Published in final edited form as:
Nat Genet. 2009 November ; 41(11): 1176–1178. doi:10.1038/ng.454.

T (Brachyury) gene duplication confers major susceptibility to familial chordoma

Xiaohong R. Yang[1,†], **David Ng**[1,†], **David A. Alcorta**[2], **Norbert J. Liebsch**[3], **Eamonn Sheridan**[4], **Sufeng Li**[2,5], **Alisa M. Goldstein**[1], **Dilys M. Parry**[1,‡], and **Michael J. Kelley**[2,5]

[1]Genetic Epidemiology Branch, Division of Cancer Epidemiology & Genetics, National Cancer Institute, NIH, DHHS, Bethesda, MD

[2]Department of Medicine, Duke University Medical Center, Durham, NC

[3]Department of Radiation Oncology, Massachusetts General Hospital, Boston, MA

[4]St. James University Hospital, Leeds, UK

[5]Durham Veterans Affairs Medical Center, Durham, NC

Abstract

Using high-resolution array-CGH, we identified unique duplications of a region on 6q27 in four multiplex families with ≥ 3 cases of chordoma, a cancer of presumed notochordal origin. The duplicated region contains only the *T* gene (Brachyury), which plays an important role in notochord development and is expressed in most sporadic chordomas. Our findings highlight the need to include screening for complex genomic rearrangements in searches for cancer susceptibility genes.

Genomic rearrangements, including rare copy number variations (CNVs), contribute significantly to disease susceptibility in sporadic genomic disorders and some inherited Mendelian diseases[1-6]. Although genomic rearrangements are one of the most common somatic genetic events in human cancers, germline CNVs, particularly inherited duplications that are the major mechanism for disease susceptibility, have seldom been reported in familial cancer syndromes. We provide here evidence that duplications of the *T* gene confer major susceptibility to familial chordoma.

Chordoma is a rare bone cancer that is believed to originate from notochordal remnants [7]. We previously reported genetic linkage to chromosome 7q33 in three multiplex chordoma families (Families 1-3) but not in a fourth chordoma family (Family 4, Supplementary Figure 1) [8,9]. Recently, we evaluated a new chordoma case in Family 1 (individual 43, Supplementary Figure 1). On magnetic resonance (MR) images the mass in her posterior nasopharynx was virtually identical in its location, appearance and signal characteristics to the clival chordoma that had been removed from her affected father (individual 20). However, she did not inherit the disease-related 7q33 haplotype from her father. Consequently, evidence of linkage to 7q33 in this

‡Corresponding author..
†These authors contributed equally to this work.

Author contributions:
XRY wrote the first draft of the paper; XRY and DN analyzed the data; DN, SL, MJK, and DAA performed the laboratory studies including genotyping, sequencing, and breakpoint evaluation; DN designed and conducted qPCR analyses; DMP initiated the clinical study and evaluated the chordoma families with DN; DMP, ES, and NJL collected the families; XRY, DN, MJK, DMP, and AMG planned the work and interpreted the results. All authors contributed to the final version of the paper.

Competing financial interests:
The authors declare no competing financial interests.

family decreased substantially (Supplementary Table 1a). This result led to a further search for chordoma susceptibility loci. In a subsequent genome-wide linkage scan using independent SNPs (Illumina 2.25K) in chordoma families 1-4, we identified six new regions with suggestive evidence for linkage. Fine mapping of the six candidate regions with microsatellite markers (STRs) further improved lod scores on 6q25-27 (3.04 in Family 1, Supplementary Table 1b), whereas the non-6q candidate regions became less interesting (lod scores 0.5-0.8). All affected individuals in Family 1, including individual 43, shared a common 6q disease-related haplotype between D6S972 and D6S503 (~11Mb). Only one unaffected individual (#44, 2 years old at MR evaluation) shared the 6q disease haplotype. Similarly, all chordoma cases and obligate gene carriers in Family 3 and the two sisters affected with chordoma in Family 4 shared a common haplotype in the 6q region. The minimal disease locus region defined by the three families was ~5 Mb. Family 2 did not show consistent evidence for linkage to this region.

The minimal disease locus region contains several biologically relevant genes that may be important in chordoma and cancers in general. Among them, the *T* gene, which encodes Brachyury, is of particular interest. Brachyury is a tissue-specific transcription factor expressed in the nucleus of notochord cells [10] and is essential for proper development and maintenance of the notochord [11]. Brachyury is specifically expressed in chordomas but not in a wide variety of non-neoplastic tissues or in 42 other types of neoplasms, including chondrosarcomas [12]. Its expression in chordomas mimics expression in the embryonic notochord. Given its obvious biological relevance, we selected the *T* gene as our top candidate gene for follow-up. We sequenced the coding region and adjacent splice sites of the nine exons and 5kb upstream and downstream of the coding region of the *T* gene in DNA from ten affected individuals from Families 1, 2, 4, and 6 (Families 1-4 were included in the linkage analyses; Families 6-8 were newly examined smaller chordoma families, Supplementary Figure 1, Table 1). We did not find any sequence variants consistent with a disease-causing mutation. We also sequenced 20 other candidate genes that reside in the minimal disease locus region, including *Dapper* and *PRS6KA2*, but did not find any disease-related mutations in them.

CNVs have recently been recognized as a significant source of genetic variation that can contribute to disease susceptibility [1]. Therefore, we conducted a genome-wide search for CNVs using a whole genome human array-CGH chip (Nimblegen 385K; average probe spacing, 7 kb). We analyzed blood-derived genomic DNA from eleven chordoma cases and two spouses selected from our seven chordoma families. We identified duplicated regions located on 6q27 in seven affected individuals from four families: Family 1 (#18, #21), Family 3 (#3), Family 4 (#1, #3), and Family 8 (#1 and #2) (Supplementary Table 2). The sizes of the duplicated regions ranged from 52 kb in Family 4 to 489 kb in Family 3. The duplications were not detected in individuals with chordoma from Families 2, 6, and 7, two unrelated spouse controls (from Families 1 and 2) or sixteen individuals from melanoma-prone families whose blood DNA was analyzed using the same assay (Supplementary Table 2). The duplicated regions in all 4 families contain only one known gene, *T*, and within the gene, there were no previously reported CNVs.

To validate the finding, we developed a quantitative PCR (qPCR) assay targeting the 3' end of exon 6 in the *T* gene and screened all 65 individuals (21 affected with chordoma) with DNA available in the 7 families for copy number changes. qPCR analyses confirmed the duplications in all affected subjects and obligate carriers in Families 1, 3, 4, and 8 (Table 1, Supplementary Figure 2A). 6q duplications were not observed in members of the other chordoma families or in 100 unrelated healthy controls (200 meioses) (Supplementary Figure 2A, 2B). The aggregation of chordoma in the three families without the duplications, each with 2 chordoma cases, may result from alterations of other susceptibility genes, another mutational mechanism targeting the *T* gene, or clustering of sporadic chordoma patients.

To confirm the *T* gene duplications and to better define the breakpoints of the amplicons, we analyzed genomic DNA from seven individuals with chordoma (#1 and #5 in Family 1, #3 in Family 3, #3 in Family 4, #1 in Family 7, #1 in Family 8) and an unaffected individual in Family 6, using a Nimblegen custom-made fine-tiling CGH array specifically targeting the 6q27 region (average probe spacing, 4 bp). Duplications of the *T* gene were clearly demonstrated in chordoma cases from Families 1, 3, 4, and 8 (Figure 1). In contrast, no duplication was observed in the two "controls" (a chordoma case from Family 7 that did not carry the duplication and an unaffected individual from Family 6). Using the predicted breakpoints defined by the fine-tiling arrays, we were able to amplify junction fragments from three of the four families (Families 1, 4, and 8). The genomic rearrangement in Family 3 is more complex than in the other families, as evidenced by the CGH segmentation analysis (Figure 1). Families 1, 4, and 8 revealed similar tandem duplications and the duplication sizes were consistent with observed array-CGH changes (124,756 bp in Family 1, 97,284 bp in Family 4, and approximately 173 kb in Family 8, Supplementary Figure 3). Amplification of junction fragments was confirmed in all individuals of Families 1, 4, and 8 who had increased copy number of the *T* gene by qPCR (Supplementary Figure 2A). In contrast, we were not able to amplify from controls the junction fragments that were observed in Families 1, 4 (200 unrelated controls), and 8 (100 unrelated controls).

Sequence analysis of the junction fragments identified the breakpoints (Families 1 and 4) or breakpoint region (Family 8). Bioinformatics analysis revealed that the breakpoints/breakpoint region were located at or near repetitive SINE and LINE elements, but they did not contain any low-copy repeats (LCRs) (Supplementary Figure 3). In Family 1, a single base pair was shared between the telomeric and centromeric sequences at the breakpoint junction (Supplementary Figure 3a). In Family 4, the telomeric and centromeric sequences were separated by a 5-bp insertion (Supplementary Figure 3b). In Family 8, the junction fragment was located within a 306 bp region of high (90%) homology formed by the fusion of two ALuY elements located at 6q27 (Supplementary Figure 3c). A precise breakpoint could not be located in this fused region as the junction fragment gradually transitioned from the telomeric ALuY (chr6:166,580,862-166,581,165) to the centromeric ALuY (chr6:166,406,935-166,407,246) sequence. The tandem duplication in Family 8 most likely resulted from an ALuY mediated non-allelic homologous recombination (NAHR). The junction fragments in Family 1 and Family 4 share features that seem to be consistent with non-homologous end joining (NHEJ), although the underlying mechanism could be more complex, as suggested in recent publications[13].

In summary, we have identified *T* as a major susceptibility gene for familial chordoma using combined genetic linkage and high-resolution array-CGH analyses. This approach has enabled us to identify a susceptibility gene in a linkage region that did not reveal disease-associated mutations by sequencing. Our findings suggest that screening for complex genomic rearrangements that co-segregate with disease in families may provide a powerful alternative to traditional gene-mapping approaches.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Deborah Zametkin for her outstanding skills as a research nurse; Drs. Gladys M. Glenn and Mary L. McMaster for their careful clinical evaluations of patients and family members; Dr. Nicholas J. Patronas for reviewing the MR images; Drs. Andrew Bergen and Stephen Chanock for SNP genotyping; Kashif Haque and Dr. Dennis Maeder for technical assistance, and, especially, the patients and their families for their participation. This work was supported by the intramural research program of the National Cancer Institute, Division of Cancer Epidemiology and Genetics,

National Institutes of Health, and by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Biomedical Laboratory Research and Development program, and a grant from the Chordoma Foundation (M.J.K.).

References

1. McCarroll SA, Altshuler DM. Nat. Genet 2007;39:S37–S42. [PubMed: 17597780]
2. Inoue K, Lupski JR. Annu. Rev. Genomics Hum. Genet 2002;3:199–242. [PubMed: 12142364]
3. Lupski JR. Nat. Genet 2007;39:S43–S47. [PubMed: 17597781]
4. Padiath QS, et al. Nat. Genet 2006;38:1114–1123. [PubMed: 16951681]
5. Le MC, et al. Nat. Genet 2006;38:1372–1374. [PubMed: 17072318]
6. Lee JA, Lupski JR. Neuron 2006;52:103–121. [PubMed: 17015230]
7. Malawer, MM.; Link, MP.; Donaldson, SS. Cancer: principles and practice of oncology in *Sarcomas of bone*. Devita, VT.; Hellman, S.; Rosenberg, SA., editors. Lippincott-Raven; Philadelphia: 1997. p. 1844-1846.
8. Kelley MJ, et al. Am. J. Hum. Genet 2001;69:454–460. [PubMed: 11452362]
9. Yang XR, et al. Int. J. Cancer 2005;116:487–491. [PubMed: 15818627]
10. Kispert A, Koschorz B, Herrmann BG. EMBO J 1995;14:4763–4772. [PubMed: 7588606]
11. Kispert A, Herrmann BG. Dev. Biol 1994;161:179–193. [PubMed: 8293872]
12. Vujovic S, et al. J. Pathol 2006;209:157–165. [PubMed: 16538613]
13. Chanda B, et al. Hum. Mol. Genet 2008;17:3446–3458. [PubMed: 18694899]

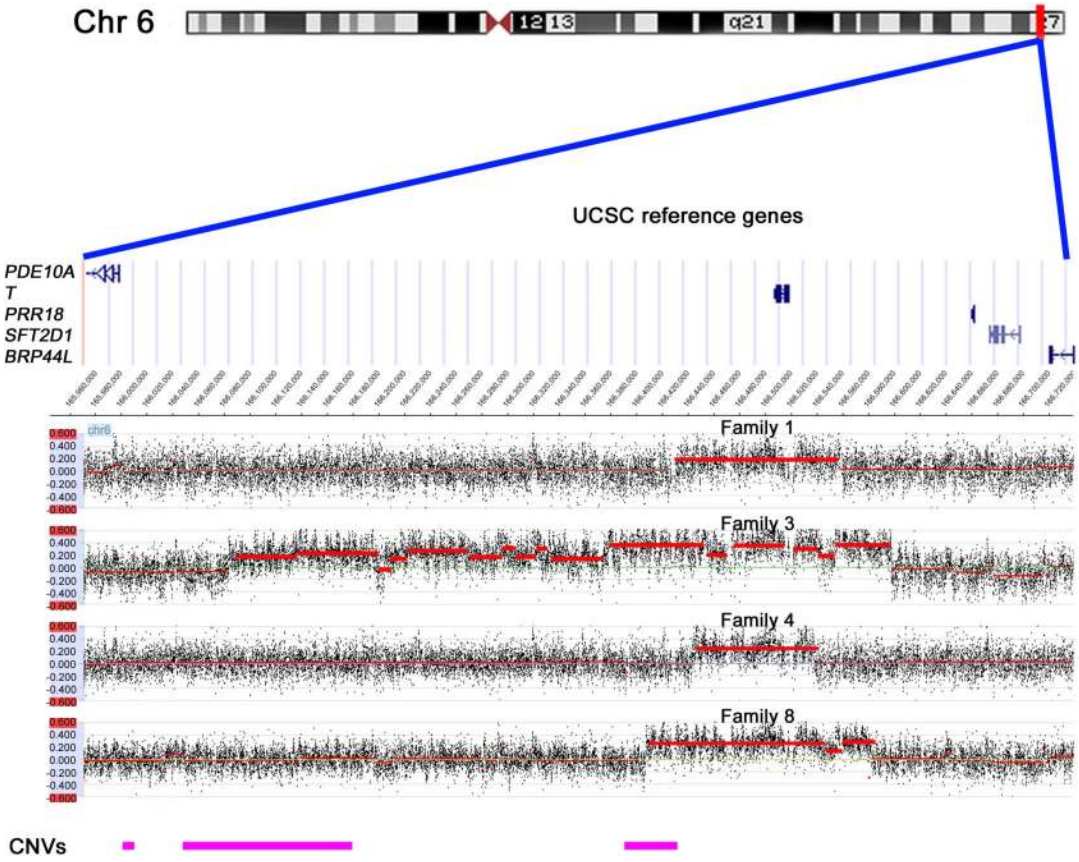

Figure 1.

Duplications of 6q27 region in chordoma cases from Families 1, 3, 4, and 8 by the custom fine-tiling CGH array. Each point represents an average of 40 oligonucleotide probes. Duplications in chordoma cases from the four families were determined by elevated log2 ratios which are shown by red horizontal bars. The segmental duplication in Family 3 is complex; probe intensities suggest a short genomic interval with normal copy number at approximately chr6:166,200,000; segments toward the telomeric end of the rearrangement have intensities suggesting triplication. The commonly amplified region in all four families is about 100 kb in size and contains only one gene, the *T* gene. There were no previously reported CNVs, as shown by pink bars, in the *T* gene region.

Nat Genet. Author manuscript; available in PMC 2010 July 12.

Table 1

Clinical features and 6q27 duplication status of subjects in seven multiplex chordoma families.

Subject	Tumor type	Tumor site	Age at diagnosis (years)	6q27 duplication by array-CGH	Fold change by qPCR (±1 SD)***
Fam 1					
1	Chordoma	Clivus	68	Yes	1.56 [1.48-1.66]
5	Chordoma	Clivus	56	Yes	1.57 [1.49-1.65]
8	Chordoma*	Clivus	46	ND**	1.62 [1.58-1.67]
9	Chordoma	Sacrum	28	ND	1.42 [1.41-1.43]
10	Chordoma	Clivus	39	ND	1.68 [1.60-1.77]
18	Chordoma	Nasopharynx	30	Yes	1.72 [1.67-1.77]
20	Chordoma	Clivus	35	ND	1.51 1.42-1.59]
21	Chordoma	Clivus	20	Yes	1.69 [1.66-1.73]
28	Chordoma	Clivus	16	ND	1.53 [1.46-1.59]
43	Chordoma*	Clivus	6	ND	1.47 [1.43-1.53]
Fam 2					
1	Chordoma	Clivus	46	No	0.93 [0.87-1.00]
3	Astrocytoma	Spinal cord	17	ND	1.00 [0.96-1.03]
7	Chordoma	Clivus	12	No	0.98 [0.95-1.02]
Fam 3					
3	Chordoma	Clivus	40	Yes	1.79 [1.74-1.84]
7	Chordoma	Clivus	21	ND	2.22 [2.08-2.36]
Fam 4					
1	Chordoma	Clivus	26	Yes	1.50 [1.46-1.55]
2	Astrocytoma	Cerebellum	12	ND	0.91 [0.86-0.98]
3	Chordoma	Clivus	7	Yes	1.47 [1.42-1.51]
Fam 6					
1	Chordoma	Clivus	46	No	1.02 [0.96-1.09]
Fam 7					
1	Chordoma	Sacrum	37	No	1.01 [0.94-1.09]
Fam 8					

Subject	Tumor type	Tumor site	Age at diagnosis (years)	6q27 duplication by array-CGH	Fold change by qPCR (±1 SD)***
1	Chordoma	Clivus	32	Yes	1.89 [1.82-1.96]
2	Chordoma	Clivus	5	Yes	1.67 [1.60-1.74]
3	Chordoma	Clivus	8	ND	1.83 [1.65-2.03]

* Chordoma diagnosed based on magnetic resonance imaging findings.

** ND = not determined.

*** SD = standard deviation.