Characterization of genetic variants in the $T$ gene in familial and sporadic chordoma

$4^{th}$ International Chordoma Research Workshop
Session I: Genetics and Genomics of Chordoma

Rose Yang, Ph.D., M.P.H.
Genetic Epidemiology Branch/DCEG
National Cancer Institute/NIH/DHHS
Chordoma is mostly sporadic; multiple families with $\geq 2$ cases suggesting genetic predisposition

The goal of our study: to identify susceptibility genes for familial chordoma

Improve our understanding of biology underlying familial and sporadic chordoma
Chordoma Family 1
Other chordoma families
6q27 gain in 4 chordoma families

Yang et al. Nat Genet, 2009
Presneau N et al. J Pathology, 2011

Copy number change of the T locus is common in sporadic chordoma tumors (N=170).
- ~50% showed gain of the T locus
- ~7% with T amplification

Knockdown of T reduced chordoma cell proliferation

No germline CNVs among 40 patients
Common $T$ variant associated with chordoma

- Pillay N et al. Nature Genet., 2012

- A common germline SNP in $T$ (rs2305089) was associated with chordoma.
  - The only exonic NS variant in >1 chordoma case
  - Variant allele A present in all 23 cases
  - Common (~47% in HapMap CEU)
  - OR=6.1, $P=4.4 \times 10^{-9}$ (40 cases, 358 controls)
  - Validated in another case-control analysis (20 cases, 363 controls, combined OR=5.3, $P=4.6 \times 10^{-12}$)

- No $T$ germline duplication in 22 sporadic cases
Finding additional susceptibility genes for familial chordoma

• Exome sequencing
  – Nimblegen v2/v3 exome capture, Illumina HighSeq, >80% coverage at 15x

• More than 800 subjects with cancers were sequenced using the same platform

• Sequenced germline DNA from 16 chordoma cases/carriers
  – 11 from 4 families; 5 young sporadic cases
Finding disease-related variants

• Non-synonymous

• Rare
  – <5 in Exome Variant Server (up to 4,300 Europeans)
  – <1% in dbSNP or 1k genome
  – <2 in internal controls (>800 subjects with other cancers)

• Same variant shared by all cases/carriers in one family and different variants in same genes shared by multiple cases across families
Finding disease-related variants

• Challenge: finding disease-causing variants
  – Database search: functional relevance, regulatory regions, somatic changes in tumors, etc.

• Technical validation/Co-segregation with disease

• Sequence genes in sporadic cases

• Functional study
Finding disease-related variants

• Found one rare variant in $T$ in a single case
  – In a young sporadic case with multi-centric chordoma
  – Observed once in ESP, not in other databases or internal controls
  – Located in a potential splice site
  – Validated by Sanger sequencing
  – Found in one parent and one of two unaffected siblings

• Function analysis, the variant did not increase the transcription activity by $T$
Characterization of $T$ variants

- **Germline $T$ duplication**
  - One new family with 2 chordoma cases
  - ~100 sporadic chordoma cases (any age, any primary site, identified in US and Canada)
  - qPCR (exon 6 of $T$)

- **Exonic single nucleotide variants in $T$**
  - All exons in $T$ sequenced using Ion Torrent
  - 115 subjects in chordoma families (39 cases/carriers, 76 unaffected family members and spouses)
  - ~100 sporadic chordoma cases
Germline $T$ amplification in sporadic chordoma
Germline $T$ duplication in a new family

Family 9

Germline $T$ dup/amp: 2% (2/100) in sporadic cases
~50% (5/9) in chordoma families
rs2305089 in familial and sporadic chordoma

Reference allele: C    Variant allele: T

<table>
<thead>
<tr>
<th>Population</th>
<th>Freq of T</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESP European</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>Controls in chordoma families</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>Cases/carriers in chordoma families</td>
<td>0.72</td>
<td>P=0.0005 vs. controls</td>
</tr>
<tr>
<td>Sporadic cases</td>
<td>0.77</td>
<td>P&lt;0.0001 vs. ESP</td>
</tr>
</tbody>
</table>

OR=2.76 (1.15, 6.63) comparing familial cases to controls using conditional logistic regression; OR=3.17 (2.24, 4.49) comparing sporadic cases to ESP.
Summary of $T$ variants

- The variant allele in rs2305089 was associated with chordoma, however, it is very common
  - Limited value in predicting risk

- The variant identified in a single case by exome sequencing was not observed in any other cases

- 3 additional novel variants were identified that occurred only in one or two cases
  - Unknown significance
  - Analyses in process to evaluate these novel and other common variants
Acknowledgements

• **DCEG Collaborators:**
  – Dilys Parry
  – Alisa Goldstein
  – David Ng
  – Kevin Jacobs
  – Ji He
  – Meredith Yeager
  – Joseph Boland
  – Laurie Burdett
  – Stephen Chanock

• **Duke University**
  – Michael Kelley
  – David Alcorta
  – Sufeng Li

• **Other Institutions:**
  – Eamonn Sheridan
    (St. James Univ, UK)
  – Norbert Liebsch
    (Mass General Hospital)

• **Roche Nimblegen**