

Cytogenetic genotype-phenotype studies: Improving genotyping, phenotyping and data storage

I. Feenstra H.G. Brunner C.M.A. van Ravenswaaij

Radboud University Nijmegen Medical Centre, Department of Human Genetics, Nijmegen (The Netherlands)

Manuscript received 17 February 2006; accepted in revised form for publication by A. Geurts van Kessel, 2 May 2006.

Abstract. High-resolution molecular cytogenetic techniques such as genomic array CGH and MLPA detect sub-microscopic chromosome aberrations in patients with unexplained mental retardation. These techniques rapidly change the practice of cytogenetic testing. Additionally, these techniques may improve genotype-phenotype studies of patients with microscopically visible chromosome aberrations, such as Wolf-Hirschhorn syndrome, 18q deletion syndrome and 1p36 deletion syndrome. In order to make the most of high-resolution karyotyping, a similar accuracy of phenotyping is needed to allow researchers and clinicians to make optimal use of the recent advances. International

agreements on phenotype nomenclature and the use of computerized 3D face surface models are examples of such improvements in the practice of phenotyping patients with chromosomal anomalies. The combination of high-resolution cytogenetic techniques, a comprehensive, systematic system for phenotyping and optimal data storage will facilitate advances in genotype-phenotype studies and a further deconstruction of chromosomal syndromes. As a result, critical regions or single genes can be determined to be responsible for specific features and malformations.

Copyright © 2006 S. Karger AG, Basel

New molecular cytogenetic techniques like array-based comparative genomic hybridization (array CGH) (Pinkel et al., 1998; Speicher and Carter, 2005) and Multiplex Ligation-dependent Probe Amplification (MLPA) (Schouten et al., 2002) can be used to search for submicroscopic chromosome aberrations in patients with unexplained mental retardation (Koolen et al., 2004; de Vries et al., 2005). In addition, these techniques will improve genotype-phenotype studies of patients with microscopically visible chromosomal imbalances by precisely determining the genomic region affected. The exact determination of breakpoints needed for genotype-phenotype studies used to be very time-consuming and only feasible for rather common cyto-

genetic syndromes. Examples are the determination of the Wolf-Hirschhorn syndrome critical region on chromosome 4p and the cat-cry region on chromosome 5p in Cri du Chat syndrome (Niebuhr, 1978b; Wright et al., 1997). An overview of current cytogenetic and molecular techniques used in clinical cytogenetics is given in Table 1.

Nowadays, size and location of chromosome aneuploidies can be determined with very high accuracy by tiling path BAC arrays, oligonucleotide arrays or SNP arrays in one single test run (Barrett et al., 2004; Slater et al., 2005; Vissers et al., 2005).

As high-resolution genotyping is rapidly becoming routine, phenotyping with an equally high accuracy is needed to fully benefit from the advantages of these new techniques.

In this report the impact of new techniques on genotype-phenotype studies is reviewed on the basis of various chromosomal syndromes. Advantages and limitations of the new approaches will be discussed, as will the need for sophisticated phenotyping and data collection.

Request reprints from Ilse Feenstra

Radboud University Nijmegen Medical Centre
Department of Human Genetics 849, P.O. Box 9101
NL-6500 HB Nijmegen (The Netherlands)
telephone: +31 24 361 39 46; fax: +31 24 366 87 74
e-mail: i.feenstra@antrg.umcn.nl

Deconstructing chromosomal syndromes

With the use of new molecular techniques, various chromosome syndromes have been analyzed in detail. Whereas in some a single gene appeared to be responsible for most of the phenotypic features, for other syndromes an increasing number of critical regions for specific clinical features can be determined. In this section we first describe the detection of critical regions and some candidate genes in a num-

ber of microscopically visible chromosome disorders. Subsequently, some examples are given of submicroscopic aberrations in which single genes appear to play a major role in the phenotypes of patients.

Cri du Chat syndrome (5p-)

Cri du Chat syndrome (CDC, OMIM 123450) was first described by Lejeune and co-workers in 1963 (Lejeune et al., 1963). The syndrome is caused by a partial deletion of the short arm of chromosome 5 and is characterized by a high-pitched cat-like cry, microcephaly, facial dysmorphism and mental retardation (Niebuhr, 1978a). Chromosome analysis showed different deletion sizes, but no clear association between deletion size and the clinical features could be demonstrated (Miller et al., 1969). In 1978, Niebuhr made an attempt to locate the genetic segment responsible for the clinical features of Cri du Chat syndrome by investigating 35 individuals with a 5p- karyotype (Niebuhr, 1978b). He concluded that the typical features of this syndrome were probably caused by a deletion of the midportion of the 5p15 segment, more specifically 5p15.2. This region is shown in the schematic overview in Fig. 1.

These findings have subsequently been confirmed (Overhauser et al., 1994; Church et al., 1995; Gersh et al., 1995; Mainardi et al., 2001).

Recently, Zhang and co-authors applied the new array CGH technique to analyze genomic DNA of 94 patients

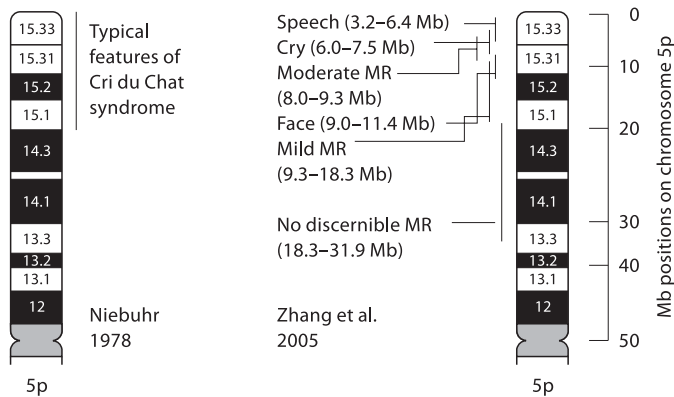


Fig. 1. Schematic overview of the clinical features of Cri du Chat syndrome and the associated critical regions on chromosome 5p. Array CGH was used in the study shown on the right, resulting in a significant refinement of the critical regions. MR = Mental retardation.

Table 1. Overview of techniques used in clinical cytogenetics

| Technique | Resolution; deletion sizes to be detected | Detectable level of mosaicism | Detection of balanced aberrations | Turn-around time | Additional requirements | Relative estimated costs |
|-----------------------------------|--|---|-----------------------------------|------------------|--|--|
| Routine karyotyping | ≥ 5–10 Mb | Depending on number of cells examined; ≥ 1% | Possible | 3–10 days | – Experienced personnel for correct interpretation | Low |
| FISH | 100 kb | ≥ 20% | Possible | 1–7 days | – Clinical indication of possible loci responsible – Labour intensive | High; depending on number of annual investigations |
| Multicolour FISH/SKY | 2–3 Mb | 10% | Possible | 1–7 days | – Clinical indication of suspected loci responsible | High |
| Comparative genomic hybridisation | ≥ 3–10 Mb | ≥ 50% | Not possible | 5–7 days | – Experienced personnel – Labour intensive | High |
| MLPA | ~ 0.1 kb | ≥ 40% | Not possible | 1–4 days | – Clinical indication of possible loci responsible | Low; depending on number of annual investigations |
| BAC array (3–32 K) | Depending on number of clones; 100 kb–1 Mb | Depending on size of aberration and array coverage; ≥ 30% | Not possible | 1–4 days | – Sophisticated equipment – Standardized storage system – Thorough statistical support | High |
| Oligonucleotide array | Depending on number of clones; 1–250 kb | Depending on size of aberration and array coverage; currently unknown | Not possible | 1–4 days | – Sophisticated equipment – Standardized storage system – Thorough statistical support | Very high |
| SNP array (100–500 K) | Depending on number of clones; 10–250 kb | Depending on size of aberration and array coverage; currently unknown | Not possible | 1–4 days | – Sophisticated equipment – Standardized storage system – Thorough statistical support | Very high |

with known deletions of 5p (Zhang et al., 2005). A detailed clinical description was available for all patients. The authors were able to define three critical regions for the cry, speech delay, and facial dysmorphology on 5p15.31, 5p15.32→p15.33 and 5p15.2→p15.31, respectively. Moreover, they concluded that there were three adjacent regions on chromosome 5p that have differential effects on the level of mental retardation (MR) if deleted. A distal 1.2-Mb deletion in 5p15.31 produces moderate MR, whereas isolated deletions of more proximal located regions result in mild or no discernable MR. In Fig. 1 an overview of critical regions associated with the different clinical features is provided.

Wolf-Hirschhorn syndrome (4p-)

In 1965, groups led by Wolf and Hirschhorn each described a patient with a deletion of the short arm of chromosome 4 presenting with growth delay, mental retardation, and congenital anomalies suggestive of a midline fusion defect (Hirschhorn et al., 1965; Wolf et al., 1965). Numerous case-reports on similar patients followed.

One of the first studies in which the investigators tried to localize the segment of chromosome 4p associated with the clinical features of Wolf-Hirschhorn syndrome (WHS, OMIM 194190) was published in 1981 (Wilson et al., 1981). Giemsa-banding (GTG) was performed in 13 patients. The authors concluded that the critical region involved in WHS is within 4p16, the most distal band of the p-arm (Fig. 2).

However, a terminal deletion could not be detected in all patients displaying the clinical features of WHS by routine karyotyping. The contribution of new molecular cytogenetic techniques such as fluorescence in situ hybridisation (FISH), enabled the diagnosis of WHS in patients with submicroscopic interstitial or terminal deletions or subtle unbalanced translocations (Altherr et al., 1991; Johnson et al., 1994).

A preliminary phenotypic map of chromosome 4p16 was put forward in 1995. A systematic genotype-phenotype analysis was performed in 11 patients with chromosome 4p deletions and/or rearrangements (Estabrooks et al., 1995). It was suggested that specific regions within 4p16 correlated with different clinical features.

In 1997 the WHS critical region (WHSCR) was refined to 165 kb with FISH using a series of landmark cosmids from a collection of WHS patient-derived cell lines (Wright et al., 1997; see Fig. 2). The WHSCR is a gene-rich region and contains, among others, the *FGFR3* gene, which is mutant in achondroplasia and other skeletal dysplasias.

Another gene designated as Wolf-Hirschhorn Syndrome Candidate 1 (*WHSC1*) was described in 1998 (Stec et al., 1998). This 25 exon gene was found to be expressed ubiquitously in early development and to undergo complex alternative splicing and differential polyadenylation. It encodes a 136-kDa protein containing four domains also present in other developmental proteins. It is expressed preferentially in rapidly growing embryonic tissues, in a pattern corre-

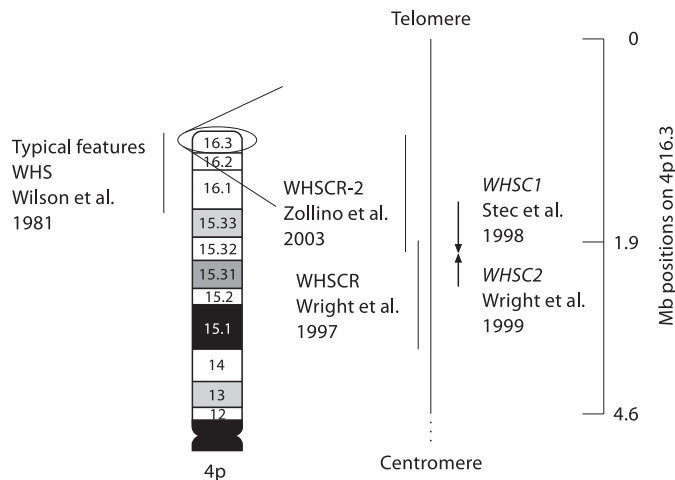


Fig. 2. Schematic overview of the deconstruction of the Wolf-Hirschhorn syndrome in critical regions and candidate genes. WHS = Wolf-Hirschhorn syndrome; WHSCR = WHS critical region; WHSCR-2 = WHS critical region 2; *WHSC1* = WHS candidate gene 1; *WHSC2* = WHS candidate gene 2.

sponding to the affected organs in WHS patients. The nature of the protein motifs, the expression pattern, and its mapping to the critical region led the authors to propose *WHSC1* as a good candidate gene for WHS. A second candidate gene (*WHSC2*) was identified one year later (Wright et al., 1999). The location of both candidate genes is depicted in Fig. 2.

In 2000 an Italian group reported the cytogenetic, molecular, and clinical findings in 16 WHS patients (Zollino et al., 2000). Submicroscopic deletions ranging from 2.8 to 4.4 Mb were detected in four patients. In one patient, no molecular deletion could be detected within the WHSCR. The precise definition of the cytogenetic defect permitted an analysis of genotype/phenotype correlations in WHS, leading to the proposal of a set of minimal diagnostic criteria. Deletions of less than 3.5 Mb resulted in a mild phenotype in which major malformations were absent. The authors proposed a 'minimal' WHS phenotype in which the clinical manifestations are restricted to the typical facial appearance, mild mental and growth retardation, and congenital hypotonia.

In 2003, the same group reported their findings in eight patients carrying a 4p16.3 microdeletion (Zollino et al., 2003). The WHSCR was fully preserved in one patient with a 1.9-Mb deletion, in spite of a typical WHS phenotype. Therefore, the authors proposed a new critical region, WHSCR2, a 300-kb interval located distally from the known WHSCR1 (Fig. 2). Furthermore, for the purpose of genetic counseling, they recommended dividing the WHS phenotype into two distinct clinical entities, i.e., a 'classical' and a 'mild' form, which are usually caused by cytogenetically visible and submicroscopic deletions, respectively. Another patient with a 1.9-Mb subtelomeric deletion was described in 2005, which supports the proposed WHSCR2 (Rodriguez et al., 2005).

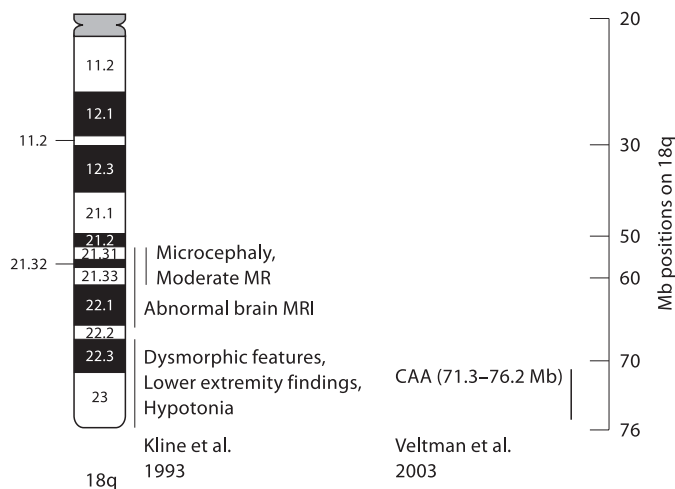


Fig. 3. Overview of the short arm of chromosome 18 and the critical regions defined for distinctive clinical features. MR = Mental retardation; CAA = congenital aural atresia.

A Belgian group reported six additional patients with an atypical 4p16.3 deletion, of whom five patients showed a (very) mild form of WHS and one patient had no clinical signs of WHS (Van Buggenhout et al., 2004). By means of a contiguous 4pter BAC array, the sizes and breakpoints were physically mapped and four terminal deletions (range 0.4–3.81 Mb) and two interstitial deletions (1.55 and 1.7 Mb) were revealed. This study enabled further refinement of the phenotypic map of this region, suggesting hemizygosity of *WHSC1* to cause the typical WHS facial appearance.

In summary, although molecular analysis allows a more detailed view of the WHS critical regions, the exact contribution of each of the proposed critical regions to the WHS phenotype still remains to be determined.

De Grouchy syndrome (18q-)

The 18q deletion syndrome (OMIM 601808) was described first in 1964 by de Grouchy et al. (1964). Most 18q cases are associated with terminal deletions and the phenotype of this syndrome is mainly characterized by mental retardation, hypotonia, short stature, ear anomalies and a flat midface. A first preliminary phenotypic map based on seven patients with deletions of 18q21.3 or 18q22.2 to 18qter was published in 1993 (Kline et al., 1993). In Fig. 3 an overview of clinical features and associated chromosome regions is provided.

Moreover, a substantial percentage of 18q- patients has congenital aural atresia (CAA), leading to hearing loss (Cody et al., 1999).

By applying array CGH, a critical region for CAA was mapped on 18q22.3→q23 (Veltman et al., 2003). The authors used a 670-kb resolution chromosome 18-specific BAC array to analyse genomic DNA of 20 patients with CAA. Of these, 18 patients had a microscopically visible 18q

deletion. In two patients, a submicroscopic 18q deletion was detected which allowed the mapping of CAA to a region of 5 Mb located in 18q22.3→q23 (Fig. 3).

2q deletions

Since 1988 a few patients with a deletion of 2q32→q33 have been described (Miyazaki et al., 1988; Palmer et al., 1990; Kreuz and Wittwer, 1993; Vogels et al., 1997).

BAC array and FISH analyses were used to delineate the deletion size to a critical region of 8.1 Mb in four patients (Van Buggenhout et al., 2005). Three patients displayed psychiatric and behavioural problems (hyperactivity, aggressiveness, anxiety) and shared a commonly deleted region of 0.5 Mb just proximal of the proximal deletion breakpoint of the fourth patient, who lacked behavioural problems. Within this region two genes are located that could cause the behavioural phenotype.

1p36 deletions

This relatively common chromosome aberration has been known for less than ten years as it was discovered only in 1997 (Shapira et al., 1997). With an estimated prevalence of one in 5,000 live births, monosomy of 1p36 (OMIM 607872) is the most common terminal deletion syndrome (Shaffer and Lupski, 2000). Because of the variability in deletion size, parental origin and clinical presentation, it has been proposed that monosomy 1p36 is a contiguous gene syndrome in which haploinsufficiency of functionally unrelated genes leads to the phenotypic features (Wu et al., 1999).

In 2003, a first physical map of 1p36 deletions was published (Heilstedt et al., 2003). First, DNA samples of 61 patients were screened with 25 microsatellite markers for the most distal part of 1p36. Then, a contig of 99 overlapping large-insert clones of this 10.5-Mb region was used to further refine the deletion size. Furthermore, clinical phenotypes of 30 patients were carefully defined. The authors proposed critical regions for hypotonia (2.2-Mb region from the telomere), a large fontanel (2.2 Mb), hearing loss (2.5 Mb), cardiomyopathy (3.1 Mb), hypothyroidism (4.1 Mb) and clefting (4.1 Mb). Because the terminal region of 1p36 is gene rich, no candidate genes could be determined.

In the same year, this group published their data using a dedicated 1p36 array CGH (Yu et al., 2003). This array was designed by using the previously assembled contig, consisting of 97 clones from 1p36, supplemented by clones for the subtelomeric regions of all chromosomes and clones for both sex-chromosomes. Genomic DNA of twenty-five patients with well-defined 1p36 deletions was studied and the array results agreed with the previously determined deletion sizes and breakpoint locations as detected by FISH and microsatellite analyses.

Recently, a tiling resolution BAC array covering 99.5% of the euchromatic parts of chromosome 1 has been applied to

study six patients with a 1p36 deletion phenotype. In all patients a 1p36 deletion was confirmed, with sizes ranging from 2 to 10 Mb. Remarkably, in two clinically similar patients two non-overlapping deletions were detected. Therefore, the authors concluded that the 1p36 phenotype is a consequence of distinct and non-overlapping deletions having a positional effect rather than being a true contiguous gene deletion syndrome (Redon et al., 2005).

Cytogenetic microdeletion syndromes and the impact of single genes

In a number of (micro)deletion syndromes, the molecular determination of breakpoints together with a comparison of clinical features has resulted in such small critical regions that single genes appear to be responsible for the (majority of) phenotypic features.

An example is Smith-Magenis Syndrome (SMS, OMIM 182290), characterized by behavioural problems, speech delay, psychomotor and growth retardation and distinct craniofacial anomalies (Smith et al., 1998). About 75% of the SMS patients have a common deletion spanning 3.5 Mb in the 17p11.2 region, although deletion sizes vary from 1.5 to 9 Mb (Greenberg et al., 1991; Vlangos et al., 2003). Recently, a number of patients who fulfil the criteria for SMS but without the 17p11.2 deletion were analyzed for mutations of the *RAI1* gene, located within the central portion of the critical region for SMS, using PCR and sequencing strategies (Slager et al., 2003; Bi et al., 2004; Girirajan et al., 2005). This resulted in the identification of nine patients having *RAI1* mutations and to the conclusion that haploinsufficiency of this gene is associated with the craniofacial, behavioral and neurological symptoms of SMS.

The 22q13 deletion syndrome (OMIM 606232) is characterized by neonatal hypotonia, severe expressive language delay in combination with mild mental retardation (Prasad et al., 2000). Included in the critical region of this syndrome is the *SHANK3/ProSAP2* gene, which is preferentially expressed in the cerebral cortex and the cerebellum. DNA analysis of *SHANK3/ProSAP2* in a patient carrying a de novo balanced translocation between chromosomes 12 and 22, t(12;22)(q24.1;q13.3), revealed a disruption within exon 21 (Bonaglia et al., 2001). Since the patient displayed all 22q13.3 deletion features, the authors proposed that *SHANK3/ProSAP2* haploinsufficiency is the cause of the 22q13 deletion syndrome. This finding was supported by another group who tested 45 patients with variable sizes of 22q13 deletions, thereby confirming a deletion for the *SHANK3/ProSAP2* gene in all patients (Wilson et al., 2003). A recent study using array CGH for molecular characterization of nine patients with 22q13 aberrations identified deletion sizes ranging from 3.3 to 8.4 Mb (Koolen et al., 2005). The authors did not observe a relation between clinical features and deletion size, thereby supporting the idea that a gene in the 3.3-Mb minimal deleted region, notably *SHANK3/ProSAP2*, may be the major candidate gene in the 22q13 deletion syndrome. Another group using array CGH

Table 2. Examples of cytogenetic microdeletion syndromes in which single genes appear to be responsible for the (majority of) clinical features

| Syndrome | Chromosome location | Gene responsible |
|---------------------------|---------------------|-----------------------|
| Smith Magenis syndrome | 17p11.2 | <i>RAI1</i> |
| 22q13 deletion syndrome | 22q13.3 | <i>SHANK3/ProSAP2</i> |
| 9q34 deletion syndrome | 9q34 | <i>EHMT1</i> |
| Rubinstein-Taybi syndrome | 16p13.3 | <i>CREBBP, EP300</i> |
| Sotos syndrome | 5q35 | <i>NSD1</i> |
| DiGeorge/VCFS syndrome | 22q11.2 | <i>TBX1</i> |

reported their findings in two unrelated 22q13.3 deletion patients (Bonaglia et al., 2006), which were consistent with the concept of *SHANK3/ProSAP2* being the best candidate gene for the neurological deficits in the 22q13.3 syndrome, although patients with the same kind of *SHANK3/ProSAP2* disruption can exhibit different degrees of severity in their phenotype.

Another terminal deletion syndrome is the 9q34 subtelomeric deletion syndrome. This syndrome is characterized by severe mental retardation, hypotonia, microcephaly and a typical face with midface depression, hypertelorism, everted lower lip, cupid bow configuration of the upper lip, and a prominent chin. The minimum critical region involved is ~1.2 Mb in size and encompasses at least 14 genes (Stewart et al., 2004). In a mentally retarded patient with a typical 9qter deletion phenotype, a balanced translocation t(X;9)(p11.23;q34.3) was detected (Kleefstra et al., 2005). Sequence analysis of the breakpoints revealed a disruption of *EHMT1*, indicating that haploinsufficiency of this gene may be responsible for the 9q subtelomeric deletion syndrome. In fact, we have since analysed the *EHMT1* gene in a series of patients with clinical phenotypes suggestive of a 9qter deletion whose telomere region was intact according to FISH and MLPA. We found a de novo nonsense mutation in one such patient and a frameshift in another (Kleefstra et al., 2006). This establishes that *EHMT1* haploinsufficiency is indeed the cause of the 9qter deletion phenotype.

Other examples of cytogenetic syndromes of which the (majority of) clinical features appear to be caused by mutations in single genes are Rubinstein-Taybi Syndrome (RSTS, OMIM 180849), Sotos syndrome (OMIM 117550) and DiGeorge/VCFS Syndrome (DGS, OMIM 188400) (Petrij et al., 1995; Jerome and Papaioannou, 2001; Kurotaki et al., 2002; Yagi et al., 2003; Roelfsema et al., 2005). Furthermore, it has been described that atypical deletions may be associated with variant phenotypes (Rauch et al., 2003, 2005).

These examples illustrate how the boundary between cytogenetic deletion syndromes and single gene conditions is becoming more and more indistinct. Ultimately, we should be able to assess the phenotype contribution of each gene within known microdeletion/microduplication syndromes.

An overview of the above-mentioned syndromes and the possible genes responsible for most phenotypic features is given in Table 2.

The phenotype; how and what to describe

The phenotype, defined as the appearance (physical, biochemical and physiological) of an individual which results from the interaction of the environment and the genotype, is usually presented in scientific articles by a clinical description, sometimes accompanied by clinical photographs. Any description of clinical features of a patient is inherently subjective. It varies between independent physicians and any emphasis on specific features may reflect the background speciality of the observer.

Description of phenotypes

To overcome the bias of subjectivity, proposals have been made to standardize the phenotype description by a systematic collection of clinical information (Freimer and Sabatti, 2003; Hall, 2003; Merks et al., 2003).

A detailed proposal for the organization and standardization of clinical descriptions of human malformations has recently been made (Biesecker, 2005). The author felt that, in contrast to the enormous improvements in molecular biology, the processes and approaches of the clinical component of molecular dysmorphology have not changed substantially. He argued that the current way of collecting phenotypic information holds several weaknesses. The quality and completeness of clinical descriptions published in the medical literature depend on the authors and editors involved. Another threat is confusion in understanding the terms used by the authors, due to the existence of synonyms, various definitions for one word, and sometimes overlapping of two different terms. The author pointed out a number of criteria for an ideal standardized clinical genetics nomenclature.

Standardization of phenotype descriptions will be crucial for a 'Human Phenome Project', in which comprehensive databases are created for such systematically collected phenotypic information (Freimer and Sabatti, 2003). The authors argued that phenotypic information should be collected on different levels: molecules, cells, tissues and whole organisms.

Visualization of phenotype

In a number of cytogenetic syndromes, such as WBS or 1p36 deletion syndrome, the clinical diagnosis is primarily based on characteristic facial features. Clinical geneticists are trained in recognizing specific patterns in different syndromes and can do this relatively well (Winter, 1996). Multiple efforts have been made to implement objective, quantitative criteria and analytical techniques for craniofacial assessments (Allanson, 1997; Shaner et al., 2001). In previous decades, anthropometry, photogrammetry and cephalometry have been applied as diagnostic methods (Garn et al., 1984, 1985; Richtsmeier, 1987; DiLiberti and Olson, 1991).

More recently, computer programs have been designed to analyze and identify faces of patients with certain syndromes on the basis of specific craniofacial features. In one study, standardized photographs of 55 patients with different syndromes were analyzed in a mathematical way by comparing feature vectors at 32 facial nodes (Loos et al., 2003). Over 75% of the patients were correctly classified by the computer, whereas clinicians who were shown the same pictures achieved a recognition rate of 62%.

More recently, a large study on computer-based three-dimensional (3D) imaging of the face of 696 individuals was published (Hammond et al., 2005). This study demonstrated the potential contributions of dense surface models (DSM) in clinical training, making clinical diagnoses and objective comparisons. Such mathematical pattern recognition might improve phenotype-genotype analyses, particularly in patients with rare or atypical chromosome aberrations.

A first application of 3D face surface models in genotype-phenotype studies was demonstrated in Williams-Beuren syndrome (WBS, OMIM 194050), involving a 7q11.23 deletion (Tassabehji et al., 2005). As the typical deletion size in WBS is 1.5 Mb and contains 28 genes, a clear genotype-phenotype correlation for craniofacial features could not be made so far. In this study, a patient with a small, atypical deletion was identified and 3D surface images of this patient's face were compared with those of WBS-individuals and controls. The patient was classified as borderline WBS with mildly dysmorphic features. Chromosome analysis revealed a heterozygous deletion at 7q11.23 of ~1 Mb, resulting in reduced expression of the *GTF2IRD1* gene. In mice, homozygous loss of *Gtf2ird1* results in craniofacial abnormalities reminiscent of those seen in WBS, together with growth retardation. These observations suggest that *GTF2IRD1* plays a role in mammalian craniofacial and cognitive development. The authors suggested that cumulative dosage of TFII-I family genes explains the main phenotypes of WBS. *Gtf2ird1*-null mice and classic WBS patients have two functional copies (*in trans* and *cis*, respectively), whereas the atypical patient had three functional genes of the *GTF2IRD1/GTF2I* cluster and showed a milder WBS phenotype.

Storage of genomic and clinical data

Cytogenetic and clinical information concerning specific chromosomal disorders are continuously published in the (inter)national medical literature. Thus, systematic collection and archiving are essential.

Many of these reports have been collected in the 'Catalogue of Unbalanced Chromosome Aberrations in Man', containing around 2,000 descriptions of patients with a rare chromosome aberration (Schinzel, 2001). This catalogue provides an unprecedented resource for genotype-phenotype studies in cytogenetically visible chromosome anomalies. In order to perform searches directed towards specific chromosome aberrations and/or clinical features, a com-

puterized version is commercially available as the Zurich Cytogenetic Database, which contains cytogenetic and clinical information on more than 7,200 cases from the medical literature and references to the original papers (Schinzel, 1994).

In the past decennium, a number of internet databases have been created. These databases allow users online access to databases which are constantly being updated.

One of the databases collecting cytogenetic, molecular, and clinical information on patients with rare unbalanced chromosome aberrations is the ECARUCA database (<http://www.ecaruca.net>) (Feenstra et al., 2006). This database is based on information from the Zurich Cytogenetic Database, and is frequently supplemented by new data of patients with (sub)microscopic chromosomal aberrations by a European network of cytogenetic laboratories. In this password protected database, geneticists can search for information on more than 1,500 different chromosome aberrations of almost 4,100 cases. Frequent submission of new cases is performed by the account holders, thereby ensuring the up-to-date quality of the collection. In addition, data is centrally checked for usage of correct cytogenetic nomenclature and description of clinical features before inclusion in the database. ECARUCA is interactive, dynamic and has long-term possibilities to store molecular data. Also, parents of children whose data have been entered in the database can anonymously add follow-up information through the website.

A specialized database for submicroscopic chromosome aberrations is Decipher (<http://www.sanger.ac.uk/PostGenomics/decipher/>). This database currently contains information on 38 microdeletion/-duplication syndromes and, like ECARUCA, members are asked to actively participate in the submission of new cases.

Linking the above mentioned databases to genome browsers allows users to directly search for molecular information in their chromosome region of interest. By comparing the clinical features of different patients and their (non) overlapping aberrations at the molecular level, this can be a helpful tool in genotype-phenotype studies.

Analysing data stored in databases

Using a mathematical model, chromosome maps for specific malformation patterns based on the catalogue of unbalanced chromosome disorders and associated congenital malformations collected in the Zurich Cytogenetic Database were created (Brewer et al., 1998, 1999). The chromosomal deletion map was assembled through the analysis of 1,753 patients with a single, non-mosaic contiguous autosomal deletion and the presence of common major malformations. This resulted in 284 positive associations between specific malformations and deleted bands, distributed among 137 malformation-associated chromosome regions (MACRs).

In a second article, a chromosomal duplication map was described. Here, a total of 143 MACRs were identified, of which 21 were highly significant.

Obviously, such maps should always be interpreted with care. Although the number of cases available for analysis was high, the accuracy of breakpoints is not known since the cytogenetic analyses were mostly performed with standard karyotyping. Nonetheless, this type of analyses can point to those chromosome regions where the search for loci involved in congenital malformations is most likely to be successful. This has been abundantly proven in the case of holoprosencephaly, where at least four genes have been found based on chromosomal mapping of critical regions (Muenke et al., 1994; Brown et al., 1995, 1998; Overhauser et al., 1995; Schell et al., 1996; Wallis et al., 1999).

As more and more submicroscopic deletions and duplications are mapped, further candidate genes for specific malformations will be revealed. For instance, a recent study in 100 patients with mental retardation and malformations detected a small duplication in 5q35.1 in a patient with lobar holoprosencephaly (de Vries et al., 2005). This region contains seven known genes of which *FBXW11* is a likely candidate gene for holoprosencephaly (Koolen et al., 2006).

Conclusion and future prospects

There are many rapid advances in genotype-phenotype studies in chromosome disorders. Many improvements have been made in the field of genotyping, leading to the detection of smaller and smaller genomic deletions and duplications. Moreover, new cytogenetic molecular techniques allow for the exact determination of breakpoints in microscopically visible aberrations.

The capacity to investigate a high number of patients by undertaking automated genotyping projects is no longer the limiting step in elucidating the molecular basis of cytogenetic syndromes. Only an equally high accuracy of phenotyping will allow researchers and clinicians to make optimal use of the recent advances in genotyping. Objective, standardized descriptions and/or visualizations of the phenotype will be critical to determine the role of critical regions or candidate genes detected by new molecular techniques.

The combination of high-resolution cytogenetic techniques, a comprehensive, systematic approach for phenotyping and collection of this information in sophisticated databases will lead to advances in genotype-phenotype studies. Thereby, a further deconstruction of chromosomal syndromes in which critical regions or single genes appear to be responsible for specific features is to be expected.

Acknowledgements

The authors would like to thank Joris Veltman and Bert de Vries for critical reading of the manuscript.

References

- Allanson JE: Objective techniques for craniofacial assessment: what are the choices? *Am J Med Genet* 70:1–5 (1997).
- Altherr MR, Bengtsson U, et al: Molecular confirmation of Wolf-Hirschhorn syndrome with a subtle translocation of chromosome 4. *Am J Hum Genet* 49:1235–1242 (1991).
- Barrett MT, Scheffer A, et al: Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA. *Proc Natl Acad Sci USA* 101:17765–17770 (2004).
- Bi W, Saifi GM, et al: Mutations of RAI1, a PHD-containing protein, in nondelusion patients with Smith-Magenis syndrome. *Hum Genet* 115:515–524 (2004).
- Biesecker LG: Mapping phenotypes to language: a proposal to organize and standardize the clinical descriptions of malformations. *Clin Genet* 68:320–326 (2005).
- Bonaglia MC, Giorda R, et al: Disruption of the ProSAP2 gene in a t(12;22)(q24.1;q13.3) is associated with the 22q13.3 deletion syndrome. *Am J Hum Genet* 69:261–268 (2001).
- Bonaglia MC, Giorda R, et al: Identification of a recurrent breakpoint within the *SHANK3* gene in the 22q13.3 deletion syndrome. *J Med Genet* 43:822–828 (2006).
- Brewer C, Holloway S, et al: A chromosomal deletion map of human malformations. *Am J Hum Genet* 63:1153–1159 (1998).
- Brewer C, Holloway S, et al: A chromosomal duplication map of malformations: regions of suspected haplo- and triplolethality – and tolerance of segmental aneuploidy – in humans. *Am J Hum Genet* 64:1702–1708 (1999).
- Brown S, Russo J, et al: The 13q- syndrome: the molecular definition of a critical deletion region in band 13q32. *Am J Hum Genet* 57:859–866 (1995).
- Brown SA, Warburton D, et al: Holoprosencephaly due to mutations in *ZIC2*, a homologue of *Drosophila odd-paired*. *Nat Genet* 20:180–183 (1998).
- Church DM, Bengtsson U, et al: Molecular definition of deletions of different segments of distal 5p that result in distinct phenotypic features. *Am J Hum Genet* 56:1162–1172 (1995).
- Cody JD, Ghidoni PD, et al: Congenital anomalies and anthropometry of 42 individuals with deletions of chromosome 18q. *Am J Med Genet* 85:455–462 (1999).
- De Grouchy J, Royer P, et al: Partial deletion of the long arms of the chromosome 18. *Pathol Biol (Paris)* 12:579–582 (1964).
- de Vries BB, Pfundt R, et al: Diagnostic genome profiling in mental retardation. *Am J Hum Genet* 77:606–616 (2005).
- DiLiberti JH, Olson DP: Photogrammetric evaluation in clinical genetics: theoretical considerations and experimental results. *Am J Med Genet* 39:161–166 (1991).
- Estabrooks LL, Rao KW, et al: Preliminary phenotypic map of chromosome 4p16 based on 4p deletions. *Am J Med Genet* 57:581–586 (1995).
- Feenstra I, Fang J, et al: European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations (ECARUCA); An online database for rare chromosome abnormalities. *Eur J Med Genet* 49:279–291 (2006).
- Freimer N, Sabatti C: The human phenome project. *Nat Genet* 34:15–21 (2003).
- Garn SM, Smith BH, et al: Applications of pattern profile analysis to malformations of the head and face. *Radiology* 150:683–690 (1984).
- Garn SM, Lavelle M, et al: Quantification of dysmorphogenesis: pattern variability index, sigma z. *AJR Am J Roentgenol* 144:365–369 (1985).
- Gersh M, Goodart SA, et al: Evidence for a distinct region causing a cat-like cry in patients with 5p deletions. *Am J Hum Genet* 56:1404–1410 (1995).
- Girirajan S, Elsas LJ 2nd, et al: *RAI1* variations in Smith-Magenis syndrome patients without 17p11.2 deletions. *J Med Genet* 42:820–828 (2005).
- Greenberg F, Guzzetta V, et al: Molecular analysis of the Smith-Magenis syndrome: a possible contiguous-gene syndrome associated with del(17)(p11.2). *Am J Hum Genet* 49:1207–1218 (1991).
- Hall JG: A clinician's plea. *Nat Genet* 33:440–442 (2003).
- Hammond P, Hutton TJ, et al: Discriminating power of localized three-dimensional facial morphology. *Am J Hum Genet* 77:999–1010 (2005).
- Heilstedt HA, Ballif BC, et al: Physical map of 1p36, placement of breakpoints in monosomy 1p36, and clinical characterization of the syndrome. *Am J Hum Genet* 72:1200–1212 (2003).
- Hirschhorn K, Cooper HL, et al: Deletion of short arms of chromosome 4–5 in a child with defects of midline fusion. *Humangenetik* 1:479–482 (1965).
- Jerome LA, Papaioannou VE: DiGeorge syndrome phenotype in mice mutant for the T-box gene, *Tbx1*. *Nat Genet* 27:286–291 (2001).
- Johnson VP, Altherr MR, et al: FISH detection of Wolf-Hirschhorn syndrome: exclusion of D4F26 as critical site. *Am J Med Genet* 52:70–74 (1994).
- Kleefstra T, Smid M, et al: Disruption of the gene *Euchromatin Histone Methyl Transferase1 (EHMTase1)* is associated with the 9q34 subtelomeric deletion syndrome. *J Med Genet* 42:299–306 (2005).
- Kleefstra T, Brunner HG, et al: Loss of function mutations in euchromatin histone methyl transferase 1 (EHMT1) cause the 9q34 subtelomeric deletion syndrome. *Am J Hum Genet* 79:370–377 (2006).
- Kline AD, White ME, et al: Molecular analysis of the 18q- syndrome – and correlation with phenotype. *Am J Hum Genet* 52:895–906 (1993).
- Koolen DA, Nillesen WM, et al: Screening for subtelomeric rearrangements in 210 patients with unexplained mental retardation using multiplex ligation dependent probe amplification (MLPA). *J Med Genet* 41:892–899 (2004).
- Koolen DA, Reardon W, et al: Molecular characterisation of patients with subtelomeric 22q abnormalities using chromosome specific array-based comparative genomic hybridisation. *Eur J Hum Genet* 13:1019–1024 (2005).
- Koolen DA, Herbergs J, et al: Holoprosencephaly and preaxial polydactyly associated with a 1.24 Mb duplication encompassing *FBXW11* at 5q35.1. *J Hum Genet* 51:721–726 (2006).
- Kreuz FR, Wittwer BH: Del(2q) – cause of the wrinkly skin syndrome? *Clin Genet* 43:132–138 (1993).
- Kurotaki N, Imaizumi K, et al: Haploinsufficiency of *NSDI* causes Sotos syndrome. *Nat Genet* 30:365–366 (2002).
- Lejeune J, Lafourcade J, et al: 3 Cases of partial deletion of the short arm of chromosome 5. *C R Hebd Seances Acad Sci* 257:3098–3102 (1963).
- Loos HS, Wiczorek D, et al: Computer-based recognition of dysmorphic faces. *Eur J Hum Genet* 11:555–560 (2003).
- Mainardi PC, Perfumo C, et al: Clinical and molecular characterisation of 80 patients with 5p deletion: genotype-phenotype correlation. *J Med Genet* 38:151–158 (2001).
- Merks JH, van Karnebeek CD, et al: Phenotypic abnormalities: terminology and classification. *Am J Med Genet A* 123:211–230 (2003).
- Miller DA, Warburton D, et al: Clustering in deleted short-arm length among 25 cases with a Bp-chromosome. *Cytogenetics* 8:109–116 (1969).
- Miyazaki K, Yamanaka T, et al: Interstitial deletion 2q32.1→q34 in a child with half normal activity of ribulose 5-phosphate 3-epimerase (RPE). *J Med Genet* 25:850–851 (1988).
- Muenke M, Gurrieri F, et al: Linkage of a human brain malformation, familial holoprosencephaly, to chromosome 7 and evidence for genetic heterogeneity. *Proc Natl Acad Sci USA* 91:8102–8106 (1994).
- Niebuhr E: The Cri du Chat syndrome: epidemiology, cytogenetics, and clinical features. *Hum Genet* 44:227–275 (1978a).
- Niebuhr E: Cytological observations in 35 individuals with a 5p- karyotype. *Hum Genet* 42:143–156 (1978b).
- Overhauser J, Huang X, et al: Molecular and phenotypic mapping of the short arm of chromosome 5: sublocalization of the critical region for the cri-du-chat syndrome. *Hum Mol Genet* 3:247–252 (1994).
- Overhauser J, Mitchell HF, et al: Physical mapping of the holoprosencephaly critical region in 18p11.3. *Am J Hum Genet* 57:1080–1085 (1995).
- Palmer CG, Heerema N, et al: Deletions in chromosome 2 and fragile sites. *Am J Med Genet* 36:214–218 (1990).
- Petrij F, Giles RH, et al: Rubinstein-Taybi syndrome caused by mutations in the transcriptional coactivator CBP. *Nature* 376:348–351 (1995).
- Pinkel D, Segraves R, et al: High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* 20:207–211 (1998).
- Prasad C, Prasad AN, et al: Genetic evaluation of pervasive developmental disorders: the terminal 22q13 deletion syndrome may represent a recognizable phenotype. *Clin Genet* 57:103–109 (2000).
- Rauch A, Beese M, et al: A novel 5q35.3 subtelomeric deletion syndrome. *Am J Med Genet A* 121:1–8 (2003).
- Rauch A, Zink S, et al: Systematic assessment of atypical deletions reveals genotype-phenotype correlation in 22q11.2. *J Med Genet* 42:871–876 (2005).
- Redon R, Rio M, et al: Tiling path resolution mapping of constitutional 1p36 deletions by array-CGH: contiguous gene deletion or 'deletion with positional effect' syndrome? *J Med Genet* 42:166–171 (2005).
- Richtsmeier JT: Comparative study of normal, Crouzon, and Apert craniofacial morphology using finite element scaling analysis. *Am J Phys Anthropol* 74:473–493 (1987).
- Rodriguez L, Zollino M, et al: The new Wolf-Hirschhorn syndrome critical region (WHSCR-2): a description of a second case. *Am J Med Genet A* 136:175–178 (2005).
- Roelfsema JH, White SJ, et al: Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in both the *CBP* and *EP300* genes cause disease. *Am J Hum Genet* 76:572–580 (2005).

- Schell U, Wienberg J, et al: Molecular characterization of breakpoints in patients with holoprosencephaly and definition of the HPE2 critical region 2p21. *Hum Mol Genet* 5:223–229 (1996).
- Schinzel A: Zurich Cytogenetics Database (1994).
- Schinzel A: Catalogue of Unbalanced Chromosome Aberrations in Man. (De Gruyter, Berlin 2001).
- Schouten JP, McElgunn CJ, et al: Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* 30:e57 (2002).
- Shaffer LG, Lupski JR: Molecular mechanisms for constitutional chromosomal rearrangements in humans. *Annu Rev Genet* 34:297–329 (2000).
- Shaner DJ, Peterson AE, et al: Soft tissue facial resemblance in families and syndrome-affected individuals. *Am J Med Genet* 102:330–341 (2001).
- Shapira SK, McCaskill C, et al: Chromosome 1p36 deletions: the clinical phenotype and molecular characterization of a common newly delineated syndrome. *Am J Hum Genet* 61:642–650 (1997).
- Slager RE, Newton TL, et al: Mutations in *RAI1* associated with Smith-Magenis syndrome. *Nat Genet* 33:466–468 (2003).
- Slater HR, Bailey DK, et al: High-resolution identification of chromosomal abnormalities using oligonucleotide arrays containing 116,204 SNPs. *Am J Hum Genet* 77:709–726 (2005).
- Smith AC, Dykens E, et al: Behavioral phenotype of Smith-Magenis syndrome (del 17p11.2). *Am J Med Genet* 81:179–185 (1998).
- Speicher MR, Carter NP: The new cytogenetics: blurring the boundaries with molecular biology. *Nat Rev Genet* 6:782–792 (2005).
- Stec I, Wright TJ, et al: *WHSCI*, a 90 kb SET domain-containing gene, expressed in early development and homologous to a *Drosophila dysmorphia* gene maps in the Wolf-Hirschhorn syndrome critical region and is fused to *IgH* in t(4;14) multiple myeloma. *Hum Mol Genet* 7:1071–1082 (1998).
- Stewart DR, Huang A, et al: Subtelomeric deletions of chromosome 9q: a novel microdeletion syndrome. *Am J Med Genet A* 128:340–351 (2004).
- Tassabehji M, Hammond P, et al: *GTF2IRD1* in craniofacial development of humans and mice. *Science* 310:1184–1187 (2005).
- Van Buggenhout G, Melotte C, et al: Mild Wolf-Hirschhorn syndrome: micro-array CGH analysis of atypical 4p16.3 deletions enables refinement of the genotype-phenotype map. *J Med Genet* 41:691–698 (2004).
- Van Buggenhout G, Van Ravenswaaij-Arts C, et al: The del(2)(q32.2q33) deletion syndrome defined by clinical and molecular characterization of four patients. *Eur J Med Genet* 48:276–289 (2005).
- Veltman JA, Jonkers Y, et al: Definition of a critical region on chromosome 18 for congenital aural atresia by array CGH. *Am J Hum Genet* 72:1578–1584 (2003).
- Vissers LE, Veltman JA, et al: Identification of disease genes by whole genome CGH arrays. *Hum Mol Genet* 14 Spec No. 2:R215–223 (2005).
- Vlangos CN, Yim DK, et al: Refinement of the Smith-Magenis syndrome critical region to approximately 950 kb and assessment of 17p11.2 deletions. Are all deletions created equally? *Mol Genet Metab* 79:134–141 (2003).
- Vogels A, Haegeman J, et al: Pierre-Robin sequence and severe mental retardation with chaotic behaviour associated with a small interstitial deletion in the long arm of chromosome 2 (del(2)(q33.1q33.3)). *Genet Couns* 8:249–252 (1997).
- Wallis DE, Roessler E, et al: Mutations in the homeodomain of the human *SIX3* gene cause holoprosencephaly. *Nat Genet* 22:196–198 (1999).
- Wilson HL, Wong AC, et al: Molecular characterization of the 22q13 deletion syndrome supports the role of haploinsufficiency of *SHANK3/PROSAP2* in the major neurological symptoms. *J Med Genet* 40:575–584 (2003).
- Wilson MG, Towner JW, et al: Genetic and clinical studies in 13 patients with the Wolf-Hirschhorn syndrome [del(4p)]. *Hum Genet* 59:297–307 (1981).
- Winter RM: What's in a face? *Nat Genet* 12:124–129 (1996).
- Wolf U, Reinwein H, et al: Deficiency on the short arms of a chromosome No. 4. *Humangenetik* 1:397–413 (1965).
- Wright TJ, Ricke DO, et al: A transcript map of the newly defined 165 kb Wolf-Hirschhorn syndrome critical region. *Hum Mol Genet* 6:317–324 (1997).
- Wright TJ, Costa JL, et al: Comparative analysis of a novel gene from the Wolf-Hirschhorn/Pitt-Rogers-Danks syndrome critical region. *Genomics* 59:203–212 (1999).
- Wu YQ, Heilstedt HA, et al: Molecular refinement of the 1p36 deletion syndrome reveals size diversity and a preponderance of maternally derived deletions. *Hum Mol Genet* 8:313–321 (1999).
- Yagi H, Furutani Y, et al: Role of *TBX1* in human del22q11.2 syndrome. *Lancet* 362:1366–1373 (2003).
- Yu W, Ballif BC, et al: Development of a comparative genomic hybridization microarray and demonstration of its utility with 25 well-characterized 1p36 deletions. *Hum Mol Genet* 12:2145–2152 (2003).
- Zhang X, Snijders A, et al: High-resolution mapping of genotype-phenotype relationships in Cri du Chat syndrome using array comparative genomic hybridization. *Am J Hum Genet* 76:312–326 (2005).
- Zollino M, Di Stefano C, et al: Genotype-phenotype correlations and clinical diagnostic criteria in Wolf-Hirschhorn syndrome. *Am J Med Genet* 94:254–261 (2000).
- Zollino M, Lecce R, et al: Mapping the Wolf-Hirschhorn syndrome phenotype outside the currently accepted WHS critical region and defining a new critical region, WHSCR-2. *Am J Hum Genet* 72:590–597 (2003).