Supplementary Material for

X-linked hydrocephalus genes: their proximity to telomeres and high A+T content compared to Parkinson's disease

Madeline Hart[1], Joshua Conrad[1], Emma Barrett[1], Kaitlyn Legg[1],
Gabrielle Ivey[1], Peter H.U. Lee[2,3], Yun Yung[4], Joon W. Shim[1*]

Correspondence to: shim@marshall.edu

This file includes:

Figure S1 to S10;
Tables S1a-c to S11a-c

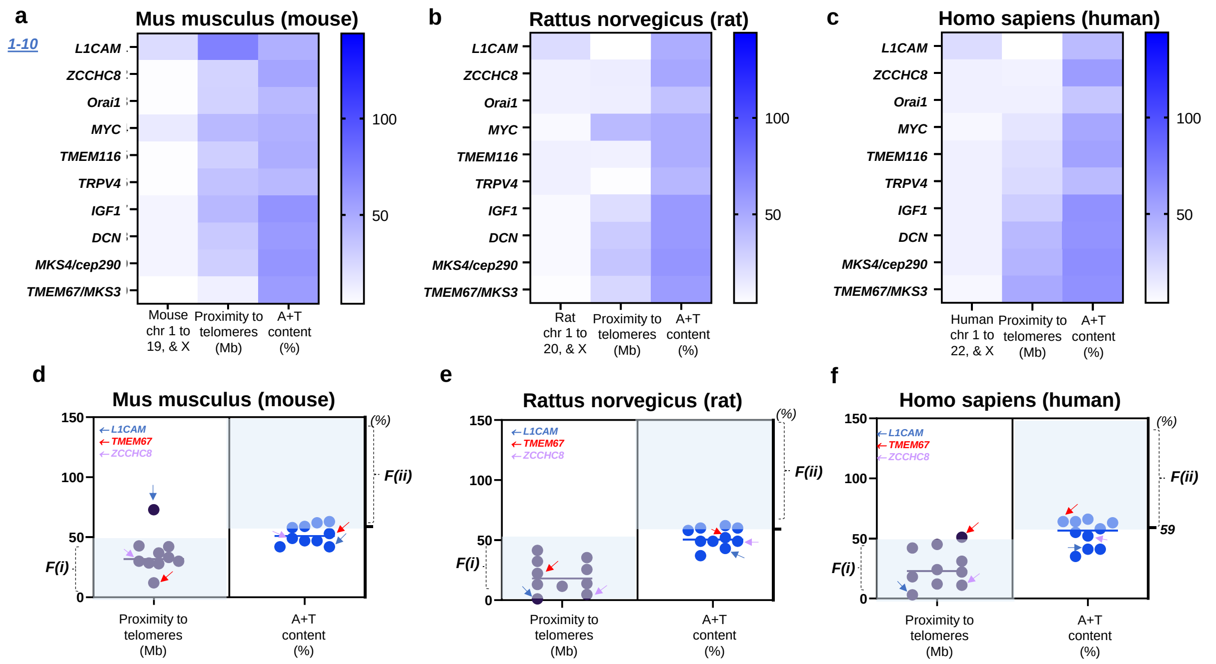

Figure S1. Relative mutability of L1CAM driven by proximity to telomeres (a) The color-coded map showing L1CAM and other genes by chromosome numbers (chr 1 to 19 and X), proximity to telomeres (Mb), and A+T content (%) in mus musculus (mouse). (b) The color-coded map showing L1CAM and other genes by chromosome numbers (chr 1 to 20 and X), proximity to telomeres (Mb), and A+T content (%) in rattus norvegicus (rat). (c) The color-coded map showing L1CAM and other genes by chromosome numbers (chr 1 to 22 and X), proximity to telomeres (Mb), and A+T content (%) in homo sapiens (human) (d) Scatter plots displaying relative mutability of ten mouse genes by proximity to telomeres (left) and A+T content (right). Shaded regions with light colors (blue) indicate genes satisfying either of the two factors. Genes outside the shaded regions at n=1 by F(i); n=7 by F(ii)) are relatively less mutable. (e) Scatter plots displaying relative mutability of ten rat genes by proximity to telomeres (left) and A+T content (right). Shaded regions with light colors (blue) indicate genes satisfying either of the two factors. Genes outside the shaded regions at n=0 by F(i); n=7 by F(ii)) are relatively less mutable. (f) Scatter plots displaying relative mutability of the same ten genes by proximity to telomeres (left) and A+T content (right) in human chromosomes. Shaded regions with light colors (blue) indicate genes satisfying either of the two factors. Genes outside the shaded regions at n=1 by F(i); n=5 by F(ii)) are relatively less mutable. Note that all human genes except one satisfied proximity to telomeres, F(i), while only a half (5 of 10) met F(ii).

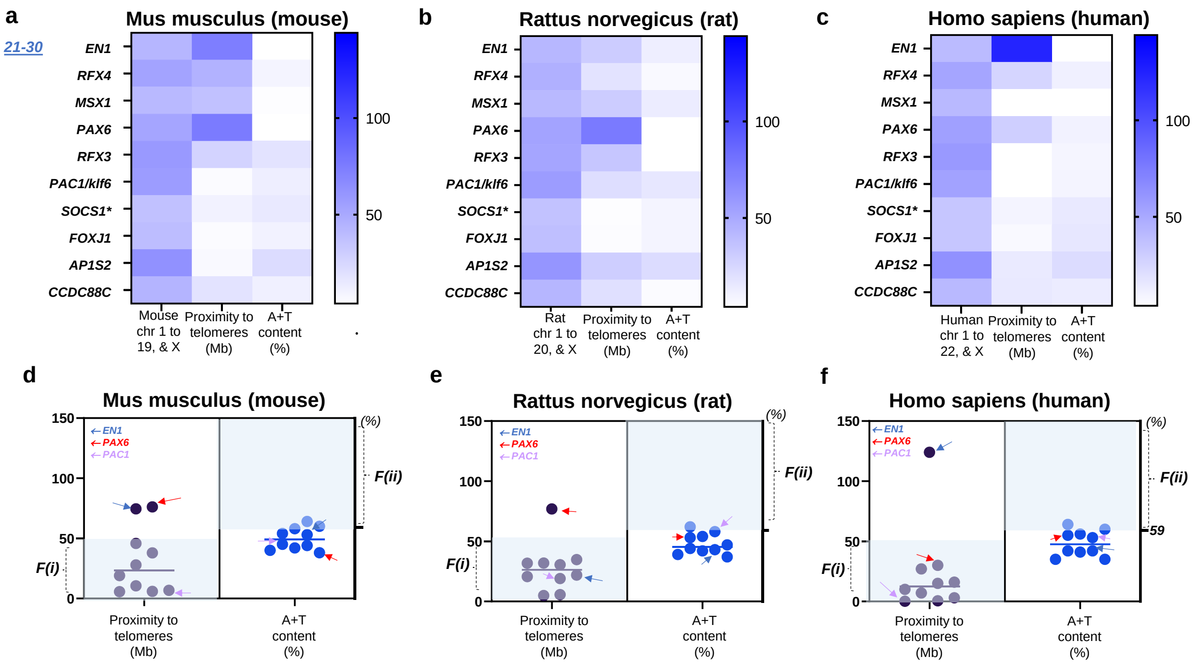

Figure S2. Relative mutability of AP1S2 driven by the two factors, F(i) and F(ii) (a) The color-coded map showing AP1S2 and other genes by chromosome numbers (chr 1 to 19 and X), proximity to telomeres (Mb), and A+T content (%) in mus musculus (mouse). (b) The color-coded map showing AP1S2 and other genes by chromosome numbers (chr 1 to 20 and X), proximity to telomeres (Mb), and A+T content (%) in rattus norvegicus (rat). (c) The color-coded map showing AP1S2 and other genes by chromosome numbers (chr 1 to 22 and X), proximity to telomeres (Mb), and A+T content (%) in homo sapiens (human) (d) Scatter plots displaying relative mutability of ten mouse genes by proximity to telomeres (left) and A+T content (right). Shaded regions with light colors (blue) indicate genes satisfying either of the two factors. Genes outside the shaded regions at n=2 by F(i); n=7 by F(ii)) are relatively less mutable. (e) Scatter plots displaying relative mutability of ten rat genes by proximity to telomeres (left) and A+T content (right). Shaded regions with light colors (blue) indicate genes satisfying either of the two factors. Genes outside the shaded regions at n=1 by F(i); n=9 by F(ii)) are relatively less mutable. (f) Scatter plots displaying relative mutability of the same ten genes by proximity to telomeres (left) and A+T content (right) in human chromosomes. Shaded regions with light colors (blue) indicate genes satisfying either of the two factors. Genes outside the shaded regions at n=1 by F(i); n=9 by F(ii)) are relatively less mutable. Note that all human genes except one satisfied proximity to telomeres, F(i), while only two genes (2 of 10) met F(ii) at A+T content > 59%.

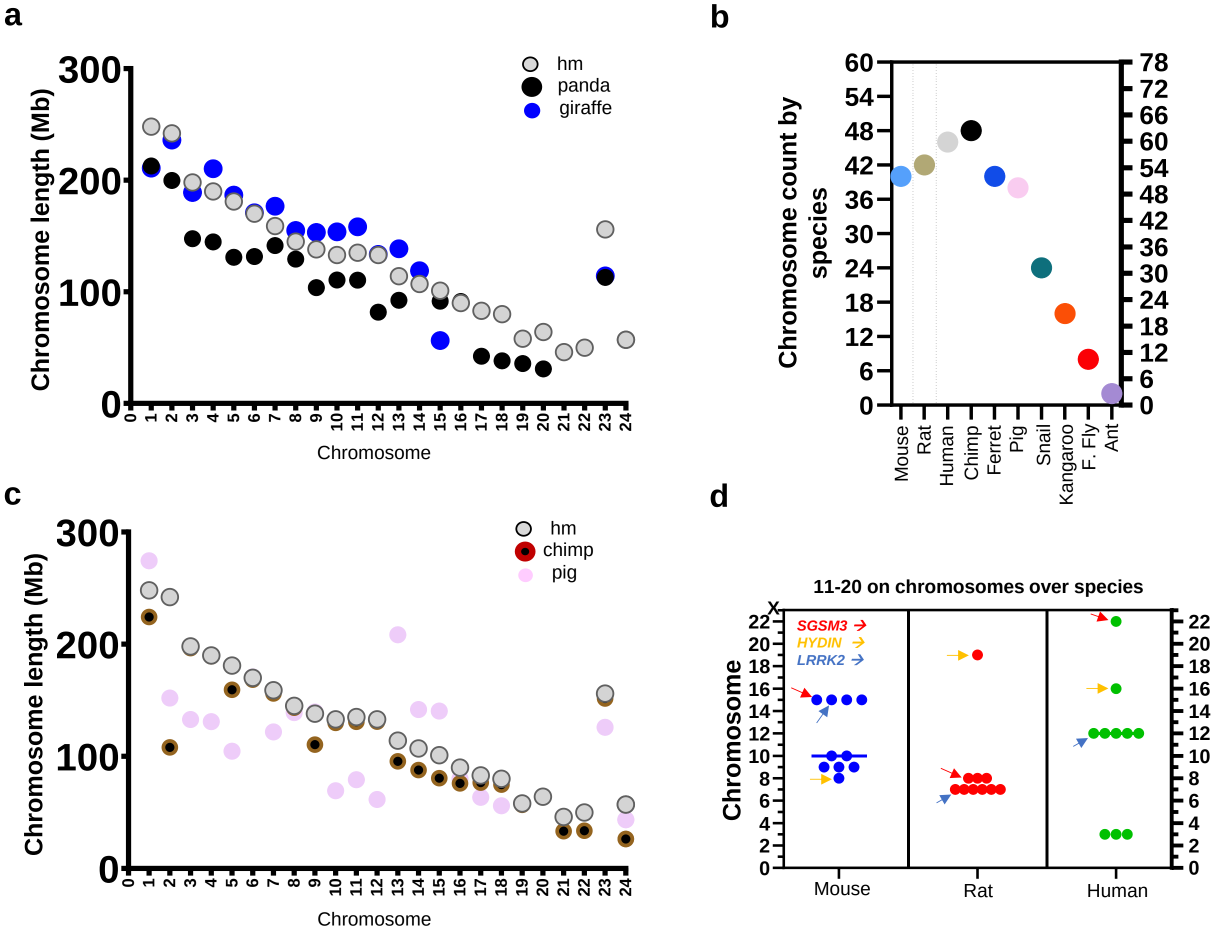

Figure S3. Chromosome length and count in different species. (a) Scatter plots showing chromosome length of three species with respect to each chromosome (chr) of chr 1 to X or Y; humans (hm), panda, and giraffe. Note that taller animals do not always show longer chromosomes. In chr 15, pandas and humans, which are smaller in body size, have longer nucleotide sizes than giraffes. x-axis is pair-wise; humans showing 23 pairs from chr 1 to chr X (plotted at 23 of x-axis) or Y (plotted at 24 of x-axis) (Morton, 1991). Note that panda and giraffe have 21 and 15 pairs of chromosomes, respectively. (b) Chromosome count of frequently used laboratory animals (mice & rats) and humans as compared to other species such as chimpanzees (chimp) to Ants. Note that chimp (48) and pig (38) have somewhat similar numbers to humans at 46 chromosomes (23 pairs) while snails are fewer at 24. (c) Scatter plots showing chromosome length of three species with respect to each chromosome of humans (hm), chimpanzee, and pig. Note that taller animals do not always show longer chromosomes. In chr 1, pigs, which are shorter in body size, have longer nucleotide sizes than humans and chimps. x-axis is pairwise; humans showing 23 pairs from chr 1 to chr X (plotted at 23 of x-axis) or Y (plotted at 24 of x-axis). (d) Chromosome number of ten (a set of #11-20) genes associated with CH over three species of mice, rats, and humans. Note that all three genes of SGSM3, HYDIN, and LRRK2 are located on different autosomes depending on species.

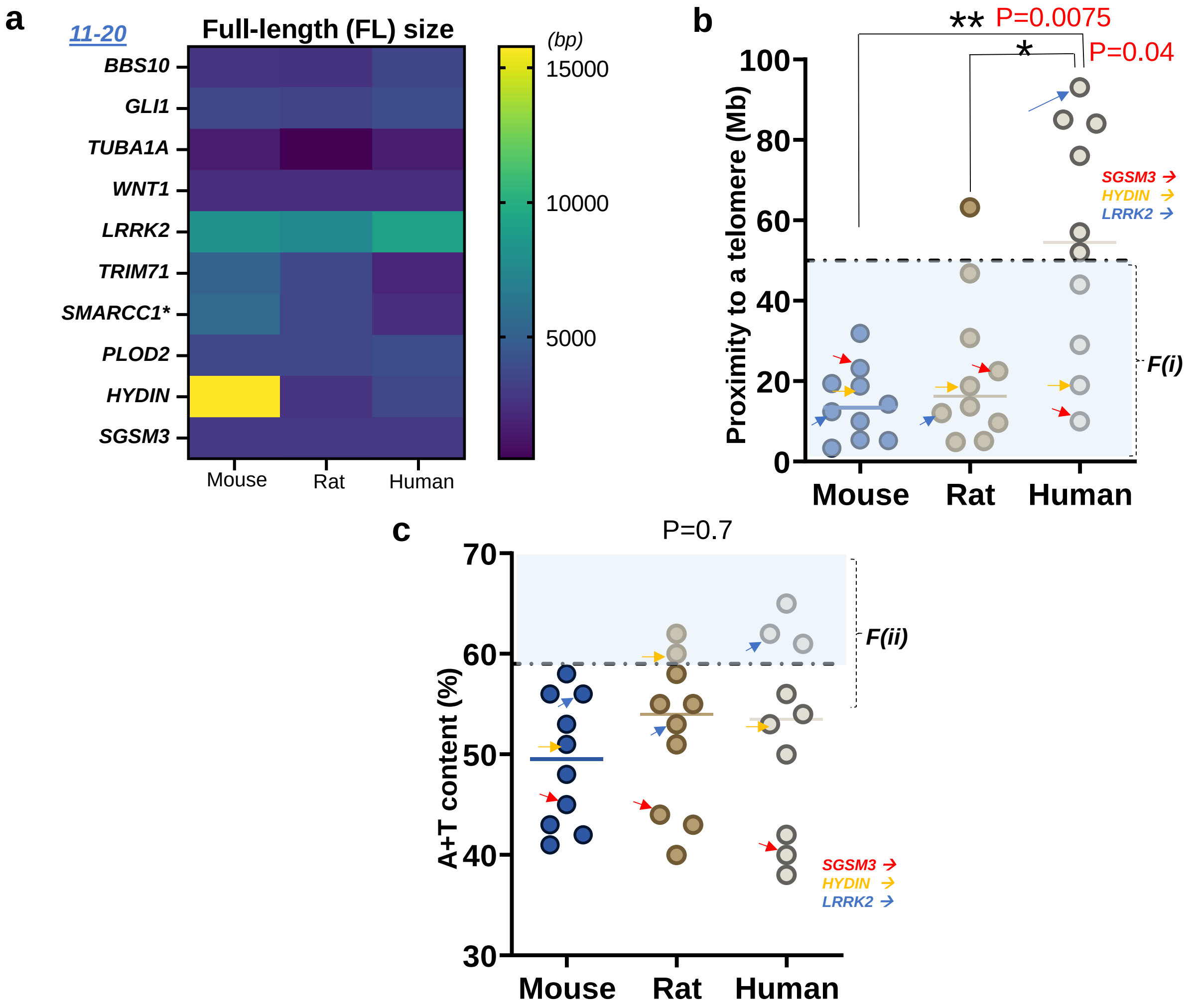

Figure S4. Chromosomal characteristics of two factors in mice, rats, and humans. (a) The full-length (FL) size of ten genes (a set of #11-20) over three species. Note that WNT1, PLOD2, and SGSM3 show consistent molecular sizes in three species while other genes slightly differ in humans as compared to rodents (b) Scatter plot summarizing ten genes (#11-20) causing CH with proximity to telomeres over mouse, rat, and human chromosomes. A horizontal dotted line indicates 50 Mb. F(i) and F(ii) indicate the first and second factor, respectively. Note that LRRK2 indicated by arrows in blue no longer satisfied F(i) in humans. *, $P < 0.05$; **, $P < 0.01$ by Kruskal-Wallis test. (c) Scatter plot showing A + T content of ten genes (#11-20) over three species demonstrates 2 (rat) or 3 (human) among 10 genes meeting the second factor, F(ii), associated with high mutation rate. Arrows in red, orange, and blue indicate SGSM3, HYDIN, and LRRK2, respectively (throughout b-c).

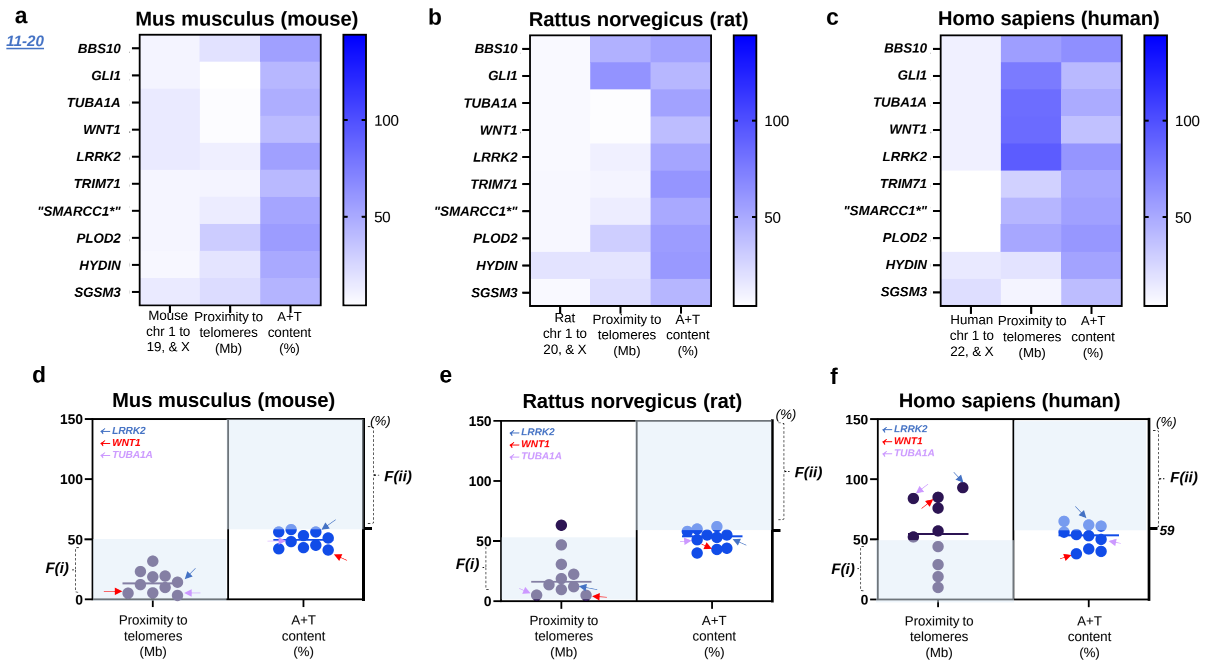

Figure S5. Relative mutability of 10 human genes by two factors. (a) The color-coded map showing chromosomal location, proximity to a telomere, and A+T content of Mus musculus (mouse) BBS10 and 9 other genes in chromosomes (chr 1 to 19 and X), proximity to telomeres (Mb), and A+T content (%). (b) the color-coded map showing chromosomal location, proximity to a telomere, and A+T content of Rattus norvegicus (rat) BBS10 and 9 other genes in chromosomes (chr 1 to 20 and X), proximity to telomeres (Mb), and A+T content (%). (c) the color-coded map showing chromosomal location, proximity to a telomere, and A+T content of Homo sapiens (human) BBS10 and 9 other genes in chromosomes (chr 1 to 22 and X), proximity to a telomere (Mb), and A+T content (%). (d) Scatter plots exhibiting relative mutability of ten mouse genes by proximity to telomeres (left) and A+T content (right). Shaded regions with light blue indicate genes satisfying either of the two factors. Genes outside the shaded regions at n = 0 by F(i); n = 7 by F(ii)) are relatively less mutable. (e) Scatter plots displaying relative mutability of ten rat genes by proximity to a telomere (left) and A+T content (right). Shaded regions with light blue indicate genes satisfying either of the two factors. Genes outside the shaded regions at n = 1 by F(i); n = 7 by F(ii)) are relatively less mutable. (f) Scatter plots indicating relative mutability of the same ten genes by proximity to a telomere (left) and A+T content (right) in human chromosomes. Shaded regions with light blue indicate human genes satisfying either of the two factors. Genes outside the shaded regions at n = 6 by F(i); n = 7 by F(ii)) are relatively less mutable. Note that unlike mice and rats, only 5 of 10 human genes satisfied proximity to telomeres, F(i).

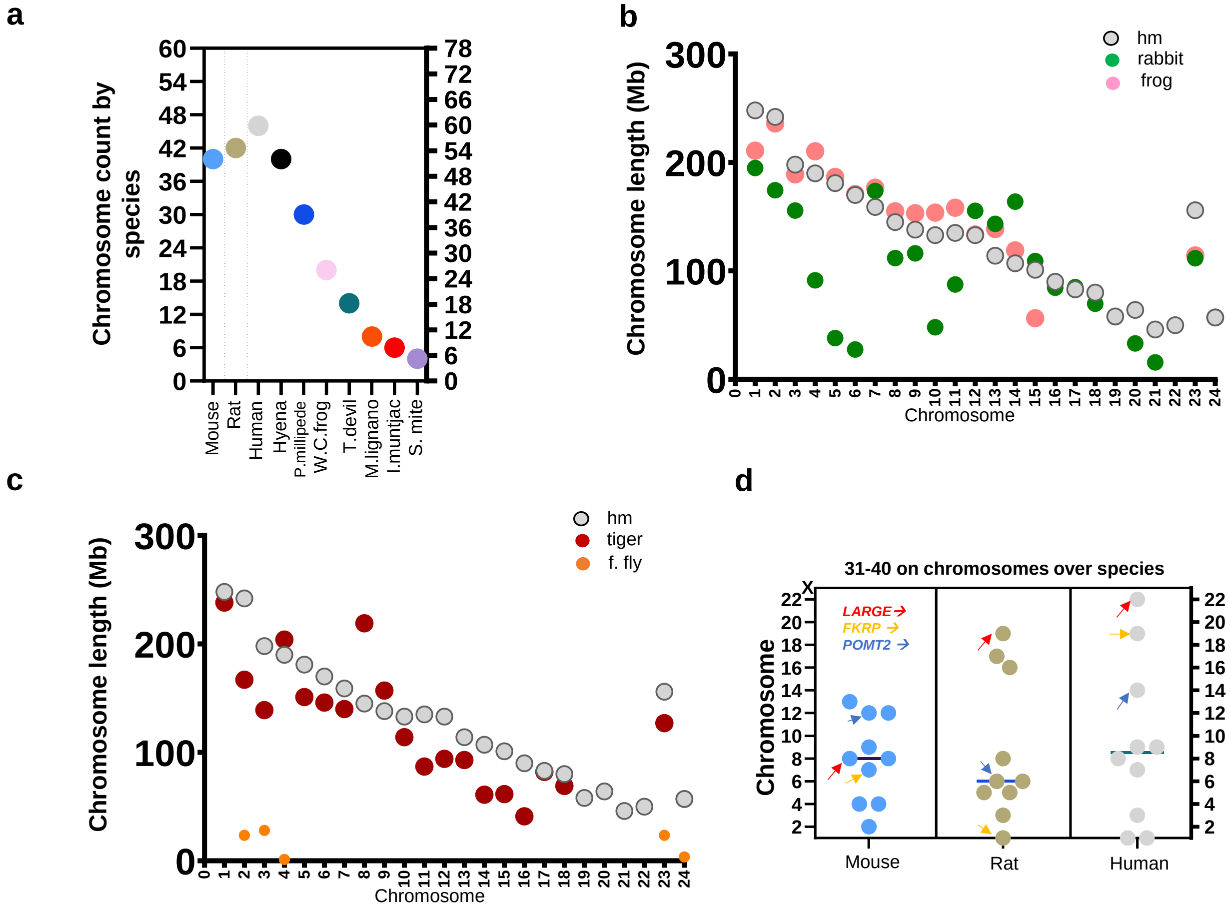

Figure S6. Species differences in chromosomes and relative mutability of 10 other human genes associated with CH by two factors. (a) Chromosome count of frequently used laboratory animals (mice & rats) and humans as compared to other species such as Hyena to Spider Mite (S. Mite). Pill millipede, P. millipede; Western clawed frog, W.C. frog; Tasmanian devil, T. devil; Indian muntjac, I. muntjac. Note that hyena have somewhat similar numbers to humans at 40 chromosomes (20 pairs) while S. Mites are fewer at 4. (b) Scatter plots showing chromosome length of three species with respect to each chromosome of humans (hm), rabbits, and (African clawed) frogs. Note that taller animals do not always show longer chromosomes. In chr 4, (African clawed) frogs, which are shorter in body size, have longer nucleotide sizes than humans and rabbits. x-axis is pairwise; humans showing 23 pairs from chr 1 to chr X (plotted at 23 of x-axis) or Y (plotted at 24 of x-axis); rabbits having 22 pairs; African clawed frogs showing 18 pairs. (c) Scatter plots showing chromosome length of three species with respect to each chromosome of humans (hm), tigers, and fruit flies. Note that taller animals do not always show longer chromosomes. In chr 4, tigers, which are shorter in body size, have longer nucleotide sizes than humans. x-axis is pairwise; humans showing 23 pairs from chr 1 to chr X (plotted at 23 of x-axis) or Y (plotted at 24 of x-axis); tigers having 19 pairs; fruit flies showing 4 pairs. (d) Chromosome number of ten (a set of #31-40) genes associated with CH over three species of mice, rats, and humans. Note that three genes of LARGE, FKRP, and POMT2 are located on different autosomes depending on species.

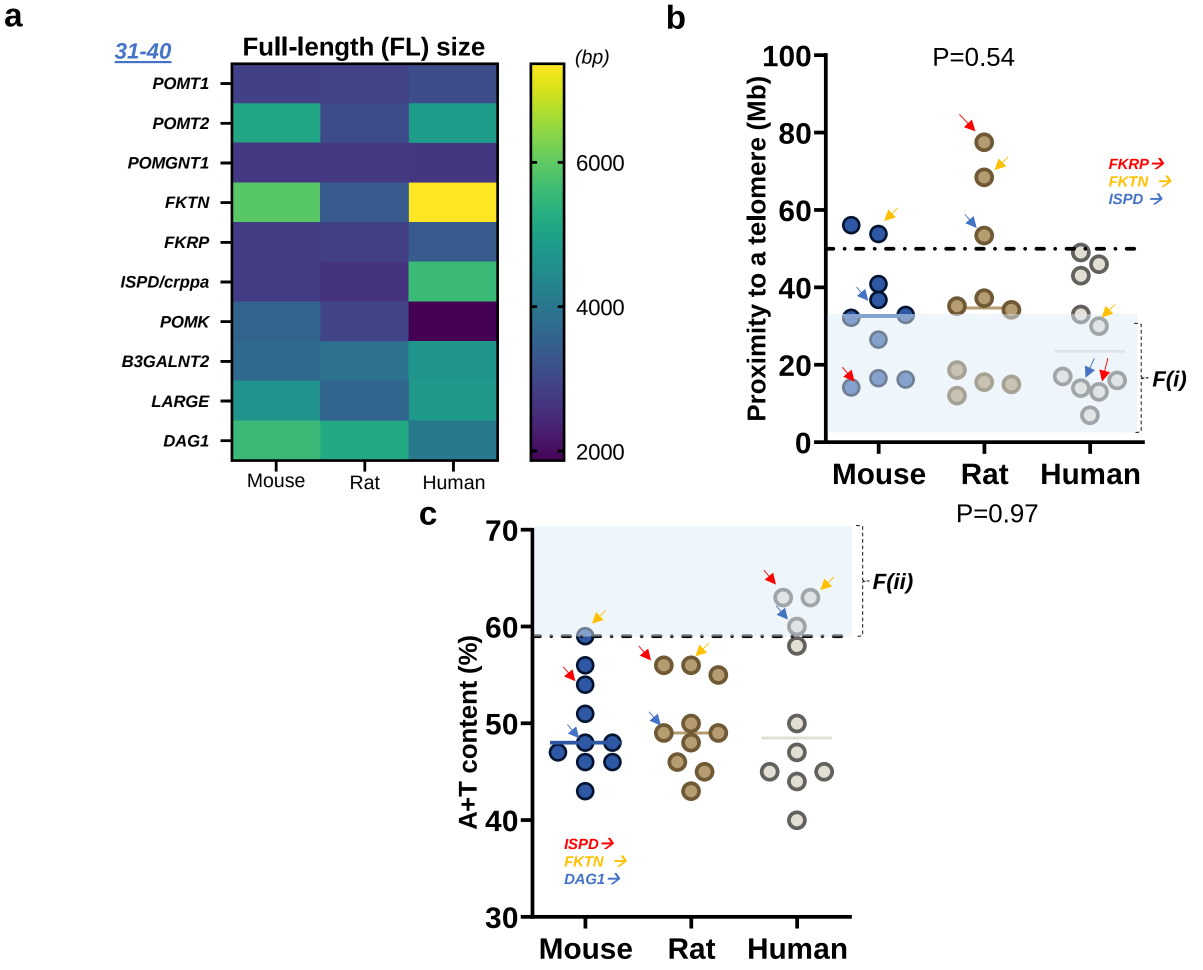

Figure S7. Nucleotide size and two factors in rodent and human chromosomes. (a) The nucleotide size of ten genes (a set of #31-40) associated with CH over three species. Note that POMT1, POMGNT1, and FKRP show consistent molecular sizes in three species while other genes slightly differ in humans as compared to rodents (b) Scatter plot demonstrating ten genes (#31-40) with proximity to telomeres over mouse, rat, and human chromosomes. A horizontal dotted line indicates 50 Mb. F(i) indicates the first factor or proximity to telomeres. Note that FKRP, FKTN, and ISPD indicated by arrows in red, yellow, and blue no longer satisfied F(i) in rats evolved in a more mutable way in humans, satisfying F(i) at <50 Mb. (c) Scatter plot showing A + T content of ten genes (#31-40) over three species displays 1 (mouse) or 3 (human) among 10 genes satisfying the second factor, F(ii), associated with high mutation rate. Arrows in red, orange, and blue indicate ISPD, FKTN, and DAG1, respectively. Statistical analysis by Kruskal-Wallis test (b-c).

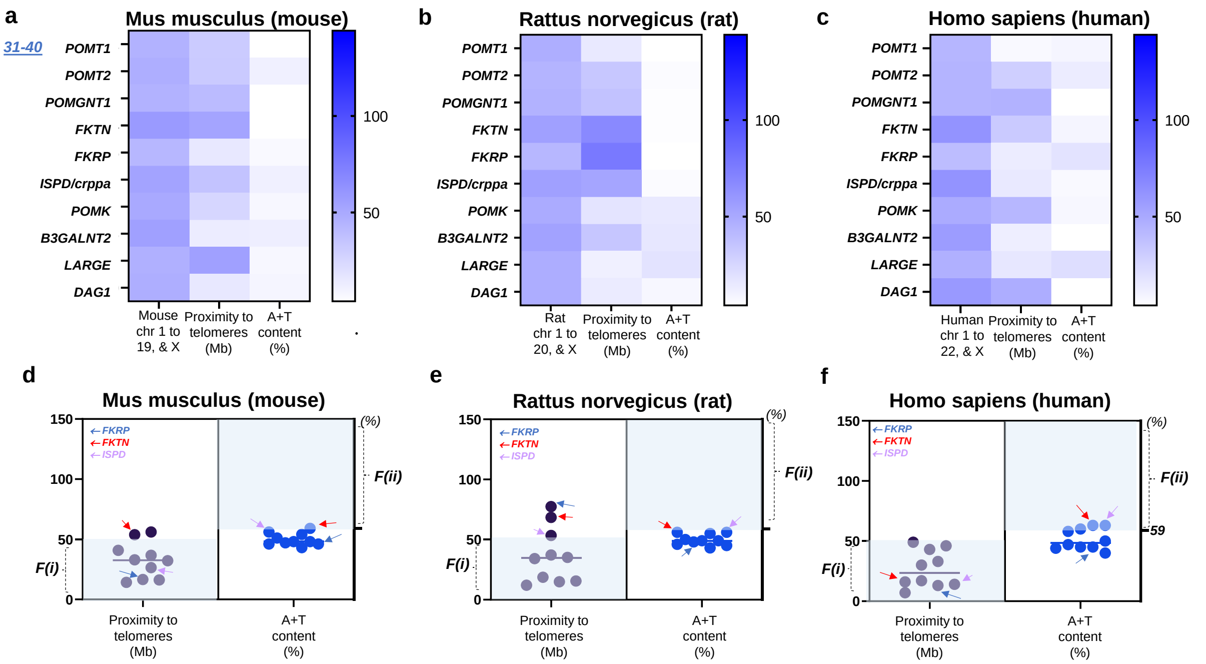

Figure S8. Relative mutability of 10 genes by two factors over rodent and human chromosome. (a) The color-coded map showing chromosomal location, proximity to a telomere, and A+T content of Mus musculus (mouse) POMT1 and 9 other genes in chromosomes (chr 1 to 19 and X), proximity to telomeres (Mb), and A+T content (%). (b) the color-coded map showing chromosomal location, proximity to a telomere, and A+T content of Rattus norvegicus (rat) POMT1 and 9 other genes in chromosomes (chr 1 to 20 and X), proximity to telomeres (Mb), and A+T content (%). (c) the color-coded map showing chromosomal location, proximity to a telomere, and A+T content of Homo sapiens (human) POMT1 and 9 other genes in chromosomes (chr 1 to 22 and X), proximity to a telomere (Mb), and A+T content (%). (d) Scatter plots summarizing relative mutability of ten mouse genes associated with CH by proximity to telomeres (left) and A+T content (right). Shaded regions with light blue indicate genes meeting either of the two factors. Genes outside the shaded regions at n = 2 by F(i); n = 8 by F(ii)) are relatively less mutable. (e) Scatter plots showing relative mutability of ten rat genes associated with CH by proximity to a telomere (left) and A+T content (right). Shaded regions with light blue indicate genes meeting either of the two factors. Genes outside the shaded regions at n = 3 by F(i); n = 7 by F(ii)) are relatively less mutable. (f) Scatter plots displaying relative mutability of the same ten genes by proximity to a telomere (left) and A+T content (right) in human chromosomes. Shaded regions with light blue indicate human genes satisfying either of the two factors. Genes outside the shaded regions at n = 1 by F(i); n = 6 by F(ii)) are relatively less mutable. Note that unlike mice and rats, all (10 of 10) human genes satisfied proximity to telomeres, F(i).

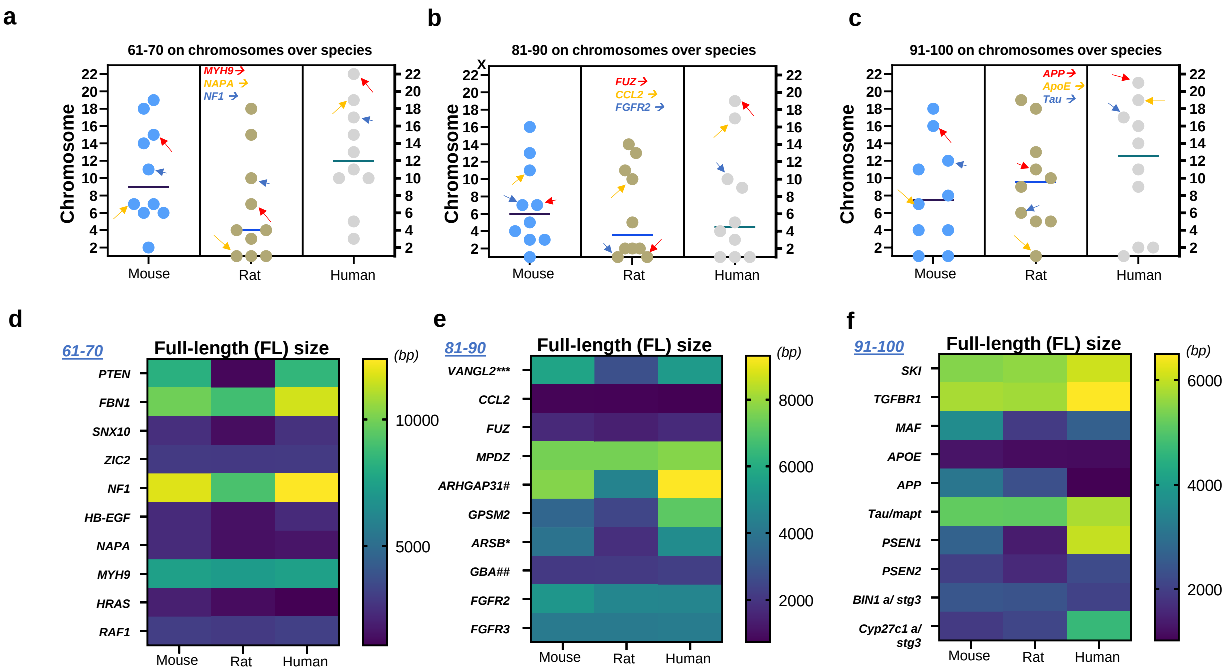

Figure S9. Chromosome numbers and nucleotide sizes of thirty genes associated with CH over three species. (a) Chromosome number of ten (a set of #61-70) genes associated with CH over three species of mice, rats, and humans. Note that three genes of MYH9, NAPA, and NF1 are located on different autosomes depending on species. (b) Chromosome number of ten (a set of #81-90) genes associated with CH over three species of mice, rats, and humans. Note that three genes of FUZ, CCL2, and FGFR2 are located on different autosomes depending on species. (c) Chromosome number of ten (a set of #91-100) genes associated with CH (#91-93) and/or Alzheimer's disease (#94-100) over three species of mice, rats, and humans. Note that three genes of APP, ApoE, and Tau are located on different autosomes depending on species. (d) The nucleotide size of ten genes (a set of #61-70) associated with CH over three species. Note that ZIC2, MYH9, HRAS, and RAF1 show consistent molecular sizes in three species while other genes slightly differ in humans as compared to rodents (e) The nucleotide size of ten genes (a set of #81-90) associated with CH over three species. Note that CCL2, MPDZ, GBA, and FGFR3 show consistent molecular sizes in three species while other genes slightly differ in humans as compared to rodents (f) The nucleotide size of ten genes (a set of #91-100) associated with CH and/or Alzheimer's disease over three species. Note that APOE show consistent molecular sizes in three species while other genes slightly differ in humans as compared to rodents

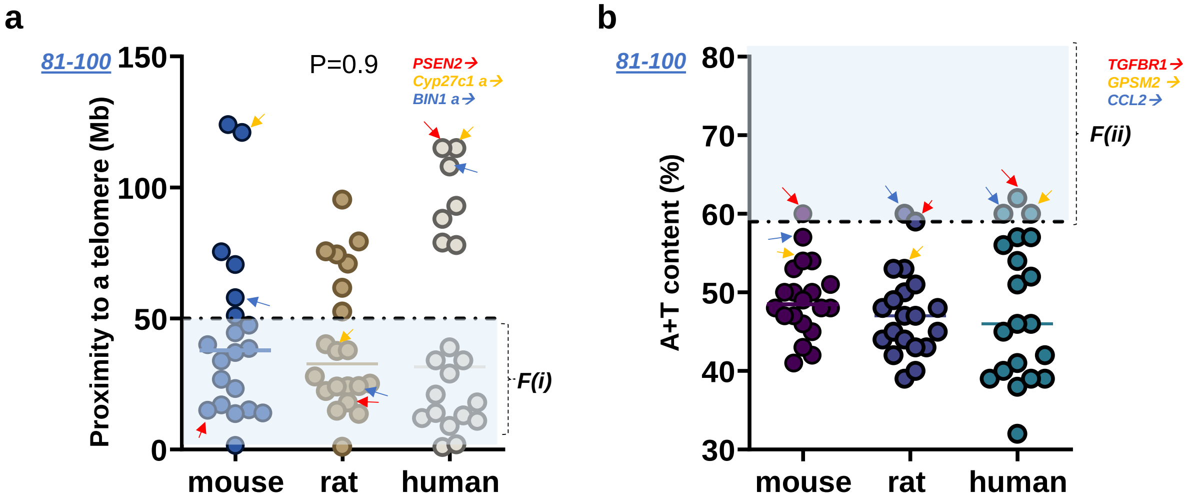

Figure S10. Proximity to telomeres and A+T content of twenty genes associated with CH and/or Alzheimer's disease over three species. (a) Scatter plot demonstrating twenty genes (#81-100) associated with CH and/or Alzheimer's disease with proximity to telomeres over mouse, rat, and human chromosomes. A horizontal dotted line indicates 50 Mb. F(i) indicates the first factor or proximity to telomeres. Note that PSEN2, Cyp28c1a, and BIN1a indicated by arrows in red, yellow, and blue no longer met F(i) in humans as compared to either mouse or rat chromosome. (b) Scatter plot displaying twenty genes (#81-100) associated with CH and/or Alzheimer's disease with A + T content over mouse, rat, and human chromosomes. A horizontal dotted line indicates 59 %. F(ii) indicates the second factor or high A + T content at >59%. Note that TGFBR1, GPSM2, and CCL2 indicated by arrows in red, yellow, and blue satisfied F(ii) in humans while the same no longer applies to mouse or rat chromosome

Table S1a. Two factor characteristics of mouse genes associated with fPD

Mouse gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
1	PARK1 or PARK4	6	60	149	89	29,27	56	1208
2	PARK2	17	11	95	84	26,24	50	3216
3	PARK3	NA						
4	PARK5	5	66	151	85	21,23	44	1156
5	PARK6	4	138	157	19	22,22	44	2774
6	PARK7	4	150	157	7	25,19	44	920
7	PARK8	15	91	104	13	29,26	55	8231
8	PARK9	4	140	157	17	18,19	37	3908
9	PARK10	4	106	157	51	27,26	53	10601
10	PARK11	1	87	195	108	29,22	51	5743
11	PARK13	6	83	150	67	21,22	43	1739
12	PARK14	15	79	104	25	22,19	41	3106
13	PARK15	10	85	131	46	23,25	48	1886
14	PARK17	8	85	130	45	30,29	60	3196
15	PARK18	16	20	98	78	26,20	46	5358
16	GBA	3	89	160	71	23,24	47	5906
17	NR4A2	2	56	182	126	26,24	50	3172

Table S1b. Two factor characteristics of rat genes associated with fPD

Rat gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
1	PARK1 or PARK4	4	89.74	182.687	92.947	28,27	55	1150
2	PARK2	1	48.8	260.522	211.722	24,20	44	1564
3	PARK3	NA						
4	PARK5	14	41	104	63	22,23	45.1	1029
5	PARK6	5	150	156	6	22,18	40	1146
6	PARK7	5	161	166	5	26,21	47	1054
7	PARK8	7	122	135	13	28,25	53	8084
8	PARK9	5	153	166	13	21,19	40	3905
9	PARK10	5	121	166	45	27,26	53	10593
10	PARK11	9	88	114	26	29,21	50	5900
11	PARK13	4	115	183	68	21,23	44	1607
12	PARK14	7	110	135	25	23,20	43	3100
13	PARK15	7	17	135	118	25,23	48	1819
14	PARK17	19	21	54	33	30,29	60	3196
15	PARK18	11	80	85	5	25,20	45	5421
16	GBA	2	174	249	75	25,24	49	2605
17	NR4A2	3	41	169	128	27,25	52	2582

Table S1c. Two factor characteristics of human genes associated with fPD

Human gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
1	PARK1 or PARK4	4	q22.1	q35.1	100	27,30	57	7111
2	PARK2	6	q26	q27	9	27,26	53	4178
3	PARK3	2	p13	p25.3	72	20,22	42	1432
4	PARK5	2	p13.3	p25.3	72	20,22	42	1432
5	PARK6	1	p36.12	p36.32	20	20,23	43	2657
6	PARK7	1	p36.32	p36.32	7	27,28	55	1127
7	PARK8	12	q12	q24.32	93	31,31	62	9239
8	PARK9	1	p36.21	p36.32	16	17,20	37	3996
9	PARK10	1	p32.2	p36.32	55	28,29	57	10800
10	PARK11	2	q37.1	q37.3	10	28,27	55	7928
11	PARK13	2	p12	p25.3	74	20,23	43	1785
12	PARK14	22	q13.1	q13.32	12	20,20	40	3299
13	PARK15	22	q12.3	q13.32	18	23,28	51	2060
14	PARK17	16	q12.1	q24.2	44	28,31	59	6776
15	PARK18	3	q26.32	q29	14	24,21	45	5078
16	GBA	1	q21.3	q44	92	22,24	46	2291
17	NR4A2	2	q24.1	q37.3	84	26,27	53	3475

Table S2a. Relative mutability of ten human genes (#1-10) associated with CH

Human gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
1	L1CAM	X	q28	q28	3	21,20	41	5099
2	ZCCHC8	12	q24.31	q24.32	11	30,28	58	4113
3	Orai1	12	q24.31	q24.32	12	16,19	35	1496
4	MYC	8	q24.13	q24.3	18	26,26	52	3721
5	TMEM116	12	q24.13	q24.32	22	25,30	55	1543
6	TRPV4	12	q24.11	q24.32	24	18,23	41	3228
7	IGF1	12	q23.1	q24.32	31	31,33	64	7277
8	DCN	12	q21.33	q24.32	42	30,33	63	6850
9	MKS4/cep290	12	q21.33	q24.32	45	41,25	66	7824
10	TMEM67/MKS3	8	q22.2	q24.3	51	30,34	64	4678

Table S2b. Relative mutability of ten rat genes (#1-10) associated with CH

Rat gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
1	L1CAM	X	151623	152,453	1	25,24	49	5079
2	ZCCHC8	12	32892	46669	13.777	27,25	52	2818
3	Orai1	12	33548	46669	13.121	20,17	37	1184
4	MYC	7	93598	135012	41.414	25,24	49	2355
5	TMEM116	12	35089	46669	11.58	23,26	49	1357
6	TRPV4	12	41997	46669	4.672	20,23	43	3211
7	IGF1	7	22361	0	22.361	30,30	60	1530
8	DCN	7	32321	0	32.321	33,27	60	1639
9	MKS4/cep290	7	35399	0	35.399	39,23	62	8232
10	TMEM67/MKS3	5	25589	0	25.589	26,32	58	4035

Table S2c. Relative mutability of ten mouse genes (#1-10) associated with CH

Mouse gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
1	L1CAM	X	72924	0	72.924	25,22	47	5288
2	ZCCHC8	5	123859	151758	27.899	26,27	53	4286
3	Orai1	5	123168	151758	28.59	19,23	42	1871
4	MYC	15	61862	104073	42.211	24,23	47	2399
5	TMEM116	5	121656	151758	30.102	21,28	49	1297
6	TRPV4	5	114796	151758	36.962	20,22	42	3247
7	IGF1	10	87772	130530	42.758	30,33	63	7125
8	DCN	10	97354	130530	33.176	32,27	59	1829
9	MKS4/cep290	10	100411	130530	30.119	39,23	62	8188
10	TMEM67/MKS3	4	12090	0	12.09	27,31	58	3452

Table S3a. Relative mutability of ten human genes (#11-20) associated with CH

Human gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
11	BBS10	12	q21.1	q24.32	57	30,35	65	3568
12	GLI1	12	q13.2	q24.32	76	21,21	42	3972
13	TUBA1A	12	q13.11	q24.32	84	23,27	50	1672
14	WNT1	12	q13.11	q24.32	85	17,21	38	2422
15	LRRK2	12	q12	q24.32	93	31,31	62	9239
16	TRIM71	3	P22.5	P26.1	29	24 29	53%	2002
17	SMARCC1*	3	P23.5	P26.1	44	25 31	56%	2344
18	PLOD2	3	Q23.5	Q29	52	30 31	61%	4009
19	HYDIN	16	Q22.2	Q24.2	19	27 27	54%	3720
20	SGSM3	22	Q13.1	Q13.31	10	19 21	40%	2986

Table S3b. Relative mutability of ten rat genes (#11-20) associated with CH

Rat gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
11	BBS10	7	46754	0	46.754	26,29	55	2667
12	GLI1	7	63169	0	63.169	22,21	43	3501
13	TUBA1A	7	130116	135012	4.896	23,32	55	482
14	WNT1	7	129942	135012	5.07	18,22	40	2360
15	LRRK2	7	122987	135012	12.025	27,26	53	7581
16	TRIM71	8	114250	123900	9.65	35,27	62	3769
17	SMARCC1*	8	110214	123900	13.686	29,22	51	3670
18	PLOD2	8	93167	123900	30.733	28,30	58	3672
19	HYDIN	19	38583	57337	18.754	34,26	60	2756
20	SGSM3	7	112544	135012	22.468	22,22	44	2854

Table S3c. Relative mutability of ten mouse genes (#11-20) associated with CH

Mouse gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
11	BBS10	10	111137	130530	19.393	27,29	56	2753
12	GLI1	10	127177	130530	3.353	21,22	43	3662
13	TUBA1A	15	98851	104073	5.222	22,26	48	1591
14	WNT1	15	98691	104073	5.382	17,24	41	2371
15	LRRK2	15	91700	104073	12.373	28,28	56	8231
16	TRIM71	9	114393	124359	9.966	19,23	42	5197
17	SMARCC1*	9	110069	124359	14.29	27,26	53	5684
18	PLOD2	9	92490	124359	31.869	28,30	58	3656
19	HYDIN	8	111336	130127	18.791	27,24	51	15783
20	SGSM3	15	80896	104073	23.177	22,23	45	3013

Table S4a. Relative mutability of ten human genes (#21-30) associated with CH

Human gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
21	EN1 **	2	Q14.2	Q36.3	124	20 21	41%	2427
22	RFX4	12	Q23.3	Q24.22	27	27 28	53%	3345
23	MSX1	4	P16.2	P16.3	3	22 20	42%	1940
24	PAX6	11	P13	P15.5	30	31 25	56%	2500
25	RFX3	9	P24.2	P24.2	0.5	30 30	60%	6274
26	PAC1/klf6	10	P15.3	P15.3	0.2	26 29	55%	4590
27	SOCS1	16	P13.13	P13.3	10	14 21	35%	1216
28	FOXJ1	17	Q25.2	Q25.3	7	19 16	35%	2587
29	AP1S2	23	P22.2	P22.3	15	30 34	64%	2119
30	CCDC88C	14	Q32.11	Q32.33	16	24 18	42%	7519

Table S4b. Relative mutability of ten rat genes (#21-30) associated with CH

Rat gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
21	EN1 **	13	31756	0	31.756	22,21	43	2678
22	RFX4	7	19015	0	19.015	25,22	47	2052
23	MSX1	14	72964	104886	31.922	21,21	42	1894
24	PAX6	3	92157	169034	76.877	30,24	54	2191
25	RFX3	1	225709	260522	34.813	29,24	53	2441
26	PAC1/klf6	17	64548	86533	21.985	27,31	58	4206
27	SOCS1	10	4884	0	4.884	15,22	37	1190
28	FOXJ1	10	101570	107211	5.641	21,18	39	2427
29	AP1S2	23	30598	0	30.598	31,31	62	1838
30	CCDC88C	6	120289	140994	20.705	25,19	44	7575

Table S4c. Relative mutability of ten mouse genes (#21-30) associated with CH

Mouse gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
21	EN1 **	1	120535	195154	74.619	21,23	44	2523
22	RFX4	10	84742	130530	45.788	26,28	54	3274
23	MSX1	5	37981	0	37.981	21,21	42	1931
24	PAX6	2	105528	181755	76.227	27,26	53	4207
25	RFX3	19	27995	0	27.995	28,32	60	9187
26	PAC1/klf6	13	5920	0	5.92	28,30	58	4217
27	SOCS1	16	10603	0	10.603	15,23	38	1220
28	FOXJ1	11	116226	121973	5.747	22,18	40	2599
29	AP1S2	23	162716	169476	6.76	30,34	64	3327
30	CCDC88C	12	100995	120092	19.097	26,19	45	7377

Table S5a. Relative mutability of ten human genes (#31-40) associated with CH

Human gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
31	POMT1	9	Q34.12	Q34.3	7	20 24	44%	3123
32	POMT2	14	Q24.3	Q31.1	30	20 25	45%	4875
33	POMGNT1	1	P34.1	P36.33	46	22 23	45%	2719
34	FKTN	9	Q31.2	Q34.3	33	30 33	63%	7364
35	FKRP	19	Q13.32	Q13.43	14	18 22	40%	3411
36	ISPD/crppa	7	P21.2	P22.3	16	30 33	63%	5592
37	POMK	8	P11.21	P23.3	43	23 27	50%	1868
38	B3GALNT2	1	P42.2	Q44	13	27 31	58%	4719
39	LARGE	22	Q12.3	Q13.33	17	23 24	47%	4816
40	DAG1	3	P21.31	P26.3	49	31 29	60%	4074

Table S5b. Relative mutability of ten rat genes (#31-40) associated with CH

Rat gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
31	POMT1	3	15538	0	15.538	22,26	48	2929
32	POMT2	6	106794	140994	34.2	21,24	45	3097
33	POMGNT1	5	129644	166875	37.231	23,23	46	2745
34	FKTN	5	68396	0	68.396	28,28	56	3402
35	FKRP	1	77486	0	77.486	20,23	43	2860
36	ISPD/crppa	6	53397	0	53.397	27,29	56	2645
37	POMK	16	66101	84729	18.628	25,25	50	2967
38	B3GALNT2	17	51377	86533	35.156	27,28	55	3936
39	LARGE	19	12048	0	12.048	24,25	49	3634
40	DAG1	8	108955	123900	14.945	23,26	49	5169

Table S5c. Relative mutability of ten mouse genes (#31-40) associated with CH

Mouse gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
31	POMT1	2	32145	0	32.145	22,24	46	2884
32	POMT2	12	87194	120092	32.898	21,27	48	5082
33	POMGNT1	4	116017	156860	40.843	23,23	46	2736
34	FKTN	4	53765	0	53.765	28,31	59	5924
35	FKRP	7	16550	0	16.55	19,24	43	2817
36	ISPD/crppa	12	36739	0	36.739	26,28	54	2818
37	POMK	8	26484	0	26.484	24,27	51	3601
38	B3GALNT2	13	14173	0	14.173	27,29	56	3716
39	LARGE	8	74080	130127	56.047	23,24	47	4671
40	DAG1	9	108141	124359	16.218	22,26	48	5591

Table S6a. Relative mutability of ten human genes (#41-50) associated with CH

Human gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
41	TBX1	22	Q11.21	Q13.33	31	20 16	36%	1538
42	B3GNT1*	11	Q13.2	Q25	69	20 24	44%	2007
43	B3GALTL	13	Q12.3	Q34	83	28 32	60%	4215
44	LAMB1	7	Q22.3	Q36.3	49	27 25	52%	5650
45	ALG13	X	Q23	Q28	45	32 35	67%	2718
46	DNAH5	5	P15.2	P15.33	13	30 27	57%	15781
47	DYX1C1	15	Q21.3	Q26.3	46	34 27	61%	1993
48	CENPF	1	Q41	Q44	34	36 24	60%	10290
49	IFT172	2	P23.2	P25.3	27	26 24	50%	5395
50	KIAA0586	14	Q23.1	Q32.33	49	30 31	61%	8330

Table S6b. Relative mutability of ten rat genes (#41-50) associated with CH

Rat gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
41	TBX1	11	82419	86241	3.822	18,16	34	1855
42	B3GNT1*	1	202345	260522	58.177	22,25	47	1980
43	B3GALTL	12	5346	0	5.346	26,28	54	4429
44	LAMB1	6	47902	0	47.902	25,24	49	5822
45	ALG13	X	107906	152453	44.547	33,30	63	1625
46	DNAH5	2	79255	0	79.255	28,27	55	15796
47	DYX1C1	8	73711	123900	50.189	31,25	56	2012
48	CENPF	13	101229	106807	5.578	35,23	58	10069
49	IFT172	6	25121	0	25.121	25,22	47	5382
50	KIAA0586	6	89725	140994	51.269	31,24	55	4983

Table S6c. Relative mutability of ten mouse genes (#41-50) associated with CH

Mouse gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
41	TBX1	16	18409	0	18.409	17,19	36	1786
42	B3GNT1*	11	22810	0	22.81	26,29	55	2699
43	B3GALTL	5	149686	151758	2.072	27,29	56	4570
44	LAMB1	12	31379	0	31.379	25,24	49	5784
45	ALG13	X	143157	169476	26.319	32,31	63	1640
46	DNAH5	15	28472	0	28.472	28,26	54	15616
47	DYX1C1	9	72880	124359	51.479	32,24	56	1978
48	CENPF	1	189420	195154	5.734	34,24	58	11130
49	IFT172	5	31448	0	31.448	25,23	48	5440
50	KIAA0586	6	71852	0	71.852	29,23	52	2777

Table S7a. Relative mutability of ten human genes (#51-60) associated with CH

Human gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
51	MKS1	17	Q22	Q25.3	24	22 25	46%	2343
52	TMEM216**	11	Q12.1	Q25	74	20 28	48%	1287
53	CC2D2A	4	P15.32	P16.3	15	31 24	56%	1682
54	CRB2	9	Q33.3	Q34.3	15	14 22	36%	5613
55	HYLS1	11	Q24.1	Q25	10	32 26	58%	1829
56	KIF7	15	Q26.1	Q26.3	12	20 17	37%	4567
57	GLI3	7	P14.1	P22.3	42	24 26	50%	8405
58	CEP83	12	Q22	Q24.33	39	36 26	62%	3686
59	OFD1	X	P22.2	P22.33	13	33 25	58%	3612
60	ZIC3	X	Q26.3	Q28	19	24 29	53%	4001

Table S7b. Relative mutability of ten rat genes (#51-60) associated with CH

Rat gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
51	MKS1	10	72667	107211	34.544	25,23	48	2480
52	TMEM216**	1	207201	260522	53.321	22,34	56	1310
53	CC2D2A	14	67435	104886	37.451	30,26	56	7779
54	CRB2	3	21564	0	21.564	14,24	38	3852
55	HYLS1	8	33922	0	33.922	29,27	56	1733
56	KIF7	1	133658	260522	126.864	23,18	41	4651
57	GLI3	17	49709	86533	36.824	25,21	46	5027
58	CEP83	7	29389	0	29.389	35,23	58	3000
59	OFD1	X	28056	0	28.056	33,25	58	3561
60	ZIC3	X	136134	152453	16.319	20,20	40	1236

Table S7c. Relative mutability of ten mouse genes (#51-60) associated with CH

Mouse gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
51	MKS1	11	87754	121973	34.219	24,23	47	2516
52	TMEM216**	19	10533	0	10.533	21,32	53	1101
53	CC2D2A	5	43898	0	43.898	30,24	54	5463
54	CRB2	2	37689	0	37.689	18,24	42	6393
55	HYLS1	9	35481	0	35.481	28,29	57	1643
56	KIF7	7	79365	144995	65.63	23,18	41	4522
57	GLI3	13	15904	0	15.904	26,26	52	8427
58	CEP83	10	94626	130530	35.904	35,23	58	3006
59	OFD1	X	165223	169476	4.253	31,27	58	4453
60	ZIC3	X	57081	0	57.081	25,30	55	4035

Table S8a. Relative mutability of ten human genes (#61-70) associated with CH

Human gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
61	PTEN	10	Q23.31	Q26.3	46	30 34	64%	8515
62	FBN1	15	Q21.1	Q26.3	49	27 28	55%	11609
63	SNX10	10	P15.2	P22.3	26	31 34	64%	2665
64	ZIC2	13	Q32.3	Q34	14	20 23	43%	2963
65	NF1	17	Q11.2	Q25.3	50	29 30	59%	12381
66	HB-EGF	5	Q31.3	Q35.3	41	24 27	51%	2358
67	NAPA	19	Q13.32	Q13.43	11	21 22	43%	1662
68	MYH9	22	Q12.3	Q13.33	11	25 18	43%	7451
69	HRAS	11	P15.5	P15.5	0.5	20 17	37%	1070
70	RAF1	3	P25.2	P26.3	12	24 26	50%	3191

Table S8b. Relative mutability of ten rat genes (#61-70) associated with CH

Rat gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
61	PTEN	1	230696	260522	29.826	33,27	60	1212
62	FBN1	3	112750	169034	56.284	25,24	49	8942
63	SNX10	4	80667	0	80.667	29,28	55	1423
64	ZIC2	15	99581	101769	2.188	21,24	45	2903
65	NF1	10	64539	107211	42.672	27,28	55	9132
66	HB-EGF	18	28116	0	28.116	25,27	52	1549
67	NAPA	1	76805	0	76.805	24,22	46	1504
68	MYH9	7	109424	135012	25.588	26,19	45	7208
69	HRAS	1	196299	260522	64.223	23,22	45	1392
70	RAF1	4	148740	102640	46.1	24,26	50	2970

Table S8c. Relative mutability of ten mouse genes (#61-70) associated with CH

Mouse gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
61	PTEN	19	32803	61420	28.617	28,32	60	8229
62	FBN1	2	125348	181755	56.407	25,24	49	9900
63	SNX10	6	51567	0	51.567	29,32	61	2567
64	ZIC2	14	122717	125139	2.422	20,25	45	2967
65	NF1	11	79472	121973	42.501	27,30	57	11847
66	HB-EGF	18	36648	0	36.648	24,28	52	2392
67	NAPA	7	15851	0	15.851	24,23	47	2402
68	MYH9	15	77726	104073	26.347	25,19	46	7439
69	HRAS	7	140773	144995	4.222	20,24	44	2043
70	RAF1	6	115653	149588	33.935	24,25	49	3084

Table S9a. Relative mutability of ten human genes (#71-80) associated with CH

Human gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
71	BRAF	7	Q34	Q36.3	19	28 32	60%	6459
72	NRAS	1	P13.3	P36.33	114	29 35	64%	4326
73	FLNA	X	Q28	Q28	0.1	21 18	39%	8533
74	RPS6KA3	X	P22.12	P22.33	20	29 33	62%	7987
75	PIK3	19	P13.11	P13.3	18	18 19	37%	3980
76	AKT3	1	Q44	Q44	5	29 31	60%	7281
77	CCND2	12	P13.32	P13.33	4	25 30	55%	6493
78	PIK3CA	3	Q26.32	Q29	19	32 34	66%	9259
79	TBC1D7	6	P23	P25.3	13	23 28	51%	1282
80	VANGL1**	1	P13.1	P36.33	115	25 31	56%	8671

Table S9b. Relative mutability of ten rat genes (#71-80) associated with CH

Rat gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
71	BRAF	4	68510	0	68.51	27,31	58	10138
72	NRAS	2	190593	249053	58.46	28,26	54	1326
73	FLNA	X	152034	152,453	0.419	24,23	47	8412
74	RPS6KA3	X	35623	0	35.623	30,32	62	7032
75	PIK3	1	82933	0	82.933	22,24	46	1984
76	AKT3	13	89225	106807	17.582	27,29	56	6854
77	CCND2	4	159989	182687	22.698	25,24	49	1414
78	PIK3CA	2	115249	0	115.249	30,30	60	6418
79	TBC1D7	17	21548	0	21.548	23,26	49	1366
80	VANGL1**	2	189637	249053	59.416	23,25	48	2597

Table S9c. Relative mutability of ten mouse genes (#71-80) associated with CH

Mouse gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
71	BRAF	6	39702	0	39.702	28,30	58	9728
72	NRAS	3	102975	159745	56.77	27,31	58	4470
73	FLNA	X	73293	0	73.293	24,23	47	8347
74	RPS6KA3	X	158151	169476	11.325	29,31	60	7320
75	PIK3	7	27340	0	27.34	22,25	47	4375
76	AKT3	1	177091	195154	18.063	27,31	58	6789
77	CCND2	6	127131	149588	22.457	24,29	53	5772
78	PIK3CA	3	32520	0	32.52	29,29	58	6987
79	TBC1D7	13	43324	0	43.324	24,25	49	1248
80	VANGL1**	3	102112	159745	57.633	24,27	51	4214

Table S10a. Relative mutability of ten human genes (#81-90) associated with CH

Human gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
81	VANGL2***	1	Q23.2	Q44	88	19 29	38%	5354
82	CCL2	17	Q12	Q13.3	34	29 31	60%	741
83	FUZ	19	Q13.33	Q13.43	9	18 21	39%	1728
84	MPDZ	9	P23	P24.3	13	30 27	57%	7734
85	ARHGAP31#	3	Q13.32	Q29	79	27 24	51%	9307
86	GPSM2	1	P13.3	P36.33	108	31 29	60%	7152
87	ARSB*	5	Q14.1	Q15.33	78	25 29	54%	4852
88	GBA##	1	Q21.3	Q44	93	22 24	46%	2291
89	FGFR2	10	Q26.13	Q26.3	12	26 26	52%	4624
90	FGFR3	4	P16.3	P16.3	1	18 21	39%	4301

Table S10b. Relative mutability of ten rat genes (#81-90) associated with CH

Rat gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
81	VANGL2***	13	84489	106807	22.318	19,24	43	2813
82	CCL2	10	67007	107211	40.204	30,30	60	780
83	FUZ	1	95384	0	95.384	20,22	42	1448
84	MPDZ	5	95920	166875	70.955	28,25	53	7516
85	ARHGAP31#	11	62151	86241	24.09	25,20	45	4566
86	GPSM2	2	196375	249053	52.678	28,25	53	2516
87	ARSB*	2	25162	0	25.162	22,22	44	1862
88	GBA##	2	174615	249053	74.438	23,25	48	2159
89	FGFR2	1	184850	260522	75.672	26,24	50	4654
90	FGFR3	14	77002	104886	27.884	21,24	45	4203

Table S10c. Relative mutability of ten mouse genes (#81-90) associated with CH

Mouse gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
81	VANGL2***	1	171856	195154	23.298	21,27	48	5714
82	CCL2	11	81928	121973	40.045	28,29	57	806
83	FUZ	7	44550	0	44.55	20,22	42	1684
84	MPDZ	4	81361	156860	75.499	28,25	53	7526
85	ARHGAP31#	16	38533	0	38.533	26,22	48	7722
86	GPSM2	3	108629	159745	51.116	29,25	54	3510
87	ARSB*	13	94079	120883	26.804	23,27	50	3989
88	GBA##	3	89116	159745	70.629	22,26	48	2142
89	FGFR2	7	129868	144995	15.127	25,26	51	5223
90	FGFR3	5	33894	0	33.894	21,24	45	4219

Table S11a. Relative mutability of ten human genes (#91-100) associated with CH

Human gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
91	SKI	1	P36.33	P36.33	2	18 21	39%	6083
92	TGFBR1	9	Q22.32	Q34.3	39	28 34	62%	6492
93	MAF	16	Q23.2	Q24.3	11	21 19	40%	2669
94	APOE	19	44 Mb	58	14	18,14	32	1166
95	APP	21	25Mb	46	29	23,19	42	1018
96	Tau/mapt	17q	45 Mb	83	18	22,23	45	5811
97	PSEN1	14q	73Mb	107	34	26,31	57	6018
98	PSEN2	11	226Mb	247	21	21,25	46	2249
99	BIN1 a/ stg3	2q	127Mb	242	115	22,19	41	2070
100	Cyp27c1 a/ stg3	2q	127mb	242	115	29,27	56	4685

Table S11b. Relative mutability of ten rat genes (#91-100) associated with CH

Rat gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
91	SKI	5	165782	166875	1.093	20,24	44	5584
92	TGFBR1	5	61710	0	61.71	26,33	59	5724
93	MAF	19	43713	57337	13.624	18,21	39	1945
94	APOE	1	79357	0	79.357	24,16	40	1172
95	APP	11	24236	0	24.236	26,21	47	2340
96	Tau/mapt	10	89236	107211	17.975	24,24	48	5149
97	PSEN1	6	103375	140994	37.619	24,27	51	1407
98	PSEN2	13	91993	106807	14.814	23,24	49	1628
99	BIN1 a/ stg3	18	24067	0	24.067	23,20	43	2399
100	Cyp27c1 a/ stg3	9	76294	114175	37.881	23,24	47	2150

Table S11c. Relative mutability of ten mouse genes (#91-100) associated with CH

Mouse gene	gene ID	chr*	gene locus	telomere locus	gene to tel. (Mb)	A, T ** (%)	A + T (%)	FL*** (bp)
91	SKI	4	155306	156860	1.554	22,24	46	5481
92	TGFBR1	4	47414	0	47.414	27,33	60	5779
93	MAF	8	116433	130127	13.694	26,28	54	3642
94	APOE	7	19M	143 M	124	23,18	41	1288
95	APP	16	84m	98m	14	27,23	50	3152
96	Tau/mapt	11	104	121	17	24,25	49	5164
97	PSEN1	12	83	120	37	25,25	50	2706
98	PSEN2	1	180	195	15	22,25	47	2026
99	BIN1 a/ stg3	18	32	90	58	23,20	43	2468
100	Cyp27c1 a/ stg3	1	74	195	121	24,23	47	1936

References for supplementary material

Morton, N.E. (1991). Parameters of the human genome. *Proc Natl Acad Sci U S A* 88**,** 7474-7476.