Electronic Supplementary Material (ESI) for Metallomics.

Supplementary data 1 – description to excel file

Sheet name: RAW-data- contains all CNAs computed by CHAMP pipeline for each sample.

Column A – sample ID

Column B – chromosome nr

Columns C and D – CNAs start and end

Column E – segment mean, measurement of the significance for given CNA to be deleted/amplified; for details see [1].

Sheet name: sorted0.33: list of CNAs meeting a significance level of "seg.mean" of 0.33 (for details see [1, 2]).

Column A – sample name

Column B – chromosome nr

Columns C and D – CNAs start and end

Column E – segment mean for "significant (above and below 0.33)" CNAs

Sheet name: CNA frequency in each sample and exposure data

Column A – sample name

Column B – CNA frequency

Column C – U-As, measurement of the arsenic in urine

Column D – samples classification to low-U-As=1 and high-U-As=2

Sheet name: Mostr_common_CNA_output– contains results of the overlapping of all CNAs from each sample (most common CNAs).

Column A, B and C – coordinates of the most common CNA,

Column D – number of samples with specific CNA,

Column E – length of the specific overlap common in all the sample harboring specific overlap

Column F – proportion of the sample where given overlap was positive – amplified.

Sheet name: FRQ_over_chromosomes– contains the frequency count of the CNA for each chromosome.

1. T. J. Morris, L. M. Butcher, A. Feber, A. E. Teschendorff, A. R. Chakravarthy, T. K. Wojdacz and S. Beck, *Bioinformatics*, 2014, 30, 428-430.
2. A. Feber, P. Guilhamon, M. Lechner, T. Fenton, G. A. Wilson, C. Thirlwell, T. J. Morris, A. M. Flanagan, A. E. Teschendorff, J. D. Kelly and S. Beck, *Genome biology*, 2014, 15, R30.