Karyomapping: a universal method for genome wide analysis of genetic disease based on mapping crossovers between parental haplotypes

Alan H Handyside, Gary L Harton, Brian Mariani, et al.

J Med Genet 2010 47: 651-658 originally published online October 25, 2009
doi: 10.1136/jmg.2009.069971

Karyomapping: a universal method for genome wide analysis of genetic disease based on mapping crossovers between parental haplotypes

Alan H Handyside,[1,2] Gary L Harton,[3] Brian Mariani,[3] Alan R Thornhill,[1,4] Nabeel Affara,[5] Marie-Anne Shaw,[2] Darren K Griffin[4]

[1]London Bridge Fertility, Gynaecology and Genetics Centre, London, UK
[2]Faculty of Biological Sciences, University of Leeds, Leeds, UK
[3]Genetics & IVF Institute, Fairfax, Virginia, USA
[4]Department of Biosciences, University of Kent, Canterbury, UK
[5]Department of Pathology, University of Cambridge, Cambridge, UK

Correspondence to
Professor Alan H Handyside, London Bridge Fertility, Gynaecology and Genetics Centre, One St Thomas Street, London SE1 9RY, UK; ahandyside@thebridgecentre.co.uk

Received 1 June 2009
Accepted 29 September 2009
Published Online First 25 October 2009

ABSTRACT

The use of genome wide single nucleotide polymorphism (SNP) arrays for high resolution molecular cytogenetic analysis using a combination of quantitative and genotype analysis is well established. This study demonstrates that by Mendelian analysis of the SNP genotypes of the parents and a sibling or other appropriate family member to establish phase, it is possible to identify informative loci for each of the four parental haplotypes across each chromosome and map the inheritance of these haplotypes and the position of any crossovers in the proband. The resulting 'karyomap', unlike a karyotype, identifies the parental and grandparental origin of each chromosome and chromosome segment and is unique for every individual being defined by the independent segregation of parental chromosomes and the pattern of non-recombinant and recombinant chromosomes. Karyomapping, therefore, enables both genome wide linkage based analysis of inheritance and detection of chromosome imbalance where either both haplotypes from one parent are present (trisomy) or neither are present (monosomy/deletion). The study also demonstrates that karyomapping is possible at the single cell level following whole genome amplification and, without any prior patient or disease specific test development, provides a universal linkage based methodology for preimplantation genetic diagnosis readily available worldwide.

INTRODUCTION

The use of DNA microarrays and comparative genomic hybridisation (array CGH) or single nucleotide polymorphism (SNP) arrays is now well established for high resolution molecular cytogenetic analysis of duplications, deletions and other unbalanced abnormalities[1–3] and has revealed extensive copy number variation throughout the genome.[4] Both of these approaches rely primarily on quantitation at known loci distributed across the genome. SNP genotyping, however, has the advantage that general features of the proband's genotype, such as loss of heterozygosity, can be analysed in conjunction with quantitation to distinguish between, for example, monosomy and uniparental isodisomy. Here we demonstrate that by Mendelian analysis of the SNP genotypes of the parents and a sibling or other appropriate family member to establish phase, it is possible to identify informative loci for each of the four parental haplotypes across each chromosome and map the inheritance of these haplotypes and the position of any crossovers in the proband. The resulting 'karyomap', unlike a karyotype, identifies the parental and grandparental origin of each chromosome and chromosome segment and is unique for every individual being defined by the independent segregation of parental chromosomes and the pattern of non-recombinant and recombinant chromosomes. Karyomapping, therefore, enables genome wide linkage based analysis of inheritance of a broad range of genetic abnormalities, identified as pre-existing in one or both parents, from structural chromosomal abnormalities and copy number variants (CNVs) to single gene defects. In addition, karyomapping identifies trisomies of meiotic origin, by the presence of both haplotypes from one parent in one or more segments of the chromosome, and monosomies and deletions, by the absence of either haplotype from that parent (figure 1), exclusively on the basis of the genotype without any need for quantitative analysis.

For standard applications such as prenatal diagnosis, conventional quantitative and genotype analysis of SNP array data and karyomapping are not mutually exclusive. Together, they provide comprehensive information on the nature and parental origin of any cytogenetic abnormalities, which can be combined with linkage based analysis of any inherited condition if required. For preimplantation genetic diagnosis (PGD), which is limited to analysis of single or small numbers of cells biopsied from each embryo following in vitro fertilisation (IVF),[5] karyomapping is ideal. Efficient methods for whole genome amplification by multiple displacement amplification (MDA) are now available[6–8] and a single universal methodology with a broad range of applications, which is widely available, eliminates the need for labour intensive and costly test development. The only requirements are that DNA is available from both parents and an appropriate family member of known disease status and the position of the relevant gene or other abnormality is known.

To demonstrate the use of karyomapping for linkage based diagnosis of a single gene defect, we have analysed two families segregating mutant alleles of the cystic fibrosis transmembrane receptor (*CFTR*) causing cystic fibrosis (CF) (figure 2). In the first family, the karyomaps of five children confirmed the known CF status of four of them using one of the affected children for linkage. Furthermore, using MDA to amplify the whole genome of single cells from one of the children, we

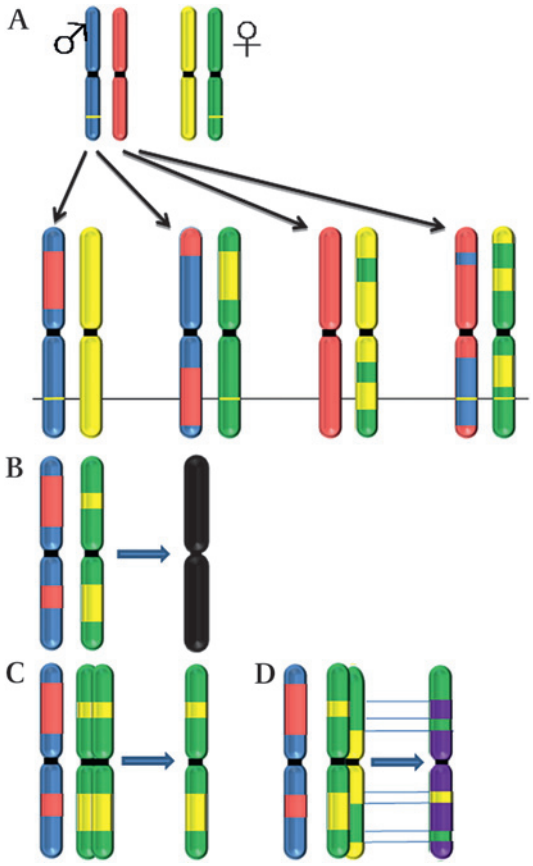

Figure 1 Karyomapping for genome wide analysis of chromosome aneuploidy and inheritance. (A) Linkage based diagnosis of inheritance of single gene defects. The two pairs of parental chromosomes (top row) are each colour coded to represent the two parental haplotypes inherited from the grandparents; the two paternal chromosomes in blue and red (left) and the two maternal chromosomes in yellow and green (right). Using Mendelian analysis of informative single nucleotide polymorphism (SNP) genotype combinations in the parents and establishing phase by reference to the genotype of one of the children, the parental haplotypes and the position of any crossovers for each paternal (left) and maternal (right) chromosomes inherited by the four children is identified and represented as a karyomap (second row). Linkage based diagnosis of the inheritance of, for example, an autosomal recessive single gene defect is then possible by comparing the parental haplotypes flanking the position of the gene (thin yellow band) in each child with those present in a child of known status. In this example, from left to right, unaffected carrier (paternal mutant allele), unaffected carrier (maternal allele), unaffected homozygote, affected homozygote. (B–D) Detection of chromosome aneuploidy. (B) Monosomy is detected by the absence of either of the haplotypes from one parent (black). (C) Trisomy caused by duplication of one chromosome or mitotic non-disjunction cannot be detected by karyomapping alone since the SNP genotype of both chromosomes is identical. (D) Trisomies resulting from the inheritance of two different products of meiosis are detected by one or more regions in which both haplotypes from one parent are present at closely adjacent informative SNP loci (purple). In this illustration, the presence of both maternal haplotypes in the pericentromeric region is consistent with a maternal meiosis I type non-disjunction error.

demonstrate that karyomapping is possible at the single cell level by limiting the analysis to only heterozygous informative SNP loci while retaining a sufficient density of loci for accurate linkage based testing. Similarly, in the second family, who had undergone IVF with PGD resulting in the birth of an unaffected child, the karyomaps of five embryos were consistent with the original diagnosis based on mutation detection alone, following

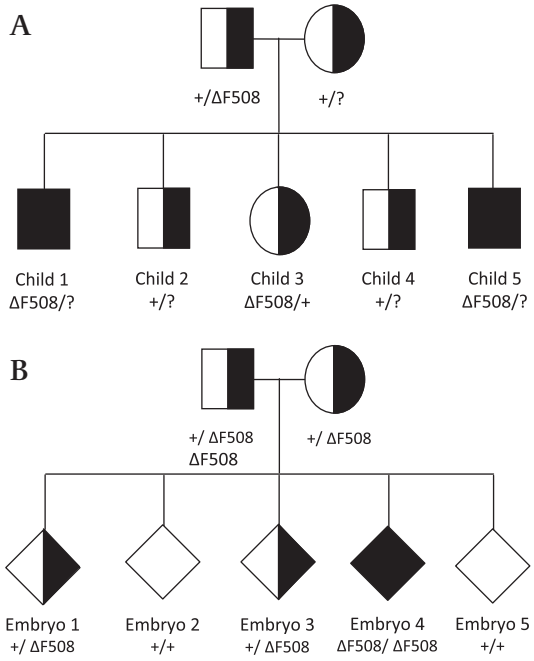

Figure 2 Pedigrees of the two families segregating mutations in *CFTR*. (A) Family 1. Both parents are carriers of mutations in the cystic fibrosis transmembrane receptor (*CFTR*) gene. The father is a carrier of the ΔF508 deletion and the mother is a carrier for an unknown minor mutation. Two of the five children are affected by cystic fibrosis and are compound heterozygotes for the two mutations. The other three (two sons and a daughter) are all carriers of one of the parental mutations as indicated. (B) Family 2. Both parents are carriers of the ΔF508 deletion and following preimplantation genetic diagnosis (PGD), five preimplantation embryos, which were not selected for transfer, were cryopreserved and subsequently consented for research. The PGD result based on analysis of one or two cells biopsied from each embryo is indicated.

MDA of either the whole embryo or small numbers of biopsied cells. In addition, a maternal aneuploidy was detected in two of the five embryos.

METHODS

Two families segregating mutant alleles of the *CFTR* causing CF were genotyped using high density, genome wide SNP arrays for karyomap analysis.

Family 1

The parents and five children of known CF status were genotyped using DNA extracted from lymphoblastoid cell lines (NIGMS Human Genetic Cell Repository, Coriell Institute, Camden, New Jersey, USA) by standard methods (figure 2A).

Family 2

In this family, both parents are known carriers of the ΔF508 deletion (p.Phe508) and have two homozygous CF affected children. Following IVF and cleavage stage biopsy of one or two single cells for PGD using standard methods,[9] two unaffected blastocyst stage embryos were transferred 5 days post-insemination resulting in the normal delivery of a healthy girl. All remaining embryos were cryopreserved and subsequently donated for research with the patients' informed consent. Three blastocyst stage and two arrested cleavage stage embryos of known CF status (figure 2B) were thawed for reanalysis by karyomapping. Three to 10 trophectoderm cells were biopsied from each blastocyst and the biopsies, one corresponding biopsied embryo and the two arrested embryos (each with

approximately two intact cells), were lysed and prepared for SNP genotyping together with parental DNA samples.

Whole genome amplification

Single lymphoblastoid cells and embryo samples were washed in phosphate buffered saline supplemented with 5% fetal bovine serum (Invitrogen-Gibco, Carlsbad, California, USA) and transferred in a minimal volume to 2.5 µl lysis buffer (0.2 N NaOH, 0.05 M DTT) in PCR tubes, covered with PCR grade oil (Sigma-Aldrich, St Louis, Missouri, USA), incubated at 65°C for 10 min, and then kept on ice before adding 2.5 µl neutralisation buffer. Isothermal MDA was then carried out according to manufacturer's instructions (Repli-G, Qiagen Inc, Valencia, California, USA) by adding 45 µl master mix and incubating at 30°C for 10 h, followed by 3 min at 65°C.

Genotyping

All DNA samples including MDA products were quantified by dsDNA staining (PicoGreen, Molecular Probes) and spectrofluorometry (Berthold Twinkle LB 970) and adjusted to 100 ng/µl in 15 µl. SNP genotyping was then carried out according to the manufacturer's instructions (HumanCNV370 Infinium-II Quad and Duo; Illumina, Inc, San Diego, CA, USA). Processed arrays were scanned and the image data analysed and converted to genotype data (BeadStudio Software Suite, v3.1, Illumina, Inc). The genotype data were finally exported as an Excel compatible file for karyomap analysis.

Karyomapping

A Visual Basic for Applications macro was developed to process and analyse the SNP genotype data (excluding the additional SNP loci used to quantify CNVs) and construct karyomaps in Microsoft Excel as follows:

1. The parental genotype combinations at each SNP locus are analysed to identify informative loci in which one parent is homozygous and the other heterozygous.
2. The genotype of one of the siblings is selected as a reference (for single gene defect diagnosis this will typically be an individual of known disease status) and used to establish phase at heterozygous informative loci, thereby defining the two parental haplotypes in both parents.
3. The genotypes of each of the other siblings are compared to the reference genotype at informative loci and the paternal and maternal haplotypes present at successive loci identified. (For karyomapping of single or small numbers of cells following whole genome amplification, only heterozygous informative SNPs were analysed to avoid errors caused by allele dropout (ADO).)
4. Karyomaps of the paternal and maternal copies of each chromosome pair at successive informative SNP loci are displayed as vertical columns of cells colour coded according to the parental haplotype present.
5. The position of crossovers in the paternal and maternal chromosomes of each of the siblings is compared and all of the karyomaps corrected for crossovers in the reference genotype identified because the crossover occurs at the same position in every other sibling. (In setting the phase of the parental haplotypes using, for example, one of the parents' children or embryos, any crossovers in this reference genotype artificially creates an apparent crossover in all siblings at that position.)

RESULTS

Family 1

SNP genotype call rates for the two parents and five children of known CF status (figure 2A) averaged 99.6% and the proportion of informative loci on chromosome 1, for example, was 41.5% with an average (SD) interval of 45.7±95.1 Kb (table 1). Using child 1 (CF affected) as a reference to establish phase and linkage, karyomaps of each of the other siblings were clear and unambiguous, and the parental haplotypes identified for each chromosomal segment were highly consistent with only a few (0.02%) isolated contradictory genotypes at informative loci (figure 3, table 1). Overall, there were an average 28±5.4 and 47.2±7.3 crossovers for the paternal and maternal chromosomes, respectively, and 75.2±9.4 in total for the region of the genome covered by the SNP loci analysed. Furthermore, the CF status of each of the children ascertained by comparing the parental haplotypes at the *CFTR* locus (including five paternal and 10 maternal intragenic informative SNP loci) with child 1, confirmed their known status (figures 2A and 3).

SNP genotyping call rates for five single cells from child 3, following whole genome amplification, averaged 92.7%. However, the incidence of contradictory genotypes at informative SNP loci increased significantly, for example, averaging 7.5% for chromosome 1 (table 1). To avoid errors caused by ADO, which is known to be prevalent after whole genome amplification of single cells,[9] we therefore restricted karyomap analysis to only those informative SNP loci which were genotyped as heterozygous in the sample being analysed. This restored the consistent pattern of parental haplotypes and well defined crossovers observed following analysis of genomic DNA from child 3, analysing all informative SNP loci. Also, the incidence of contradictory genotypes was considerably reduced (0.13%) (table 1). Restricting karyomap analysis to these heterozygous informative SNP loci reduced the number of loci that can be analysed and as a consequence increased the intervals between them. Nevertheless, average intervals between these loci ranged from 64.8–106.4 Kb for the autosomes, with only a small minority of intervals exceeding 1 Mb (data not shown).

Family 2

SNP genotype call rates for the two parents averaged 99.6% and 38.4% of SNP loci on chromosome 1 were informative.

Table 1 Distribution of informative loci and the incidence of contradictory calls on chromosome 1 (family 1)

Chromosome 1	Child 1–5		Single cells 1–5 (Child 3)	
	All informative SNP loci	Heterozygous informative SNP loci only	All informative SNP loci	Heterozygous informative SNP loci only
Average SNP call rate (%)	99.6		92.7	
Number of informative SNP loci (% total)	9712 (41.5)	4855 (20.8)	7736 (33.1)	2416 (10.3)
Average SNP interval ±SD (Kb)	45.7±95.1	91.0±192.8	57.4±114.1	182.6±317.8
Median interval (Kb)	17.1	25.0	20.9	56.6
Maximum interval (Mb)	2.35	3.2	2.95	3.90
Minimum interval (bp)	11	11	11	11
Average distribution of intervals between loci (%)				
<50 Kb	7551 (77.8)	3211 (66.2)	5539 (71.6)	1144 (47.5)
<1 Mb >50 Kb	2144 (22.1)	1579 (32.9)	2178 (28.2)	1189 (49.3)
>1 Mb	13 (0.13)	42 (0.9)	15 (0.2)	78 (3.25)
% of contradictory calls	0.02	0.02	7.5	0.13

SNP, single nucleotide polymorphism.

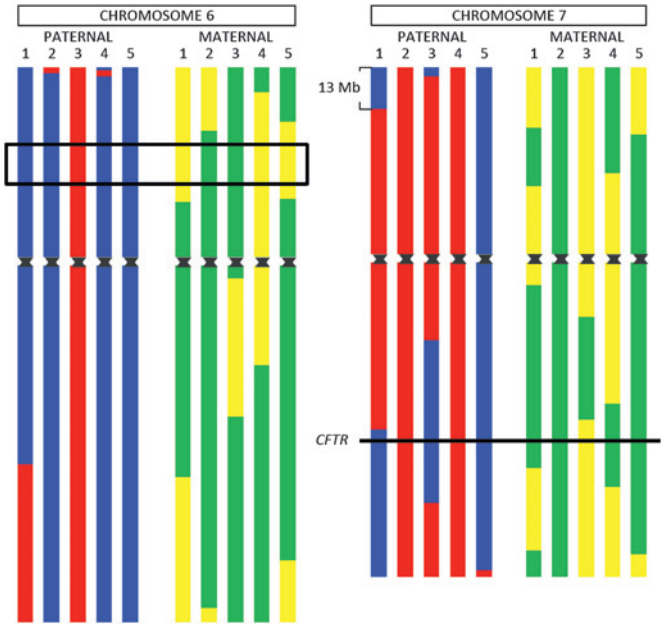

Figure 3 Karyomaps of chromosomes 6 and 7 for the five children in family 1. The consecutive series of informative single nucleotide polymorphism (SNP) loci across the length of each chromosome is represented in the karyomap by a continuous vertical column, coloured according to the parental haplotype identified. Where the SNP genotype indicates the same parental haplotype as that in the reference used to establish phase, the column is coloured blue or yellow, and where the opposite haplotype is present, the column is coloured red or green for the paternal and maternal chromosomes, respectively. The position of the centromere is indicated by the grey segment. Note that the length of the karyomap is determined by the number of informative SNPs on that chromosome and does not represent the physical distance between the SNP loci. The paternal and maternal chromosomes are grouped on the left and right, respectively, to facilitate comparison between siblings (child 1—5). Note that the grandparental origin of each chromosome can be determined by examining the parental haplotype in the pericentromeric region. The position of the cystic fibrosis transmembrane receptor (*CFTR*) located on the long arm of chromosome 7 (7q31.2) is indicated by a horizontal line. The extent of the human leucocyte antigen (HLA) region (approximately 24—35.7 Mb) tested for preimplantation genetic diagnosis (PGD) with HLA matching is indicated by the box on the chromosome 6 panel.

Following whole genome amplification, the call rates for biopsies from three blastocyst stage embryos (3—10 cells), one corresponding biopsied embryo and two arrested embryos in which only approximately two cells remained intact in each case, all of which had been cryopreserved subsequent to cleavage stage biopsy and PGD, were lower—averaging 92.8%. Using embryo 1, previously identified by PGD as an unaffected heterozygous carrier, as the reference genotype, the karyomaps for each of the other embryos clearly identified the parental haplotypes and positions of crossovers and consistent with the original PGD result in each case (figure 4). Furthermore, karyomapping identifies the origin of the mutant allele as maternal in both of the carriers (embryos 1 and 3). In this family, this is less definitive since both parents carry the same mutation (p.Phe508) and no linkage based testing was attempted originally.

The karyomaps of the five embryos also revealed two aneuploidies of maternal origin, a monosomy for chromosome 6 (biopsy 2) and a trisomy for chromosome 9 (biopsy 3) (figure 4). With the trisomy 9, the distribution of regions in which both maternal haplotypes were detected, on both arms distal from the centromere, is consistent with non-disjunction

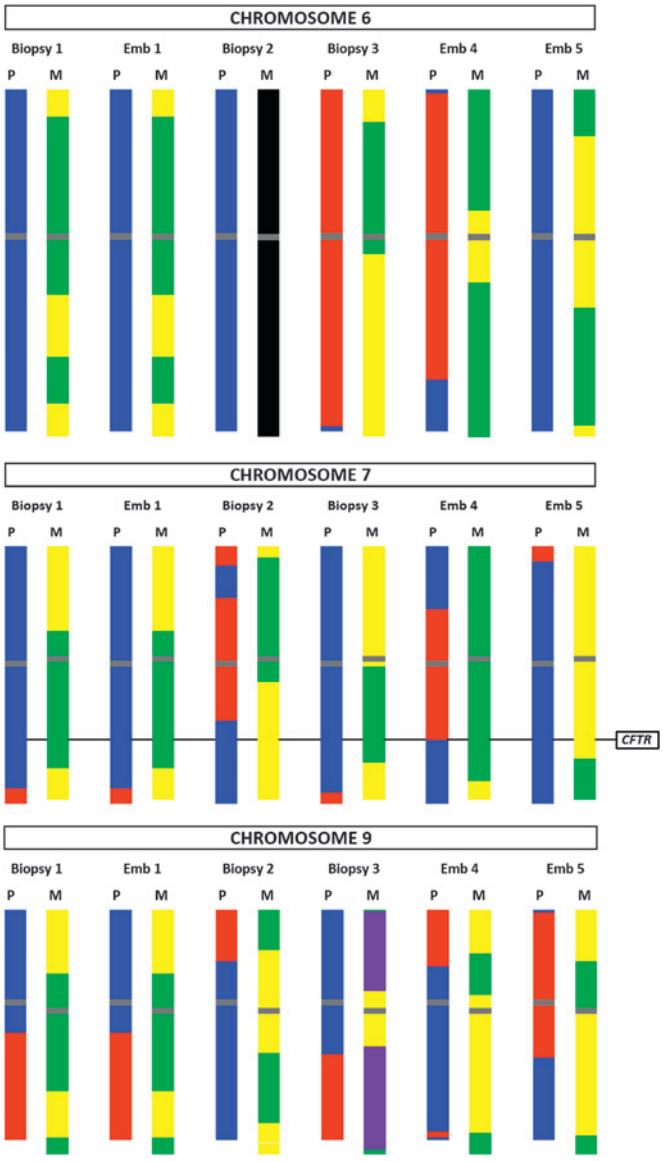

Figure 4 Karyomaps of chromosomes 6, 7 and 9 for embryos 1—5 (family 2). Karyomaps representing the parental haplotypes identified in the biopsied cells or embryos for the paternal (left) and maternal (right) chromosomes are paired in each case. Where the SNP genotype indicates the same parental haplotype as that in the reference used to establish phase, the column is coloured blue or yellow, and where the opposite haplotype is present, the column is coloured red or green for the paternal and maternal chromosomes, respectively. The position of the centromere is indicated by the grey segment. In addition, if both or neither parental haplotypes from one parent are identified, the column is coloured purple or black, respectively. The position of the cystic fibrosis transmembrane receptor (*CFTR*) located on the long arm of chromosome 7 (7q31.2) is indicated by a horizontal line.

in meiosis II, though other mechanisms are possible (see Discussion).

DISCUSSION

By Mendelian analysis of a series of closely spaced informative SNP loci across each chromosome within a family, karyomapping enables highly accurate, genome wide linkage based analysis of inheritance of a broad range of genetic abnormalities from structural chromosomal abnormalities and CNVs to single gene defects identified as pre-existing in one or both parents. All that is required is that the exact position of the genetic abnormality

is known. With genomic DNA (family 1), karyomapping unambiguously confirmed the known status of each of the four siblings for mutations in the *CFTR* causing CF, using one of the affected children to establish phase and linkage to the mutations in both parents. Also, crossovers between parental haplotypes were sharply defined (figure 3) and the number and distribution of crossovers in the five children from family 1 (table 2) was broadly consistent with previous studies.[10–12] There were a few isolated contradictory SNP genotype calls at informative loci (table 1). Most of these are likely to be genotyping errors in either the reference genotype or the proband, since a double recombination between closely spaced flanking SNP loci is extremely unlikely,[13] although rare mutations or gene conversion cannot be excluded in some cases. In principle, therefore, for single gene defects, there may be no need for direct mutation detection. In practice, however, many if not most families do not fulfil the criteria for linkage testing even after extensive investigation to understand the mode of inheritance and identify the causative mutation. For example, many families present singleton cases, so that one cannot easily determine whether or not it is a new mutation or an inherited case without mutation testing. In these families, karyomapping may only be useful as an aid to investigate possible linkage.

For PGD of single gene defects, following IVF and cleavage stage embryo biopsy, knowledge of the causative mutation and/or established linkage to one or more family members are essential prerequisites because of the relatively high incidence of errors with single cell genetic analysis. For example, ADO, in which one of the parental alleles randomly fails to amplify by PCR or following whole genome amplification by MDA, is a major potential source of errors.[7 14] Tests for single gene defects, therefore, typically include one or more closely spaced flanking markers in addition to direct mutation detection and informative, highly polymorphic short tandem repeat (STR) markers would normally be preferred to biallelic SNPs.[15] Indeed, karyomap analysis of single cells genotyped following MDA, at all informative SNP loci, resulted in an unacceptably high incidence of contradictory calls, presumably caused by ADO, which often made it impossible to define the position of crossovers accurately. By limiting analysis to only those informative loci that are heterozygous, this problem is virtually eliminated (table 1). Although the average interval between these heterozygous informative SNP loci is increased (64.8–106.4 Kb for the autosomes), with few exceptions, the spacing is more than adequate for accurate linkage based diagnosis with or without direct mutation detection. This was demonstrated by karyomap analysis of samples with small numbers of cells following MDA from a PGD cycle for CF (family 2), which was consistent with the original conventional single cell analysis (figure 4). Thus a single universal linkage based methodology providing genome wide linkage markers, which is widely available, eliminates the need for labour intensive and costly patient or disease specific test development.

Another feature of karyomapping is that analysis of informative loci across each chromosome not only provides genome wide linkage markers but also enables detection of chromosome imbalance, including identifying the parent of origin, where either both (trisomy) or no (monosomy or deletion) parental haplotypes from one parent are present, exclusively on the basis of the SNP genotype without any need for quantitative analysis (figure 1). Thus, any trisomy of meiotic origin, in which the whole of the chromosomes or one or more segments of the two chromosomes are derived from the two parental haplotypes, will be detected. Furthermore, the pericentromeric or distal position of these dual haplotype segments may identify their origin as resulting from non-disjunction in meiosis I or II, respectively, although other mechanisms have been proposed.[16] Without quantitation, however, it is not possible to detect duplication of chromosomes or smaller segments such as de novo CNVs since the genotype is identical. For example, uniparental isodisomy cannot be distinguished from monosomy by karyomapping alone. By contrast, whole chromosome monosomies and segmental or even smaller deletions should be detected with a resolution only limited by the spacing between informative loci.

This ability of karyomapping to combine detection of chromosome imbalance with genome wide linkage analysis is particularly important for PGD since there is a relatively high incidence of chromosome aneuploidy and other complex structural abnormalities of meiotic and postzygotic origin in human preimplantation embryos following IVF.[17 18] This was demonstrated here, where reanalysis by karyomapping, following a conventional PGD cycle for CF, revealed that two out of five embryos had maternal aneuploidies, one monosomy 6 and one trisomy 9 (meiosis II type), and would not have established a viable pregnancy if transferred (figure 4). Recently, several randomised clinical trials of the use of PGD and fluorescent in situ hybridisation (FISH) analysis of 5–7 chromosomes have shown that live birth rates are not increased in women of advanced maternal age in which the incidence of aneuploidy is significantly increased.[19] This may be partly explained by the limited number of chromosomes tested and errors caused by chromosomal mosaicism which arises through postzygotic malsegregation or loss of chromosomes,[20] so that the biopsied single cell is not representative of the remaining cells in the embryo. Amplification bias during whole genome amplification from single cells by either PCR or MDA methods has, until recently, made the use of DNA microarray based CGH for comprehensive aneuploidy testing unreliable.[21–23] However, several groups have now reported accurate detection of aneuploidies by conventional quantitative analysis of SNP array data.[24] Karyomapping is exclusively based on genotype analysis and has several advantages. These include genome wide analysis for all chromosomes, identifying the parent of origin, which may be clinically significant, and, for trisomies, providing direct evidence for the inheritance of two chromosomes from one parent even if only a single cell is analysed. In this event, although trisomy rescue is a theoretical possibility, there is a risk of uniparental disomy, and unlike monosomies, which are generally not viable, some trisomies result in viable but abnormal pregnancy.

A major application of karyomapping for PGD is for carriers of balanced structural chromosome abnormalities, mainly reciprocal and Robertsonian translocations, which can cause infertility or repeated miscarriage. Karyomapping should in these cases detect both the duplicated and deleted segments of any unbalanced combination in the embryo resulting from fertilisation with an unbalanced gamete. Furthermore, unlike FISH based methods with combinations of chromosome specific centromeric and telomeric probes,[25] karyomapping should also distinguish between normal and balanced embryos by identifying the parental haplotypes across the breakpoints. Another important application is for human leucocyte antigen (HLA) matching, with or without single gene defect testing if required, with the aim of isolating cord blood stem cells at birth for transplantation to an existing child with a serious blood related illness.[26 27] Here the provision of genome wide linkage markers for the HLA region on chromosome 6 as well as any other locus for single gene defect analysis is a significant advantage and

Table 2 The number and location of crossovers detected in the autosomes and X chromosomes of child 1–5 (family 1)

		Paternal								Maternal							
Chromosome		1	2	3	4	5	Average	±SD	Range	1	2	3	4	5	Average	±SD	Range
1	p	2	2	0	1	2	1.4	0.89	0–2	3	0	3	2	0	1.6	1.52	0–3
	q	2	1	1	1	1	1.2	0.45	1–2	3	2	2	2	2	2.2	0.45	2–3
	All	4	3	1	2	3	2.6	1.14	1–4	6	2	5	4	2	3.8	1.79	2–6
2	p	0	1	1	0	2	0.8	0.84	0–2	1	1	2	0	0	0.8	0.84	0–2
	q	0	0	1	4	0	1	1.73	0–4	1	0	1	2	1	1	0.71	0–2
	All	0	1	2	4	2	1.8	1.48	0–4	2	1	3	2	1	1.8	0.84	1–3
3	p	1	0	1	1	0	0.6	0.55	0–1	1	3	1	1	1	1.4	0.89	1–3
	q	1	0	1	0	1	0.6	0.55	0–1	1	1	2	2	1	1.4	0.55	1–2
	All	2	0	2	1	1	1.2	0.84	0–2	2	4	3	3	2	2.8	0.84	2–4
4	p	0	1	2	0	1	0.8	0.84	0–2	0	0	1	0	2	0.6	0.89	0–2
	q	0	1	2	1	1	1	0.71	0–2	3	2	2	0	0	1.4	1.34	0–3
	All	0	2	4	1	2	1.8	1.48	0–4	3	2	3	0	2	2	1.22	0–3
5	p	1	0	0	1	1	0.6	0.55	0–1	0	1	1	1	2	1	0.71	0–2
	q	1	1	1	0	1	0.8	0.45	0–1	1	2	1	2	2	1.6	0.55	1–2
	All	2	1	1	1	2	1.4	0.55	1–2	1	3	2	3	4	2.6	1.14	1–4
6	p	0	1	0	2	0	0.6	0.89	0–2	1	1	0	1	2	1	0.71	0–2
	q	1	0	0	0	0	0.2	0.45	0–1	1	1	2	1	1	1.2	0.45	1–2
	All	1	1	0	2	0	0.8	0.84	0–2	2	2	2	2	3	2.2	0.45	2–3
7	p	1	0	1	0	0	0.4	0.55	0–1	2	0	0	1	1	0.8	0.84	0–2
	q	1	0	2	0	1	0.8	0.84	0–2	3	0	2	2	1	1.6	1.14	0–3
	All	2	0	3	0	1	1.2	1.30	0–3	5	0	2	3	2	2.4	1.82	0–5
8	p	0	1	1	0	1	0.6	0.55	0–1	0	1	1	2	1	1	0.71	0–2
	q	1	0	0	2	1	0.8	0.84	0–2	2	0	0	1	2	1	1.00	0–2
	All	1	1	1	2	2	1.4	0.55	1–2	2	1	1	3	3	2	1.00	1–3
9	p	1	0	0	1	0	0.4	0.55	0–1	1	1	2	1	2	1.4	0.55	1–2
	q	1	0	1	0	1	0.6	0.55	0–1	1	2	2	1	2	1.6	0.55	1–2
	All	2	0	1	1	1	1	0.71	0–2	2	3	4	2	4	3	1.00	2–4
10	p	0	1	0	0	1	0.4	0.55	0–1	2	1	1	2	2	1.6	0.55	1–2
	q	0	0	0	1	0	0.2	0.45	0–1	0	2	2	3	2	1.8	1.10	0–3
	All	0	1	0	1	1	0.6	0.55	0–1	2	3	3	5	4	3.4	1.14	2–5
11	p	1	1	1	0	1	0.8	0.45	0–1	1	1	1	0	1	0.8	0.45	0–1
	q	0	0	1	1	1	0.6	0.55	0–1	1	1	1	0	1	0.8	0.45	0–1
	All	1	1	2	1	2	1.4	0.55	1–2	2	2	2	0	2	1.6	0.89	0–2
12	p	0	1	0	0	0	0.2	0.45	0–1	1	1	0	0	0	0.4	0.55	0–1
	q	1	1	2	2	1	1.4	0.55	1–2	2	1	2	3	0	1.6	1.14	0–3
	All	1	2	2	2	1	1.6	0.55	1–2	3	2	2	3	0	2	1.22	0–3
13	p	Not analysed								Not analysed							
	q	1	1	2	3	2	1.8	0.84	1–3	0	2	3	2	0	1.4	1.34	0–3
	All	1	1	2	3	2	1.8	0.84	1–3	0	2	3	2	0	1.4	1.34	0–3
14	p	Not analysed								Not analysed							
	q	1	1	1	2	1	1.2	0.45	1–2	2	1	2	1	2	1.6	0.55	1–2
	All	1	1	1	2	1	1.2	0.45	1–2	2	1	2	1	2	1.6	0.55	1–2
15	p	Not analysed								Not analysed							
	q	1	1	1	2	2	1.4	0.55	1–2	1	2	2	1	0	1.2	0.84	0–2
	All	1	1	1	2	2	1.4	0.55	1–2	1	2	2	1	0	1.2	0.84	0–2
16	p	0	1	0	1	0	0.4	0.55	0–1	0	0	2	1	2	1	1.00	0–2
	q	0	1	0	0	0	0.2	0.45	0–1	1	1	1	2	1	1.2	0.45	1–2
	All	0	2	0	1	0	0.6	0.89	0–2	1	1	3	3	3	2.2	1.10	1–3
17	p	1	0	0	0	0	0.2	0.45	0–1	2	1	1	0	1	1	0.71	0–2
	q	0	0	1	3	2	1.2	1.30	0–3	2	1	2	2	1	1.6	0.55	1–2
	All	1	0	1	3	2	1.4	1.14	0–3	4	2	3	2	2	2.6	0.89	2–4
18	p	0	1	0	1	1	0.6	0.55	0–1	1	0	0	0	1	0.4	0.55	0–1
	q	0	0	1	0	0	0.2	0.45	0–1	1	1	2	1	1	1.2	0.45	1–2
	All	0	1	1	1	1	0.8	0.45	0–1	2	1	2	1	2	1.6	0.55	1–2
19	p	0	1	0	1	1	0.6	0.55	0–1	0	1	0	0	0	0.2	0.45	0–1
	q	1	1	1	1	1	1	0.00	1	1	1	1	1	0	0.8	0.45	0–1
	All	1	2	1	2	2	1.6	0.55	1–2	1	2	1	1	0	1	0.71	0–2
20	p	0	0	0	1	0	0.2	0.45	0–1	1	0	1	1	0	0.6	0.55	0–1
	q	1	1	0	0	0	0.4	0.55	0–1	1	1	2	0	0	0.8	0.84	0–2
	All	1	1	0	1	0	0.6	0.55	0–1	2	1	3	1	0	1.4	1.14	0–3

Continued

Table 2 Continued

Chromosome		Paternal 1	2	3	4	5	Average	±SD	Range	Maternal 1	2	3	4	5	Average	±SD	Range
21	p	0	0	0	0	0	0	0.00		0	0	0	0	0	0	0.00	
	q	1	0	1	1	0	0.6	0.55	0–1	0	0	2	2	0	0.8	1.10	0–2
	All	1	0	1	1	0	0.6	0.55	0–1	0	0	2	2	0	0.8	1.10	0–2
22	p	Not analysed								Not analysed							
	q	0	1	1	2	2	1.2	0.84	0–2	1	0	1	1	1	0.8	0.45	0–1
	All	0	1	1	2	2	1.2	0.84	0–2	1	0	1	1	1	0.8	0.45	0–1
X (non- PAR)	p			0			0	0	0	2	1	1	1	1	1.2	0.45	0–2
	q			0			0	0	0	0	2	4	0	3	1.8	1.79	0–4
	All			0			0	0	0	2	3	5	1	4	3	1.58	1–5
Total		23	23	28	36	30	**28**	5.43	23–36	48	40	59	46	43	**47.2**	7.26	40–59

avoids the need for developing a patient specific test with multiple markers in both these regions. Using family 1 as an example, in the 11.7 Mb interval between D6S299 and D6S1645, two of the most distal and proximal STR markers currently used for PGD analysis,[28] for the five single cells from child 3, there were an average of 79 (range 66–98) and 119 (range 86–115) heterozygous informative SNP loci for the paternal and maternal chromosomes, respectively, far exceeding the 6–10 STR markers normally used.

For PGD, there is only limited time available for genetic analysis between embryo biopsy, normally at cleavage stages early on day 3 following insemination, and transfer on days 4–6. Preliminary results with shortened protocols indicate that whole genome amplification, SNP array processing and data analysis for karyomapping can be achieved within 24–48 h, which would avoid the need to cryopreserve biopsied embryos (unpublished observations). Compared with standard methods, the use of arrays is relatively expensive, particularly as multiple embryos may need to be tested. However, expensive labour intensive test development is eliminated. Also, with improvements in embryo culture and cryopreservation by vitrification, which involves ultra rapid cooling, it is now possible to biopsy only those viable embryos able to develop to the later blastocyst stage, cryopreserve them pending genetic analysis, and then thaw selected embryos for transfer in a later reproductive cycle. Using this strategy and testing for aneuploidy by conventional CGH, survival of vitrified blastocysts following thawing was close to 100% and clinical pregnancy rates were 69% per embryo transfer.[29]

Use of genome wide SNP genotyping or other arrays for clinical applications raises challenging clinical and ethical issues. Because karyomapping analysis is genome wide, there is a strong likelihood of its use in postnatal or adult counselling for multifactorial disorders. While our knowledge of the genes contributing to common complex disorders is incomplete, knowledge of the major loci contributing to a disorder will inform our views on risk. Karyomapping, however, simply uses SNP genotype information anonymously to map inheritance from one generation to the next across the genome. It cannot per se identify de novo changes at individual SNP loci and, if deemed necessary, it would be straightforward to blind or remove some or all of the genotype data from clinical records. With rapid advances in new sequencing technologies, including quantitative single molecule methods, the cost of sequencing targeted regions or the whole genome of individuals is steadily decreasing.[30] If this becomes possible for routine clinical application, complete data on inherited and de novo alterations will be available. The challenge will then be to interpret and counsel patients on the significance of a multitude of rare variants. For single cell analysis, however, the extent and accuracy of sequence information may be limited without further improvements in whole genome amplification.

Acknowledgements AHH is grateful to Paul Williams, The Bridge Centre Ltd, for encouraging and supporting this work. GLH wishes to thank the staff of the Genetics & IVF Institute, especially Kasie Heeres and Khoa Tran, for their hard work, dedication and very gifted hands in the laboratory. DKG is supported by a BBSRC Career Development Fellowship BB/E024211/1.

Funding Self funded.

Competing interests A patent application has been submitted by AHH and The Bridge Centre Ltd (PCT/GB2006/004221).

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Van den Veyver IB,** Patel A, Shaw CA, Pursley AN, Kang SH, Simovich MJ, Ward PA, Darilek S, Johnson A, Neill SE, Bi W, White LD, Eng CM, Lupski JR, Cheung SW, Beaudet AL. Clinical use of array comparative genomic hybridization (aCGH) for prenatal diagnosis in 300 cases. *Prenat Diagn* 2009;**29**:29–39.
2. **Lu XY,** Phung MT, Shaw CA, Pham K, Neil SE, Patel A, Sahoo T, Bacino CA, Stankiewicz P, Kang SH, Lalani S, Chinault AC, Lupski JR, Cheung SW, Beaudet AL. Genomic imbalances in neonates with birth defects: high detection rates by using chromosomal microarray analysis. *Pediatrics* 2008;**122**:1310–18.
3. **Rauch A,** Ruschendorf F, Huang J, Trautmann U, Becker C, Thiel C, Jones KW, Reis A, Nurnberg P. Molecular karyotyping using an SNP array for genomewide genotyping. *J Med Genet* 2004;**41**:916–22.
4. **Redon R,** Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, González JR, Gratacòs M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. Global variation in copy number in the human genome. *Nature* 2006;**444**:444–54.
5. **Thornhill AR,** Handyside AH. Human embryo biopsy procedures. In: Gardner DK, Weissman A, Howles CM, Shoham Z, eds. *'Textbook of Assisted Reproductive Technologies: Laboratory and Clinical Perspectives'* 3rd edn. Informa Healthcare, 2008:191–206.
6. **Dean FB,** Hosono S, Fang L, Wu X, Faruqi AF, Bray-Ward P, Sun Z, Zong Q, Du Y, Du J, Song W, Kingsmore SF, Egholm M, Lasken RS. Comprehensive human genome amplification using multiple displacement amplification. *Proc Natl Acad Sci U S A* 2002;**99**:5261–6.
7. **Handyside AH,** Robinson MD, Simpson RJ, Omar MB, Shaw MA, Grudzinskas JG, Rutherford A. Isothermal whole genome amplification from single and small numbers of cells: a new era for preimplantation genetic diagnosis of inherited disease. *Mol Hum Reprod* 2004;**10**:767–72.
8. **Schowalter KV,** Fredrickson J, Thornhill AR. Efficient isothermal amplification of the entire genome from single cells. *Methods Mol Med* 2007;**132**:87–99.
9. **Handyside AH,** Lesko JG, Tarin JJ, Winston RM, Hughes MR. Birth of a normal girl after in vitro fertilization and preimplantation diagnostic testing for cystic fibrosis. *N Engl J Med* 1992;**327**:905–9.
10. **Kong A,** Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K. A high-resolution recombination map of the human genome. *Nat Genet* 2002;**31**:241–7.
11. **Tease C,** Hulten MA. Inter-sex variation in synaptonemal complex lengths largely determine the different recombination rates in male and female germ cells. *Cytogenet Genome Res* 2004;**107**:208–15.

12. **Tease C,** Hartshorne GM, Hulten MA. Patterns of meiotic recombination in human fetal oocytes. *Am J Hum Genet* 2002;**70**:1469—79.
13. **Broman KW,** Weber JL. Characterization of human crossover interference. *Am J Hum Genet* 2000;**66**:1911—26.
14. **Ray PF,** Winston RM, Handyside AH. Reduced allele dropout in single-cell analysis for preimplantation genetic diagnosis of cystic fibrosis. *J Assist Reprod Genet* 1996;**13**:104—6.
15. **Fiorentino F,** Biricik A, Karadayi H, Berkil H, Karlikaya G, Sertyel S, Podini D, Baldi M, Magli MC, Gianaroli L, Kahraman S. Development and clinical application of a strategy for preimplantation genetic diagnosis of single gene disorders combined with HLA matching. *Mol Hum Reprod* 2004;**10**:445—60.
16. **Hulten MA,** Patel SD, Tankimanova M, Westgren M, Papadogiannakis N, Jonsson AM, Iwarsson E. On the origin of trisomy 21 Down syndrome. *Mol Cytogenet* 2008;**1**:21.
17. **Delhanty JD,** Harper JC, Ao A, Handyside AH, Winston RM. Multicolour FISH detects frequent chromosomal mosaicism and chaotic division in normal preimplantation embryos from fertile patients. *Hum Genet* 1997;**99**:755—60.
18. **Vanneste E,** Voet T, Le Caignec C, Ampe M, Konings P, Melotte C, Debrock S, Amyere M, Vikkula M, Schuit F, Fryns JP, Verbeke G, D'Hooghe T, Moreau Y, Vermeesch JR. Chromosome instability is common in human cleavage-stage embryos. *Nat Med* 2009;**15**:577—83.
19. **Mastenbroek S,** Scriven P, Twisk M, Viville S, Van der Veen F, Repping S. What next for preimplantation genetic screening? More randomized controlled trials needed? *Hum Reprod* 2008;**23**:2626—8.
20. **Chatzimeletiou K,** Morrison EE, Prapas N, Prapas Y, Handyside AH. Spindle abnormalities in normally developing and arrested human preimplantation embryos in vitro identified by confocal laser scanning microscopy. *Hum Reprod* 2005;**20**:672—82.
21. **Le Caignec C,** Spits C, Sermon K, De Rycke M, Thienpont B, Debrock S, Staessen C, Moreau Y, Fryns JP, Van Steirteghem A, Inge L, Joris RV. Single-cell chromosomal imbalances detection by array CGH. *Nucleic Acids Res* 2006;**34**:e68.
22. **Fiegler H,** Geigl JB, Langer S, Rigler D, Porter K, Unger K, Carter NP, Speicher MR. High resolution array-CGH analysis of single cells. *Nucleic Acids Res* 2007;**35**:e15.
23. **Hu DG,** Guan XY, Hussey N. Gender determination and detection of aneuploidy in single cells using DNA array-based comparative genomic hybridization. *Methods Mol Med* 2007;**132**:135—51.
24. **Wells D,** Alfarawati S, Fragouli E. Use of comprehensive chromosomal screening for embryo assessment: microarrays and CGH. *Mol Hum Reprod* 2008;**14**:703—10.
25. **Scriven PN,** Handyside AH, Ogilvie CM. Chromosome translocations: segregation modes and strategies for preimplantation genetic diagnosis. *Prenat Diagn* 1998;**18**:1437—49.
26. **Verlinsky Y,** Rechitsky S, Schoolcraft W, Strom C, Kuliev A. Preimplantation diagnosis for Fanconi anemia combined with HLA matching. *JAMA* 2001;**285**:3130—3.
27. **Van de Velde H,** De Rycke M, De Man C, De Hauwere K, Fiorentino F, Kahraman S, Pennings G, Verpoest W, Devroey P, Liebaers I. The experience of two European preimplantation genetic diagnosis centres on human leukocyte antigen typing. *Hum Reprod* 2009;**24**:732—40.
28. **Fiorentino F,** Kahraman S, Karadayi H, Biricik A, Sertyel S, Karlikaya G, Saglam Y, Podini D, Nuccitelli A, Baldi M. Short tandem repeats haplotyping of the HLA region in preimplantation HLA matching. *Eur J Hum Genet* 2005;**13**:953—8.
29. **Fragouli E,** Katz-Jaffe M, Alfarawati S, Stevens J, Colls P, Goodall NN, Tormasi S, Gutierrez-Mateo C, Prates R, Schoolcraft WB, Santiago M, Dagan W. Comprehensive chromosome screening of polar bodies and blastocysts from couples experiencing repeated implantation failure. *Fertil Steril* 2009.
30. **Voelkerding KV,** Dames SA, Durtschi JD. Next-generation sequencing: from basic research to diagnostics. *Clin Chem* 2009;**55**:641—58.