*Figure S8. Box plot for the scores the first six principal components. The black line represents the median; top and bottom edges of the box are the upper and lower quartiles, whiskers extend to upper and lower quartiles plus and minus 1.5× the IQR. Kruskal-Wallis test with Dunn–Bonferroni correction for multiple comparisons test (confidence level 95%, * = $p < 0.05$, **= $p < 0.01$) reveal a significant difference between the H and the other two groups.*