1814T

panelcn. MOPS reaches clinical standards as a CNV detection tool for targeted panel sequencing data. *V. Haunschmid[1], G. Povysil[1], J. Vogt[2], K. Wimmer[2], G. Klambauer[1], S. Hochreiter[1].* 1) Institute of Bioinformatics, Johannes Kepler University Linz, Linz, Austria; 2) Division of Human Genetics, Medical University Innsbruck, Innsbruck, Austria.

Targeted panel sequencing is becoming increasingly important as a cost-effective strategy to identify disease-causing variants in clinical and research applications. While various copy number variation (CNV) detection methods exist for whole-genome and whole-exome sequencing data, highly accurate methods for panel sequencing data that are suitable for clinical purposes are still missing. The challenges with this kind of data are the small size and number of target regions as well as their uneven coverage. For clinical applications a method should furthermore be able to detect both short CNVs affecting only single exons or even just parts thereof as well as longer CNVs that affect multiple exons or even an entire gene. We present panelcn. MOPS for copy number detection which extends our previously developed method cn. MOPS to targeted panel sequencing data. The method is well suited for this type of data since it can estimate technical and biological characteristics influencing the read counts of each targeted region by a mixture of Poissons model. The design of the count windows, the read counting procedure, the parameters of the model and the segmentation algorithm have been optimized for targeted panel sequencing. cn. MOPS supplies integer copy numbers together with probabilities which inform users about the reliability of the copy number estimates. We have tested panelcn. MOPS on simulated and real sequencing data. On 240 simulated data sets, that resembled the characteristics of targeted panel sequencing data, panelcn. MOPS has reached an average accuracy of 99. 96%. The real sequencing data was enriched with the TruSight cancer panel that targets 94 cancer predisposition genes including NF1/2, BRCA1/2 and APC. panelcn. MOPS detected 100% of CNVs known from previous MLPA analyses without any false positives. The size of the CNVs ranged from an 80bp deletion starting in the intron and affecting only part of one exon over duplications of several exons to deletions of 350kb affecting the entire gene. These results show that CNVs in targeted panel sequencing data can accurately be predicted with panelcn. MOPS. Consequently additional biotechnologies to detect CNVs, such as MLPA, can be omitted in order to reduce time and costs.

1815F

***i*PsychCNV: A robust method for copy number variation detection on dried blood spots.** *J. H. Thygesen[1,2], M. Bertalan[1], S. Weinsheimer[1], W. Mazin[1], T. Sparsø[1], T. Werge[1].* 1) Mental Health Centre, Sct. Hans and, Roskilde, DK, Research Institute of Biological Psychiatry, Psychiatric Center Sct. Hans, Roskilde, Denmark; 2) Neuroscience of Mental Health Department, Division of Psychiatry, University Collage London.

Dried blood spot (DBS) has been collected in Denmark for over 30 years, offering uncountable possibilities for population genetics studies. However, DBS genomic analysis offers unusual challenges, which current methods for copy-number variation (CNV) detection are not designed to handle. Existing methods predict large number of false positive CNVs on DBS data, making association analysis infeasible. Here we describe a novel methodological approach, *i*PsychCNV, the first tool designed to predict and analyze CNVs from DBS genomic data obtained via Illumina SNP array. *i*PsychCNV outperforms the widely used algorithm PennCNV on three datasets with different source of genomic DNA: whole blood, DBS and mock data (simulating DBS). Direct comparison of matched whole blood vs. DBS data from four samples reveals that PennCNV and *i*PsychCNV have a ratio of 0. 75% and 45% of true positives, respectively. To evaluate the methods specificity and sensitivity, we generated mock data that simulates 800 different CNV combinations found in Infinium PsychArray BeadChip (Illumina) from DBS data. On mock data, PennCNV has poor performance with 0. 65 area under the ROC curve (AUC), when compared to *i*PsychCNV which has excellent performance of 0. 92 ROC AUC. Traditional CNV prediction methods perform poorly because they rely mostly on Log R ratio signal, which on DBS data can have three times higher standard deviation than observed from whole blood. *i*PsychCNV takes full advantage of B allele frequency distribution, whereby false positive CNVs are reduced. The *i*Psych project includes 80,000 DBS samples from five psychiatric diseases and controls; therefore program functions are designed to manage large datasets, like searching for CNV hotspots, which summarize genomic regions that are more relevant for a specific disease. A support vector machine (SVM) model can be constructed using selected variables from CNV hotspot regions, increasing true positive calls. *i*PsychCNV can improve existing CNV call from other programs by using B allele frequency, hotspots and SVM. *i*PsychCNV is a R package easily installed, implemented and efficiently executed using multiple cores. It performs well for different datasets, offering a robust alternative to the existing CNV prediction methods. *i*PsychCNV is publicly available on Github as an open source project: https://github. com/mbertalan/iPsychCNV.