Global analysis of uniparental disomy using high density genotyping arrays

S Bruce, R Leinonen, C M Lindgren, K Kivinen, K Dahlman-Wright, M Lipsanen-Nyman, K Hannula-Jouppi, J Kere

Background: Uniparental disomy (UPD), the inheritance of both copies of a chromosome from a single parent, has been identified as the cause for congenital disorders such as Silver-Russell, Prader-Willi, and Angelman syndromes. Detection of UPD has largely been performed through labour intensive screening of DNA from patients and their parents, using microsatellite markers.

Methods: We applied high density single nucleotide polymorphism (SNP) microarrays to diagnose whole chromosome and segmental UPD and to study the occurrence of continuous or interspersed heterodisomic and isodisomic regions in six patients with Silver-Russell syndrome patients who had maternal UPD for chromosome 7 (matUPD7).

Results: We have devised a new high precision and high-throughput computational method to confirm UPD and to localise segments where transitions of UPD status occur. Our method reliably confirmed and mapped the matUPD7 regions in all patients in our study.

Conclusion: Our results suggest that high density SNP arrays can be reliably used for rapid and efficient diagnosis of both segmental and whole chromosome UPD across the entire genome.

Uniparental disomy (UPD) occurs when a child receives both copies of a particular chromosome (or part of a chromosome) from only one parent, thus distorting the fundamental concept of biparental inheritance. UPD may involve two copies of the same chromosome (isodisomy) or one copy of the contributing parent’s pair of chromosomes (heterodisomy). Meiotic recombination events can result in a mixed UPD with interspersed regions of heterodisomy and isodisomy along the chromosome. UPD can result from the fertilisation of aneuploid gametes, with either gametic complementation or trisomic rescue as the mechanism. A non-disjunction event occurring during meiosis I or II is likely to be the original error in these instances. In addition, somatic events such as mitotic recombination (causing segmental UPDs) or duplication of a viable chromosome to compensate for an inherited dysfunctional chromosome can also result in UPD. Imprinting is defined as the bias in the expression of certain genes depending on the parent of origin, and is generally associated with parent specific methylation patterns of DNA established in the germline. If imprinted genes reside on the UPD chromosome, distortion of their expression patterns will occur and may cause specific phenotypes or syndromes, as is the case with patUPD1 and Beckwith-Wiedemann syndrome (MIM #130650), matUPD15 and Prader-Willi syndrome (MIM #176270), and patUPD15 and Angelman syndrome (MIM #105830). Furthermore, recessive, normally non-penetrant alleles in isodisomic regions may cause recessive diseases.

Maternal UPD of chromosome 7 (matUPD7) is found in around 10% of patients with Silver-Russell syndrome (SRS; MIM #180860), suggesting that imprinted genes play a role in its aetiology. SRS is a congenital syndrome manifested by intrauterine and postnatal growth retardation and typical dysmorphic features. A segmental matUPD reported for 7q31-qter and cytogenetic abnormalities for the 7p11-p14 region found in SRS patients delineate two potential regions for imprinted candidate genes for SRS. A recessive allele explaining the association between matUPD7 and SRS is unlikely, as a common isodisomic region in matUPD7s has not been found.

Detection of UPD has largely been performed through labour intensive screening with microsatellite markers of DNA from patients and their parents. Other means of identifying UPD are methylation specific PCR or bisulfite sequencing of known imprinted genes on the chromosome of interest. In this study, we set out to explore the new generation of hybridisation based single nucleotide polymorphism (SNP) genotyping methods, such as the Affymetrix GeneChip human mapping 10K array (HMA10K), as a simple tool for studying UPD. We present statistical means for diagnosing UPD, including a new multipoint mapping method that can be used to identify isodisomic and heterodisomic regions on the autosomal chromosomes of the genome. The regions are identified by calculating binomial probabilities with 95% confidence intervals (CI) against genomic background, under the assumption that only parts of the genotyped regions are affected by UPD. We specifically studied previously diagnosed SRS patients with matUPD7 and one patient with suspected matUPD7, and found that segments of isodisomy and heterodisomy can be mapped with great precision and certainty. We also provide further evidence for the involvement of imprinted genes in the SRS phenotype of matUPD7 patients.

Materials and Methods

Study subjects

Patients and/or parents provided written informed consent, after which blood samples were obtained. The study was approved by the ethics review board of the Hospital for Children and Adolescents, University of Helsinki, Finland.

Six parent–offspring trios were included in the study, of which five had previously been reported to have matUPD7 and one was a suspected case based on analysis of 12 chromosome 7 microsatellite markers, which had shown paternal alleles missing for eight markers (patient designated

Abbreviations: HMA10K, human mapping 10K array; SNP, single nucleotide polymorphism; SRS, Silver-Russell syndrome; UPD, uniparental disomy

See end of article for authors’ affiliations

Correspondence to: Dr Ju Kere, Department of Biosciences at Novum, CBT, 7th Floor, Karolinska Institutet, S-14157 Huddinge, Sweden; juha.kere@biosci.ki.se

Received 24 February 2005
Revised form received 25 April 2005
Accepted for publication 5 May 2005

Methods: We applied high density single nucleotide polymorphism (SNP) microarrays to diagnose whole chromosome and segmental UPD and to study the occurrence of continuous or interspersed heterodisomic and isodisomic regions in six patients with Silver-Russell syndrome patients who had maternal UPD for chromosome 7 (matUPD7).

Results: We have devised a new high precision and high-throughput computational method to confirm UPD and to localise segments where transitions of UPD status occur. Our method reliably confirmed and mapped the matUPD7 regions in all patients in our study.

Conclusion: Our results suggest that high density SNP arrays can be reliably used for rapid and efficient diagnosis of both segmental and whole chromosome UPD across the entire genome.

U

niparental disomy (UPD) occurs when a child receives both copies of a particular chromosome (or part of a chromosome) from only one parent, thus distorting the fundamental concept of biparental inheritance. UPD may involve two copies of the same chromosome (isodisomy) or one copy of the contributing parent’s pair of chromosomes (heterodisomy). Meiotic recombination events can result in a mixed UPD with interspersed regions of heterodisomy and isodisomy along the chromosome. UPD can result from the fertilisation of aneuploid gametes, with either gametic complementation or trisomic rescue as the mechanism. A non-disjunction event occurring during meiosis I or II is likely to be the original error in these instances. In addition, somatic events such as mitotic recombination (causing segmental UPDs) or duplication of a viable chromosome to compensate for an inherited dysfunctional chromosome can also result in UPD. Imprinting is defined as the bias in the expression of certain genes depending on the parent of origin, and is generally associated with parent specific methylation patterns of DNA established in the germline. If imprinted genes reside on the UPD chromosome, distortion of their expression patterns will occur and may cause specific phenotypes or syndromes, as is the case with patUPD1 and Beckwith-Wiedemann syndrome (MIM #130650), matUPD15 and Prader-Willi syndrome (MIM #176270), and patUPD15 and Angelman syndrome (MIM #105830).

Furthermore, recessive, normally non-penetrant alleles in isodisomic regions may cause recessive diseases.

Maternal UPD of chromosome 7 (matUPD7) is found in around 10% of patients with Silver-Russell syndrome (SRS; MIM #180860), suggesting that imprinted genes play a role in its aetiology. SRS is a congenital syndrome manifested by intrauterine and postnatal growth retardation and typical dysmorphic features. A segmental matUPD reported for 7q31-qter and cytogenetic abnormalities for the 7p11-p14 region found in SRS patients delineate two potential regions for imprinted candidate genes for SRS. A recessive allele explaining the association between matUPD7 and SRS is unlikely, as a common isodisomic region in matUPD7s has not been found.

Detection of UPD has largely been performed through labour intensive screening with microsatellite markers of DNA from patients and their parents. Other means of identifying UPD are methylation specific PCR or bisulfite sequencing of known imprinted genes on the chromosome of interest. In this study, we set out to explore the new generation of hybridisation based single nucleotide polymorphism (SNP) genotyping methods, such as the Affymetrix GeneChip human mapping 10K array (HMA10K), as a simple tool for studying UPD. We present statistical means for diagnosing UPD, including a new multipoint mapping method that can be used to identify isodisomic and heterodisomic regions on the autosomal chromosomes of the genome. The regions are identified by calculating binomial probabilities with 95% confidence intervals (CI) against genomic background, under the assumption that only parts of the genotyped regions are affected by UPD. We specifically studied previously diagnosed SRS patients with matUPD7 and one patient with suspected matUPD7, and found that segments of isodisomy and heterodisomy can be mapped with great precision and certainty. We also provide further evidence for the involvement of imprinted genes in the SRS phenotype of matUPD7 patients.

Materials and Methods

Study subjects

Patients and/or parents provided written informed consent, after which blood samples were obtained. The study was approved by the ethics review board of the Hospital for Children and Adolescents, University of Helsinki, Finland.

Six parent–offspring trios were included in the study, of which five had previously been reported to have matUPD7 and one was a suspected case based on analysis of 12 chromosome 7 microsatellite markers, which had shown paternal alleles missing for eight markers (patient designated

Abbreviations: HMA10K, human mapping 10K array; SNP, single nucleotide polymorphism; SRS, Silver-Russell syndrome; UPD, uniparental disomy

See end of article for authors’ affiliations

Correspondence to: Dr Ju Kere, Department of Biosciences at Novum, CBT, 7th Floor, Karolinska Institutet, S-14157 Huddinge, Sweden; juha.kere@biosci.ki.se

Received 24 February 2005
Revised form received 25 April 2005
Accepted for publication 5 May 2005

Methods: We applied high density single nucleotide polymorphism (SNP) microarrays to diagnose whole chromosome and segmental UPD and to study the occurrence of continuous or interspersed heterodisomic and isodisomic regions in six patients with Silver-Russell syndrome patients who had maternal UPD for chromosome 7 (matUPD7).

Results: We have devised a new high precision and high-throughput computational method to confirm UPD and to localise segments where transitions of UPD status occur. Our method reliably confirmed and mapped the matUPD7 regions in all patients in our study.

Conclusion: Our results suggest that high density SNP arrays can be reliably used for rapid and efficient diagnosis of both segmental and whole chromosome UPD across the entire genome.

U
matUPD7 VII) (table 1). This patient was born after induced labour for intrauterine growth retardation at 34 gestational weeks to a 30 year old mother and 31 year old father of average height. Her birth length was 39 cm (−3.7 SD) and weight 1425 g (−3.1 SD). During the pregnancy, slightly elevated alpha fetoprotein levels were followed up by amniocentesis, which gave a normal female karyotype. Otherwise, the pregnancy was uneventful. The child has two siblings with normal growth and development.

At 4 months of age, the patient was suspected to have SRS because of her severe growth retardation (−5.2 SD) and typical dysmorphic features, including leanness (weight in proportion to height −16%), relative macrocephaly (head circumference −2 SD), prominent ears, a large fontanelle, and clindactyly. As she grew older, she continued to show slight dysmorphic SRS-like features, and similar speech and circumflex

### Genotyping
We used the Affymetrix Mapping10K_Xba142 assay, which enables the simultaneous genotyping of over 10,000 SNPs using only 250 ng of genomic DNA (Affymetrix Inc., Santa Clara, CA, USA). The genotyping was carried out according to the GenChip mapping 10K assay protocol, using the column cleanup option, available on the Affymetrix website (www. affymetrix.com). The chips were scanned with the GeneChip Scanner 3000 (Affymetrix Inc.) and genotypes were analysed using GDAS software version 3.0 (Affymetrix Inc.).

### Data analysis
The genotype calls for the matUPD7 trios were analysed with PedCheck 1.1 to identify inconsistencies of mendelian inheritance. To explore regions of isodisomy and heterodisomy on the autosomal chromosomes of the genome was applied to all parent offspring trios. Isodisomy was defined as long stretches of homozygous genotypes $g_i$, and heterodisomy was defined as long stretches of identical genotypes $g_i$ between child and parent. Only markers with a successful genotype for all individuals of the trio were included in the analysis. In our test, the trial was successful if an SNP had the genotype $g_i$ or $g_0$ with the constant genomic probabilities of success $f_0$ or $f_i$, respectively. We required all trials in the heterodisomic and isodisomic regions to be successful and the binomial probabilities could be expressed simply as $f_0^n$ and $f_i^n$. P values were Bonferroni corrected by dividing $f_0^n$ and $f_i^n$ by the total number of putatively isodisomic and heterodisomic regions, respectively. The 95% confidence interval was defined by excluding $c/2$ SNPs from both ends of the UPD regions, where $c$ satisfies the condition $f_0^n<0.05$ or $f_i^n < 0.05$. We used the statistics software R for binomial probability calculation and for the creation of the graphs.

### Power estimation
We estimated the power of our method by calculating the minimum number of SNPs for the UPD regions to be detected using HMA10K with 95% power at $p=0.05$. Firstly, we calculated the minimum number of mendelian errors ($c$) in chromosomes without uniparental disomy, for the cumulative binomial probability of observing at least $c$ mendelian errors to be $<0.05$. This was performed using the mendelian error frequency $f_0$ calculated from the number of mendelian errors in chromosomes 1–6 and 8–22 in the matUPD7 I, II, III, IV, and VII trios. Secondly, we calculated the minimum number of SNPs in the UPD regions to give rise to at least $c$ mendelian errors with 95% probability, using the mendelian error frequency $f_0$ calculated from the number of mendelian errors in chromosome 7 in the matUPD7 I, II, III, IV, and VII trios. The minimum genetic length of the UPD regions was calculated by multiplying the minimum number of SNPs minus 1 by the average distance between the SNPs on each chromosome. As the significance of a UPD region is dependent on the number of SNPs it contains, an increased marker density is likely to improve the power of detection substantially. Therefore, the same estimation was performed for the reported average distance between the SNPs on the HMA10K array.

### Table 1 Patients included in the study and the intervals of recombination on chromosome 7

<table>
<thead>
<tr>
<th>SRS patient*</th>
<th>UPD†</th>
<th>Sex</th>
<th>SNP Location (bp)</th>
<th>Recomb region begins</th>
<th>Recomb region ends</th>
<th>Region length (Mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>matUPD7 I</td>
<td>isodisomy</td>
<td>M</td>
<td>rs718939 41639971</td>
<td>n1006106 42903835</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>matUPD7 II</td>
<td>iso/hetero</td>
<td>M</td>
<td>rs1074977 138899924</td>
<td>n11181740 14053098</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>matUPD7 III</td>
<td>iso/hetero</td>
<td>F</td>
<td>rs721317 23093041</td>
<td>n723699 27371748</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>matUPD7 IV</td>
<td>iso/hetero</td>
<td>F</td>
<td>rs919198 97373871</td>
<td>n20171799 102815675</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>matUPD7 VII</td>
<td>isodisomy</td>
<td>F</td>
<td>rs281900 144571391</td>
<td>n1406288 146506171</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>matUPD7q31-</td>
<td>iso/hetero</td>
<td>q-</td>
<td>rs727505 124016584</td>
<td>n1419702 124799900</td>
<td>0.8</td>
<td></td>
</tr>
</tbody>
</table>

Recomb, recombination, iso, isodisomy; hetero, heterodisomy. *Five of the patients have been previously reported: matUPD7 I, III, IV, and V in Hannula et al. and matUPD7q31-qter in Hannula et al. Patient matUPD7 VII was first reported in this study.

www.jmedgenet.com
RESULTS

Genotyping and mendelian inheritance analysis

We genotyped 18 individuals, including three parent-offspring trios of matUPD7 SRS cases with a suggested occurrence of interspersed heterodisomic and isodisomic regions (matUPD7 II, matUPD7 III, and matUPD7 IV), one trio with the patient having isodisomy covering the whole chromosome 7 (matUPD7 I), one trio with a matUPD7 case with reported segmental isodisomy for 7q31-qter, and one trio where matUPD7 was suspected (matUPD7 VII, unpublished) (table 1). The SNP genotype call rate and detection rate were 91.5% and 95.2%, respectively. The concordance rate of 99.9% (4 discordant genotypes out of 10204) was estimated by genotyping Affymetrix reference DNA, and comparing the generated genotype calls with the reference genotype calls downloaded from the Affymetrix website. The PedCheck analysis identified 461 mendelian errors (out of 166611 genotypes), of which 77% occurred on chromosome 7, strongly suggesting the presence of UPD. To test the significance of this finding, a p value was calculated for each patient as a cumulative binomial probability of seeing at least the observed number of mendelian errors on each autosomal chromosome by chance, under the assumption that all markers are equally likely to give rise to mendelian errors. A significant deviation from the expected number of mendelian errors with a p value < 10^-10 was observed on chromosome 7 for all matUPD7 I, matUPD7 II, matUPD7 III, matUPD7 IV, matUPD7 VII, and matUPD7q31-qter.

Multipoint UPD mapping

Our generated genotype data were subsequently analysed using a strategy to explore regions of isodisomy, defined as long stretches of homozygous genotypes, and heterodisomy, defined as long stretches of genotypes that are identical between child and parent. Using this approach, we could confirm the reported and suspected matUPD7 regions for all six tested patients with UPD, using highly significant, Bonferroni corrected p values and the number of coinciding mendelian errors in these regions (fig 1A). Furthermore, for the patients matUPD7 II, III, and IV we could delineate the isodisomic and heterodisomic regions along chromosome 7 with high confidence. No region of isodisomy common for all the matUPD7 cases studied at this high resolution could be identified, supporting the notion that imprinted gene(s) and not a recessive disease allele contribute to the SRS phenotype in these patients (fig 1A). Isodisomy in the centromeric region of matUPD7 II and IV suggested that the UPD had arisen from a meiosis II error, whereas heterodisomy in the centromeric region of matUPD7 III suggested a meiosis I error. Isodisomy throughout the whole chromosome 7 in matUPD7 I and VII indicated a postfertilisation rather than a meiotic error. Although there was no full chromosomal coverage by the 95% confidence interval regions, the predicted regions outside the confidence intervals still support UPD, but there is uncertainty in determining the exact recombination sites.

Three observations are critical for detailed interpretation of the results. Firstly, if the UPD region resides at the end of the chromosome, it is likely to extend all the way to the telomeric repeat. Secondly, in the case of overlapping isodisomic and heterodisomic regions, if the first region ends at marker m1 and the second region starts at marker m2 then the meiotic recombination must have occurred between markers m1-1 and m2+1.

Figure 1 Isodisomy and heterodisomy regions in matUPD7. (A) Location of the interspersed regions of isodisomy and heterodisomy on chromosome 7. Patients (from top to bottom): matUPD7 I, matUPD7 II, matUPD7 III, matUPD7 IV, matUPD7 VII, and matUPD7q31-qter. The Bonferroni corrected p values are shown above the 95% CI (boxes) for each region. The distributions of the mendelian errors on chromosome 7 are indicated for each individual, where each bar represents one error. The dashed vertical line indicates the centromeric position. (B) The delineation of recombination regions in overlapping regions of isodisomy and heterodisomy. The recombination is defined as occurring between m1-1 and m1+1, where m1 corresponds to the end of the isodisomic region and m2 corresponds to the start of the heterodisomic region. (C) The delineation of recombination regions in segmental UPDs, where m1 denotes the start of the isodisomic region and m2 denotes the most proximal mendelian error supporting the UPD region. The recombination is defined as occurring between m1-1 and m2+1.
mendelian error is candidate genes discussed by Hannula the NCBI Map Viewer human build 35.1. In particular, the www.jmedgenet.com frequency of mendelian errors the probability to detect UPD regions, given the frequency, was a putative heterodisomic region on chromosome 2 (rs1419702) and thus the somatic recombination could be mapped to a distinct region of 783 kb between the markers (rs1419702) and D7S686 (at position 117–125 Mb), further narrowing down the potential SRS candidate gene region to 7q31.33-qter. This region includes 419 genes, according to narrowing down the potential SRS candidate gene region to 7q31.33-qter. This region includes 419 genes, according to

<table>
<thead>
<tr>
<th>Chr</th>
<th>Chr. length (Mb)</th>
<th>No. of SNPs</th>
<th>Average distance between SNPs (Mb)</th>
<th>Minimum no. of SNPs in UPD region*</th>
<th>Minimum UPD region length† (Mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>245</td>
<td>977</td>
<td>0.32</td>
<td>63</td>
<td>19.76</td>
</tr>
<tr>
<td>2</td>
<td>242</td>
<td>962</td>
<td>0.29</td>
<td>63</td>
<td>17.87</td>
</tr>
<tr>
<td>3</td>
<td>199</td>
<td>813</td>
<td>0.28</td>
<td>51</td>
<td>14.18</td>
</tr>
<tr>
<td>4</td>
<td>191</td>
<td>816</td>
<td>0.27</td>
<td>51</td>
<td>13.73</td>
</tr>
<tr>
<td>5</td>
<td>181</td>
<td>780</td>
<td>0.26</td>
<td>51</td>
<td>12.99</td>
</tr>
<tr>
<td>6</td>
<td>171</td>
<td>793</td>
<td>0.25</td>
<td>51</td>
<td>12.30</td>
</tr>
<tr>
<td>7</td>
<td>159</td>
<td>585</td>
<td>0.31</td>
<td>51</td>
<td>15.45</td>
</tr>
<tr>
<td>8</td>
<td>146</td>
<td>556</td>
<td>0.30</td>
<td>51</td>
<td>15.05</td>
</tr>
<tr>
<td>9</td>
<td>138</td>
<td>544</td>
<td>0.28</td>
<td>51</td>
<td>13.98</td>
</tr>
<tr>
<td>10</td>
<td>135</td>
<td>610</td>
<td>0.25</td>
<td>51</td>
<td>12.67</td>
</tr>
<tr>
<td>11</td>
<td>134</td>
<td>644</td>
<td>0.25</td>
<td>51</td>
<td>12.26</td>
</tr>
<tr>
<td>12</td>
<td>132</td>
<td>545</td>
<td>0.27</td>
<td>51</td>
<td>13.71</td>
</tr>
<tr>
<td>13</td>
<td>114</td>
<td>492</td>
<td>0.26</td>
<td>51</td>
<td>13.08</td>
</tr>
<tr>
<td>14</td>
<td>106</td>
<td>401</td>
<td>0.30</td>
<td>51</td>
<td>15.02</td>
</tr>
<tr>
<td>15</td>
<td>100</td>
<td>335</td>
<td>0.36</td>
<td>38</td>
<td>13.14</td>
</tr>
<tr>
<td>16</td>
<td>88</td>
<td>259</td>
<td>0.39</td>
<td>38</td>
<td>14.39</td>
</tr>
<tr>
<td>17</td>
<td>78</td>
<td>188</td>
<td>0.46</td>
<td>38</td>
<td>17.03</td>
</tr>
<tr>
<td>18</td>
<td>76</td>
<td>346</td>
<td>0.25</td>
<td>38</td>
<td>9.10</td>
</tr>
<tr>
<td>19</td>
<td>63</td>
<td>98</td>
<td>0.43</td>
<td>38</td>
<td>15.75</td>
</tr>
<tr>
<td>20</td>
<td>62</td>
<td>222</td>
<td>0.30</td>
<td>38</td>
<td>11.19</td>
</tr>
<tr>
<td>21</td>
<td>47</td>
<td>197</td>
<td>0.28</td>
<td>38</td>
<td>10.49</td>
</tr>
<tr>
<td>22</td>
<td>49</td>
<td>82</td>
<td>0.66</td>
<td>38</td>
<td>24.44</td>
</tr>
</tbody>
</table>

Table 2 The minimum physical length of UPD regions detectable by HMA10K and HMA100K

Chr., chromosome; *95% power and 0.05 significance level.

The average meiotic recombination region length was 3.5 Mb and the variance 4.3 Mb. Thirdly, if a segmental UPD region starts at marker m1 and the most adjacent marker with a mendelian error is m2, then the recombination must occur in the region between markers m1− and m2 (fig 1C).

In the case of matUPD7q31-qter, the first marker of the UPD region coincided with the first mendelian error (rs1419702) and thus the somatic recombination could be mapped to a distinct region of 783 kb between the markers rs1419702 (m1−) and rs1419702 (m2) (at position 124–125 Mb on 7q31.33). Our estimate improves the precision 10 fold compared with the previously reported 8.7 Mb recombination region between microsatellite markers D7S633 and D7S686 (at position 117–125 Mb), further narrowing down the potential SRS candidate gene region to 7q31.33-qter. This region includes 419 genes, according to the NCBI Map Viewer human build 35.1. In particular, the candidate genes discussed by Hannula et al9 are still present within the segment, including the 7q32 imprinting cluster.

No evidence for UPD in other parts of the genome was found in our analysis, supporting the hypothesis of the matUPD7 as the cause of the SRS phenotype in these patients. The only instance where a genomic region turned up as a possible UPD, defined using the Bonferroni corrected p value, was a putative heterodisomic region on chromosome 7 (table 2). By using HMA100K, our method can detect UPD regions extending at least 15% of the chromosome length.

Estimating power of detection

We estimated the power of our method by calculating the probability to detect UPD regions, given the frequency of mendelian errors $f_e$ in the UPD regions, and $f_o$ in the non-UPD regions. Based on the number of mendelian errors in matUPD7 I, II, III, IV and VII trios, the $f_e$ and $f_o$ frequencies are 0.12 and 0.001, respectively. We calculated the minimum number of SNPs and the minimum physical length of UPD regions that can be detected in HMA10K and HMA100K with 95% power and at a significance level of 0.05 (table 2). For HMA10K, the minimum physical length varied from 9.1 Mb on chromosome 18 to 24.4 Mb on chromosome 22, indicating that HMA10K has sufficient power to detect whole chromosome UPD for all chromosomes. Notably, for the most relevant chromosomes to screen systematically for UPD (6, 7, 11, 14, and 15), our method can detect all UPD regions extending at least 15% of the chromosome length. For HMA100K, the minimum physical distance varied from 1.8 Mb on chromosome 21 to 4.7 Mb on chromosome 1 (table 2). By using HMA100K, our method can detect UPD regions on chromosomes 6, 7, 11, 14, and 15 if they extend at least 5% of the chromosome length.

**DISCUSSION**

In this study, we show that UPD can be statistically confirmed with SNP data from high density microarrays by making sensible assumptions about its multipoint genotypic appearance. This was exemplified by diagnosing both chromosome wide and segmental UPD, and by localising interspersed regions of heterodisomy and isodisomy in patients with matUPD7. The proposed UPD analysis approach has an advantage over the currently used methods such as microsatellite genotyping, methylation specific PCRs, and bisulphite sequencing, as it is rapid and reliable, and the genomewide result means it has wide application.

The significance of the UPD regions is determined by the number of consecutive SNPs meeting the UPD criteria. Consequently, shorter regions of UPD appear less significant than longer, and regions with a lower SNP density appear less significant than those with higher density. In this study, all UPD regions had very low p values, but if many recombinations have occurred or if the marker density is low, the UPD regions may lack significant Bonferroni corrected p values. This underlines the importance of using mendelian errors to detect chromosomes affected by UPD. If
non-significant isodisomic and heterodisomic regions are strongly supported by mendelian errors, or are adjacent to other UPD regions, they are excellent candidates to represent real UPD. Furthermore, our method is based on the assumption that all cells have UPD, which is not always the case. Beckwith-Wiedemann patients with high levels of mosaicism for patUPD11p would be more difficult to detect with the automated genotype calling of HMA10K. However, the discrepancy between cell rate and detection rate, which has been described as the consequence of mixing DNA, should, when it occurs in a specific chromosomal region, alert the investigator to consider the possibility of mosaic UPD.

It should be noted that the regions we define as isodisomic cannot be distinguished from putative deletions at the genotype type. Isolated isodisomic regions should therefore be confirmed by other methods such as karyotyping or signal intensity analyses. Recently, it has been reported that the HMA10K array can be used for molecular karyotyping based on signal intensity. Several of the congenital syndromes associated with UPD are similarly associated with chromosomal defects, such as deletions of the 15q11-q13 region in Prader-Willi and Angelman syndromes, and duplications of 7p11-p14 in SRS.2 This underlines the potential for using the HMA10K array when investigating developmental defects on a genomewide level, where both a copy number assessment on the hybridisation signal level and a UPD analysis on the genotype level can be performed using the same technique. Our power estimation suggests that this approach would be even more powerful using the new HMA100K array, enabling a very good coverage of the genome, and thus opening the possibility of discovering new functional aberrations associated with congenital syndromes.

Previous algorithms for multipoint UPD mapping were generated for microsatellite data from the chromosome of interest and are based on calculating genetic maps or calculating the probability distribution of recombination events based on parent-offspring genotype patterns. These algorithms have been implemented in studies of recombination patterns contributing to UPD in more general attempts to understand the mechanisms leading to non-disjunction, as has been done for matUPD15. We believe that our method will also serve well in such studies, as the transition between heterodisomy and isodisomy pinpoints the recombination loci with high accuracy. We strongly encourage using our proposed three step approach for exploring UPD: (a) considering the mendelian error statistics, (b) considering the result from the novel UPD algorithm and (c) if it is of relevance for the study, using the output from this algorithm to define the regions where recombination events have occurred.

CONCLUSION

We have shown that high density SNP marker arrays with whole genome coverage are useful for studying UPD in both clinical and biological settings.

ACKNOWLEDGEMENTS

We thank M Rönnholm at the Affymetrix core facility Novum, Karolinska Institutet for expert technical help. This work was supported by the Sigrid Jusélius Foundation, Academy of Finland, the Ulla & Gustaf of Ulgas Foundation, the Alex & Eva Wallström Foundation, and the Swedish Research Council. We are grateful to the Wallenberg Consortium North for supporting the Affymetrix core facility at NOVUM.

REFERENCES