Human Molecular Genetics, 1999, Vol. 8, No. 9 1769–1777

Point mutations throughout the GLI3 gene cause Greig cephalopolysyndactyly syndrome

Martha Kalff-Suske, Anja Wild, Juliane Topp, Martina Wessling, Eva-Maria Jacobsen, Dorothea Bornholdt, Hartmut Engel, Heike Heuer, Cora M. Aalfs[1], Margreet G. E. M. Ausems[2], Rita Barone[3], Andreas Herzog[4], Peter Heutink[5], Tessa Homfray[6], Gabriele Gillessen-Kaesbach[7], Rainer König[8], Jürgen Kunze[9], Peter Meinecke[10], Dietmar Müller[11], Renata Rizzo[3], Sibylle Strenge[12], Andrea Superti-Furga[13] and Karl-Heinz Grzeschik[+]

Medizinisches Zentrum für Humangenetik, Philipps-Universität Marburg, D-35037 Marburg, Germany, [1]Afdeling Klinische Genetica, Academisch Medisch Centrum, Universiteit van Amsterdam, NL-1105 AZ Amsterdam, The Netherlands, [2]Klinisch Genetisch Centrum Utrecht, NL-3501 CA Utrecht, The Netherlands, [3]Clinica Pediatrica, Universita di Catania, I-95125, Catania, Italy, [4]D-06526 Sangerhausen, Germany, [5]Klinische Genetica, Erasmus Universiteit Rotterdam, NL-3000 DR Rotterdam, The Netherlands, [6]South West Thames Regional Genetics Service, St George's Hospital Medical School, London SW17 ORE, UK, [7]Institut für Humangenetik, Universitätsklinikum Essen, D-45122 Essen, Germany, [8]Institut für Humangenetik, Johann Wolfgang Goethe-Universität, D-60590 Frankfurt, Germany, [9]Institut für Humangenetik, Virchow-Klinikum, Humboldt-Universität, D-13353 Berlin, Germany, [10]Abteilung Medizinische Genetik, Altonaer Kinderkrankenhaus, D-22763 Hamburg, Germany, [11]Säuglingsklinik, Klinikum Chemnitz, D-09009 Chemnitz, Germany, [12]Institut für Humangenetik, Universität Leipzig, D-04103 Leipzig, Germany and [13]Kinderspital Zürich, Universitäts-Kinderklinik, CH-8032 Zürich, Switzerland

Received April 27, 1999; Revised and Accepted June 8, 1999

Greig cephalopolysyndactyly syndrome, characterized by craniofacial and limb anomalies (GCPS; MIM 175700), previously has been demonstrated to be associated with translocations as well as point mutations affecting one allele of the zinc finger gene GLI3. In addition to GCPS, Pallister–Hall syndrome (PHS; MIM 146510) and post-axial polydactyly type A (PAP-A; MIM 174200), two other disorders of human development, are caused by GLI3 mutations. In order to gain more insight into the mutational spectrum associated with a single phenotype, we report here the extension of the GLI3 mutation analysis to 24 new GCPS cases. We report the identification of 15 novel mutations present in one of the patient's GLI3 alleles. The mutations map throughout the coding gene regions. The majority are truncating mutations (nine of 15) that engender prematurely terminated protein products mostly but not exclusively N-terminally to or within the central region encoding the DNA-binding domain. Two missense and two splicing mutations mapping within the zinc finger motifs presumably also interfere with DNA binding. The five mutations identified within the protein regions C-terminal to the zinc fingers putatively affect additional functional properties of GLI3. In cell transfection experiments using fusions of the DNA-binding domain of yeast GAL4 to different segments of GLI3, transactivating capacity was assigned to two adjacent independent domains (TA_1 and TA_2) in the C-terminal third of GLI3. Since these are the only functional domains affected by three C-terminally truncating mutations, we postulate that GCPS may be due either to haploinsufficiency resulting from the complete loss of one gene copy or to functional haploinsufficiency related to compromised properties of this transcription factor such as DNA binding and transactivation.

INTRODUCTION

Polydactylies are caused by disturbances of anterior/posterior patterning during limb development. The Greig cephalopolysyndactyly syndrome (GCPS; MIM 175700) is a rare autosomal dominant disorder affecting limb and craniofacial development in humans. GCPS-affected individuals are characterized by post-axial and pre-axial polydactyly as well as syndactyly of hands and feet, macrocephaly, a broad nasal root with mild hypertelorism and a prominent forehead (1). Inter- and intrafamilial variability of the phenotype commonly is observed (2). Haploinsufficiency of the human zinc finger protein gene *GLI3*, located on chromosome 7p13, has been considered as a cause for the GCPS phenotype, because deletions eliminating this region as well as translocations interrupting the gene were detected in GCPS patients (3,4). *GLI3*

[+]To whom correspondence should be addressed. Tel: +49 6421 286232; Fax: +49 6421 288920; Email: grzeschi@mailer.uni-marburg.de

forms, together with *GLI1* and *GLI2*, a gene family characterized by multiple regions of sequence similarity, with the central DNA-binding domain composed of five zinc finger motifs showing the highest degree of identity (5). The relative order and relative location of the homologies within each of the proteins is maintained.

By assigning GLI3 a role as a potential developmental regulator (4), the way was paved for intensive studies of its gene family in humans and a broad range of model organisms (6). Most of our current understanding of the role of GLI family proteins derives from the analysis of Cubitus interruptus (Ci), the single Gli homolog in *Drosophila melanogaster* (7). The Ci protein fulfills within the Hedgehog (Hh) developmental pathway multiple tasks as a transcriptional activator or repressor translating Hh signals into anterior/posterior positional information. In the absence of the Hh signal, Ci is part of a cytoplasmic complex with the protein kinase Fused (Fu) and with Suppressor of fused [Su(fu)], anchored at the microtubules through the kinesin-related protein Costal-2 (Cos-2). This association leads to targeting of Ci to the proteasome where it is cleaved to release an N-terminally truncated form which appears to enter the nucleus and act as transcriptional repressor (7,8). In contrast, the reception of the Hh signal leads to activation of Fu, which triggers the dissociation of Su(fu) and Ci, possibly through Su(fu) degradation. It also opposes the inhibitory activity of Cos-2 by releasing it from the microtubules (9). Consequently, Ci processing is reduced, full-length Ci accumulates and the transcription of Hh target genes is activated, presumably by full-size Ci protein (10,11). Homologous genes acting in a similar mode in various animals suggest that this pathway is one of the basic, highly conserved tools used to generate pattern during development (6). However, the situation in vertebrates is complicated by the existence of the three paralogous GLI family members (Gli1, Gli2 and Gli3) which might share all or part of the functions assigned to Ci. Studies of expression patterns during limb development of several vertebrate model organisms indicated that Gli1 might act preferentially as transcriptional activator, close to Sonic hedgehog signal release, whereas Gli3, expressed at more anterior sites, possibly functions as a repressor of target genes (12–14). These observations suggested that the limb phenotype in GCPS might result from an impairment of the repressor capacity of GLI3, possibly located in its N-terminal segment.

To contribute to the understanding of the role of human *GLI3* during limb development, we analyzed mutations of this gene in polydactyly syndromes. Previously, we associated two point mutations with GCPS. The nonsense mutation Q496X generates a stop codon truncating the protein in the C–H link of the first zinc finger, and a missense mutation P707S maps to a highly conserved putative phosphorylation site C-terminally of the zinc finger domain (ZFD) (15). Two other human developmental disorders, Pallister–Hall syndrome (PHS; MIM 146510) and post-axial polydactyly type A (PAP-A; MIM 174200), whose single overlapping feature with GCPS is polydactyly at the posterior side of the limbs were also attributed to GLI3 point mutations. In two families with autosomal dominant PHS (16) and a large PAP-A family (17), frameshift mutations were found that result in GLI3 proteins truncated C-terminally of the ZFD. The small number of identified truncation mutations appeared to fall into categories with respect to known and presumed functions of the GLI3 protein. The hypothesis was derived that in GCPS, N-terminal GLI3 protein moieties without a DNA-binding domain would be unable to function as a transcriptional repressor. Frameshift mutations in PHS and PAP-A truncating the protein after the ZFD would leave the DNA-binding and N-terminal functions intact (18). This, however, did not take into account the second, more C-terminal, point mutation we described in GCPS (P707S).

To gain more insight into the mutational spectrum in GCPS and the corresponding molecular lesions of GLI3, we have extended our mutation analysis to 24 new cases. Here, we report the identification of 15 novel mutations distributed throughout the coding *GLI3* gene regions implying that impairment of functions other than DNA-binding may cause GCPS. In order to determine which functional properties may be affected by the C-terminal mutations we observe in GCPS, we have analyzed the potential of different segments of this DNA-binding factor to act as transcriptional activator, a function predicted by the dual role of Ci (6,7). Two independent domains of GLI3 appear to have retained the potential to activate target genes. In contrast to the conclusions drawn from expression patterns in vertebrate embryonal tissues, both our functional studies and the observed C-terminal mutations suggest that an impairment of the activating capacity of GLI3 might be involved in the etiology of GCPS.

RESULTS

Novel *GLI3* mutations detected in GCPS

We have screened PCR-derived fragments spanning the complete coding region and the exon–intron boundaries of the 15 exon *GLI3* gene in DNA from 24 unrelated GCPS patients for mutations. PCR products that exhibited altered banding pattern in the single strand conformation analysis (SSCA) were compared with probes of unaffected family members and 100 control individuals. Two SSCA variants detected in the intervening sequences appeared in control individuals, as well, and were considered to represent polymorphisms [c.368–19G→A or IVS3–19G→A and c.1242+12C→G or IVS8+12C→G (data not shown)]. Several previously described (15,19) as well as new polymorphisms were detected in exons II, V, IX and XV. The following nucleotide exchanges represent novel polymorphic markers: c.39G→A, c.1320T→G and c.4609C→T. While the first polymorphism is a wobble polymorphism (K13K), the two others lead to the amino acid exchanges D440E and R1537C. These coding region polymorphisms are present at frequencies of 1, 5 and 12%, respectively in the control population. In nine cases, no mutation was identified in the coding *GLI3* sequences from karyotypically normal GCPS patients. Absence of major deletions within the segment of chromosome 7q13 carrying the *GLI3* gene was ascertained by fluorescence *in situ* hybridization (FISH) with yeast artificial chromosome (YAC) clone 32ID10 or detection of heterozygosity at coding region polymorphic sites (data not shown).

In 15 cases, a causative mutation within the *GLI3* gene could be identified. The specific alterations observed in the patients studied are listed in Table 1. These mutations are all present in a heterozygous state. The majority are truncating mutations, including three nonsense and six frameshift mutations. Additional changes include missense and splicing mutations. Several functional domains encoded by the *GLI3* gene are possibly

Table 1. *GLI3* mutations in GCPS patients identified in this study

Cases[a]	Nucleotide position of mutation	Amino acid position	Type of mutation
A (3)	c.473+1G→A; IVS4+1G→A		Donor splice site
B	c.706G→T	E236X	Nonsense
C	c.924delC	M309X	Frameshift at codon 308
D	c.931delA	L346X	Frameshift at codon 311
E	c.1095–1096insA	E411X	Frameshift at codon 366
F (2)	c.1497+1G→C; IVS10+1G→C		Donor splice site
G (1)	c.1497+2T→G; IVS10+2T→G		Donor splice site
H (1)	c.1543T→G	C515G	Missense
I (2)	c.1559G→A	C520Y	Missense
J (3)	c.1627G→T	E543X	Nonsense
K (2)	c.2374C→T	R792X	Nonsense
L	c.2424A→G	I808M	Missense
M (2)	c.3503delG	L1205X	Frameshift at codon 1168
N (4)	c.4291–4292insG	N1435X	Frameshift at codon 1431
O (3)	c.4359delA	V1487X	Frameshift at codon 1453

Numbering of nucleotide and amino acid positions is according to ref. 42.
For the position of the mutations relative to the *GLI3* gene structure, see Figure 3.
[a]For familial cases, the number of analyzed individuals affected with the mutation is indicated in parentheses.

affected, given the size of the deduced prematurely terminated proteins or the location of the detected amino acid exchanges.

Mutations in the N-terminal part of GLI3 including the ZFD

The nonsense mutation E236X (case B) and the three frameshift mutations at codon 308 (M309X; case C), codon 311 (L346X; case D) and codon 366 (E411X; case E) should remove most regions of the 1580 amino acid wild-type protein. The mutation involving the G of the invariant GT sequence of the donor splice site at position +1 of intron 4 (case A; IVS4+1G→A) may also fall within this category.

Within the DNA-binding domain composed of five zinc fingers extending from amino acid 462 to 645 encoded by exons X–XIII of the *GLI3* gene, two missense mutations were mapped. They involve the first and the second cysteine residue of the second zinc finger, respectively (case H, C515G; and case I, C520Y). A single truncating mutation, the nonsense mutation E543X (case J), leads to a protein with only the first two zinc fingers. Two splicing mutations involving the invariant GT of the 5' splice site of intron 10 were identified in independent familial cases that exhibit inter- as well as intrafamilial variations in the expression of the GCPS traits (case F, IVS10+1G→C; and case G, IVS10+2T→G).

Mutations in the C-terminal moiety of GLI3

Mutations located C-terminally to the ZFD were detected in a sporadic case and in four families in which the affected individuals exhibit the typical variable expressivity of GCPS. The missense mutation I808M (case L) leaves the protein intact, exchanging only a single amino acid. A nonsense mutation R792X (case K) truncates GLI3 after the DNA-binding domain. Three frameshift mutations remove sequentially larger parts of the C-terminus. A frameshift due to a single nucleotide deletion at codon 1168 gives rise to 37 altered amino acids before premature termination (case M; L1205X). A single nucleotide insertion of G between the *GLI3* cDNA positions 4291 and 4292 was detected in all analyzed affected members of family N. This mutation results in a frameshift creating four mutant amino acid residues following residue 1430 (N1435X). Similarly, a single nucleotide deletion in codon 1453 produces a mutant protein (V1487X) with 34 changed amino acids in consequence of the frameshift in the three affected individuals of case O.

Identification of two autonomous transactivation domains in the C-terminal part of GLI3

In order to obtain experimental evidence for a possible role of GLI3 as transcriptional activator which might be compromised by *GLI3* mutations, we examined the capacity of different segments of GLI3 to direct GAL4-binding site-dependent transcriptional activation. Fusion constructs were transfected together with a constant amount of GAL4-dependent luciferase reporter into non-small cell lung cancer NCI-H661 cells that express GLI3 endogeneously. Linear concentration dependence was seen within a range of 10–1000 ng of co-transfected expression plasmid (data not shown). The LUC activities obtained upon co-transfection of 100 ng of expression plasmids are given in Figure 1. The GAL4 domain alone caused a minimal increase in luciferase activity, which was assigned a value of 1. Fusion of the entire moiety of GLI3 C-terminal to the ZFD (GAL4–GLI3$^{626-1580}$) resulted in a 5.6-fold induction of LUC activity. In contrast, a fusion of the N-terminus of GLI3 without the ZFD (GAL4–GLI3$^{18-428}$) decreases transcriptional activation through the GAL4-binding sites below the level obtained with GAL4 alone. LUC induction was also completely lost upon removal of 559 C-terminal amino acids leaving central portions of the GLI3 protein (regions 3, 4 and 5) fused to the heterologous DNA-binding domain in GAL4–

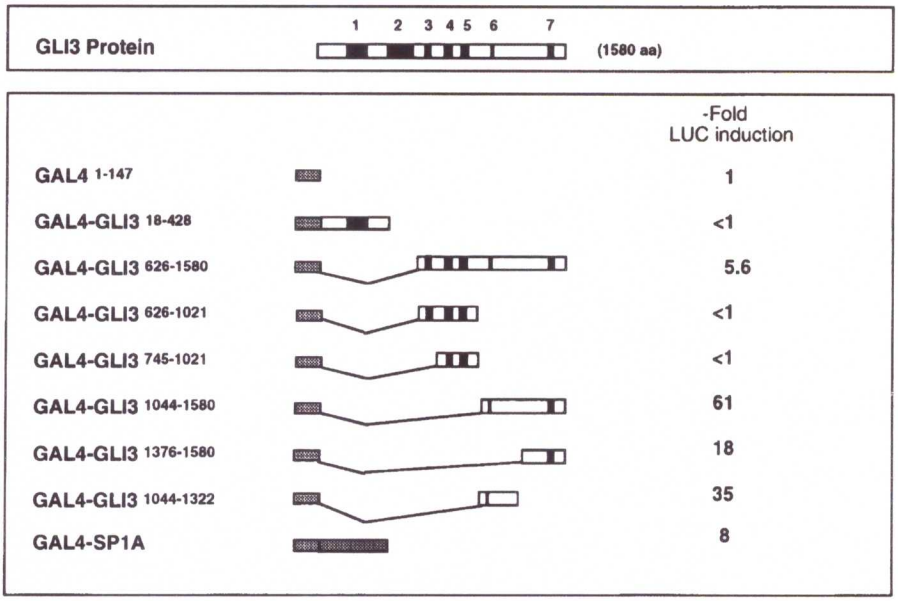

Figure 1. Deletion analysis of GLI3 linked to the GAL4 DNA-binding domain reveals two independent C-terminal transactivation domains. The top panel gives a schematic representation of the GLI3 protein (1580 amino acids). The regions of similarity between members of the human GLI family as defined in ref. 5 are shown as filled boxes with arabic numerals identifying them above. Region 2 includes the ZFD. In the bottom panel, the DNA-binding domain of GAL4 (amino acids 1–147) given as stippled bars was fused with various segments of GLI3 as indicated. The names of the resulting constructs and the GLI3 residues they contain are on the left. Expression plasmids (0.1 μg each) were co-transfected with the GAL4-dependent LUC reporter gene G5E1bLUC (1 μg) into NCI-H661 cells and were tested for stimulation of LUC activity. Firefly LUC activities were corrected for transfection efficiency as measured by *Renilla* LUC. Values obtained for each expression construct are given relative to the value obtained for GAL4 alone, set arbitrarily at 1. Values were averaged from at least two independent sets of transfection experiments, with deviations <20%.

GLI3$^{626-1021}$ and GAL4$^{725-1021}$. However, direct fusion of a segment encompassing the C-terminal part in GAL4–GLI3$^{1044-1580}$ created a potent transactivator yielding a >60-fold LUC induction. Co-transfection of a GAL4 fusion construct of a strong transactivation domain of another zinc finger transcription factor, the glutamine-rich A domain of Sp1, for comparison, resulted in 8-fold activation of the GAL4-dependent LUC reporter activity in NCI-H661 cells.

To map the transactivation domain more closely, fusion constructs containing adjacent segments of this GLI3 region were examined. The construct containing solely the region of residues 1376–1580 of GLI3 activated GAL4-binding site-driven LUC activity 18-fold. This transcription activation domain encoded within the C-terminal 204 amino acid residues of GLI3 is called TA_1. Direct fusion of the preceding residues 1044–1322 of GLI3 to the GAL4 DNA-binding domain showed a 35-fold LUC induction. This stretch of 278 amino acids within GLI3, called TA_2, harbors a second autonomous region with transactivation potential.

DISCUSSION

Structural comparison of the human GLI proteins reveals six regions of similarity besides the highly conserved ZFD (5) (Fig. 2). Consistent with the generally modular structure of transcription factors, these regions may reflect functions common to the gene family. Whereas the role of the ZFD in binding to specific DNA sequences has been analyzed extensively for vertebrate GLI proteins (20–22), the identification and characterization of the other functional domains is only beginning and is guided mainly by comparison with the *Drosophila* homolog Ci as illustrated in Figure 2. A repression activity has been ascribed to the N-terminal parts of Ci and GLI3. It is not clear, however, whether the GLI3 region numbered 1 that has conserved sequence similarity to Ci is involved in repression.

The Ci domains responsible for its post-translational modifications including steps governing subcellular compartmentalization of the full-length and the proteolytically processed form of the *Drosophila* protein have been localized C-terminally to the ZFD (8) (Fig. 2). While the processed Ci form acts as a transcriptional repressor, the full-length protein appears to activate target genes. Protein kinase A phosphorylation has been involved in the regulation of the activity and proteolysis of Ci (23). Initially, full-length GLI3 expression constructs were reported to exert only negative but not positive transcriptional regulation on artificial GLI-binding sites (24). Our analysis of specific fragments of the GLI3 protein to study potential activation capacities yields evidence for two adjacent but independently acting domains of GLI3 located around regions 6 and 7 which appear to share the function of activating target genes with Ci (Fig. 2). The identification of transactivation potential within the most C-terminal domain TA_1 (amino acids 1376–1580) supports a notion of Ruppert *et al.* (5). On the basis of primary sequence, these authors predicted an α-helical region in GLI3 between amino acids 1494 and 1512 corresponding to the seventh region of sequence similarity between human GLI1 and GLI3. This structure has similarity to well-established acidic activation domains such as that of

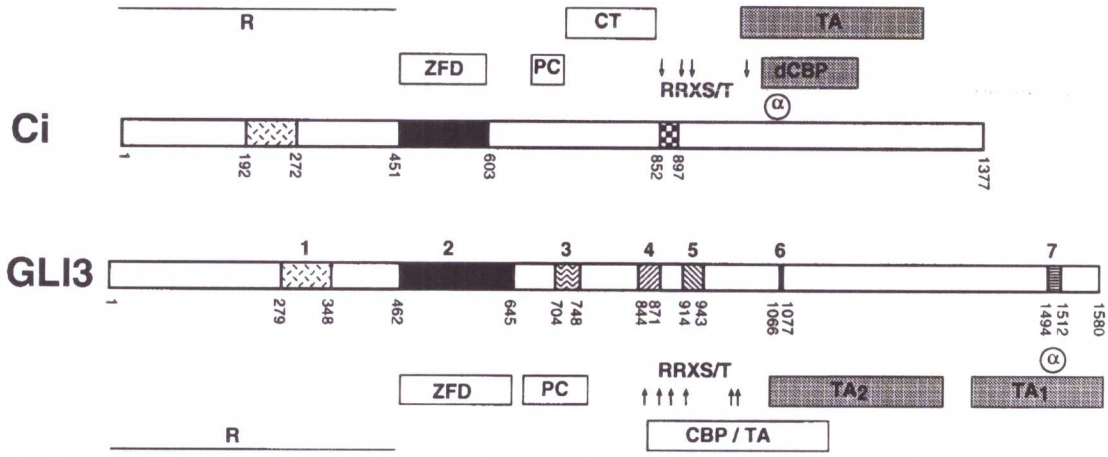

Figure 2. Sequence and functional homologies of *Drosophila* Ci and human GLI3 proteins. The Ci protein and GLI3 proteins are depicted as bars. The boxes within GLI3 represent the seven regions of similarity between human GLI proteins originally defined by Ruppert *et al.* (5). These regions are identified by arabic numerals with region 2 including the ZFD. The amino acid positions delimiting these regions are indicated. Two regions of Ci (accession no. X54360.1) besides the central ZFD show sequence conservation with the human GLI3 protein as indicated. An N-terminal region shows similarity to the region numbered 1 in GLI3, whereas a C-terminal region corresponds to the GLI3 residues between the C-terminus of region 4 and the N-terminus of region 5 (amino acids 872–913). The currently identified structural and functional features are indicated above each protein. For Ci, the repression and activation regions (R and TA) (10), an acidic α-helical region (α) (26,29), the dCBP-binding region (dCBP) (43), the protein kinase A phosphorylation sites (RRXS/T) (23), and the regions for proteolytic cleavage and for cytoplasmic tethering (PC and CT) (8) are given. For GLI3, the two independent transactivation domains TA_1 (amino acids 1376–1580) and TA_2 domains (amino acids 1044–1322) identified in this study, the transactivation and CBP-binding regions (TA/CBP) reported by Dai *et al.*, (27) the consensus protein kinase A phosphorylation sites clustered between amino acids 846 and 1006, and the location of the site presumed to be responsible for protein cleavage (PC) (27) are indicated. An acidic α-helical region overlaps with region 7.

herpes simplex VP16 (25). Experimental evidence that the corresponding α-helical domain acts as an activation domain was obtained recently for the paralogous gene *GLI1* (26). The presence of a transactivation activity mediated by an acidic type domain in both GLI1 and GLI3 thus appears to provide them with a common mechanism to increase transcription of genes targeted by their ZFDs.

In this study, we demonstrate the existence of a second independent domain capable of transcriptional activation within the region encompassing amino acids 1044–1322 (TA_2) (Fig. 2). Consistent with our data, Dai *et al.* (27) very recently reported the identification of a specific domain of GLI3 with transcriptional activation potential (amino acids 827–1132). This region overlapped with a CBP-binding module (amino acids 827–1180), considered to foster transcriptional activation of target genes. Combining the results of these authors (transactivation and CBP binding: amino acids 827–1132) with our findings (transactivation: amino acids 1044–1322) allows prediction of the location of the minimal sequence requirements for CBP-mediated transactivation to residues 1044–1132. This region contains the motif PSI[S/T]EN conserved among GLI1, GLI2 and GLI3 [sequence similarity region 6 (5)] embedded within different sequence environments. Interestingly, CBP is reported to bind to GLI3 but not to GLI1 (27).

In recent reports, activation of GLI3 target genes was also shown in transfection experiments with full-length GLI3 expression constructs [mGli1 (27) and PTCH1 (28)]. The question of whether both TA domains are active on all target genes or only on a subset or under defined conditions remains to be addressed. It is not known whether the TA domain identified in Ci can also be split in two autonomous subdomains. In contrast to GLI3, where the acidic region compatible with the formation of an α-helix is located at the very C-terminus (marked α in Fig. 2), a similar structure is found in the *Drosophila* Ci protein N-terminal to the CBP-binding region (26,29). Thus, besides obvious parallels in the activation function of GLI3 and Ci, these observations point to possible differences in the mechanism. The presence of a domain mediating transcriptional repression postulated in the N-terminal part of the GLI proteins (Fig. 2) (26,27) is neither ruled out nor confirmed by our results because the observed expression levels of the N-terminal fusion constructs were too low for accurate quantitation.

Proof for *GLI3* being responsible for GCPS is provided by the identification of heterozygous splice, missense, nonsense and frameshift mutations of this gene in patients with the characteristic phenotype (Fig. 3). Observing sites of mutations dispersed throughout the whole *GLI3* coding sequence, we show that GCPS cannot, exclusively, be associated with loss of the DNA-binding domain, as predicted previously (18). Truncating lesions scattered over *GLI3* are induced by at least 10 out of 16 mutations. In four cases, so far, GCPS appears to be caused by missense mutations affecting different regions of the gene. The nature of most alterations indicates loss of all or some functions of the protein. However, with the exception of mutations affecting the DNA-binding ZFD, it is not apparent which function of GLI3 might be impaired.

The group of C-terminal mutations observed so far is quite heterogeneous, most probably affecting several of the putative

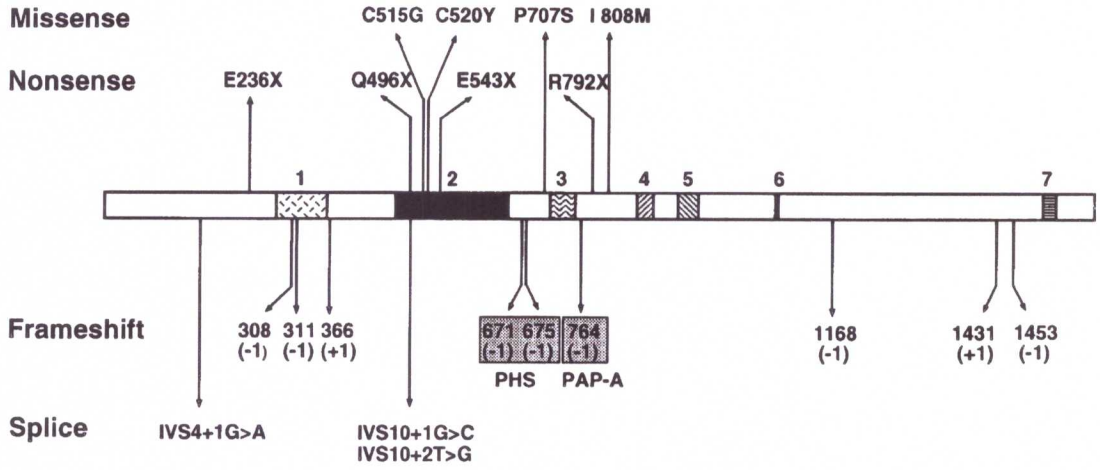

Figure 3. Summary of the presently known *GLI3* mutations in GCPS, PHS and PAP-A. The GLI3 protein is drawn schematically as in Figure 2. The nature of the mutations is indicated on the left. Their position within the protein is indicated by numbers specifying the codons or amino acid residues, given as the single letter code. The number of bases inserted or deleted in frameshift mutations is given in parentheses. For splicing mutations, the altered position within the intervening sequence (IVS) is given. Data are from this report (Table 1), ref. 15 (GCPS: Q496X, P707S); ref. 16 [PHS: 671(−1) and 675(−1)] and ref. 17 [PAP-A: 764(−1)].

functional properties of GLI3 described above. The nonsense mutation R792X leaves the DNA-binding domain intact. Premature termination occurs between regions 3 and 4. According to the interpretation proposed by Biesecker (18), this should result in PAP-A. However, the phenotype resulting from this mutation clearly includes GCPS symptoms not listed for PAP-A. Assuming that human GLI3 is subject to intracellular compartmentalization and/or post-translational processing as suggested recently (27,28), mutations affecting a putative site that tethers full-length GLI3 in the cytoplasm might cause the constitutive release of a transcriptional repressor form that is able to translocate to the nucleus. The GCPS mutants I808M from this work and P707S described by Wild *et al.* (15) and the known PAP-A mutation (17) may be due to an functional impediment of one or both these processes. However, experimental evidence for the retention of these functions, as analyzed in Ci, is still lacking for GLI3.

The missense mutation I808M falls slightly N-terminal to the region containing six putative protein kinase A phosphorylation sites [RRXS/T consensus (30)] clustered between amino acids 846 and 1006 as indicated in Figure 2. The primary sequence immediately surrounding the site mutated in I808M is conserved between GLI3 proteins of human (5; accession no. M57609), mouse (31; accession no. X95255 and *Xenopus* (13; accession no. U42461) but differs considerably from GLI1 and GLI2, suggesting that this region may be critical for a GLI3-specific property.

Three frameshifts truncate segments of different extensions from the C-terminus. With the identification of the TA domains of GLI3, the four frameshift mutations (R792X, L1205X, V1487X and N1435X) may now be functionally explained. These frameshift mutations completely or partially remove the two TA domains. Even mutants that retain TA_2 display a phenotype. This domain, although able to activate promoters independently of TA_1 through heterologous recognition sequences such as GAL4-binding sites, might activate natural target genes only in concert with TA_1. The C-terminally truncated mutant proteins, provided that they are stable, support the notion that the activation domains are required for proper function of the normal GLI3 protein. In this case, activation would constitute an essential role for GLI3. Alternatively, the mutant protein with an intact zinc finger but lacking the TA domains may influence wild-type protein expressed from the non-affected allele and/or other GLI factors in a dominant-negative manner by the occupation of their binding sites through mutant GLI3 proteins. The potential of C-terminally truncated GLI3 for repression should not be affected.

The majority of identified mutations map within the N-terminus and the central ZFD of GLI3 (Fig. 3). Having lost the capacity to bind DNA, they might behave as null mutants, compatible with the proposed role of haploinsufficiency in this disorder. The nonsense mutation E236X and the three frameshift mutations M309X, L346X and E411X should result in loss of most of the functionally important regions including the DNA-binding domain. If the splicing mutation involving the G of the invariant GT sequence of the donor splice site at position +1 of intron 4 (c.473+1G→A; IVS4+1G→A) should lead to a premature translational stop, a severely truncated protein ensues. Exon skipping, which is the preferred pattern of aberrant splicing when the 5' splice site is disrupted (32,33), would introduce the chain-terminating amber codon at position 215. While it is tempting to attribute the molecular defect to the out-of-frame deletion of exon IV, alternative splicing patterns cannot be ruled out.

Within the DNA-binding domain composed of five zinc fingers (Fig. 3), missense mutations involve the first and the second cysteine residue of the second zinc finger, respectively (C515G and C520Y). The absence of one of the cysteine residues in the finger motif is expected to compromise the tetrahedral coordination of the zinc atom. In addition to the nonsense mutation Q496X we described previously (15), we have now detected a second mutation truncating the ZFD. The nonsense mutation E543X occurs within the H–C link (the amino acid sequence connecting the histidine of one finger to the cysteine

of the next) between the second and third zinc finger, leading to a protein that lacks three of the five zinc finger motifs, probably unable to bind DNA specifically. Crystallographic analyses of GLI1 have revealed that while zinc finger 1 does not contact the DNA, fingers 2–5 bind in the major groove of the helix, with fingers 4 and 5 making extensive base contacts in the 9 bp consensus GLI recognition site (21).

The mutations involving the invariant GT of the 5' splice site of intron 10 (IVS10+1G→C and IVS10+2T→G) may also result in a mutant GLI3 with reduced or defective DNA binding depending on the adopted aberrant splicing pattern. Exon skipping would cause an in-frame deletion of exon X that encodes the first zinc finger. While this finger is not involved directly in DNA contacts, it is known to form extensive protein–protein interactions with finger 2 (21) and may influence the stability or specificity of the recognition site binding. Alternative splicing patterns may lead to premature termination. Altogether, translation products without a functional DNA-binding domain could not fulfill any of the tasks assigned to GLI proteins in Hh signaling, not even the function of a transcriptional repressor expected to be mediated by sequences in the most N-terminal domain. In addition, the mutations within the ZFD may interfere with other putative functions of GLI3. Recently, Smad proteins that have a role in transforming growth factor-β signaling have been shown to interact with a region of the murine Gli3 protein adjacent to and partly overlapping the ZFD (34).

Our *GLI3* mutation screen extended to a larger number of GCPS cases demonstrates that this phenotype is not only caused by mutations that impair solely the DNA-binding activity, as hypothesized by Bieseker (18). Instead, it seems that GCPS involves a larger spectrum of functions, specifically those relating to transcriptional activation by this factor. We identified in 15 of 24 cases mutations within the structural regions of the gene. Any attempt at a phenotype–genotype correlation for GCPS symptoms and GLI3 mutations is bound to be complicated by intrafamilial and even intraindividual phenotypic variability. Mouse mutants on a uniform genetic background presumably might be helpful to resolve this issue through the detection of modifying genes. The mouse *extra toes* mutation (Xt), a deletion of part of Gli3, originally described by Johnson (35), exhibits considerable phenotypic variability of the affected feet in Xt/+ outcrossed to CB mice (the F_1 of CBA/Gr and C57BL/Gr). In addition to stochastic events, phenotypic heterogeneity might be attributed to modifying interaction partners, in particular within the Hh signaling cascade, and to paralogy.

The unresolved cases may be attributed in part to the detection rate of the applied screening method; however, other explanations need to be taken into consideration. Mutations affecting the proper *GLI3* mRNA level required for normal temporal and spatial development might cause the phenotype but would remain unnoticed by the present mutation search. In addition, it cannot be excluded that phenotypic manifestations of GLI1 and GLI2 structural or regulatory mutants are coincident or overlapping in nature to GCPS. Both specific and overlapping functions and expression of murine Gli2 and Gli3 during development recently have been shown by analyzing knockout mice (36). Mutations in other *GLI* genes (notably *GLI2* on the basis of higher degree of similarity to *GLI3*) might give rise to phenotypes that share some or all characteristics with GCPS.

Besides the six GLI3 mutations located C-terminally to the ZFD found in GCPS cases [five from this study and P707S (15)], two were found in familial PHS cases (16) and a single one in a large family with PAP-A (17), as depicted in Figure 3. Additional GLI3 mutations were detected recently in pre-axial polydactyly type IV and post-axial polydactyly type B (U. Radhakrishna, in preparation). On comparing the GLI3 mutations observed in these different syndromes, no simple obvious genotype–phenotype correlation emerges. One might speculate that these syndromes are phenotypic subtypes of GCPS associated with mutations affecting specific functions within the various tasks of GLI3 and/or its expression pattern. The mutations known so far obviously do not saturate the *GLI3* gene for the detection of all functionally important sites. Therefore, a further extension of the mutation analysis in GCPS as well as in other polydactyly syndromes is promising.

MATERIALS AND METHODS

Subjects

The patients with GCPS analyzed here were clinically examined at the referring institutions and included in the study after informed consent was obtained. The probands show all or some of the typical manifestions associated with the syndrome, including post-axial polysyndactyly of the hands, pre-axial polydactyly of the hands and feet as well as syndactylies and mild craniofacial abnormalities allowing unambiguous distinction of GCPS from other GLI3-associated syndromes such as PAP-A and PHS. Cases A, F, G, H, I, J, K, M, N and O (Table 1) are familial with classical GCPS. While the family history is available for two generations in families A, F, G, H, K, M and O, the pedigrees of families I, J and N extend over four, five and three generations, respectively. Expressivity varies considerably within families. In case F, the affected son shows mental retardation in addition to a GCPS phenotype. The analyzed proband in case A (20-year-old male) manifests, in addition to GCPS, gynecomasty and elevated 17α-hydroxyprogesterone levels. Cases C, D (37) and L represent sporadic cases with phenotypically inconspicuous parents. In two cases, the family history is either incompletely documented (B) or unknown (E). Cytogenetic analysis including FISH analysis with the 32ID10 YAC probe from within the *GLI3* gene (38) was performed to screen for microdeletions when living cells were available. DNA samples from GCPS patients and from 100 control individuals from the German population were purified using standard methods.

Exon amplification and SSCA

The 15 exons and the corresponding exon–intron boundaries of the *GLI3* gene were amplified by PCR using primers as described previously (15) with the following modifications for amplification primers: ExVIrev*, 5'-GCCATTTCCCAAGACTC-3'; ExVIIrev1*, 5'-GCTGAAGAGCTGCTACGG-3'; ExXIfor*, 5'-TGATGAATACGTTTCCATTTG-3'; ExXIrev*, 5'-AAGGACCCAAGTGTGCCTG-3'; ExXIIrev*, 5'-CCTTATGCAAGCTCCATGCC-3'; ExXIIIrev1*, 5'-GACCTGGACTGTGAATGGCTG-3'; ExXVrev12*, 5'-CTTGGTAGATGTT-

GATGTGTG-3'; ExXVfor15*, 5'-CTATGACCAAACCGT-GGGC-3'; and ExXVrev16*, 5'-GATTTCCGTTGGTT-GCAGTC-3'. SSCA was performed according to two protocols. Seven cases were screened by resolving [α-^{32}P]dCTP-labeled exon amplification products as described previously (15). For all subsequent cases, conditions were adapted for exon amplification using 'Ready-To-Go' PCR beads (Amersham Pharmacia Biotech). PCR products (3.5 µl) were diluted in glycerol-containing buffer prior to denaturation, and subsequently resolved on 12% acrylamide gels (49:1) at 10°C (250 V) and 20°C (150 V) for 16 h. Gels were silver stained. Sequencing of allele-specific and heterozygote DNA templates was essentially as described previously (15).

Correction of the size calculated for GLI3

In the course of sequencing exon XV–16 PCR products, the thymine residue at cDNA position 4646 (numbering according to ref. 5; accession no. M57609) was found to be absent not only from patient DNAs but also from wild-type genomic DNAs as well as from the original *GLI3* cDNA clone (5). This single nucleotide shift in codon 1549 of *GLI3* engenders the translation of a 1580 amino acid protein with 32 altered C-terminal residues compared with the originally deduced protein of 1596 amino acids. This change to the original GLI3 sequences has been communicated to the GenBank/EMBL databases.

Plasmid constructions

In the GAL4–GLI3 expression constructs, the 147 N-terminal codons of the yeast transcription factor GAL4 were fused to various segments of GLI3. Expression of the fusion proteins is driven by the SV40 promoter. The constructs containing C-terminal GLI3 segments were generated as follows. The GLI3 fragments in question were isolated from pGLI3-bs2 (5) using the specified restriction enzymes, fused to the appropriate *Eco*RI linkers depending on the reading frame and ligated to the dephosphorylated 3.4 kb *Eco*RI-cut pGAL4-Sp1A (39). For pGAL4-GLI3$^{626–1580}$, the 3123 bp *Afl*III–*Eco*RI *GLI3* fragment was fused to 12mer *Eco*RI linkers; for pGAL4-GLI3$^{626–1021}$, the 1178 bp *Afl*III–*Bss*HI fragment was also linked to 12mers; for pGAL4-GLI3$^{745–1021}$, the 825 bp *Cla*I–*Bss*HI fragment was ligated to 10mer linkers; for pGAL4-GLI3$^{1044–1580}$, the 1860 bp *Bss*HI–*Cla*I fragment was used in connection with 8mers; and for pGAL4-GLI3$^{1376/1580}$, the 876 *Hin*dIII–*Eco*RI fragment was fused to 12mers prior to vector ligation. The pGAL4-GLI3$^{1044–1322}$ expression plasmid was obtained by *Kpn*I–*Xba*I restriction of pGAL4-GLI3$^{1044–1580}$ followed by Klenow fill-in and religation reactions. The GAL4 fusion construct containing the GLI3 segment N-terminal to the zinc fingers (amino acids 18–428) consists of an in-frame fusion of GAL4 sequences to the 1228 bp *Eco*RI–*Bst*EII fragment of pGLI3-bs2 in pGAL4-Sp3 (40). Maintenance of the correct *GLI3* reading frame in the fusion constructs was confirmed by sequencing.

The reporter plasmid G5E1bLUC is a derivative of G5E1bCAT (41) and was constructed as follows. A 130 bp *Hin*dIII (with Klenow-filled overhang)–*Bam*HI fragment containing five GAL4-binding sites fused to the E1a TATA box were isolated from G5E1b CAT and inserted into the *Sma*I–*Bgl*II-cut pGL3 basic vector (Promega). The pRL-SV40 plasmid (Promega) was used to assess transfection efficiencies.

Cell culture, transfections and luciferase assays

The human non-small cell lung cancer cell line, NCI-H661, was purchased from the American Type Culture Collection. The cells were propagated as monolayers in RPMI medium supplemented with 10% fetal calf serum. Cells were transfected by a lipofection method using a total of 2 µg of DNA. Expression plasmids (100 ng) with carrier DNA (pBSII; Stratagene) to make up 1 µg were co-transfected with 1 µg of firefly luciferase reporter plasmid and 25 ng of pRL-SV40 *Renilla* luciferase transfection efficiency control plasmid (Promega). DNA was incubated with 5 µl of Lipofectin reagent and Opti-MEM (Life Technologies) according to the manufacturer's instructions and added to 35 mm plates containing 1.5×10^5 cells. Opti-MEM medium was removed after 5 h and the incubation continued for 42 h prior to lysate preparation. The sequential assays of firefly and *Renilla* luciferases were performed according to the specifications of the manufacturer (Promega).

ABBREVIATIONS

GCPS, Greig cephalopolysyndactyly syndrome; IVS, intervening sequence; LUC, luciferase; PAP-A, postaxial polysyndactyly type A; PHS, Pallister–Hall syndrome; SSCA, single strand conformation analysis; TA, transcription activation; ZFD, zinc finger domain.

ACKNOWLEDGEMENTS

We are grateful to the patients and their families for their consent to this study, to our many medical colleagues for clinical, radiological and cytological evaluation of the referred patients, and to M. Koch (Marburg) for providing the DNA from control individuals. We wish to thank A. Hinney, A. Schelbert and G. Suske for advice and discussions. This work was supported by the Alfred und Ursula Kulemann-Stiftung (M.K.-S.) and the Deutsche Forschungsgemeinschaft (Gr 373/20-2).

REFERENCES

1. Greig, D.M. (1926) Oxycephaly. *Edinburgh Med. J.*, **33**, 189–218.
2. Gollop, T.R. and Fontes, L.R. (1985) The Greig cephalopolysyndactyly syndrome: report of a family and review of the literature. *Am J. Med. Genet.*, **22**, 59–68.
3. Wagner, K., Kroisel, P. and Rosenkranz, W. (1990) Molecular and cytogenetic analysis in two patients with microdeletions of 7p and Greig syndrome: hemizygosity for *PGAM2* and *TCRG* genes. *Genomics*, **8**, 487–491.
4. Vortkamp, A., Gessler, M. and Grzeschik, K.-H. (1991) *GLI3* zinc finger gene interrupted by translocations in Greig syndrome families. *Nature*, **352**, 539–540.
5. Ruppert, J.M., Vogelstein, B., Arheden, K. and Kinzler, K.W. (1990) *GLI3* encodes a 190-kilodalton protein with multiple regions of GLI similarity. *Mol. Cell. Biol.*, **10**, 5408–5415.
6. Ruiz i Altaba, A. (1997) Catching a Gli-mpse of Hedgehog. *Cell*, **90**, 193–196.
7. Ingham, P.W. (1998) Transducing Hedgehog: the story so far. *EMBO J.*, **17**, 3505–3511.
8. Aza-Blanc, P., Ramírez-Weber, F.-A., Laget, M.-P., Schwartz, C. and Kornberg, T.B. (1997) Proteolysis that is inhibited by Hedgehog targets Cubitus interruptus protein to the nucleus and converts it to a repressor. *Cell*, **89**, 1043–1053.
9. Monnier, V., Dusillol, F., Alves, G., Lamour-Isnard, C. and Plessis, A. (1998) Suppressor of Fused links Fused and Cubitus interruptus on the Hedgehog signalling pathway. *Curr. Biol.*, **8**, 583–586.

10. Alexandre, C., Jacinto, A. and Ingham, P.W. (1996) Transcriptional activation of hedgehog target genes in *Drosophila* is mediated directly by the Cubitus interruptus protein, a member of the GLI family of zinc finger DNA-binding proteins. *Genes Dev.*, **10**, 2003–2013.
11. Ohlmeyer, J.T. and Kalderon, D. (1998) Hedgehog stimulates maturation of Cubitus interruptus into a labile transcriptional activator. *Nature*, **396**, 749–753.
12. Marigo, V., Johnson, R., Vortkamp, A. and Tabin, C. (1996) Sonic hedgehog differentially regulates expression of Gli and Gli3 during limb development. *Dev. Biol.*, **180**, 273–283.
13. Marine, J.C., Bellefroid, E.J., Pendeville, H., Martial, J.A. and Pieler, T. (1997) A role for *Xenopus* Gli-type zinc finger proteins in the early embryonic patterning of mesoderm and neuroectoderm. *Mech. Dev.*, **63**, 211–225.
14. Büscher, D. and Rüther, U. (1998) Expression profile of Gli family members and Shh in normal and mutant mouse limb development. *Dev. Dyn.*, **211**, 88–98.
15. Wild, A., Kalff-Suske, M., Vortkamp, A., Bornholdt, D., König, R. and Grzeschik, K.-H. (1997) Point mutations in human GLI3 cause Greig syndrome. *Hum. Mol. Genet.*, **6**, 1979–1984.
16. Kang, S., Graham, J.M.Jr, Olney, A.H. and Biesecker, L.G. (1997) GLI3 frameshift mutations cause autosomal dominant Pallister–Hall syndrome. *Nature Genet.*, **15**, 266–268.
17. Radhakrishna, U., Wild, A., Grzeschik, K.-H. and Antonarakis, S.E. (1997) Mutation in GLI3 in postaxial polydactyly type A. *Nature Genet.*, **17**, 269–271.
18. Biesecker, L.G. (1997) Strike three for GLI3. *Nature Genet.*, **17**, 259–260.
19. Kang, S., Rosenberg, M., Ko, V.D. and Biesecker, L.G. (1997) Gene structure and allelic expression assay of the human *GLI3* gene. *Hum. Genet.*, **101**, 154–157.
20. Kinzler, K.W. and Vogelstein, B. (1999) The *GLI* gene encodes a nuclear protein which binds specific sequences in the human genome. *Mol. Cell. Biol.*, **10**, 634–642.
21. Pavletich, N.P. and Pabo, C.O. (1993). Crystal structure of a five-finger GLI–DNA complex: new perspectives on zinc fingers. *Science*, **261**, 1701–1707.
22. Vortkamp, A., Gessler, M. and Grzeschik, K.-H. (1995) Identification of optimized target sequences for the GLI3 zinc finger protein. *DNA Cell Biol.*, **14**, 629–634.
23. Chen, Y., Gallaher, N., Goodman, R. and Smolik, S.M. (1998) Protein kinase A directly regulates the activity and proteolysis of cubitus interruptus. *Proc. Natl Acad. Sci. USA*, **95**, 2349–2354.
24. Sasaki, H., Hui, C., Nakufuku, M. and Kondoh, H. (1997) A binding site for Gli proteins is essential for *HNF-3β* floor plate enhancer activity in transgenics and can respond to Shh *in vitro*. *Development*, **124**, 1313–1322.
25. Triezenberg, S.J., Kingsbury, R.C. and McKnight, S.L. (1988) Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Dev.*, **2**, 718–729.
26. Yoon, J.W., Liu, C.Z., Yang, J.T., Swart, R., Iannaccone, P. and Walterhouse, D. (1998) GLI activates transcription through a herpes simplex viral protein 16-like activation domain. *J. Biol. Chem.*, **273**, 3496–3501.
27. Dai, P., Akimaru, H., Tanaka, Y., Maekawa, T., Nakafuku, M. and Ishii, S. (1999) Sonic hedgehog-induced activation of the *GLI1* promoter is mediated by GLI3. *J. Biol. Chem.*, **274**, 8143–8152.
28. Shin, S.H., Kogerman, P., Lindstrom, E., Toftgard, R. and Biesecker, L.G. (1999) GLI3 mutations mimic *Drosophila* Cubitus interruptus protein functions and localization. *Proc. Natl Acad. Sci. USA*, **96**, 2880–2884.
29. Orenic, T.V., Slusarski, D.C., Kroll, K.L. and Holmgren, R.A. (1990) Cloning and characterization of the segment polarity gene cubitus interruptus Dominant of *Drosophila*. *Genes Dev.*, **4**, 1053–1067.
30. Kemp, B.E. and Pearson, R.B. (1990) Protein kinase recognition sequence motifs. *Trends Biochem. Sci.*, **15**, 342–346.
31. Thien, H., Büscher, D. and Rüther, U. (1996) Cloning and sequence analysis of the murine *Gli3* cDNA. *Biochim. Biophys. Acta*, **1307**, 267–269.
32. Talerico, M. and Berget, S.M. (1990) Effect of 5' splice site mutations on splicing of the preceding intron. *Mol. Cell. Biol.*, **10**, 6299–6305.
33. Nakai, K. and Hiroshi, S. (1994) Construction of a novel database containing aberrant splicing mutations of mammalian genes. *Gene*, **141**, 171–177.
34. Liu, F., Massagué, J. and Ruiz i Altaba, A. (1998) Carboxy-terminally truncated GLI3 proteins associate with Smads. *Nature Genet.*, **20**, 325–326.
35. Johnson, D.R. (1967) *Extra-toes*: a new mutant gene causing multiple abnormalities in the mouse. *J. Embryol. Exp. Morphol.*, **17**, 543–581.
36. Mo, R., Freer, A.M., Zinyk, D.L., Crackower, M.A., Michaud, J., Heng, H.H., Chik, K.W., Shi, X.M., Tsui, L.C., Cheng, S.H., Joyner, A.L. and Hui, C. (1997) Specific and redundant functions of Gli2 and Gli3 zinc finger genes in skeletal patterning and development. *Development*, **124**, 113–123.
37. Kunze, J. and Kaufmann, H.J. (1985) Greig cephalopolysyndactyly syndrome: report of a sporadic case. *Helv. Paediatr. Acta*, **40**, 489–495.
38. Vortkamp, A., Gessler, M., Le Paslier, D., Elaswarapu, R., Smith, S. and Grzeschik, K.-H. (1994) Isolation of a yeast artificial chromosome contig spanning the Greig cephalopolysyndactyly syndrome gene region. *Genomics*, **22**, 563–568.
39. Southgate, C.D. and Green, M.R. (1991) The HIV-1 Tat protein activates transcription from an upstream DNA-binding site: implications for Tat function. *Genes Dev.*, **5**, 2496–2507.
40. Dennig, J., Beato, M. and Suske, G. (1996) An inhibitor domain in Sp3 regulates its glutamine-rich activation domains. *EMBO J.*, **15**, 5659–5667.
41. Lillie, J.W. and Green, M.R. (1989) Transcription activation by the adenovirus E1a protein. *Nature*, **338**, 39–44.
42. Antonarakis, S.E. and the Nomenclature Working Group (1998) Recommendations for a nomenclature system for human gene mutations. *Hum. Mutat.*, **11**, 1–3.
43. Akimura, H., Chen, Y., Dai, P., Hou, D.-X., Nonaka, M., Smolik, S.M., Armstrong, S., Goodman, R.H. and Ishii, S. (1997) *Drosophila* CBP is a co-activator of cubitus interruptus in hedgehog signalling. *Nature*, **386**, 735–738.