
Christopher Schroeder1,+,*, Marc Sturm1,+, Andreas Dufke1, Ulrike Mau-Holzmann1, Thomas Eggermann2, Sven Poths1, Olaf Riess1 and Michael Bonin1
1Department of Medical Genetics, University of Tübingen, Calwerstr. 7, 72076 Tübingen, Germany
2 Institute of Human Genetics, RWTH Aachen, Pauwellsstr. 30, 52074 Aachen, Germany
† both authors contributed equally to this work

ABSTRACT
Summary: UPDtool is a computational tool for detection and classification of uniparental disomy (UPD) in trio SNP-microarray experiments. UPDs are rare events of chromosomal malsegregation and describe the condition of two homologous chromosomes or homologous chromosomal segments that were inherited from one parent. The occurrence of UPD can be of major clinical relevance. Though high-throughput molecular screening techniques are widely-used, detection of UPDs and especially the subclassification remains complex. We developed UPDtool to detect and classify UPDs from SNP-microarray data of parent-child-trios. The algorithm was tested using five positive controls including both iso- and heterodisomic segmental UPDs and 30 trios from the HapMap project as negative controls. With UPDtool we were able to correctly identify all occurrences of non-mosaic UPD within our positive controls, whereas no occurrence of UPD was found within our negative controls. In addition, the chromosomal breakage points could be determined more precisely than by microsatellite analysis. Our results were compared to both the gold standard, microsatellite analysis, and SNP trio, another program available for UPD-detection. UPDtool is platform-independent, light-weight and flexible. Because of its simple input format UPDtool may also be used with other high-throughput technologies (e.g. Next-Generation-Sequencing).
Availability and implementation: UPDtool executables, documentation and examples can be downloaded from http://www.uni-tuebingen.de/uni/thk/de/f-genomik-software.html.
Contact: christopher.schroeder@med.uni-tuebingen.de, marc.sturm@med.uni-tuebingen.de
Supplementary information: Supplementary data are available at Bioinformatics online.

1 INTRODUCTION
In a normal diploid karyotype (46,XX or 46,XY, respectively) each chromosome pair consists of one maternal and one paternal inherited homologue. Uniparental disomy (UPD) is characterized by the presence of a chromosome pair or homologous chromosomal segment that was inherited from solely one parent. UPDs can be divided into maternal and paternal by origin of the homologous chromosomes. Further subclassification is based on the information whether both homologous chromosomes originate from the same (isodisomy) or both (heterodisomy) chromosomes of one parent. Also mixtures of both subtypes (combined iso- and heterodisomy) are possible. The mechanisms leading to UPD are complex and were discussed in detail in a recent review (Yamazawa, et al., 2010). Generally, UPD can affect each chromosome and the incidence is estimated to be around 1:3500 of live births. Phenotypes associated with UPD, due to either imprinting-disturbance or unmasking recessive mutations, have been described for nearly all autosomes (Gardner, et al., 2011). The pathogenesis of UPD is determined by both epigenetic imprinting (e.g. Silver-Russell syndrome (chr. 7, maternal), Prader-Willi syndrome (chr. 15, maternal), Beckwith-Wiedemann syndrome (chr. 11, paternal) and Angelman-syndrome (chr 15, paternal)) as well as unmasking of autosomal-recessive diseases (e.g. isodisomy of chromosome 9 in Leigh syndrome (Tiranti, et al., 1999)). The clinical findings of two of our positive controls with UPD(7) were published previously (Eggermann, et al., 2008).

Microsatellite analysis as well as methylation specific tests as a genomewide screening tool for UPDs are laborious, expensive and imprecise due to the limited number of markers that are available per chromosome. In contrast, the widely applied high-throughput genotyping technologies such as whole-genome high-density SNP-microarrays or Next-Generation-Sequencing provide hundreds of thousands genotypes in one experiment. Due to homozygosity by descendants microsatellite analysis as well as UPD-detection based on genotypes may be of limited informative value in consanguineous families. Especially trio experiments are thought to render UPD-detection possible (Altug-Teber, et al., 2005; Bruce, et al., 2005). Genotype information from children and their parents allows detection of inheritance errors, also called Mendelian errors (ME). Causes for unexpected calls can be found in genotyping errors but also in UPD-regions. Contrary to UPD regions, genotyping errors are evenly distributed over all chromosomes and occur at a low rate, e.g. < 0.1 % in an Affymetrix 50K array (Saunders, et al., 2007). Yet, detection of UPD and especially heterodisomic
branches to maternal and paternal UPD. The detection is based on the fraction of SNPs that are identical to the mother's alleles.

---

2 METHOD AND RESULTS

We developed a novel algorithm for UPD detection which is briefly summarized in the following. An UPD stretch contains SNPs of same inheritance that can be both informative (i.e. MEs) and non-informative (i.e. inheritance of both parents cannot be excluded). Each ME is used as starting point for a putative UPD stretch and extended to both sides. This extension is stopped by the occurrence of SNPs of different inheritance. To mitigate the effect of random genotype detection errors adjacent stretches with the same inheritance mode are joined. Finally, all stretches with less than a given number of mendelian errors will be removed. The mode of inheritance (maternal, paternal) depends on the pattern of MEs found. For further subclassification all UPD stretches are split into stretches of isodisomy and heterodisomy by the fraction of homo-
zygous and heterozygous SNPs that are acquired using a sliding window approach. This additional information helps to detect loss of heterozygosity (LOH) that is indicative for an isodisomy, whereas a large fraction of genotypes that could only be inherited from one parent indicates a heterodisomy. UPDtool does not evaluate allelic ratios and is therefore not optimized for the detection of mosaicism. The whole workflow and an example of a combined hetero- and isodisomy are depicted in Figure 1. For further details on workflow and file formats see the documentation of this software package.

To test the performance of UPDtool we obtained microarray data (Genome-Wide Human SNP Array 6.0, Affymetrix) from 5 trios that were previously characterized by microsatellite analysis (s. Supplementary Table 1, positive controls) and 30 HapMap trios (negative controls). All positive controls were processed using the standard protocols given by the manufacturer followed by a basic analysis with GATK (Genome analysis toolkit v. 4.1.1, Affymetrix). Copy-number-variations (CNV) were ruled out prior to analysis, since CNVs might interfere with UPD-detection. For further analysis with UPDtool the genotypes were exported and converted to the input format using the UPDconverter tool.

We were able to detect all non-mosaic UPDs present within our positive controls (s. Supplementary Table 1). The breakpoints between hetero- and isodisomy could be determined more precisely than by microsatellite analysis. Furthermore we were able to confirm the rate of MEs found in the literature within our reference cohort (rate of MEs: 0.00036 per patient, SD +/- 0.00013). All MEs found within the reference cohort were distributed evenly over all chromosomes and were evenly inherited from both parents. The analysis of one trio took approximately 60 seconds. As an internal control we were able to safely detect previously identified deletions that may mimic isodisomic stretches down to 100 kb (~ 20 MEs). Detection of small indels is considered to be a measure for sensitivity representing the resolution of the algorithm used. The overall results of UPDtool are comparable to those of SNPtrio and microsatellite analysis (s. Supplementary Table 1). However, SNPtrio divided at least in our hands heterodisomic regions into multiple short stretches of non-specified and isodisomic UPD.

3 CONCLUSION

UPDtool was written in C# and we provide a platform-independent executable for windows and linux, both 32 and 64 bit. The only requirement of UPDtool is an installation of the .NET 4.0 or Mono framework. For generation of images R has to be installed and added to the PATH environment variable. A data converter for the widely used Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara) is provided. Additionally, UPDtool comes with a tool that enables batch analysis of larger cohorts.

ACKNOWLEDGEMENTS

We thank Ms. Hirtreiter (Affymetrix) for providing the affymetrix hapmap trios.

Conflict of Interest: none declared.

REFERENCES


