Array-CGH analysis indicates a high prevalence of genomic rearrangements in holoprosencephaly: an updated map of candidate loci.

Claude Bendavid, Lucie Rochard, Christèle Dubourg, Jonathan Seguin, Isabelle Gicquel, Laurent Pasquier, Jaqueline Vigneron, Annie Laquerrière, Pascale Marcorelles, Corinne Jeanne-Pasquier, et al.

Human Mutation

Array-CGH analysis indicates a high prevalence of genomic rearrangements in Holoprosencephaly: an updated map of HPE candidate loci.

Journal:	*Human Mutation*
Manuscript ID:	draft
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Bendavid, Claude; Faculté de Médecine, Université de Rennes , IFR140, UMR 6061 CNRS, Institut de Génétique et Développement de Rennes, Equipe Génétique des Pathologies liées Au Développement Rochard, Lucie; Faculté de Médecine, Université de Rennes 1,IFR140, UMR 6061 CNRS, Institut de Génétique et Développement de Rennes, Equipe Génétique des Pathologies liées Au Développement DUBOURG, Christèle; Laboratoire de Génétique Moléculaire, CHU PONTCHAILLOU; Faculté de Médecine, Université de Rennes1, IFR 140, UMR 6061 CNRS, Institut de Génétique et Développement de Rennes, Equipe Génétique des Pathologies liées Au Développement Seguin, Jonathan; Faculté de Médecine, Université de Rennes1, IFR 140, UMR 6061 CNRS, Institut de Génétique et Développement de Rennes, Equipe Génétique des Pathologies liées Au Développement Gicquel, Isabelle; Faculté de Médecine, Université de Rennes1, IFR 140, UMR 6061 CNRS, Institut de Génétique et Développement de Rennes, Equipe Génétique des Pathologies liées Au Développement PASQUIER, Laurent; Faculté de Médecine, Université de Rennes 1,IFR140, UMR 6061 CNRS, Institut de Génétique et Développement de Rennes, Equipe Génétique des Pathologies liées Au Développement; Hopital Sud, Service de génétique Clinique VIGNERON, Jacqueline; Maternité régionale A. Pinard, Service de Médecine et réanimation néonatales génétique LAQUERRIERE, Annie; Université de Rouen Marcorelles, Pascale; CHU Brest, Hôpital Morvan, Anatomie Pathologique Jeanne-Pasquier, Corinne; CHU Côte de Nacre, Anatomie et cytopathologie Rouleau, Caroline; Hopital Lapeyronnie, CHU Montpellier, Service d'anatomie et cytopathologie Jaillard, Sylvie; CHU RENNES, Laboratoire de Cytogénétique Mosser, Jean; Ouest Génopole, Transcriptomic platform; Faculté de Médecine, Université de Rennes , IFR140, UMR 6061 CNRS, Institut de Génétique et Développement de Rennes, Equipe Oncogénèse et régulation ODENT, SYLVIE; Faculté de Médecine, Université de Rennes 1,IFR140, UMR 6061 CNRS, Institut de Génétique et Développement de Rennes, Equipe Génétique des Pathologies liées Au Développement; CHU RENNES, Service de génétique Clinique DAVID, Véronique; Faculté de Médecine, Université de Rennes1, IFR 140, UMR 6061 CNRS, Institut de Génétique et Développement de Rennes, Equipe Génétique des Pathologies liées Au Développement; Laboratoire de Génétique Moléculaire, CHU Pontchaillou
Key Words:	holoprosencephaly, CGH array, rearrangements, candidate loci

Array-CGH analysis indicates a high prevalence of genomic rearrangements in Holoprosencephaly: an updated map of HPE candidate loci.

Acknowledgments

This work was funded by PHRC interrégional 2004, GIS maladies rares (GISMR0701), FMO (Fédération des Maladies Orphelines).

We particularly thank the SOFFOET for providing several Fetus cases with complete anatomical description of cerebral defects, all clinicians who referred HPE cases and all families for giving their consent to this study.

ABSTRACT

Holoprosencephaly (HPE) is the most frequent congenital malformation of the brain. Aetiology includes karyotype anomalies, environmental factors and genic forms that can be syndromic or isolated. Non-random structural chromosomal anomalies previously compiled from chromosomal HPE predicted at least 12 different HPE loci. To date, eight HPE genes have been identified from recurrent chromosomal rearrangements or from the sequencing of genes from Nodal and SHH pathways.

Our cohort of isolated HPE presents a high genetic heterogeneity. Point mutations were found in *SHH, ZIC2, SIX3* and *TGIF* genes in about 20% of cases (with 10% in *SHH*). Submicroscopic deletions in these same genes were found in 7.5% and 4.4% presented with other subtelomeric gain or losses. Consequently molecular basis of HPE remains unknown in 70% of our cohort.

In order to detect new HPE candidate genes, we used array-CGH to refine the previous karyotype based HPE loci map. We analysed 111 HPE patients with high performance Agilent arrays made of 44K or 244K oligonucleotidic probes and found that 28 presented with submicroscopic anomalies involving known or new potential HPE loci located on different chromosomes but with poor redundancy. We observed 14 isolated deletions, 9 isolated duplications and 5 associated genomic losses and gains. Compiling these new data with frequencies of deletions in known HPE genes and of subtelomeric anomalies, we give evidence that microrearrangements could be a major molecular mechanism in HPE. Additionally, this study opens new insights on HPE candidate genes identification giving an updated HPE candidate loci map.

INTRODUCTION

Holoprosencephaly (HPE; MIM 236100) is the most common forebrain developmental anomaly in humans, resulting from a complete or partial failure of cleavage of the forebrain during development. The clinical spectrum ranges from alobar HPE (absent interhemispheric fissure) to semilobar (posterior midline separation) and lobar HPE (continuity only across the frontal cortex) generally associated with facial anomalies. HPE is a severe pathology, associated with mental retardation in all affected live newborns, with poor or symptomatic treatment (Cohen, 2006). The genetic counselling in HPE families is very complex due to the extreme phenotypical variability, the genetic heterogeneity, and a high recurrence risk (13%) in apparently sporadic cases. Chronologically, non-random structural chromosomal anomalies previously compiled from chromosomal HPE predicted at least 12 different HPE loci, and out of these 12 loci, eight genes have been really implicated in HPE with mutations found in isolated HPE: Sonic hedgehog (SHH; 7q36; HPE3)(Belloni, et al., 1996; Roessler, et al., 1996)], ZIC2 (13q32; HPE5)(Brown, et al., 1998), SIX3 (2p21; HPE2)(Wallis and Muenke, 1999), TGIF (18p11.3; HPE4)(Gripp, et al., 2000), PATCHED1 (9q22)(Ming and Muenke, 2002), TDGF1 (3p21.31)(de la Cruz, et al., 2002), FAST1 (8q34) (Ouspenskaia, et al., 2002) and GLI2 (2q24)(Roessler and Muenke, 2003). In total, these actors play a role either in the SHH pathway, or in the Nodal/Transforming Growth Factor beta (TGFβ) pathway, or as transcription factors.

Point mutations in the four major genes, *SHH*, *ZIC2*, *SIX3* and *TGIF*, were identified in 20% of our HPE patients. Among these genes, SHH appears to be the major one accounting for 50% of the identified mutations (Dubourg, et al., 2004).

Animal studies and rare human cases showed that double heterozygous mutations could be involved in HPE phenotype, introducing the multi-hit hypothesis in this developmental disorder (Ming and Muenke, 2002). This hypothesis helped in the understanding of the variable penetrance of familial mutations in the disease and led us to systematically screen the four genes, even if one mutation was initially found in the first sequenced gene. Additional work is focusing on environmental factors including low cholesterol levels (Edison, et al., 2007).

Since 2003, we also screened these genes for microrearrangements and proved for the first time the implication of gene deletions in 7.5% in holoprosencephaly (3.2% in

SHH, 2% in ZIC2, 1.6% in SIX3 and 0.65% in TGIF) (Bendavid, et al., 2006a; Bendavid, et al., 2006b). When comparing the combined mutation and deletion results observed in the foetus cohort and in the live-born children, the total proportions of gene anomalies are close (28% and 24% respectively), but the rate of point mutations is much higher in live born children than in foetuses (23% versus 14%) whereas submicroscopic deletions, which represent gross alterations, occur more frequently in foetuses (10% versus 5%) who generally have a more severe phenotype.

Because of the HPE clinical and genetic heterogeneity, and the lack of informative families, a classical positional cloning strategy based on genome wide scan is not possible. Consequently, we were led to develop a strategy based on molecular biology and cytogenetics to identify candidate regions and thus candidate genes. This approach will complete the initial karyotype based study of recurrent chromosomal abnormalities which led to the identification of the first HPE genes (Belloni, et al., 1996) and HPE loci map (Roessler and Muenke, 1998).

After screening of microdeletions or mutations in known HPE genes with QMPSF/MLPA or DHPLC plus sequencing respectively, we searched for submicroscopic rearrangements in subtelomeric chromosomal regions using the Multiplex Ligation Probe Amplification (MLPA) method (Hogervorst, et al., 2003). Indeed, subtelomeric aberrations were detected in 4,4% of our HPE patients with no known anomaly, showing either a single anomaly or an association between a deletion and a gain, these rearrangements were very heterogeneous, encompassing 10 different subtelomeric regions. Some targeted regions known to be implicated in HPE (7q encompassing the SHH gene, 18p encompassing the TGIF gene, 21q including a candidate gene, LSS (Lanosterol Synthase)), but also new regions (1p, 5q, 7p, 8p, 9q, 17q, 18q and 22q) were identified. Several samples, mainly foetal ones, consisted of an association between a duplication and a deletion in two chromosomal subtelomeres like (dup7pter; del7qter) or (dup20pter; del21qter). Moreover, rearrangements presented by foetuses generally implied known HPE loci, while those observed in live children encompassed regions not previously described (Bendavid, et al., 2007). But, even if we compile these large deletions, point mutations and deletions in the known genes, the combined rate of patients with identified molecular basis only reaches 30%, so more than 70% of the cases remain

unexplained, suggesting the involvement of many other genes in HPE (Dubourg, et al., 2007) and/or non genetic factors.

The previous data showed the growing importance of microrearrangements in a complex genetic disease where mutations in the different genes have a variable penetrance and may need a second genetic event to give the disease. In order to identify new candidate loci and thus novel candidate genes, we decided to perform a genome wide screening for submicroscopic anomalies on isolated HPE patients, using Agilent array-CGH technology. We first tested the technique on 10 patients with known alterations in *SHH*, *TGIF*, *ZIC2*, *SIX3* or subtelomeres in order to better define the size of the rearrangement and the breakpoints. Then 111 samples (64 fetuses and 47 live-born children), with no known karyotypes alterations, were hybridized. We chose these patients with no regard to their mutational status for the main HPE genes, resulting in that 18% of this group were carrying a mutation.

Confirming the growing importance of micro rearrangements, this study showed up an impressive rate of 28 patients among 111 with submicroscopic chromosomal anomalies. These defects involved known or new potential HPE loci located on different chromosomes but with poor redundancy. Added to the previous microdeletion findings in known HPE genes and subtelomeres, our data showed that microrearrangements could be the major molecular mechanism in HPE and strongly reinforce the multigenic origin in this developmental disorder.

MATERIAL AND METHODS

Patients and controls

The cohort consisted in 9 patients presenting with a known deletion already detected by qPCR or QMPSF in HPE genes (SHH, TGIF, ZIC2 and SIX3) or telomeric rearrangements shown by MLPA experiments and 111 HPE patients (64 fetuses and 47 live-born children) with normal karyotype. Out of this whole cohort, all foetuses had severe HPE (central nervous system (CNS) findings consistent with HPE) whereas 15 live-borns had severe HPE and 32 had a spectrum of HPE microsigns (midline defect without cerebral malformation, including facial clefting, single central maxillary incisor, hypotelorism).

Twenty patients with already known point mutations in the four HPE genes were also included, 20 of them with an inherited mutation and a severe phenotype whereas the transmitting parent had a microform.
For only 21 cases with positive CGH result, we had both maternal and paternal DNA that could be investigated by array-CGH, MLPA or quantitative PCR to confirm de novo alterations.

In order to simplify the interpretation of the results and better characterize CNVs, we used genomic DNA from one well-characterized normal male 46,XY and 1 well-characterized normal female 46,XX as control. These two controls were regular blood donors that gave consent (to Etablissement Français du Sang) to use anonymously their DNA for diagnosis or research purposes.

This study was approved by the Institutional Review Board, and parents of all participants gave their informed consent.

Comparative Genomic Hybridization study

Briefly, for live-born children, genomic DNA was extracted from peripheral whole blood by the Flexigene DNA kit from QIAGEN. For foetuses, DNA was extracted from tissues or cultured amniotic cells, using the QIAamp DNA minikit from QIAGEN. DNA concentration was determined using a NanoDrop Spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE).

Array-CGH analysis can be hampered by the large DNA input requirement: a minimum of 0.5 µg per sample are needed to process one array-CGH. Most of our samples were extracted from foetuses, and only a small amount of DNA is often available. So, when necessary, DNA samples were amplified using a whole genome amplification method. In this case, control DNA was also amplified in order not to co-hybridize a native DNA from a control with an amplified DNA from a patient.

DNA amplification

DNA samples were amplified using the GenomePlex WGA kit (Sigma product code WGA2) according to the protocol provided by Sigma. The amount of DNA input into WGA reactions was 50 ng.

Array-CGH

Array-CGH was performed using the Agilent Human Genome Microarray Kit 44A and 244A (Agilent Technologies, Santa Clara, CA, USA). These are high resolution 60-mer oligonucleotide based microarrays containing 44,000 or 244,000 60-mers probes respectively, spanning coding and non-coding genomic sequences with median spacing of 24 kb and 7.4 kb respectively.

When using native DNA, both patient and a sex-matched control's DNAs were separately digested with both AluI and RsaI (Promega, Madison, WI, USA) for 2 h at 37C°. Quality of digestion was controlled on a 1% agarose gel.

When patient and control DNA were amplified digestion was not necessary as the WGA method already generates small fragments, and DNA was directly labelled after purification.

Labelling

DNA concentration was re-controlled using Qubit quantification method (Invitrogen, Carlsbad, CA, USA) in order to use the same DNA input for the patient and the control labelling. According to the protocol provided by Agilent (Protocol v4.0, June 2006), the native digested or WGA2 amplified DNA were labelled by random priming using the Agilent "Genomic DNA Labelling Kit Plus". Patient DNA and control DNA were labelled with Cy3-dUTP and Cy5-dUTP respectively. Labelled products were purified by Microcon YM-30 filters (Millipore, Billerica, MA, USA). Finally, patient and control DNAs were pooled based on equimolar DNA concentration measured with a Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies, Rockland, Delaware, USA). The mix was denaturated with Human Cot I DNA, and then hybridized on microarrays at 65°C for 24 to 48 h in a hybridization oven with a rotator rack. Washing steps and acetonitrile rinsing were performed according to the Agilent protocol in an ozone free area. Arrays were analysed using the Agilent scanner G2565BA and the Agilent Feature Extraction software version 9.1(CGH-v4_9.1 protocol). The software removed outliers pixels and subtracted local background.

Normalisation of the data was achieved by linear dye normalisation and log2 ratios of the dye-normalized signals were calculated.

Bioinformatics

Data were imported into Agilent CGHanalytics software version 3.4.27 for analysis. Identification of probes with a significant gain or loss was based on cut-off values of 0.5 and -1 respectively. Based on CGHanalytics Quality control metrics (QCmetrics), the arrays included had a Derivative Log Ratio (DLR) spread score under 0,320. The DLR Spread metrics estimates the log ratio noise by calculating the spread of log ratio differences between consecutive probes along all chromosomes (Largo, et al., 2007). DNA sequence information from the software was linked to the public UCSC database (Human Genome Browser, May 2004 Assembly: Hg18). We only conserved gains or losses that encompassed at least 3 consecutive spots on the array. We didn't present in the result table the gains and losses corresponding to high frequency copy number variation (CNV).

Based on a quality score for the arrays, we classified them in two groups so that to select arrays that could enter a transversal analysis using the Nexus Copy Number software (www.Biodiscovery.com). Patient CGH files extracted from Feature extraction were uploaded into the Nexus Copy Number software. This software also gives a quality score: all arrays with a score under 0.180 (corresponding to 0.320 in CGH analytics) allowed the transversal analysis without strong interfering background (that could generate false positive gains or losses). Using this transversal analysis, we compiled all array data and gave a graphic view of the combined results (aggregate) where the frequencies of anomalies were represented by histograms. This aggregate easily pointed out small redundant rearrangements that resulted either from frequent CNV either from one anomaly of the control DNA (mirror image in most of the patients). Finally, all gains or losses could be compared to CNV listed in the database (monthly updated from TCAG) and exported in Tables (Tables 1, 2 and 3) and Figure 1.

The aim of this approach was to detect the redundancy of small regions and get a percentage value to compare with rare CNV frequencies for the same area when such CNV was already described in databases (TCAG human variation website: http://projects.tcag.ca/variation/).

RESULTS

Search for CNVs in control DNA

Male and female control DNAs were co-hybridized on a 244K microarray in order to better characterize copy number variants (CNV) that could be present even in these phenotypically normal individuals.

Gains or losses observed in the two control genomic DNA were listed in order to be taken into account in the interpretation of patients' array-CGH results (data not shown).

Array-CGH analysis in patients

120 patients (70 foetuses and 50 live-born children) were analysed either on a 44K or a 244K Agilent oligonucleotide arrays.

Out of them, 9 presented with a known deletion already detected by qPCR or QMPSF on known genes (SHH, TGIF, ZIC2 and SIX3) or subtelomeric rearrangements shown by MLPA experiments. They were analysed on 44K arrays.

A first cohort of 37 patients without any karyotypic alterations was also tested with 44K agilent array-CGH and another cohort of 74 HPE patients was secondarily tested on 244K array-CGH as soon as this platform became commercially available.

Patients with known rearrangements

Out of the 9 patients with known rearrangements analysed on 44K arrays, 7 were foetuses and 2 were live born children. 7 had severe HPE phenotype with alobar or semilobar form. These patients were known to present deletions in known genes like *SHH*, *TGIF*, *ZIC2* and *SIX3*, but array-CGH analysis could give the size of the losses which ranged from 1 gene to 30Mb, and showed that 4 of them had also gain of genomic DNA.

An overview of all imbalances is shown in Table 1.

Patients analysed on 44K arrays

A first array-CGH analysis on 37 HPE patients was performed, using oligonucleotide 44K Agilent® array. 19 cases were foetuses and 18 were live-born children. Sub microscopic chromosomal imbalances were detected in 9 (5 fetuses, 4 children) out

of them, 6 had an isolated deletions and 3 had duplications. The size of the imbalances ranged from 50 Kb to 16.9 Mb. (Table 2)

Patients analysed on 244K arrays

Another cohort of 74 HPE (45 foetuses and 29 live-born children) was then analysed, using a higher resolution array (244K Agilent®).

19 patients (12 fetuses, 7 children) out of 74 (26%) presented copy number variations not described in the CNV databases. These rearrangements are more frequently deletions as 12 patients had an isolated or associated loss. 11 patients had isolated or associated duplications. The size's range of the imbalances is very large, from 300 Kb to 16.5 Mb (Table 3).

Nexus study

Compiling 90 arrays (quality score below 0,180) in the Nexus study, the aggregate pointed out 2 regions overlapping known CNV but with a higher recurrence than what was reported in CNV databases. The first locus included two genes: *MACROD2* and *FLRT3* in 20p12.1. The second one included the gene *TPPP* in 5p15.3.

DISCUSSION

This study is the first series of HPE patients to be screened for chromosomal imbalances with high resolution oligonucleotide microarrays. In a total of 111 patients with normal karyotype, 28 were detected with chromosomal imbalances (25%). This unexpected high frequency, whatever the size, location and redundancy of the rearrangements, demonstrates that CGH is mandatory to detect submicroscopic molecular defects; consequently, adding these anomalies to the results of the classic diagnosis screening of isolated HPE, the rate of identified molecular defects could exceed for the first time 50% of cases.

Methodology

Control DNA

We decided to use genomic DNA from one normal male 46,XY and one normal female 46,XX as controls. Indeed, it seemed to us that it was easier to validate the CNVs status of these two DNA samples and thus avoid false positive or negative results in our series. A similar strategy was chosen by Carter et al (Carter, 2007),

who selected NA15510, the source of fosmid library used to confirm genome assembly during the finishing of the human genome, and NA10851, a well characterized cell line DNA from the Hap-Map collection.

CNV interference

Results need to take into account the presence of potential copy number variants (CNV) present in the normal human genome. Indeed, the recent appreciation of widespread copy number variation in genomes of healthy subjects was a significant challenge for teams that wished to use array-CGH in order to correlate chromosomal imbalances and diagnosis of constitutional disorders. Public available databases that accumulate CNV data on hundreds of healthy individuals are now available. So we systematically compared our results to the Database of Genomic Variants (DGV) (http://projects.tcag.ca/variations), which in June 2008 contained 17641 CNVs entries from 49 different published studies using different array platforms and healthy human groups; most of these CNVs range from 40 to 100 kb (Carter, 2007). The referenced CNVs could account for over 20% of the human genome. Of course, this database should be considered with great caution, especially when these CNVs are reported only in a single individual out of hundreds what should definitely not deny any morbid role for such locus. Every month, new CNV data are described in the literature, so it is very difficult to consider all these data and to appreciate their accuracy. In our Nexus study, we pointed out two loci of interest based on a higher frequency than what could be found in CNV database:

A first chromosomal region presented 4 gains and 5 losses in HPE patients, covering mostly 150 Kb, in 5p15.3. This redundant rearranged region contains a candidate gene, TPPP, expressed in adult brain and often described in Alzheimer papers.

A second one was located in 20p. Indeed, among 110 HPE patients, 4 presented rearrangement at 20p12.1, 3 losses and 1 gain. These rearrangements spanned over several closed regions, all concerning one or several exons of *MACROD2* gene (*C20orf133*) (Figure 1) whose embryonic expression pattern in the mouse orthologue is compatible with a role in holoprosencephaly (Maas, et al., 2007). This gene contains a nested gene in its third exon, called *FLRT3*, coding a fibronectin leucin-rich repeat transmembrane protein, which could be a conserved Nodal target. Indeed, loss of function in the FLRT3 gene leads to defects in ventral closure, headfold fusion and definitive endoderm migration (Maretto, et al., 2008).

A deletion of MACROD2 was previously found in one patient with Kabuki syndrome (KS) ; sequencing of 19 other KS patients did not reveal any mutation in MACROD2 discarding this gene as a unique candidate gene for KS (Maas, et al., 2007).

To confirm that the rearrangements observed in our 4 patients were not benign CNV, the different database reporting these large-scale polymorphisms were investigated and unlike the Maas paper data, we could determine that they overlapped described CNVs. Moreover, Kuniba et al recently described the results of a deletion assay for the exon 5 in MACROD2 and a mutation analysis of MACROD2 and FLRT3 among 43 patients with KS in Japan (Kuniba, et al., 2008). They also showed that 2 patients out of 18 presented copy number variations in this region, and concluded that MACROD2 and/or FLRT3 could not be the causative gene in most Japanese KS patients.

These two papers and our own analysis demonstrate that one should be very cautious about the involvement of genes in diseases without an achieved CNV database to analyze the results.

Minimal size to be considered

Currently available data suggest using a DNA size cut-off for a positive result. To avoid listing many anomalies (small CNV or background noise related to one or two consecutive spots on the array), we chose that all data presented in tables must result from three consecutive co-deleted or co-duplicated probes on the 44K or 244K arrays to produce strong evidence of any genomic defect; consequently their respective resolution can be estimated to 48kb and 14.8kb respectively.

This compromise is not perfect as this cut-off of 3 spots on the 44K array would have led us to discard a deletion located only on *SIX3* gene (Table 1) and involving only one spot, if we had not been aware of its existence based on previous specific studies (Bendavid, et al., 2006a).

Impact of array resolution

We used the oligonucleotide 44K Agilent® arrays to test 37 HPE patients (19 foetuses and 18 live-born children). Sub microscopic chromosomal imbalances were detected in 9 out of them (24%), 6 deletions and 4 duplications (Table 2). Another

cohort of 74 HPE (45 foetuses and 29 live-born children) was then analysed, using 244K Agilent® arrays. 19 patients out of 74 (26 %) presented rearrangements.

When using the high resolution arrays, the rate of detected rearrangements modestly raises (26% versus 24%). This suggests that the resolution of the 44K arrays could be sufficient for routine screening of HPE patients. On the opposite, the higher rate of small deletions detected by 244K needs more data compiling (based on Nexus software to point out the more frequent anomalies and exclude putative CNV). In routine diagnosis, the cost of the 244K array versus its modest raise of the detection rate could certainly limit the CGH approach to the 44K array.

Patient's results

Identified loci were heterogeneous in size and poorly redundant, but large anomalies were preferentially found in fetuses

Previous data suggested that deletions in known genes (SHH, SIX3, ZIC2 and TGIF) were more frequent in foetuses than in children, who, on the opposite, presented a higher rate of mutations (Bendavid, et al., 2006a; Bendavid, et al., 2006b). The present study corroborates this hypothesis as 17 chromosomal rearrangements were observed in foetuses (61%) versus 11 in children (39%).

The structural variations observed in our cohort present a wide range of sizes.

We first determined the precise size of rearrangements in samples known to be deleted for HPE genes (*SHH*, *TGIF*, *SIX3* and *ZIC2*) by qPCR, QMPSF or MLPA or by high resolution karyotype (Table 1). The patient (N°8) with a deletion of TGIF in 18p shown by quantitative PCR (Bendavid, et al., 2006a) had in fact a 10 Mb loss of telomeric genomic DNA. Deletions in 13q including *ZIC2* ranged from 1.7 Mb to 30.7 Mb, while those including *SHH* in 7q ranged from 3 Mb to 7 Mb. The severity of the phenotype seems to be correlated with the size of the deletion, if we consider that wider rearrangements are mostly found in foetuses with alobar or semi-lobar HPE, while the smaller ones (even concerning a single probe like the patient (N°1) with isolated *SIX3* deletion) were preferentially observed in live-born children, with lobar HPE accompanied with minor signs. Nevertheless, phenotype/genotype correlation is not straight as children with typical HPE do not have a higher rate of rearrangements than children with minor signs.

Parental analysis

In our series, out of 28 patients with gains or losses, both parents could be tested in 21 cases (focusing on 26 rearrangements). The aim of this parental analysis was to further modulate the role of these rearrangements in phenotype onset based on their inheritance.

Therefore, we postulated first that all anomalies found isolated (not associated with a mutation) and that raised de novo in the proband would be the more likely to be involved in the phenotype. Second, all anomalies found isolated but inherited from one parent would be part from a genetic background and may only act with a variable penetrance or with an associated factor. Finally, if found associated with a mutation and inherited from one parent, the anomaly would more likely be a very minor modulator.

In our series, 9 patients had at least one gain or loss inherited from one parent; 8 were duplications whereas only 3 were deletions (patients N° 23, 30 and 32). In 6 patients (N° 18, 21, 27, 30, 32 and 37), rearrangements were associated to mutations. Therefore, no definitive conclusion about the involvement of these different regions can be given.

On the opposite, out of the 11 patients with de novo gains or losses, we had 9 deletions and only 4 duplications. 8 involved known HPE loci: 2p (SIX3), 7qter (SHH), 13q (ZIC2), 18pter (TGIF), 20pter and 21qter. This reinforces the implication of loci like 20pter or 21qter in HPE but also gives new candidate loci with regions not previously involved in the disease: 1q, 6q, 7p, 10q, 14q and 17q. The 14q del was already reported by Kamnasaran et al in 2005 as a putative HPE loci (Kamnasaran, et al., 2005). The two overlapping 6q del are also particularly interesting as they mapped a new locus that appeared twice de novo.

It's worth noting that one patient (N°25) presents a *SHH* locus duplication. This could be associated with a SHH gain of function, what has never been described before in HPE. Nevertheless, this defect was also found in the mother (normal phenotype). SHH gains of function are usually associated with basocellular carcinoma that have not been described to date in this family. On another hand, this rearrangement may cause a *SHH* loss of function based on the modification of the chromatin environment. Further investigations on *SHH* expression in this family should be overtaken.

Update of HPE loci map

The major aim of our research is to identify candidate regions, map them (Figure 2) and thus extract new HPE candidate genes. Because of the large number of non redundant loci found and the large number of genes covered, it was not relevant to list every candidate gene that could be hypothesized out of these regions. Nevertheless, our results are a milestone in further genetic research and diagnosis in HPE. Effectively, an important bioinformatics investigation on these regions can therefore be performed, based on gene networks and known functional studies results in order to prioritize the best candidate regions and genes out of our data.

Today, only one region could be easily hypothesized as it showed a redundancy in 4 patients who presented a deletion of about 100Kb without overlap with CNVs in the TCAG database to date. This region is located in 10p12.1 and contains a candidate gene, *PATCHED3*, belonging to the PATCHED family, receptor of SHH, one of the main HPE genes. Nevertheless, we could investigate the parents for one case (N°30) and the deletion was inherited from the father, knowing that furthermore the father also transmitted a *SHH* mutation. Another case (N°27) with no parental study also had an associated *SHH* mutation. Consequently, a role for *PATCHED3* is questionable and it's difficult to consider it as a strong modulator making the fetus phenotype worse than his father's.

CGH use in routine diagnosis

For diagnosis, this study demonstrates that the CGHarray approach must be another part of the molecular routine so that to get as much potential markers of the disease as possible to help the genetic counseling. In our experience, this method has been very helpful in the identification of unbalanced subtelomeric anomalies (as MLPA for subtelomeres but with the advantage of determining the breakpoints in the same time) and led to identification of parental cryptic balanced translocations by FISH.

Finally, this study also reinforces the multi-hit hypothesis showing 16 patients with associated gains and losses or rearrangements and mutations. The patients (with severe phenotype) with known mutations inherited from one parent (with mild phenotype) plus a gain and/or a loss are a strong example of what could be considered as a genetic background helping the mutation penetrance.

Conclusion

This is the first study screening a large cohort of isolated HPE by CGHarray. Results demonstrated a high frequency of submicroscopic anomalies with yield of 25%. These anomalies are heterogeneous in size and poorly redundant, what give even more evidence of the multi and plurigenic origin of this developmental disorder. The map of these anomalies added to known candidate loci will certainly help to associate potential candidate genes from developmental research with the human disease. This study also demonstrates that CGH must be part of the molecular diagnosis algorithm to help clinicians to get more disease markers for the difficult HPE genetic counseling.

Databases

Database of Genomic Variants (http://projects.tcag.ca/variations),

Human Genome Browser, (May 2004 Assembly: Hg18).

Nexus Copy Number software (www.Biodiscovery.com)

BIBLIOGRAPHY

Belloni E, Muenke M, Roessler E, Traverso G, Siegel-Bartelt J, Frumkin A, Mitchell HF, Donis-Keller H, Helms C, Hing AV, Heng HH, Koop B, Martindale D, Rommens JM, Tsui LC, Scherer SW. 1996. Identification of Sonic hedgehog as a candidate gene responsible for holoprosencephaly. Nat Genet 14(3):353-6.

Bendavid C, Dubourg C, Gicquel I, Pasquier L, Saugier-Veber P, Durou MR, Jaillard S, Frebourg T, Haddad BR, Henry C, Odent S, David V. 2006a. Molecular evaluation of foetuses with holoprosencephaly shows high incidence of microdeletions in the HPE genes. Hum Genet 119(1-2):1-8.

Bendavid C, Dubourg C, Pasquier L, Gicquel I, Le Gallou S, Mottier S, Durou MR, Henry C, Odent S, David V. 2007. MLPA screening reveals novel subtelomeric rearrangements in holoprosencephaly. Hum Mutat 28(12):1189-97.

Bendavid C, Haddad BR, Griffin A, Huizing M, Dubourg C, Gicquel I, Cavalli LR, Pasquier L, Shanske AL, Long R, Ouspenskaia M, Odent S, Lacbawan F, David V, Muenke M. 2006b. Multicolor FISH and quantitative PCR can detect submicroscopic deletions in holoprosencephaly patients with a normal karyotype. J Med Genet.

Brown SA, Warburton D, Brown LY, Yu CY, Roeder ER, Stengel-Rutkowski S, Hennekam RC, Muenke M. 1998. Holoprosencephaly due to mutations in ZIC2, a homologue of Drosophila odd-paired. Nat Genet 20(2):180-3.

Carter NP. 2007. Methods and strategies for analyzing copy number variation using DNA microarrays. Nat Genet 39(7 Suppl):S16-21.

Cohen MM, Jr. 2006. Holoprosencephaly: clinical, anatomic, and molecular dimensions. Birth Defects Res A Clin Mol Teratol 76(9):658-73.

de la Cruz JM, Bamford RN, Burdine RD, Roessler E, Barkovich AJ, Donnai D, Schier AF, Muenke M. 2002. A loss-of-function mutation in the CFC domain of TDGF1 is associated with human forebrain defects. Hum Genet 110(5):422-8.

Dubourg C, Bendavid C, Pasquier L, Henry C, Odent S, David V. 2007. Holoprosencephaly. Orphanet J Rare Dis 2:8.

Dubourg C, Lazaro L, Pasquier L, Bendavid C, Blayau M, Le Duff F, Durou MR, Odent S, David V. 2004. Molecular screening of SHH, ZIC2, SIX3, and TGIF genes in patients with features of holoprosencephaly spectrum: Mutation review and genotype-phenotype correlations. Hum Mutat 24(1):43-51.

Edison RJ, Berg K, Remaley A, Kelley R, Rotimi C, Stevenson RE, Muenke M. 2007. Adverse birth outcome among mothers with low serum cholesterol. Pediatrics 120(4):723-33.

Gripp KW, Wotton D, Edwards MC, Roessler E, Ades L, Meinecke P, Richieri-Costa A, Zackai EH, Massague J, Muenke M, Elledge SJ. 2000. Mutations in TGIF cause holoprosencephaly and link NODAL signalling to human neural axis determination. Nat Genet 25(2):205-8.

Hogervorst FB, Nederlof PM, Gille JJ, McElgunn CJ, Grippeling M, Pruntel R, Regnerus R, van Welsem T, van Spaendonk R, Menko FH, Kluijt I, Dommering C, Verhoef S, Schouten JP, van't Veer LJ, Pals G. 2003. Large genomic deletions and duplications in the BRCA1 gene identified by a novel quantitative method. Cancer Res 63(7):1449-53.

Kamnasaran D, Chen CP, Devriendt K, Mehta L, Cox DW. 2005. Defining a holoprosencephaly locus on human chromosome 14q13 and characterization of potential candidate genes. Genomics 85(5):608-21.

Kuniba H, Tsuda M, Nakashima M, Miura S, Miyake N, Kondoh T, Matsumoto T, Moriuchi H, Ohashi H, Kurosawa K, Tonoki H, Nagai T, Okamoto N, Kato M, Fukushima Y, Naritomi K, Matsumoto N, Kinoshita A, Yoshiura KI, Niikawa N. 2008. Lack of C20orf133 and FLRT3 mutations in 43 patients with Kabuki syndrome in Japan. J Med Genet 45(7):479-80.

Largo C, Saez B, Alvarez S, Suela J, Ferreira B, Blesa D, Prosper F, Calasanz MJ, Cigudosa JC. 2007. Multiple myeloma primary cells show a highly rearranged unbalanced genome with amplifications and homozygous deletions irrespective of the presence of immunoglobulin-related chromosome translocations. Haematologica 92(6):795-802.

Maas NM, Van de Putte T, Melotte C, Francis A, Schrander-Stumpel CT, Sanlaville D, Genevieve D, Lyonnet S, Dimitrov B, Devriendt K, Fryns JP, Vermeesch JR. 2007. The C20orf133 gene is disrupted in a patient with Kabuki syndrome. J Med Genet 44(9):562-9.

Maretto S, Muller PS, Aricescu AR, Cho KW, Bikoff EK, Robertson EJ. 2008. Ventral closure, headfold fusion and definitive endoderm migration defects in mouse embryos lacking the fibronectin leucine-rich transmembrane protein FLRT3. Dev Biol 318(1):184-93.

Ming JE, Muenke M. 2002. Multiple hits during early embryonic development: digenic diseases and holoprosencephaly. Am J Hum Genet 71(5):1017-32.

Ouspenskaia MV, Karkera JP, Roessler E, Shen MM, Goldmunts E, Bowers P, Towbin J, Belmont J, Muenke M. Role of FAST1 gene in the development of holoprosencephaly (HPE) and congenital cardiac malformations in humans; 2002 Oct 2002; Baltimore, USA. American Society of Human Genetics. p N° 822.

Roessler E, Belloni E, Gaudenz K, Jay P, Berta P, Scherer SW, Tsui LC, Muenke M. 1996. Mutations in the human Sonic Hedgehog gene cause holoprosencephaly. Nat Genet 14(3):357-60.

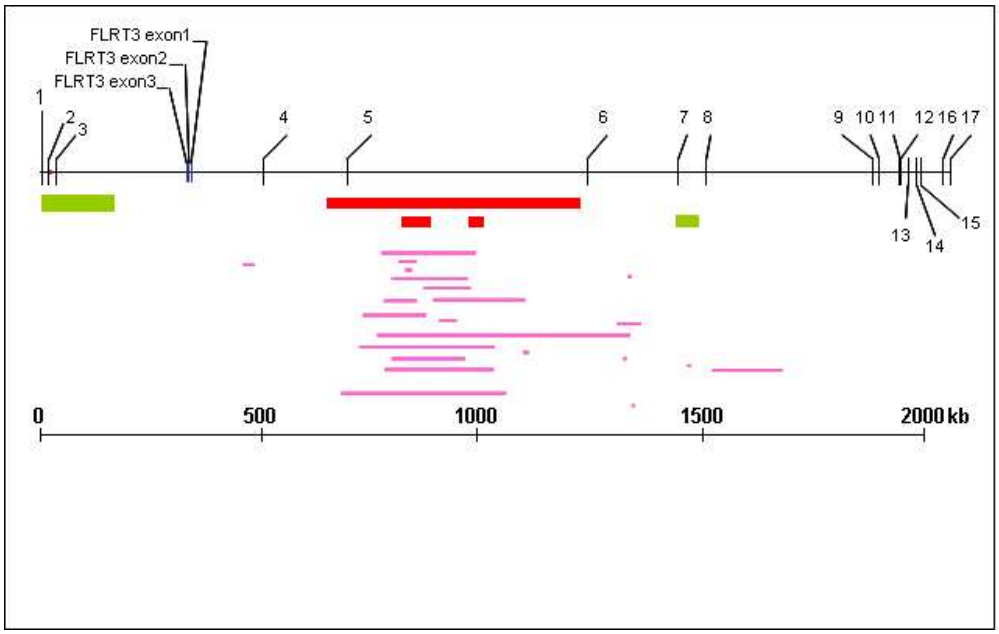
Roessler E, Muenke M. 1998. Holoprosencephaly: a paradigm for the complex genetics of brain development. J Inherit Metab Dis 21(5):481-97.

Roessler E, Muenke M. 2003. How a Hedgehog might see holoprosencephaly. Hum Mol Genet 12 Spec No 1:R15-25.

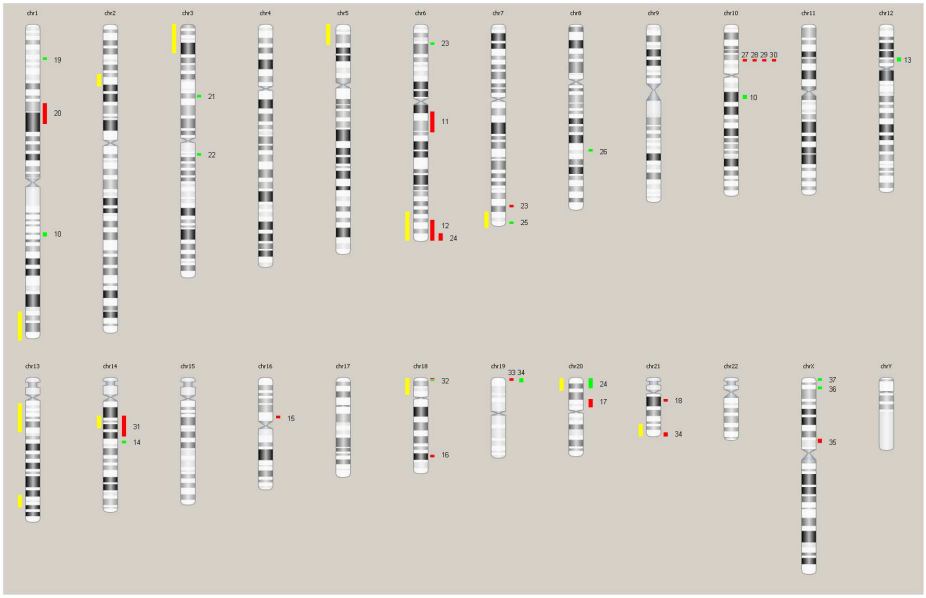
Wallis DE, Muenke M. 1999. Molecular mechanisms of holoprosencephaly. Mol Genet Metab 68(2):126-38.

Figure 1: Schematic view of the exons and introns of MACROD2 and FLRT3 genes. MACROD2 contains 17 exons and is 2057697 bp size. It contains in the third intron the FLRT3 gene. This three exons gene is 13628 bp in size. Red lines indicate deletion of patients, Green line indicate duplication. Positions of CNVs are represented by pink lines (updated in june 2008).

Figure 2: This figure represents all chromosome ideograms with: 1) in yellow, on the left of each ideogram, all related known HPE loci included in Roessler et al (1998) and 14q loci from Kamnasaran and al (2005) papers. 2) in green all gains found in our cohort 3) in red all losses found in our cohort. Gains and losses are associated with the patient numbers that can be found in tables 2, 3 and 4.



: Schematic view of the exons and introns of MACROD2 and FLRT3 genes. MACROD2 contains 17 exons and is 2057697 bp size. It contains in the third intron the FLRT3 gene. This three exons gene is 13628 bp in size. Red lines indicate deletion of patients, Green line indicate duplication. Positions of CNVs are represented by pink lines (updated in june 2008).
170x106mm (96 x 96 DPI)



This figure represents all chromosome ideograms with: 1) in yellow, on the left of each ideogram, all related known HPE loci included in Roessler et al (1998) and 14q loci from Kamnasaran and al (2005) papers. 2) in green all gains found in our cohort 3) in red all losses found in our cohort. Gains and losses are associated with the patient numbers that can be found in tables 2, 3 and 4.
508x304mm (96 x 96 DPI)

Table 1: CGH analysis in 9 patients with known defects

Patient	Phenotype	F / C	Sexe	Anomalies	Start-End	size (Mb)	sequencing
1	lobar	C		Del 2p	45022541-45025894	Only six3 gene	
2	lobar	C	M	Del 5qter	177354922-ter	3,5	
				Dup 17q	70880633-ter	7,9	
3	alobar	F		Dup 7pter	0-6167900	7	
				Del 7qter	152087784-ter	7	
4	semi lobar	F	F	Del 7q	154943585-1578473434	3	
				Dup 8p	0-409876	0.4	
5	syndromic HPE	F	F	Del 7qter	152232455-ter	6,2	
6	semi lobar	F	M	Del 13q	98323528-100026969	1,7	
7	alobar	F	M	Del 13q ter	83266096-ter	30,7	
8	semi lobar	F	M	Del 18pter	66272268-ter	10	
9	alobar cardiac anomalies	F	M	Del 20p	0-8268492	8,2	
				Dup 21q	35804023-ter	11	

Table 2: Summary of copy number changes detected by 44K array CGH (short clinical description and associated mutations in HPE genes)

Patient	Phenotype	F / C	Sexe	Anomalies	Start-End	taille (Mb)	sequencing
10	semi lobar	F	F	Dup 1q	165319168-165662668	0,3	
				Dup 10q	59312961-60086986	0,8	
11	minor signs	C	M	Del 6q	69637936-86100038	16,9	
12	Semi lobar	F	F	Del 6q	155166802-ter	15,7	
13	semi lobar	C	M	Dup 12p	26727022-27427780	0,8	
14	lobar, polymalformations	F	M	Dup14q	49150663-49200220	0,05	
15	syndromic HPE	C	F	Del 16p	29407325-30255748	0,85	
16	semi lobar	F	M	Del 18q	62405035-63364892	1	
17	alobar	F	F	Del 20p	18377837-23793823	5,4	
18	semi lobar	C	F	Del 21q	19024763-20666763	1,6	Zic2 de novo

Table 3: Summary of copy number changes detected by 244K array CGH (short clinical description and associated mutations in HPE genes)

Patient	Phenotype	F / C	Sexe	Anomalies	Start-End	Size (Mb)	sequencing
19	lobar	C	M	Dup 1p	26866731-27930229	1,06	
20	minor signs	C	F	Del 1p	61656706-77968936	16,4	
21	semi lobar	F	M	Dup 3p	54969531-55519135	0,55	SHH paternal
22	alobar	F	F	Dup3q	101837302-101896892	0,6	
23	minor signs	C	F	Del 7q	142530842-142603299	0,7	
				Dup 6p	13251543-13558381	0,3	
24	minor signs	C	M	Del 6qter	165767853-ter	5,2	
				Dup 20pter	0-7967454	8	
25	semi lobar	F	M	Dup 7q	154623987-155557785	0,93	
26	minor signs	C	F	Dup 8q	98220486-98365015	0,15	
27	minor signs	C	M	Del 10p	27656534-27753830	0,10	SHH
28	lobar	F	M	Del 10p	27645595-27754606	0,11	
29	alobar	F	F	Del 10p	27645595-27754606	0,11	
30	lobar + familial form	F	M	Del 10p	27645595-27754606	0,11	SHH paternal
31	alobar	F	F	Del 14q	29960612-46532138	16,5	
32	alobar	F	M	Del 18p	0-602022	0,6	
				Dup 18p	1707459-1834111	0,13	Zic2 *(de novo)*
33	semi lobar	F	M	Del 19p	0-402006	0,4	
34	minor signs	C	F	Dup 19p	0-1124473	1,1	
				Del 21q	44012487-ter	2,9	
35	semi lobar	F	F	Del Xp	48145179-52507792	4,4	
36	minor signs	F	F	DupXp	7542546-8091810	0,5	
37	alobar	F	M	Dup Xp	2709755-2923910	0,3	2 in Zic2

Table 4:Study of inheritance in 21 probands (out of 31) whose both parents were available for analysis. Gene dosage was performed (QMPSF or MLPA or CGH) in the parents to check if the remaniement was inherited. Out of three patients with subtelomeric gain and loss, FISH was only performed in one case and confirmed a balanced translocation in one parent.

		mutations		remaniement	
Patient	anomalies	father	mother	father	mother
10	Dup 1q	No	No	No	No
	Dup 10q	No	No	No	No
1	Del 2p	No	No	No	No
21	Dup 3p	Yes	no	Yes	No
22	Dup3q	No	No	Yes	No
23	Dup 6p	No	No	No	Yes
	Del 7q	No	No	No	Yes
11	Del 6q	No	No	No	No
12	Del 6q	No	No	No	No
3	Del 7qter	No	No	No	balanced
	Dup 7pter	No	No	No	balanced
25	Dup 7q	No	No	No	Yes
26	Dup 8q	No	No	No	Yes
30	Del 10p	Yes	No	Yes	No
13	Dup 12p	No	No	Yes	No
6	Del 13q	No	No	No	No
7	Del 13q ter	No	No	No	No
31	Del 14q	No	No	No	No
2	Dup 17q	No	No	No	No
32	Del 18p	No	No	Yes	No
	Dup 18p	No	No	Yes	No
8	Del 18pter	No	No	No	No
9	Del 20p	No	No	No	No
	Dup 21q	No	No	No	No
18	Del 21q	No	No	No	No
36	DupXp	No	No	Yes	No