Publicacions més rellevants de la línia de recerca:
Associació CNVs i malaltia

Referència: Gonzalez JR, Subirana I, Escaramis G, Peraza S, Caceres A, Estivill X, Armengol L. Accounting for uncertainty when assessing association between copy number and disease: a latent class model *BMC Bioinformatics*, **10** (2009), pp. 172+.

Abstract: BACKGROUND: Copy number variations (CNVs) may play an important role in disease risk by altering dosage of genes and other regulatory elements, which may have functional and, ultimately, phenotypic consequences. Therefore, determining whether a CNV is associated or not with a given disease might be relevant in understanding the genesis and progression of human diseases. Current stage technology give CNV probe signal from which copy number status is inferred. Incorporating uncertainty of CNV calling in the statistical analysis is therefore a highly important aspect. In this paper, we present a framework for assessing association between CNVs and disease in case-control studies where uncertainty is taken into account. We also indicate how to use the model to analyze continuous traits and adjust for confounding covariates. RESULTS: Through simulation studies, we show that our method outperforms other simple methods based on inferring the underlying CNV and assessing association using regular tests that do not propagate call uncertainty. We apply the method to a real data set in a controlled MLPA experiment showing good results. The methodology is also extended to illustrate how to analyze aCGH data. CONCLUSION: We demonstrate that our method is robust and achieves maximal theoretical power since it accommodates uncertainty when copy number status are inferred. We have made R functions freely available.

Referència: Gonzalez JR, Carrasco JL, Armengol L, Villatoro S, Jover L, Yasui Y, Estivill X. Probe-specific mixed-model approach to detect copy number differences using multiplex ligation-dependent probe amplification (MLPA) *BMC Bioinformatics*, **9** (2008), pp. 261+.

Abstract: BACKGROUND: MLPA method is a potentially useful semi-quantitative method to detect copy number alterations in targeted regions. In this paper, we propose a method for the normalization procedure based on a non-linear mixed-model, as well as a new approach for determining the statistical significance of altered probes based on linear mixed-model. This method establishes a threshold by using different tolerance intervals that accommodates the specific random error variability observed in each test sample. RESULTS: Through simulation studies we have shown that our proposed method outperforms two existing methods that are based on simple

threshold rules or iterative regression. We have illustrated the method using a controlled MLPA assay in which targeted regions are variable in copy number in individuals suffering from different disorders such as Prader-Willi, DiGeorge or Autism showing the best performace. CONCLUSION: Using the proposed mixed-model, we are able to determine thresholds to decide whether a region is altered. These threholds are specific for each individual, incorporating experimental variability, resulting in improved sensitivity and specificity as the examples with real data have revealed.

Referència: Armengol L, Villatoro S, González JR, Pantano L, García-Aragonés M, Rabionet, R, Cáceres, M, Estivill, X. Identification of Copy Number Variants Defining Genomic Differences among Major Human Groups. *PLoS ONE*, **4(9)** (2009), pp. e7230+.

Abstract: BACKGROUND: Understanding the genetic contribution to phenotype variation of human groups is necessary to elucidate differences in disease predisposition and response to pharmaceutical treatments in different human populations. METHODOLOGY/PRINCIPAL FINDIGNS: We have investigated the genome-wide profile of structural variation on pooled samples from the three populations studied in the HapMap project by comparative genome hybridization (CGH) in different array platforms. We have identified and experimentally validated 33 genomic loci that show significant copy number differences from one population to the other. Interestingly, we found an enrichment of genes related to environment adaptation (immune response, lipid metabolism and extracellular space) within these regions and the study of expression data revealed that more than half of the copy number variants (CNVs) translate into gene-expression differences among populations, suggesting that they could have functional consequences. In addition, the identification of single nucleotide polymorphisms (SNPs) that are in linkage disequilibrium with the copy number alleles allowed us to detect evidences of population differentiation and recent selection at the nucleotide variation level. CONCLUSIONS: Overall, our results provide a comprehensive view of relevant copy number changes that might play a role in phenotypic differences among major human populations, and generate a list of interesting candidates for future studies.