BIOINFORMATICS APPLICATIONS NOTE Vol. 29 no. 12 2013, pages 1562–1564
doi:10.1093/bioinformatics/btt174

Genome analysis

Advance Access publication April 4, 2013

UPDtool: a tool for detection of iso- and heterodisomy in parent–child trios using SNP microarrays

Christopher Schroeder[1,*,†], Marc Sturm[1,†], Andreas Dufke[1], Ulrike Mau-Holzmann[1], Thomas Eggermann[2], Sven Poths[1], Olaf Riess[1] and Michael Bonin[1]

[1]Department of Medical Genetics, University of Tübingen, Calwerstr. 7, 72076 Tübingen, Germany and [2]Institute of Human Genetics, RWTH Aachen, Pauwelsstr. 30, 52074 Aachen, Germany

Associate Editor: Martin Bishop

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 with both the gold standard, microsatellite analysis and SNPtrio, 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

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on November 25, 2012; revised on April 7, 2013; accepted on April 8, 2013

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 ~1:3500 of live births. Phenotypes associated with UPD, owing 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 owing 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 of genotypes in one experiment. Owing 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 UDP 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 not only 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 50 K array (Saunders *et al.*, 2007). Yet, detection of UPD and especially heterodisomic regions as well as further subclassification of UPD regions (e.g. paternal, maternal) is still not routinely carried out. To our knowledge, only SNPtrio, a webtool, is publicly available and implements trio information for UPD detection (Ting *et al.*, 2007). Among the reasons for the development of UPDtool

*To whom correspondence should be addressed.
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

were the need for a lightweight and flexible tool with simple input and output file format that might be used with different genotyping techniques (microarrays but also next-generation sequencing). UPDtool was implemented as a portable and platform-independent software solution for detection of both hetero- and isodisomic regions on tab-delimited SNP files.

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 MEs 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 homozygous and heterozygous SNPs that are acquired using a sliding window approach. This additional information helps to detect loss of heterozygosity 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 five trios that were previously characterized by microsatellite analysis (Supplementary Table S1, positive controls) and 30

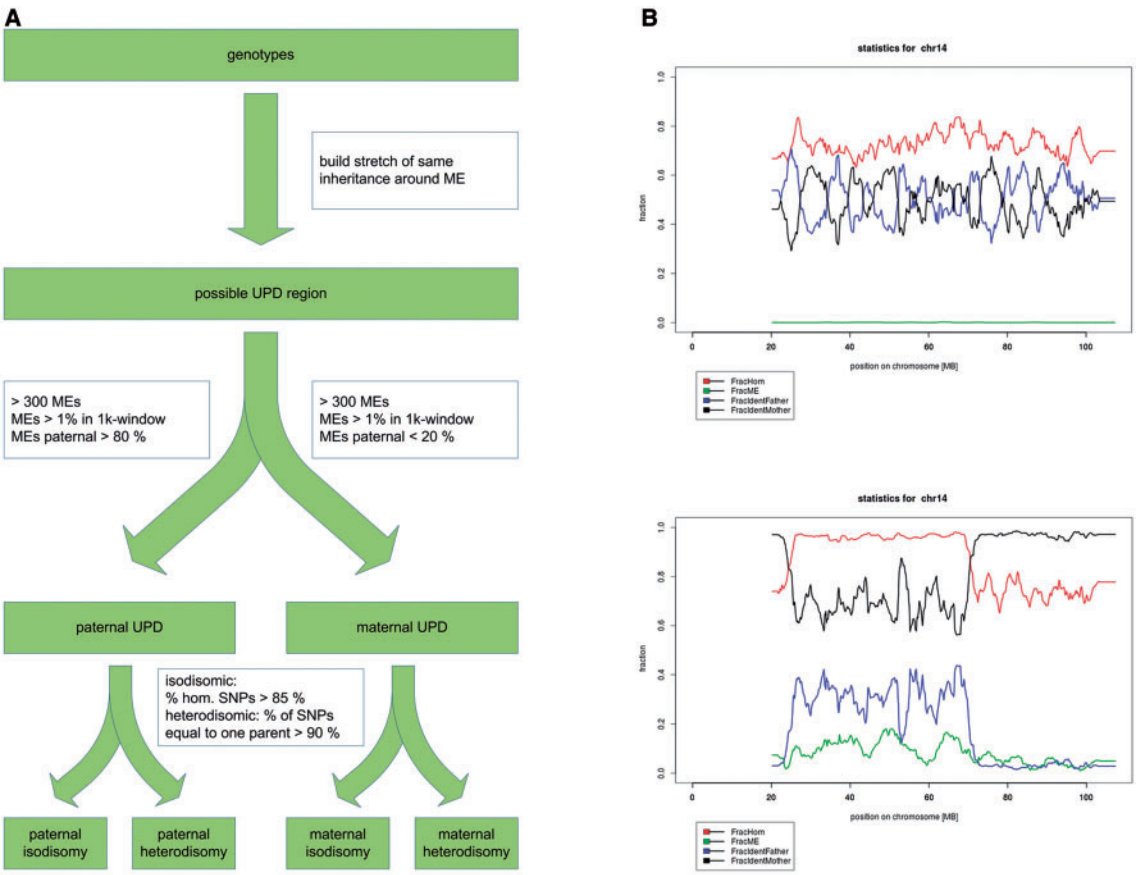

Fig. 1. Overview on the algorithm used by UPDtool to detect UPDs and the results of two exemplary samples for chromosome 14. (**A**) Part A shows the decision tree for detection and classification of UPDs. The given parameters are the standard parameters used by UPDtool. (**B**) Part B is divided into two subsections. The upper part shows the results of UPDtool for chromosome 14 in a control case, whereas the lower part illustrates some characteristic features of different UPD types. In our positive control, the fraction of MEs in a 1 k window is above 1% throughout chromosome 14. In addition, a high fraction of genotypes identical to the mother indicate a heterodisomy from 20 to 24 MB and from 71 to 106 MB. From 24 to 70 MB, a high fraction of homozygous genotypes points toward an isodisomy. Legend: FracHom = red line = fraction of genotypes in 1 k window that is homozygous, FracME = green line = fraction of MEs in a 1 k window, FracIdentFather = blue line = fraction of genotypes within a 1 k window where both alleles are identical to the fathers' alleles, FracIdentMother = black line = fraction of genotypes within a 1 k window where both alleles are identical to the mothers' alleles. All different lines start at 20.5 MB, as there are no genotypes available before this position on chromosome 14 owing to constitutional heterochromatin

HapMap trios (negative controls). All positive controls were processed using the standard protocols given by the manufacturer followed by a basic analysis with Genome analysis toolkit (v. 4.1.1, Affymetrix). Copy number variations (CNV) were ruled out before analysis, as 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 (Supplementary Table S1). 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 ~60 s. 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 with those of SNPtrio and microsatellite analysis (Supplementary Table S1). However, SNPtrio divided at least in our hands heterodisomic regions into multiple short stretches of non-specified and isodisomic UPD.

3 CONCLUSION

UPDtool is a fast, lightweight and platform-independent tool for detection and classification of UPDs. We were able to detect hetero- and isodisomies from maternal and paternal origin present in our positive controls. All UPD stretches were identified correctly, and breakpoints within a chromosome could be determined more precisely compared with microsatellite analysis. Owing to the simple input format, UPDtool may likely be used to analyze SNPs of other high-throughput platforms. Moreover, UPDtool is easy to install because it is platform independent and portable. UPDtool currently does not evaluate copy number data—for this reason, CNVs should be ruled out by another bioinformatic method before UPD detection.

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 Dr A. Hirtreiter (Affymetrix) for providing the affymetrix hapmap trios.

Conflict of Interest: none declared.

REFERENCES

Altug-Teber,O. *et al.* (2005) A rapid microarray based whole genome analysis for detection of uniparental disomy. *Hum. Mutat.*, **26**, 153–159.

Bruce,S. *et al.* (2005) Global analysis of uniparental disomy using high density genotyping arrays. *J. Med. Genet.*, **42**, 847–851.

Eggermann,T. *et al.* (2008) Segmental maternal UPD(7q) in Silver-Russell syndrome. *Clin. Genet.*, **74**, 486–489.

Gardner,R.J.M. *et al.* (2011) *Chromosome Abnormalities and Genetic Counseling*. Oxford University Press, USA.

Saunders,I.W. *et al.* (2007) Estimating genotyping error rates from Mendelian errors in SNP array genotypes and their impact on inference. *Genomics*, **90**, 291–296.

Ting,J.C. *et al.* (2007) Visualization of uniparental inheritance, Mendelian inconsistencies, deletions, and parent of origin effects in single nucleotide polymorphism trio data with SNPtrio. *Hum. Mutat.*, **28**, 1225–1235.

Tiranti,V. *et al.* (1999) Leigh syndrome transmitted by uniparental disomy of chromosome 9. *J. Med. Genet.*, **36**, 927–928.

Yamazawa,K. *et al.* (2010) Uniparental disomy and human disease: an overview. *Am. J. Med. Genet. C Semin. Med. Genet.*, **154 C**, 329–334.