

Supplemental Figure 2. Alignments between scaffolds containing more than one disagreement with the BEAT consensus sequence, the BEAT consensus, and the human reference. Scaffolds that meet these criteria were found to belong to two categories: likely paralogs, and possible assembly errors. Paralogs are indicated when the entire sequence can be reconstructed from shotgun sequence reads, while assembly errors are marked by regions of scaffold-read disagreement. All alignments to scaffolds shown were conducted with BWA-MEM (Li 2013), and used only the original reads from Perry, Reeves et al. (2012). Titles with a light blue border indicate likely paralogs, while titles with a red border indicate possible assembly error.

CDH23 Scaffold with disagreement 1

Human	CCTCTCACGGTGCTCAATGGGCCATCCTGGCCCTGGATGCAGACCAAGACATCTACGCCGGTGACCTACCAGCTGCTGGGTGCCAGAGTGGCCTTTTGACATCAACAGCAGCACCGGTG
BEAT	CCCCTCACGGTGCTCAATGGGCCATCCTGGCCCTGGATGCAGACCAGGACATCTACGCCGGTGACCTACCAGCTGCTGGGCACCCAGAGTGGCCTTTTGGACATCAACAGCAGCACTGGCG
Scaffold	<div style="display: flex; justify-content: space-between; width: 100%;"> 1112131415161718191101111121 </div>
BWA alignment to scaffold	<div style="font-family: monospace; font-size: 0.8em;"> <p>CCCCTCACGGTGCTCAATGGGCCATCCTGGCCCTGGATGCAGACCAGGACATCTACGCCGGTGACCTACCAGCTGCTGGGCAGCTCTTCCGATCTTTTGGACATCAACAGCAGCACTGGTGM</p> <p>.....C.....</p> <p>.....C.....C...T.....C.....</p> <p>.....C.....a.....</p> <p>.....G.....</p> </div>

KCNQ4 Scaffold with disagreement

Human	GCCGCGGGTACCCAGGGCAACATCTTCGCCACGTCCGCGTGCAGCATGCGCTTCTGCAGATCTGCGCATGGTGCATGGACCGCGGGCGGCACCTGGAAGCTGCTGGGCTCAGTGGTCTACGCGCATAGCAAGG
BEAT	GCTGCGGGCACACAGGGCAACATCTTCGCCACGTCTGCAGCATGCGCTTCTGCAGATCTGCGCATGCTGCGGATGGACCGGAGGGCGGCACCTGGAAGCTGCTGGGCTCGGTGGTCTACGCGCACAGCAAGG
Scaffold	<div style="display: flex; justify-content: space-between; font-size: small;"> 1 11 21 31 41 51 61 71 81 91 101 111 121 131 141 </div> <pre> GCCGCGGGTCCAGGGCAATGCTTTGCCACGTCTGCACTCCGGAGCCTGCGGTTCTGCAGATCTGCGGATGATCCGCATGGACCGGAGGGGAGGCACCTGGAAGCTCCTGGGCTCCGCCATCTGCGCCACAGCAAGGNC.....C.....C..G.....G.....CA.....C.....C..G.....C.....T..T.....TGG..A.....C.....C.....C..g.....Tgg..a.....C.....C.....C..g.....TGG..A.....a.....G..C.....C..C.....C.....t.....C.....g.....g..tg..g..C..C.....a.....a.....a.....a..a.....g.....g.....C.....A..G..G.....C.....T..A.....C..g.....G.....C.....C..g..g.....C.....C.....C..g..g.....C.....g..C.....C..g..g.....C.....g..g.....g.....C.....C.....C.....C..G..G.....C.....g.....C.....C..g.....C..g.....TGG..A.....C.....C..G..G.....C.....C.....C.....C.....TGG..A..... </pre>
BWA alignment to scaffold	

OTOF Scaffold with disagreement

Human	GCCGCTTCCTCCCTCGCTGACAAGGACCAGGGCCACTCATCCCGCACCAGGCTTGACCGGGAGCGCCTCAAGTCCTGCATGAGGGAGCTGG
BEAT	GCCGCTTCCTCCCTCGCTGACAAGGACCAGAGCCACTCGTCCCGCACCAGGCTTGACCGGGAGCGCCTCAAGTCCTGCATGAGGGAGCTGG
Scaffold	<div style="display: flex; justify-content: space-between; width: 100%;"> 1 11 21 31 41 51 61 71 81 91 </div> <p>GCCGCTTCCTCCCTCGCGGACAAGGACCAGAGCCACTCGTCCCGCACCAGGCTGGACCGGGAGCGCCTCAAGTCCTGCATGAGGGAGCTGGNNN</p>
BWA alignment to scaffold	<p>The alignment visualization shows approximately 15 reads aligned to the scaffold. The scaffold sequence is: GCCGCTTCCTCCCTCGCGGACAAGGACCAGAGCCACTCGTCCCGCACCAGGCTGGACCGGGAGCGCCTCAAGTCCTGCATGAGGGAGCTGGNNN. The reads are mostly aligned to the first 91 positions of the scaffold. Mismatches are indicated by 't' (thymine) or 'g' (guanine) characters below the alignment lines. For example, at position 51, the scaffold has 'G' and the reads have 't'. At position 61, the scaffold has 'G' and the reads have 't'. At position 71, the scaffold has 'C' and the reads have 'T'. At position 81, the scaffold has 'C' and the reads have 'T'. At position 91, the scaffold has 'N' and the reads have 't'.</p>

