

Cell, Volume 148

## Supplemental Information

### Genome Sequencing and Analysis

#### of the Tasmanian Devil

#### and Its Transmissible Cancer

Elizabeth P. Murchison, Ole B. Schulz-Trieglaff, Zemin Ning, Ludmil B. Alexandrov, Markus J. Bauer, Beiyuan Fu, Matthew Hims, Zhihao Ding, Sergii Ivakhno, Caitlin Stewart, Bee Ling Ng, Wendy Wong, Bronwen Aken, Simon White, Amber Alsop, Jennifer Becq, Graham R. Bignell, R. Keira Cheetham, William Cheng, Thomas R. Connor, Anthony J. Cox, Zhi-Ping Feng, Yong Gu, Russell J. Grocock, Simon R. Harris, Irina Khrebtukova, Zoya Kingsbury, Mark Kowarsky, Alexandre Kreiss, Shujun Luo, John Marshall, David J. McBride, Lisa Murray, Anne-Maree Pearse, Keiran Raine, Isabelle Rasolonjatovo, Richard Shaw, Philip Tedder, Carolyn Tregidgo, Albert J. Vilella, David C. Wedge, Gregory M. Woods, Niall Gormley, Sean Humphray, Gary Schroth, Geoffrey Smith, Kevin Hall, Stephen M.J. Searle, Nigel P. Carter, Anthony T. Papenfuss, P. Andrew Futreal, Peter J. Campbell, Fengtang Yang, David R. Bentley, Dirk J. Evers, and Michael R. Stratton

Table S1. Tasmanian Devil Genome Size Estimation, Related to Table 1

Method	Estimated Tasmanian devil genome size (Gb)
Number of assembled bases (contigs)	2.93
Number of assembled bases (supercontigs)	3.17
Kmer frequency analysis	3.03
Flow karyotype analysis	2.89
Relative DNA content analysis	2.92

**Table S2. *In Silico* Chromosome Assignment , Related to Table 1**

<b>Chromosome</b>	<b>Chromosome DNA content (cytogenetics) (Mb)</b>	<b>Number of assigned supercontigs</b>	<b>Assigned bases (Mb)</b>
1	571	6729	684
2	610	8381	740
3	556	7197	641
4	450	4817	487
5	341	3188	300
6	277	2844	263
X	122	2378	86.6
Unassigned		(440)	(1.23)
Total	2927	35,974 (total supercontigs) 35,534 (total assigned supercontigs)	3202.8 (Total bases) 3201.6 (Total assigned bases)

**Table S3. Coordinates of Tumour-Specific Rearrangements, Related to Figure 4**

The rearrangements listed were identified and confirmed only in 87T or 53T but not in both tumours.

<b>Sample</b>	<b>Breakpoint 1</b>		<b>Breakpoint 2</b>	
<b>87T</b>	<b>Coordinate range</b>	<b>Strand</b>	<b>Coordinate range</b>	<b>Strand</b>
	Chr4_supercontig_000000294: 363864-364141	-	Chr4_supercontig_000000294: 285688-285971	+
	Chr3_supercontig_000000393: 1780799-1781083	-	Chr3_supercontig_000000281: 21827-22022	-
	Chr3_supercontig_000001556: 850-1120	-	Chr3_supercontig_000000315: 147673-147969	+
	Chr5_supercontig_000003149: 550-825	+	Chr2_supercontig_000001430: 157997-158230	+
	Chr4_supercontig_000000654:2 4128-24412	+	Chr4_supercontig_000000229: 2437030-2437307	+
	Chr2_supercontig_000000207:1 235093-1235312	-	Chr2_supercontig_000000207: 1230368-1230615	+
	Chr1_supercontig_000000629: 186664-186891	-	Chr1_supercontig_000000629: 185915-186201	+
	Chr2_supercontig_000000207: 1218849-1219055	-	Chr2_supercontig_000000207: 1214688-1214894	+
	Chr2_supercontig_000000207: 1336201-1336430	+	Chr2_supercontig_000000207: 1335745-1336030	+
	Chr2_supercontig_000000344: 674461-674728	-	Chr2_supercontig_000000344: 672331-672615	-
	Chr1_supercontig_000000182: 1304512-1304801	+	Chr1_supercontig_000000182: 1303661-1303939	+
<b>53T</b>	<b>Coordinate range</b>	<b>Strand</b>	<b>Coordinate range</b>	<b>Strand</b>
	Chr3_supercontig_000000191: 233670-233926	-	Chr1_supercontig_000000303: 3376378-3376658	-
	Chr2_supercontig_000000364: 1366493-1366808	+	Chr2_supercontig_000000017: 367364-367641	-
	Chr5_supercontig_000001108: 870-1177	-	Chr1_supercontig_000001069: 40054-40331	-
	Chr1_supercontig_000000253: 91816-92089	-	Chr1_supercontig_000000253: 89326-89611	+
	Chrx_supercontig_000000022: 31467-31764	+	Chrx_supercontig_000000022: 21372-21685	-
	Chr2_supercontig_000000274: 784350-784641	-	Chr2_supercontig_000000274: 739815-740095	+
	Chr3_supercontig_000000130: 1756277-1756562	-	Chr3_supercontig_000000130: 1749479-1749756	+
	Chr1_supercontig_000000356: 1345071-1345388	-	Chr1_supercontig_000000101: 1147092-1147437	-

	Chr5_supercontig_000000097: 125064-125382	+	Chr3_supercontig_000000190: 1609202-1609504	-
	Chr2_supercontig_000001667: 152094-152382	+	Chr2_supercontig_000000584: 111379-111671	+
	Chr3_supercontig_000000115: 261665-261938	-	Chr3_supercontig_000000115: 260047-260297	+
	Chr1_supercontig_000000104: 1148801-1149123	-	Chr1_supercontig_000000104: 1145329-1145615	+
	Chr5_supercontig_000000152: 1991912-1992185	-	Chr2_supercontig_000000061: 2384483-2384759	-
	Chr5_supercontig_000000350: 74757-75032	-	Chr5_supercontig_000000350: 38473-38804	+
	Chr3_supercontig_000000377: 2627634-2627909	-	Chr3_supercontig_000000377: 2622358-2622613	+
	Chr3_supercontig_000000401: 151726-151992	+	Chr3_supercontig_000000393: 3853491-3853755	-
	Chr2_supercontig_000000345: 100261-100543	+	Chr2_supercontig_000000014: 1392972-1393238	+

**Table S4. Variation Involving Genes in DFTD, Related to Figure 4**

Table is available as an attached Excel spreadsheet.

(A) Non-synonymous single base substitutions in gene exons in 87T and 53T. These variants were found in 87T and 53T but not in either of the normal Tasmanian devil genomes.

(B) Non-synonymous single base substitutions in gene exons in 87T. Variants were found only in 87T but not in 53T or either of the normal Tasmanian devil genomes.










(C) Non-synonymous single base substitutions in gene exons in 53T. Variants were found only in 53T but not in 87T or either of the normal Tasmanian devil genomes.

















(D) Indels in gene exons in 87T and 53T. Variants were found in both 87T and 53T or either 87T or 53T but not in either of the normal Tasmanian devil genomes.

(E) Structural variants involving genes in DFTD. Coordinates and gene IDs are listed for structural variants involving genes that were identified in both 87T and 53T but not in either of the normal Tasmanian devil genomes. As confirmation has not been performed for the whole set, this list may include some false positives.

**Table S5. Genotyping Summary, Related to Figure 5**

The genotypes that define the coloured symbols on the phylogenetic trees in Figure 5 and Figure S4 are shown. The mitochondrial coordinates are based on DDBJ/EMBL/GenBank accession JN216828.

Symbol	Genotype	Type
	Probable ancestral genotype.	
	Chr1_supercontig_000000134:1378535C>T	Heterozygous
	Chr3_supercontig_000000336:2001045T> Chr1_supercontig_000000248:697762G>A Chr1_supercontig_000000134:1378535C>T	Heterozygous loss Heterozygous Heterozygous
	Chr3_supercontig_000000336:2001045T> Chr2_supercontig_000000158:601261C>T Chr1_supercontig_000000095:2618076T>G Chr1_supercontig_000000248:697762G>A Chr1_supercontig_000000134:1378535C>T MT:1828A>G MT:3255G>A	Heterozygous loss Heterozygous Homozygous Homozygous Heterozygous Mitochondrial Mitochondrial
	Chr3_supercontig_000000336:2001045T> Chr2_supercontig_000000158:601261C>T Chr1_supercontig_000000095:2618076T>G Chr1_supercontig_000000248:697762G>A Chr1_supercontig_000000134:1378535C>T MT:1828A>G MT:1928T>C MT:3255G>A	Heterozygous loss Heterozygous Homozygous Homozygous Heterozygous Mitochondrial Mitochondrial Mitochondrial
	Chr3_supercontig_000000336:2001045T> Chr2_supercontig_000000158:601261C>T Chr1_supercontig_000000095:2618076T>G Chr1_supercontig_000000248:697762G>A Chr1_supercontig_000000248:403816C>T Chr1_supercontig_000000134:1378535C>T MT:1828A>G MT:3255G>A	Heterozygous loss Heterozygous Homozygous Homozygous Homozygous Homozygous Mitochondrial Mitochondrial
	Chr3_supercontig_000000336:2001045T> Chr2_supercontig_000000158:601261C>T Chr1_supercontig_000000095:2618076T>G Chr1_supercontig_000000248:697762G>A Chr4_supercontig_000000200:4423976T>C Chr3_supercontig_000000295:352495G>A Chr3_supercontig_000000362:528865A>G Chr1_supercontig_000000248:403816C>T Chr1_supercontig_000000134:1378535C>T MT:1828A>G MT:3255G>A MT:2572G>A	Heterozygous loss Heterozygous Homozygous Homozygous Heterozygous Homozygous Homozygous Heterozygous Homozygous Mitochondrial Mitochondrial Mitochondrial
	Chr3_supercontig_000000336:2001045T> Chr2_supercontig_000000158:601261C>T Chr1_supercontig_000000095:2618076T>G Chr1_supercontig_000000248:697762G>A Chr1_supercontig_000000134:1378535C>T MT:1828A>G MT:3430T>A MT:4311G>A MT:3255G>A	Heterozygous loss Heterozygous Homozygous Heterozygous Heterozygous Mitochondrial Mitochondrial Mitochondrial Mitochondrial
	Chr3_supercontig_000000336:2001045T> Chr2_supercontig_000000158:601261C>T Chr1_supercontig_000000095:2618076T>G Chr1_supercontig_000000248:697762G>A Chr1_supercontig_000000134:1378535C>T MT:1828A>G MT:2731A>G MT:3430T>A	Heterozygous loss Heterozygous Homozygous Homozygous Heterozygous Mitochondrial Mitochondrial Mitochondrial

	MT:4311G>A MT:3255G>A	Mitochondrial Mitochondrial
	MT:7198T>C	Mitochondrial
	MT:12288C>T MT:7198T>C	Mitochondrial Mitochondrial
	MT:3161G>A	Mitochondrial
	MT:3161G>A MT:14209A>T	Mitochondrial
	MT:2707G>A	Mitochondrial
	MT:2707G>A MT:3679G>A	Mitochondrial
	MT:6745T>C	Mitochondrial
	MT:14114C>T	Mitochondrial
	MT:2942C>T	Mitochondrial
	MT:4266G>A	Mitochondrial
	MT:2261T>C	Mitochondrial
	Chr1_supercontig_000000095:591062C>A Chr2_supercontig_000000205:1697548C>T Chr1_supercontig_000000246:95466A>G Chr2_supercontig_000000205:4043679A>T	Homozygous Homozygous Homozygous Homozygous
	Chr1_supercontig_000000095:591062C>A Chr2_supercontig_000000205:1697548C>T Chr1_supercontig_000000246:95466A>G Chr2_supercontig_000000205:4043679A>T MT:10133C>T	Homozygous Homozygous Homozygous Homozygous Mitochondrial
	Chr3_supercontig_000000336:2001045T>- Chr2_supercontig_000000158:601261C>T Chr1_supercontig_000000095:2618076T>G Chr1_supercontig_000000248:697762G>A Chr4_supercontig_000000200:4423976T>C Chr3_supercontig_000000295:352495G>A Chr3_supercontig_000000362:528865A>G Chr1_supercontig_000000248:403816C>T Chr1_supercontig_000000134:1378535C>T MT:1828A>G MT:3255G>A MT:2572G>A MT:611C>T	Heterozygous loss Heterozygous Homozygous Homozygous Heterozygous Homozygous Homozygous Heterozygous Homozygous Mitochondrial Mitochondrial Mitochondrial Mitochondrial
	Chr3_supercontig_000000336:2001045T>- Chr2_supercontig_000000158:601261C>T Chr1_supercontig_000000095:2618076T>G Chr1_supercontig_000000248:697762G>A Chr1_supercontig_000000134:1378535C>T MT:1828A>G MT:3255G>A MT:4422G>A	Heterozygous loss Heterozygous Homozygous Homozygous Heterozygous Mitochondrial Mitochondrial Mitochondrial
	Chr1_supercontig_000000095:591062C>A Chr2_supercontig_000000205:1697548C>T Chr1_supercontig_000000246:95466A>G Chr2_supercontig_000000205:4043679A>T MT:10133C>T Chr2_supercontig_000000031:817867C>T	Homozygous Homozygous Homozygous Homozygous Mitochondrial Heterozygous