Table S1. Exome data quality

Pedigree/ Subject	Exons ≥10X Average		Exons ≥20X Average		Exons ≥50X Average		Mapped Reads	Reads in Exons	% of Mapped	# variant calls per subject	# variants used in analysis
**0328/0328 (Proband)*	226,303	98.49%	219,103	95.36%	178,648	77.75%	50,336,430	33,160,266	65.88	590,189	931,996
0328/0326 (Mother)	226,520	98.58%	220,653	96.03%	185,044	80.53%	48,457,440	31,400,649	64.80	616,927	
0328/0327 (Father)	226,414	98.54%	219,506	95.53%	180,332	78.48%	46,921,607	30,630,060	65.28	604,568	
***2009/2009 (Proband)*	188,990	84.81%	184,199	82.66%	153,447	68.86%	91,039,553	65,083,250	71.49	132,292	115,730
2009/2433 (Father)	189,279	84.94%	185,201	83.11%	164,388	73.77%	98,752,261	72,658,338	73.58	131,498	
2009/2434 (Mother)	189,145	84.88%	185,023	83.03%	160,801	72.16%	92,414,068	67,844,788	73.41	130,738	
2009/2625 (Sister)	187,117	83.97%	180,544	81.02%	117,414	52.69%	67,682,913	48,741,385	72.01	130,897	
2009/2626 (Brother)	187,653	84.12%	180,878	81.17%	122,405	54.93%	70,854,543	50,829,897	71.74	130,921	
***7621/7261 (Proband)*	193,558	86.86	182,794	82.03	96,288	43.21%	61,040,317	45,873,872	75.15	119,589	125,437
7621/8268 (Father)	194,872	87.45	187,830	84.29	125,948	56.52%	74,156,062	55,932,680	75.43	121,123	
7621/8267 (Mother)	193,290	86.74	182,549	81.92	99,096	44.47%	59,571,527	45,802,482	76.89	117,276	
7621/8198 (Sister)	194,337	87.21	187,875	84.31	125,970	56.53%	71,326,004	55,219,896	77.42	120,182	
***5762/5762 (Proband)*	190,081	85.30%	179,853	80.71%	113,023	50.72%	65,707,714	44,398,392	67.63	321,003	341,206
5762/6766 (Father)	190,638	85.55%	180,900	81.18%	119,619	53.68%	70,459,783	47,246,019	67.05	331,958	
5762/6765 (Mother)	190,549	85.51%	181,858	81.61%	127,931	57.41%	80,231,612	52,189,040	65.05	359,359	
5762/6767 (Brother)	189,970	85.25%	178,360	80.04%	110,349	49.52%	67,773,116	44,932,428	66.30	331,808	

* Exon coverage was calculated with the BEDtools coverage tool v2.25.0 (Quinlan and Hall, 2010) using the bed file associated with the exome capture kit. Only reads in which greater than 50% of the base pairs overlapping each feature (exon) were counted. Each feature that contained at least 10X-, 20X-, or 50X-coverage was counted and divided by the sum of the total features. Mapped reads were calculated with SAMtools view v1.3.1 (Li et al., 2009) by filtering out reads with the unmapped flag set.
** UCSC build hg19 was used to define genomic (intronic and exonic) and exonic regions for all known genes. The total exon range was calculated for each transcript and the transcript with the highest total exon range was considered the largest transcript. The total exon range represents the sum of the exon lengths for a transcript. In cases where a gene had multiple equal length largest transcripts, only one was used for statistical processing. Ultimately, 27,145 transcripts and 222,839 exons were utilized. These transcripts were defined using 0-based coordinates. Sambamba 0.6.6 (Tarasov et al., 2015) in conjunction with a shell script was used to procure the mean coverage and total mapped reads for each genomic and exonic region. A read was counted for mean coverage and for calculating the total mapped reads, if it passed a mapping quality threshold of ≥ 0. Duplicate reads and reads that failed quality control were not included. Regions were differentiated into 5 categories: no coverage, mean coverage $0<X<10$, mean coverage $10\leq X<20$, mean coverage $20\leq X<50$, and mean coverage ≥ 50. The percentage for each category was calculated by dividing the number of exons that fell into a specific category over the total number of exons present in the generated largest transcript reference BED file. Total mapped reads for each patient were generated by summating the read counts in the genomic regions. Percentage of mapped reads was calculated by dividing the reads in exons by the total mapped reads.

Table S2. Variants considered

Proband	Gene	Genomic Coordinates	Variant	CADD	ExAC Frequency	ExAC pLI	Inheritance Model	Variant Description
328	*UBTF*	chr17:g.42290219C>T	NM_014233.2(UBTF): c.628G>A(p.Glu210Lys)	32	NA	1.00	de novo	missense
328	*COL6A5*	chr3:g.130107975A>G chr3:g.130095164C>T chr3:g.130095493G>A	NM_153264.5(COL6A5):c.2414A>G(p.His805Arg) NM_153264.5(COL6A5):c.152C>T(p.Thr51Met) NM_153264.5(COL6A5):c.481G>A(p.Gly161Arg)	8.641 8.254 23	1.80E-2(137/7638) 1.67E-5(2/119816) 2.08E-4(22/105643)	0.00	compound heterozygous	missense missense missense
328	*MUC17*	chr7:g.100680261C>T chr7:g.100683221A>G	NM_001040105.1(MUC17): c.5564C>T(p.Thr1855Ile) NM_001040105.1(MUC17): c.8524A>G (p.Thr2842Ala)	12.39 0.061	NA 2.19E-3(266/121348)	NA	compound heterozygous	missense missense
328	*MUC5B*	chr11:g.1253994G>A chr11:g.1251805C>T	NM_002458.1(MUC5B):c.2068G>A(p.Val690Ile) NM_002458.1(MUC5B):c.1454C>T (p.Thr485Met)	16.59 15.04	3.18E-5(3/94252) 8.85E-5(9/101730)	1.00	compound heterozygous	missense missense
328	*TMEM131*	chr2:g.98375414G>A chr2:g.98377264T>C	NM_015348.1(TMEM131): c.5309C>T(p.Ala1770Val) NM_015348.1(TMEM131): c.5003A>G(p.Lys1668Arg)	15.59 21.5	5.87E-5(7/119260) NA	1.00	compound heterozygous	missense missense
328	*UBASH3A*	chr21:g.43824076C>A chr21:g.43854987G>C	NM_018961.2(UBASH3A):c.22C>A(p.Leu8Ile) NM_018961.2(UBASH3A): c.1316G>C(p.Arg439Thr)	23.1 1.38	1.79E-4(3/16768) 2.47E-5(3/121356)	0.00	compound heterozygous	missense missense
328	*SSX5*	chrX:g.48054229T>C	NM_021015.3(SSX5):c.254A>G(p.Glu85Gly)	NA	2.29E-5(2/87186)	0.00	X-linked	missense
2009	*UBTF*	chr17:g.42290219C>T	NM_014233.2(UBTF): c.628G>A(p.Glu210Lys)	32	NA	1.00	de novo	missense
2009	*KLC4*	chr6:43034797C>G chr6:43039337G>A	NM_201523.2(KLC4): c.909C>T (p.Ser303Ser) NM_201523.2(KLC4): c.1342G>A (p.Glu448Lys)	16.96 25.8	6.02E-4(73/121170) 5.97E-4(72/120606)	0.24	compound heterozygous	missense
2009	*FBXO21*	chr12:117595754C>T	NM_033624.2(FBXO21): c.1462G>A(p.Val488Ile)	22.4	1.24E-4(15/121336)	0.99	de novo	missense
7621	*UBTF*	chr17:g.42290219C>T	NM_014233.2(UBTF): c.628G>A(p.Glu210Lys)	32	NA	1.00	de novo	missense
7621	*MUC4*	chr3:g. 195510146G>C	NM_004532.5(MUC4):c.83-4820C>G	10.53	NA	NA	de novo	intronic
7621	*MUC4*	chr3:g.195510181T>C	NM_004532.5(MUC4):c.83-4855A>G	4.216	8.15E-4(17/20862)	NA	de novo	intronic
7621	*MUC4*	chr3:g.195510182G>A	NM_004532.5(MUC4):c.83-4856G>A	NA	1.86E-3(38/20456)	NA	de novo	intronic
7621	*DBF4*	chr7:g.87530095G>A	NM_006716.3(DBF4):c.826G>A(p.Asp276Asn)	18.02	1.65E-5(2/121212)	1.00	de novo	missense
7621	*KRT85*	chr12:g.52757964G>A	NM_002283.3(KRT85):c.691-17C>T	3.067	NA	0.00	de novo	intronic
7621	*DNAH9*	chr17:g.11696996A>G	NM_001372.3(DNAH9):c.8223+15A>G	1.89	7.09E-4(85/119934)	0.00	de novo	intronic
7621	*HMCN1*	chr1:g.185878454C>T chr1:g.186105795C>T	NM_031935.2(HMCN1):c.622-15C>T NM_031935.2(HMCN1):c.13313-5C>T	5.71 7.612	NA 2.73E-3(331/121066)	0.00	compound heterozygous	intronic intronic
7621	*CACNA2D3*	chr3:g.54925426C>T chr3:g.54925493C>A	NM_018398.2(CACNA2D3): c.2195C>T (p.Thr732Met) NM_018398.2(CACNA2D3):c.2246+16C>A	34 10.28	7.62E-3(919/120586) 8.93E-3(1071/119836)	1.00	compound heterozygous	Missense intronic
7621	*POMT1*	chr9:g.134394788G>A chr9:g.134397464C>T	NM_007171.3(POMT1):c.1565G>A(p.Arg522Lys) NM_007171.3(POMT1):c.1922C>T (p.Ala641Val)	1.671 11.42	1.25E-3(26/20754) 7.58E-3(919/121230)	0.00	compound heterozygous	Missense missense
7621	*TAOK3*	chr12:g.118610323C>T chr12:g.118704579G>A	NM_001346487.1(TAOK3): c.1865G>A(p.Arg622His) NM_001346487.1(TAOK3):c.-193-15C>T	35 0.867	8.24E-5(10/121412) NA	1.00	compound heterozygous	Missense intronic
7621	*DNAH3*	chr16:g.21093012C>T chr16:g.21139071G>A	NM_017539.2(DNAH3):c.2914G>A(p.Glu972Lys) NM_017539.2(DNAH3):c.1145C>T(p.Pro382Leu)	23.5 20.3	1.65E-3(200/121388) 9.72E-3(1180/121406)	0.00	compound heterozygous	missense missense
7621	*MAST3*	chr19:g.18246622G>A chr19:g.18255981G>A	NM_015016.1(MAST3):c.1856G>A(p.Arg619Gln) NM_015016.1(MAST3):c.2894G>A(p.Arg965Gln)	22.2 26.2	NA NA	1.00	compound heterozygous	missense missense
7621	*DHX35*	chr20:g.37634873C>T chr20:g.37659385C>T	NM_021931.3(DHX35):c.1096C>T(p.Arg366*) NM_021931.3(DHX35):c.1922C>T(p.Ala641Val)	42 25.1	5.77E-5(7/121412) 5.99E-3(727/121402)	0.00	compound heterozygous	nonsense missense
7621	*C4orf47*	chr4:g.186353135G>T chr4:g.186370727G>A	NM_001114357.1(C4orf47):c.101-1G>T NM_001114357.1(C4orf47):c.882-23G>A	24.7 10.06	1.45E-3(31/21364) 4.80E-2(932/19410)	NA	compound heterozygous	Splice-site intronic
7621	*C9orf43*	chr9:g.116181419A>G chr9:g.116187302C>G	NM_001278629.1(C9orf43):c.319A>G(p.Ile107Val) NM_001278629.1(C9orf43): c.811C>G(p.Pro271Ala)	13.33 22.6	1.82E-2(2205/121024) 5.39E-2(6485/120402)	0.00	compound heterozygous	missense missense
5762	*UBTF*	chr17:g.42290219C>T	NM_014233.2(UBTF): c.628G>A(p.Glu210Lys)	32	NA	1.00	de novo	missense
5762	*PIBF1*	chr13:g.73401237G>A chr13:g73547733TTAAA>T	NM_001349655.1(PIBF1):c.896G>A(p.Arg299Gln) NM_001349655.1(PIBF1): c.1974_1977delTAAA(p.Asn658Lysfs*15)	25.7 35	1.07E-2(1264/118622) 1.09E-4(13/119262)	0.00	compound heterozygous	missense outframe deletion
5762	*HCN4*	chr15:g.73615084G>A chr15:g.73617804G>T	NM_005477.2(HCN4):c.3350C>T (p.Pro1117Leu) NM_005477.2(HCN4):c.1591-19C>A	22.9 1.243	3.16E-3(63/19948) 6.77E-3(820/121172)	0.23	compound heterozygous	missense intronic
5762	*NLRP6*	chr11:g.284242C>G	NM_138329.2(NLRP6):c.2214C>G (p.Cys738Trp)	22.9	NA	NA	de novo	missense
5762	*CTSF*	chr11:g.66333485G>C	NM_003793.3(CTSF):c.867+8C>G	4.596	8.24E-6(1/121338)	0.00	de novo	intronic

Table S3. Autistic behaviors of TGen_0328 at 4.6 years of age

Difficulty following directions at home and school
Difficulty paying attention at school
Easily distractible
Difficulty completing tasks
Restless and fidgety
Difficulty sitting still
Difficulty following rules
Frequent crying
Lacked understanding of social cues
Temper tantrums
Lack of interest in other children
Did not enjoy interacting with peers
Did not share toys and play space
Did not take turns at play
Did not pretend play
Did not tolerate changes in routine
Bothered by particular sounds

Table S4. Gene expression in fibroblasts

Tissue	***UBTF***		**pre-rRNA**		**18S**
	Total *UBTF*	*UBTF1*	Random primer cDNA	Gene specific primer cDNA	Random primer cDNA
Controls (N = 4)	1.00 ± 0.04	1.00 ± 0.02	1.00 ± 0.05	1.00 ± 0.05	1.00 ± 0.03
Patients (N = 4)	1.08 ± 0.03	1.09 ± 0.04	3.18 ± 0.67*	3.95 ± 0.60*	2.27 ± 0.11*

QRT-PCR values are referenced to normal controls and presented as means ± SEM. β-actin was used as the endogenous control. *$P < 0.05$.

Table S5. Expression of genes regulated by UBTF2 in human fibroblasts and mouse brain

Gene	*Fibroblasts*		*Ubtf*$^{+/+}$ Mice		*Ubtf*$^{+/-}$ Mice	
	Controls	Patients	Cerebral cortex	Cerebellum	Cerebral cortex	Cerebellum
Hyal1/HYAl1	1.00 ± 0.03	1.30 ± 0.09	1.00 ± 0.04	2.07 ± 0.09	1.05 ± 0.03	1.98 ± 0.08
Ppargc1a/ PPARGC1A	1.00 ± 0.04	2.27 ± 0.35*	1.00 ± 0.03	1.54 ± 0.06	1.04 ± 0.04	1.47 ± 0.06
Fancb/FANCB	1.00 ± 0.02	0.97 ± 0.07	1.00 ± 0.03	1.21 ± 0.04	0.92 ± 0.04	1.17 ± 0.05
Hist1h4b/ HIST1H4B	1.00 ± 0.03	0.60 ± 0.05*	1.00 ± 0.03	1.26 ± 0.08	0.79 ± 0.02	0.97 ± 0.05
Slc4a4/SLC4A4	1.00 ± 0.04	0.69 ± 0.06*	1.00 ± 0.05	3.64 ± 0.33	0.95 ± 0.03	3.37 ± 0.39

β-actin was used as the endogenous control. QRT-PCR values are referenced to wild-type (*Ubtf*$^{+/+}$) cerebral cortex or human fibroblast controls and presented as means ± SEM (N = 4 samples/genotype). **P* < 0.05.

Table S6. Fly lethality tables

	c155*-GAL4/c155-GAL4 X UAS-*UBTF1/CyO		
Genotype	**Observed**	**Expected**	***P* value**
c155-GAL4;*CyO*	26	23	
c155-GAL4; UAS-*UBTF1*	20	23	0.68
	c155*-GAL4/c155-GAL4 X UAS-*UBTF1*-E210K/*CyO		
c155-GAL4;*CyO*	118	59	
c155-GAL4; UAS-*UBTF1*-E210K	0	59	< 0.0001
	eye3.5*-GAL4/*eye3.5*-GAL4 X UAS-*UBTF1/CyO		
eye3.5-GAL4;*CyO*	38	24	
eye3.5-GAL4; UAS-*UBTF*	10	24	0.0052
	eye3.5*-GAL4/*eye3.5*-GAL4 X UAS-*UBTF1* E210K/*CyO		
eye3.5-GAL4;*CyO*	56	28	
eye3.5-GAL4; UAS-*UBTF1*-E210K	0	28	< 0.0001

Table S7. *Ubtf* expression in mice

Tissue	*Ubtf*$^{+/+}$		*Ubtf*$^{+/-}$	
	Total *Ubtf*	*Ubtf1*	Total *Ubtf*	*Ubtf1*
Cerebral cortex	1.00 ± 0.02	1.00 ± 0.02	0.54 ± 0.03*	0.93 ± 0.04
Cerebellum	1.63 ± 0.12	1.83 ± 0.09	0.85 ± 0.06*	1.60 ± 0.15
Liver	0.64 ± 0.02	0.61 ± 0.02	0.37 ± 0.03*	0.55 ± 0.02

β-actin was used as the endogenous control. QRT-PCR values are referenced to wild-type (*Ubtf*$^{+/+}$) cerebral cortex and presented as means ± SEM (N = 4 mice/genotype). *$P < 0.05$.

Table S8. Effects of genotype and sex on weight and behavioral measures in 3-month old mice.

	Male		Female	
	Ubtf$^{+/+}$ (n=23)	*Ubtf*$^{+/-}$ (n=20)	*Ubtf*$^{+/+}$ (n=21)	*Ubtf*$^{+/-}$ (n=22)
Weight (g)	26.2 ± 0.5	25.3 ± 0.3	19.2 ± 0.3	20.3 ± 0.4
Grip strength (g)	3.50 ± 0.06	3.56 ± 0.06	2.98 ± 0.07	3.12 ± 0.13
Grip strength/weight	0.133 ± 0.002	0.140 ± 0.003	0.151 ± 0.004	0.146 ± 0.004
Dominance tube (win & tie %)	27% & 33%	40% & 33%	12% & 48%	40%* & 48%
Cross maze score (%)	30.6 ± 3.3	30.2 ± 2.5	30.5 ± 2.7	27.6 ± 1.6
Rope climbing (s)	2.79 ± 0.15	3.61 ± 0.20*	3.02 ± 0.17	2.83 ± 0.12
Rope climbing/weight	0.107 ± 0.005	0.140 ± 0.008*	0.153 ± 0.009	0.135 ± 0.007
Open field activity				
Distance traveled (cm)	1179.2 ± 109.4	971.5 ± 55.4	1062.6 ± 77.1	1354.5 ± 74.7*
Ambulatory count	590.2 ± 67.0	461.5 ± 27.7	525.3 ± 41.3	692.5 ± 41.8*
Stereotypic Count	1580.0 ± 69.6	1417.9 ± 65.9	1391.3 ± 57.0	1574.1 ± 50.7*
Vertical count	69.0 ± 10.4	41.6 ± 4.0*	76.3 ± 33.9	61.3 ± 14.6
Jump count	17.9 ± 3.0	10.9 ± 1.8	18.6 ± 2.1	14.2 ± 2.0
Average velocity (cm/s)	35.7 ± 1.7	33.8 ± 1.8	32.2 ± 2.1	33.4 ± 1.6
Ambulatory episodes	48.4 ± 4.2	41.9 ± 2.6	46.2 ± 3.6	58.2 ± 3.3*
DigiGait^{TM}				
Propel (s) Forelimb	0.121 ± 0.002	0.128 ± 0.004	0.116 ± 0.003	0.120± 0.002
Propel (s) Hindlimb	0.177 ± 0.004	0.189 ± 0.004*	0.174 ± 0.003	0.176 ± 0.003
Stride length (cm) Forelimb	6.69 ± 0.13	6.99 ± 0.16	6.50 ± 0.10	6.72 ± 0.10
Stride length (cm) Hindlimb	6.67 ± 0.12	7.05 ± 0.15*	6.50 ± 0.11	6.75 ± 0.10*
Stride Frequency (steps/s) Forelimb	3.05 ± 0.07	2.89 ± 0.08	3.12 ± 0.05	3.03 ± 0.05
Stride Frequency (steps/s) Hindlimb	3.06 ± 0.06	2.88 ± 0.07*	3.12 ± 0.05	3.03 ± 0.05
Stance width (cm) Forelimb	1.66 ± 0.03	1.52 ± 0.02*	1.59 ± 0.03	1.57 ± 0.03
Stance width (cm) Hindlimb	2.84 ± 0.05	2.67 ± 0.04*	2.57 ± 0.04	2.65 ± 0.05
Step angle (deg) Forelimb	69.17 ± 0.75	70.25 ± 0.77	67.29 ± 0.97	68.62 ± 0.75
Step angle (deg) Hindlimb	54.87 ± 1.68	56.25 ± 0.86	56.03 ± 1.97	56.09 ± 1.51
Paw Area (cm^2) Forelimb	0.30 ± 0.01	0.31 ± 0.01	0.31 ± 0.01	0.31 ± 0.01
Paw Area (cm^2) Hindlimb	0.60 ± 0.01	0.55 ± 0.01*	0.58 ± 0.01	0.52 ± 0.01*

Ambulatory count, the total number of X + Y photo beam breaks while in ambulatory movement status. Stereotypic count, any partial-body movements that occur within the ambulatory box such as grooming, head-weaving or scratching. Vertical count, number of periods of continuous Z photo beam breaks. Jump count, the number of times that the mouse leaves the photo beam array for a period of time. Ambulatory episodes, the number of times the mouse has started moving after the resting delay has expired. Values are means ± SEM except for dominance tube. *$P < 0.05$, for effect of genotype within sex.

Table S9. Primers for genotyping, QRT-PCR, and Sanger sequencing

Primer	Sequence (5'→3')	Locus	Usage	Product (bp)
Ubtf_KO_A	tgatccctccctttctgatg	NC_000077.6 102315400 - 381	*Ubtf* ko mice genotyping	429 (WT mice with Ubtf_KO_A+B+C)
Ubtf_KO_B	tggggataggccttagagaga	NC_000077.6 102313816 - 796	*Ubtf* ko mice genotyping	628 (Floxed mice with Ubtf_KO_A+B+C)
Ubtf_KO_C	cacgggaaaacaaggtcact	NC_000077.6 102313388 - 407	*Ubtf* ko mice genotyping	529 (KO mice with Ubtf_KO_A+B+C)
Ubtf_q85F	gaacggagaagcggactg	NM_011551.6 319 - 336	QRT-PCR	
Ubtf_q85R	tccagcagagtcagcatatctt	NM_011551.6 411 - 390	QRT-PCR	93 for both *Ubtf1* and *Ubtf2* (with Ubtf_q85F)
Ubtf_q9F	agcggaaggagtacgagga	NM_011551.6 1071 - 1089	QRT-PCR	
Ubtf_q9R	tcgaacttgtccttgagctg	NM_011551.6 1197 - 1178	QRT-PCR	127 for *Ubtf1* (with Ubtf_q9F)
Slc4a4_p26F	actgtctccagtgcaagtagga	NM_018760.2 754 -775	QRT-PCR	
Slc4a4_p26R	tgtcagattcctgtgggtca	NM_018760.2 828 - 809	QRT-PCR	75 (with Slc4a4_p26R)
Fancb_p21F	tgccaagtgatgaaatggtta	NM_175027.4 1595 - 1615	QRT-PCR	
Fancb_p21R	ggcaccatgtctgctctaca	NM_175027.4 1738 - 1719	QRT-PCR	144 (with Fancb_p21F)
Hist1h4b_p81F	cgtcacctacacggagcac	NM_178193.2 294 - 312	QRT-PCR	
Hist1h4m_p81R	tagtcccttaacccccgaat	NM_178193.2 403 - 384	QRT-PCR	110 (with Hist1h4b_p81F)
Ppargc1a_p29F	tgaaagggccaaacagagag	NM_008904.2 2113 - 2132	QRT-PCR	
Ppargc1a_p29R	gtaaatcacacggcgctctt	NM_008904.2 2176 - 2157	QRT-PCR	64 (with Ppargc1a_p29F)
Hyal1_p29F	tttccagagacccccatgt	NM_008317.4 975 - 993	QRT-PCR	
Hyal1_p29R	aggctgtgctccagttcct	NM_008317.4 1071 - 1053	QRT-PCR	97 (with Hyal1_p29F)
SLC4A4_P68F	tcatggaccaacaaattacagc	NM_001098484.2 2504 -2525	QRT-PCR	
SLC4A4_P68R	cacccaaaagagatccaagtg	NM_001098484.2 2595 -2575	QRT-PCR	92 (with SLC4A4_P68F)
FANCB_P29F	tggagaaggaactagtcaccctta	NM_001018113.2 2563 - 2586	QRT-PCR	
FANCB_P29R	tgctcacttcacacctctgc	NM_001018113.2 2649 - 2630	QRT-PCR	87 (with FANCB_P29F)
HIST1H4B_P9F	aagtgctgcgggataacatc	NM_003544.2 62 - 81	QRT-PCR	
HIST1H4B_P9R	ttaaccccaccacgccta	NM_003544.2 134 - 117	QRT-PCR	73 (with HIST1H4B_P9F)
PPARGC1A_P6F	cgcagtcacaacacttacaagc	NM_001330751.1 1371 - 1392	QRT-PCR	
PPARGC1A_P6R	ggggtcatttggtgactctg	NM_001330751.1 1444 - 1425	QRT-PCR	74 (with PPARGC1A_P6F)
HYAL1_P5F	ggaagtcacagatgtatgtgcaa	NM_033159.3 927 - 949	QRT-PCR	
HYAL1_P5R	gattggggtcaccagcag	NM_033159.3 1001 - 984	QRT-PCR	75 (with HYAL1_P5F)
UBTF_Q48F	caaaaccaccgaatcacaca	NM_014233.3 434 - 453	QRT-PCR	
UBTF_Q48R	tgtcaatgtacggaacttcctc	NM_014233.3 554 - 533	QRT-PCR	121 for both *UBTF1* and *UBTF2* (with UBTF_Q48F)
UBTF_Q38F	ctcaaagtgcggccagat	NM_014233.3 939 - 956	QRT-PCR	

UBTF_Q38R	tctttttgtccgagagctgag	NM_014233.3	1020 - 1000	QRT-PCR	82 for *UBTF1* (with UBTF_Q38R)
H45S_F	gccttctctagcgatctgagag	NR_145819.1	1417 - 1438	QRT-PCR	
H45S_R	ccataacggaggcagagaca	NR_145819.1	1498 - 1479	QRT-PCR	82 (with H45S_F)
ACTB_F	atgggtcagaaggattcctatgt	NM_001101.3	223 - 245	QRT-PCR	
ACTB_R	ggtcatcttctcgcggtt	NM_001101.3	444 - 427	QRT-PCR	222 (with ACTB_F)
H18S_F	ttcgaacgtctgccctatcaa	NR_003286.2	344 - 364	QRT-PCR	
H18S_R	atggtaggcacggcgacta	NR_003286.2	393 - 375	QRT-PCR	50 (with H18S_F)

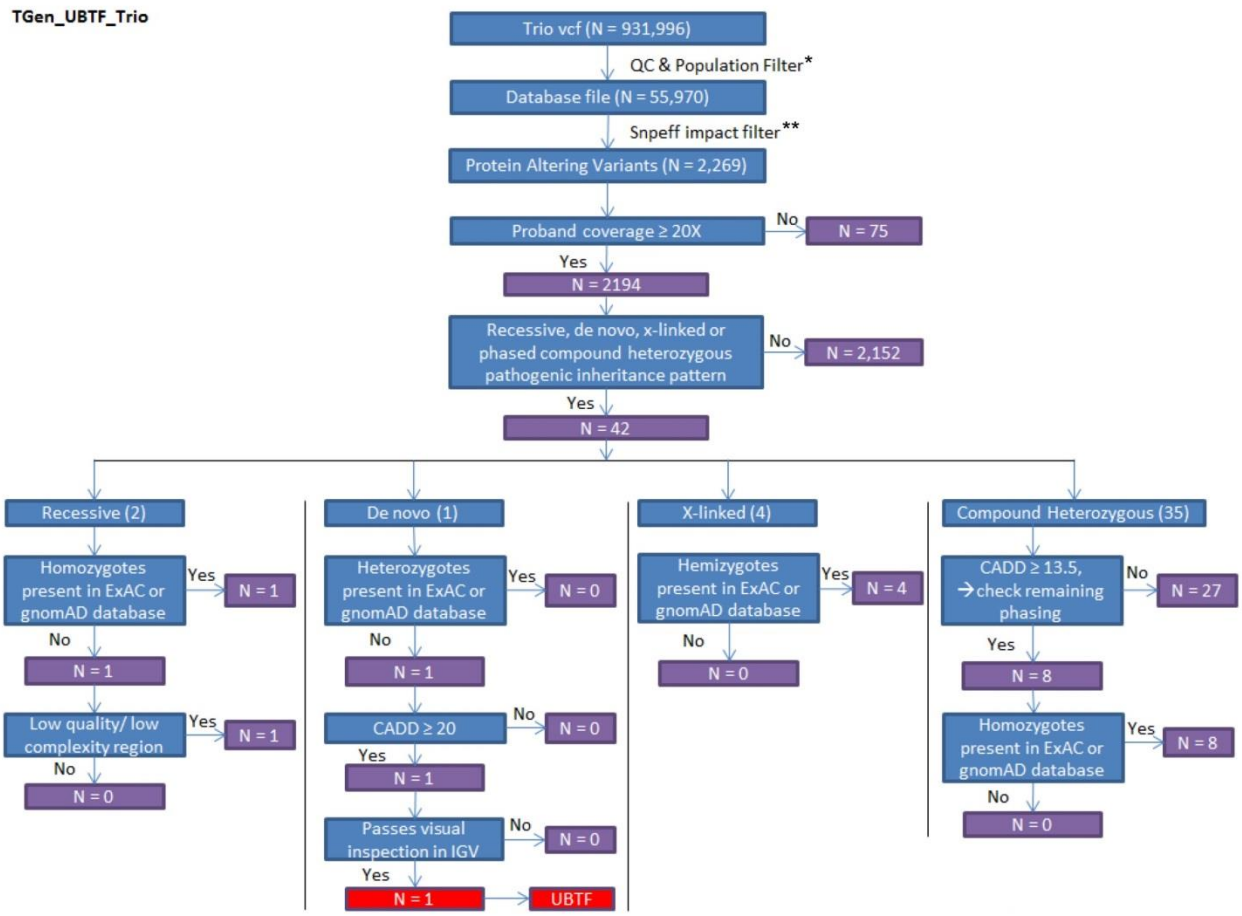


Figure S1. Translational genomics (TGen) exome analysis workflow. *GATK QualScore ≥ 300 *and* 1000 Genomes frequency ≤ 0.05. **Snpeff Variant Impact = "HIGH" + "MODERATE" and gene ≠ to "none" or "unknown."

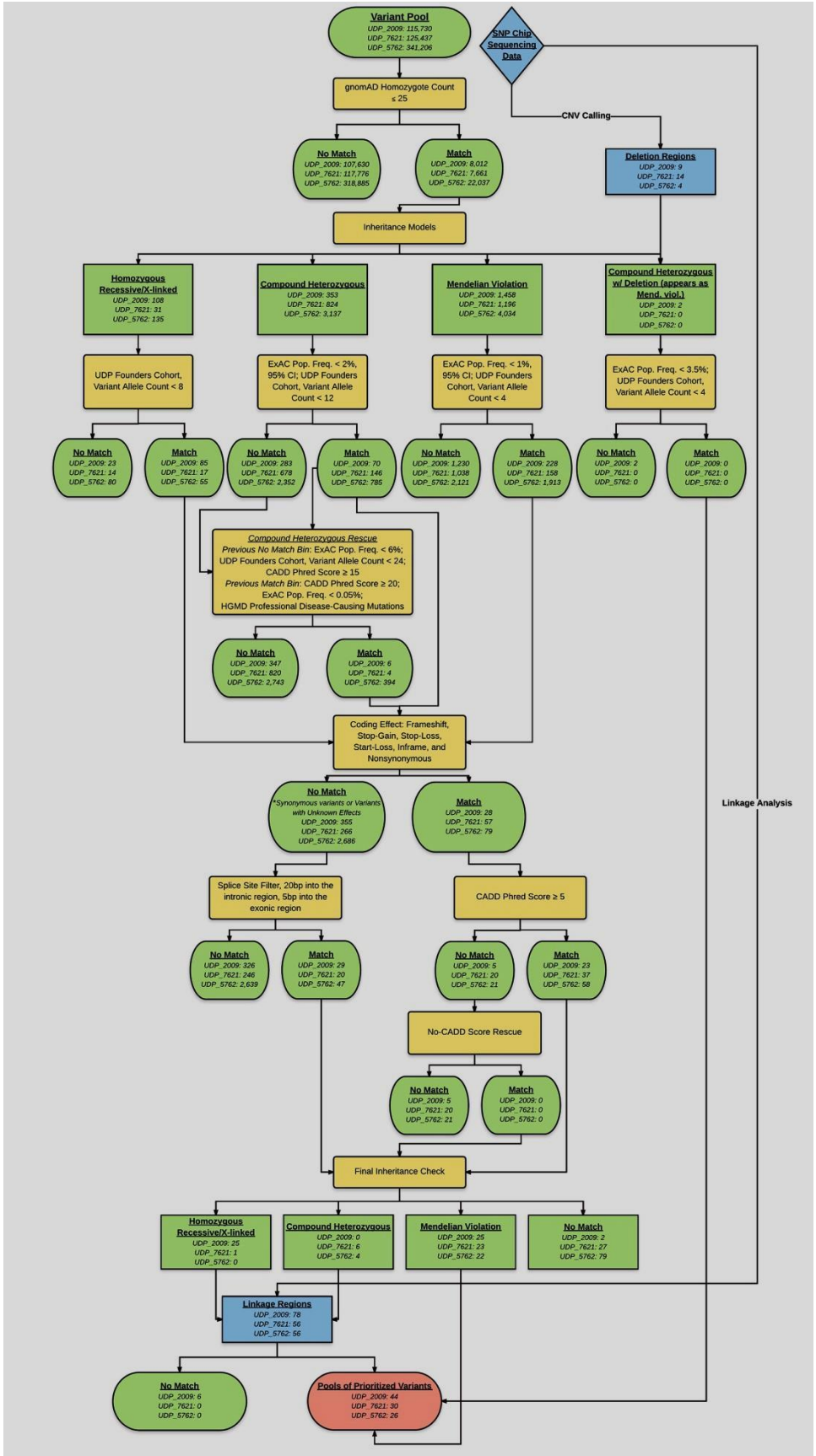
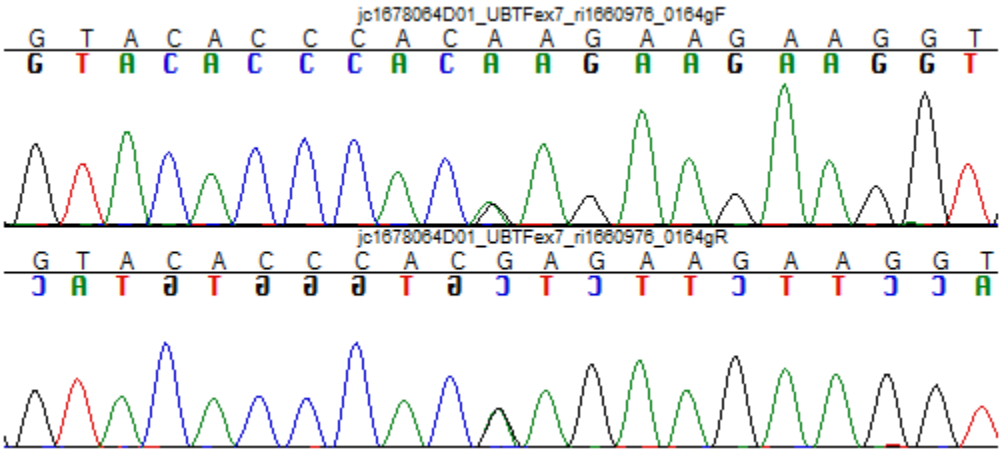
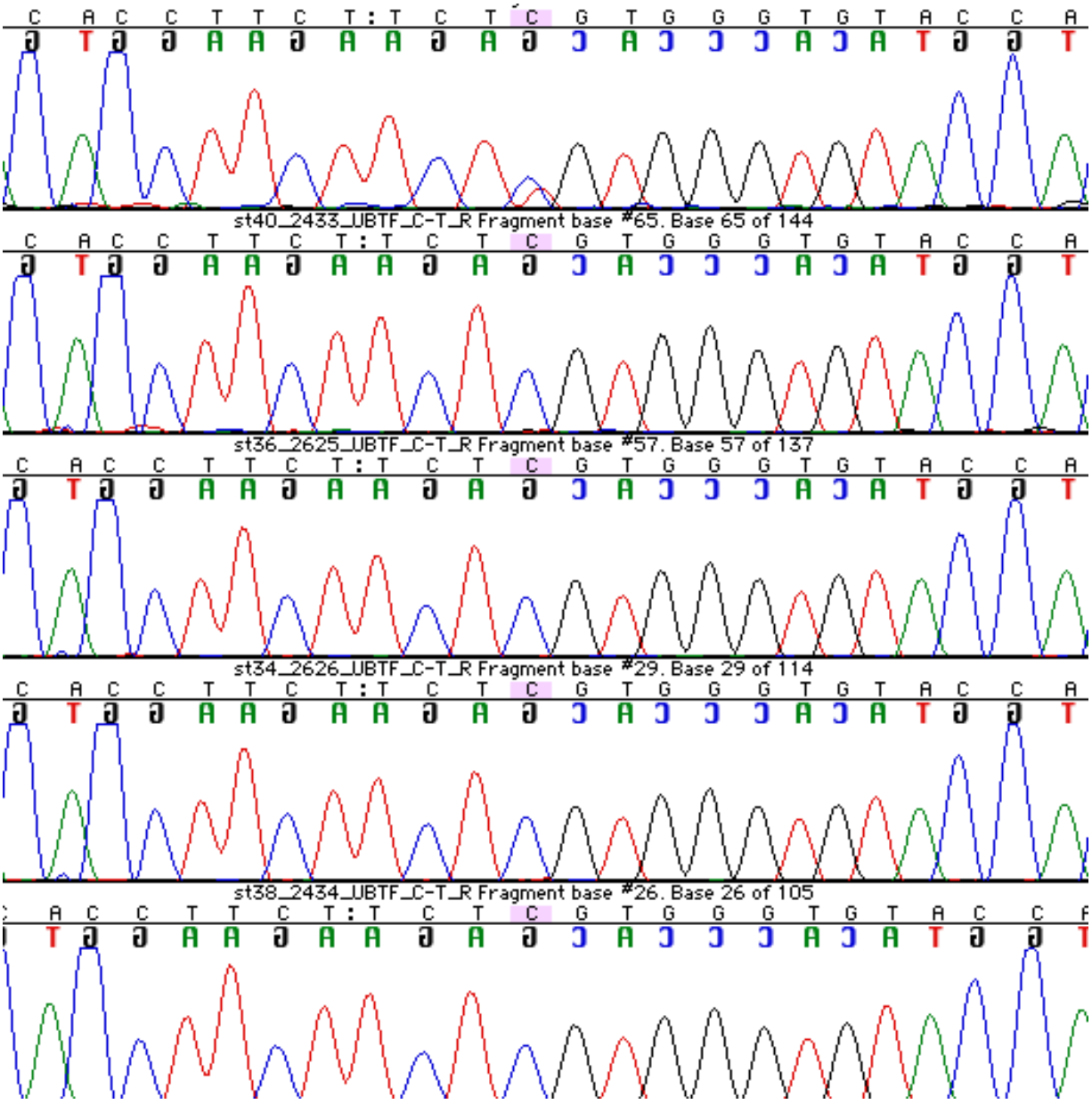

Variant Pool
UDP_2009: 115,730
UDP_7621: 125,437
UDP_5762: 341,206
SNP Chip Sequencing Data
gnomAD Homozygote Count ≤ 25
CNV Calling
No Match
UDP_2009: 107,630
UDP_7621: 117,776
UDP_5762: 318,885
Match
UDP_2009: 8,012
UDP_7621: 7,661
UDP_5762: 22,037
Deletion Regions
UDP_2009: 9
UDP_7621: 14
UDP_5762: 4
Inheritance Models
Homozygous Recessive/X-linked
UDP_2009: 108
UDP_7621: 31
UDP_5762: 135
Compound Heterozygous
UDP_2009: 353
UDP_7621: 824
UDP_5762: 3,137
Mendelian Violation
UDP_2009: 1,458
UDP_7621: 1,196
UDP_5762: 4,034
Compound Heterozygous w/ Deletion (appears as Mend. viol.)
UDP_2009: 2
UDP_7621: 0
UDP_5762: 0
UDP Founders Cohort, Variant Allele Count < 8
ExAC Pop. Freq. < 2%, 95% CI; UDP Founders Cohort, Variant Allele Count < 12
ExAC Pop. Freq. < 1%, 95% CI; UDP Founders Cohort, Variant Allele Count < 4
ExAC Pop. Freq. < 3.5%; UDP Founders Cohort, Variant Allele Count < 4
No Match
UDP_2009: 23
UDP_7621: 14
UDP_5762: 80
Match
UDP_2009: 85
UDP_7621: 17
UDP_5762: 55
No Match
UDP_2009: 283
UDP_7621: 678
UDP_5762: 2,352
Match
UDP_2009: 70
UDP_7621: 146
UDP_5762: 785
No Match
UDP_2009: 1,230
UDP_7621: 1,038
UDP_5762: 2,121
Match
UDP_2009: 228
UDP_7621: 158
UDP_5762: 1,913
No Match
UDP_2009: 2
UDP_7621: 0
UDP_5762: 0
Match
UDP_2009: 0
UDP_7621: 0
UDP_5762: 0
Compound Heterozygous Rescue
Previous No Match Bin: ExAC Pop. Freq. < 6%; UDP Founders Cohort, Variant Allele Count < 24; CADD Phred Score ≥ 15
Previous Match Bin: CADD Phred Score ≥ 20; ExAC Pop. Freq. < 0.05%; HGMD Professional Disease-Causing Mutations
No Match
UDP_2009: 347
UDP_7621: 820
UDP_5762: 2,743
Match
UDP_2009: 6
UDP_7621: 4
UDP_5762: 394
Coding Effect: Frameshift, Stop-Gain, Stop-Loss, Start-Loss, Inframe, and Nonsynonymous
No Match
*Synonymous variants or Variants with Unknown Effects
UDP_2009: 355
UDP_7621: 266
UDP_5762: 2,686
Match
UDP_2009: 28
UDP_7621: 57
UDP_5762: 79
Linkage Analysis
Splice Site Filter, 20bp into the intronic region, 5bp into the exonic region
CADD Phred Score ≥ 5
No Match
UDP_2009: 326
UDP_7621: 246
UDP_5762: 2,639
Match
UDP_2009: 29
UDP_7621: 20
UDP_5762: 47
No Match
UDP_2009: 5
UDP_7621: 20
UDP_5762: 21
Match
UDP_2009: 23
UDP_7621: 37
UDP_5762: 58
No-CADD Score Rescue
No Match
UDP_2009: 5
UDP_7621: 20
UDP_5762: 21
Match
UDP_2009: 0
UDP_7621: 0
UDP_5762: 0
Final Inheritance Check
Homozygous Recessive/X-linked
UDP_2009: 25
UDP_7621: 1
UDP_5762: 0
Compound Heterozygous
UDP_2009: 0
UDP_7621: 6
UDP_5762: 4
Mendelian Violation
UDP_2009: 25
UDP_7621: 23
UDP_5762: 22
No Match
UDP_2009: 2
UDP_7621: 27
UDP_5762: 79
Linkage Regions
UDP_2009: 78
UDP_7621: 56
UDP_5762: 56
No Match
UDP_2009: 6
UDP_7621: 0
UDP_5762: 0
Pools of Prioritized Variants
UDP_2009: 44
UDP_7621: 30
UDP_5762: 26

Figure S2. National Institutes of Health (NIH)/National Human Genome Research Institute (NHGRI)/Undiagnosed Disease Program (UDP) exome analysis workflow. The variant pools were processed with population frequency, inheritance models, and predicted deleteriousness filters. Variants that passed the filters were placed in the "Match" bins. SNP Chip sequencing data were utilized to call long regions (>150bp) with copy number variant of 0 or 1 and procure regions that were possibly disease-segregating through linkage analysis. In a gene, compound heterozygous candidates were considered if there was at least one maternally inherited variant and one paternally inherited variant. Weaker compound heterozygous candidates (ExAC Population Frequency < 6%, UDP Founders Cohort, Variant Allele Count < 24) were rescued with extremely strong compound heterozygous candidates (CADD Phred Score ≥ 20, ExAC Population Frequency < 0.05%, HGMD Professional Disease-Causing Mutations). A CADD Phred Score ≥ 15 filter was applied to weaker compound heterozygous candidates to curtail the number of rescued variants. Variants that passed the gnomAD homozygote count ≤ 25 and the ExAC population frequency filters were inspected for coding effects (frameshift, stop-gain, stop-loss, inframe, and nonsynonymous variants) and for possible splice site effects (20 bp into the intronic region and 5 bp into the exonic region). Variants with coding effects were then filtered with a CADD Phred Score ≥ 5 threshold, a threshold used for at least benign variants. Variants that did not have a CADD Phred Score annotation were rescued. A final inheritance check was conducted to categorize variants in each inheritance model. Variants that did not pass the final inheritance check were placed in the "No Match" bin as either failed compound heterozygous candidates or inheritance-unresolved candidates. Compound heterozygous and homozygous recessive/X-linked variants found in the linkage regions, Mendelian violation variants, and compound heterozygous variants with deletion variants were gathered as pools of prioritized variants for visual inspection and literature verification. Variants that passed visual inspection and literature verification were considered for further validation (Table S2).

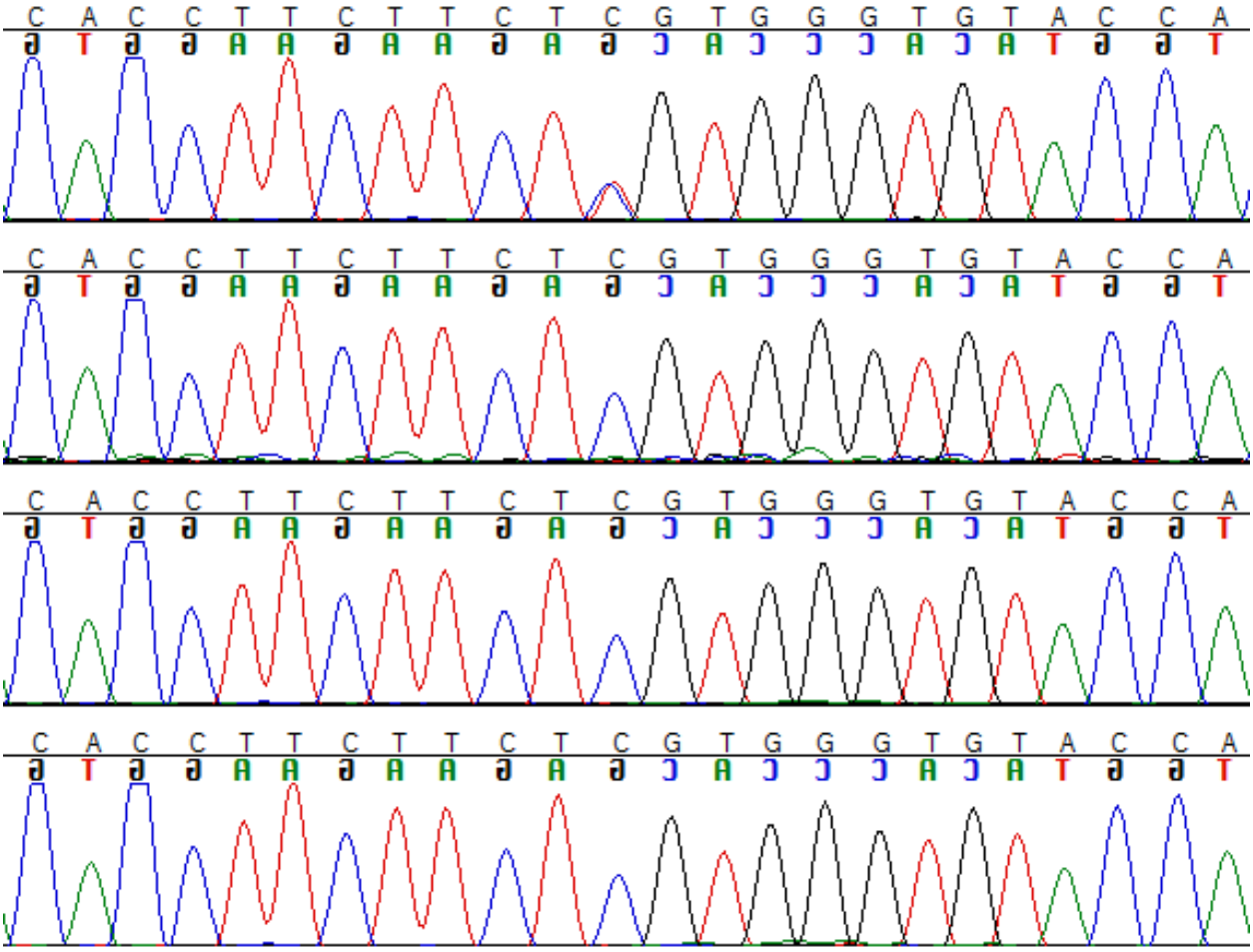
TGen_0328


jc1678064D01_UBTFex7_ri1660976_0164gF
G T A C A C C C A C A A G A A G A A G G T
jc1678064D01_UBTFex7_ri1660976_0164gR
G T A C A C C C A C G A G A A G A A G G T

UDP_2009


st40_2433_UBTF_C-T_R Fragment base #65. Base 65 of 144
st36_2625_UBTF_C-T_R Fragment base #57. Base 57 of 137
st34_2626_UBTF_C-T_R Fragment base #29. Base 29 of 114
st38_2434_UBTF_C-T_R Fragment base #26. Base 26 of 105

UDP_7621


C A C C T T C T T C T C G T G G G T G T A C C A
C A C C T T C T T C T C G T G G G T G T A C C A
C A C C T T C T T C T C G T G G G T G T A C C A
C A C C T T C T T C T C G T G G G T G T A C C A

UDP_5762

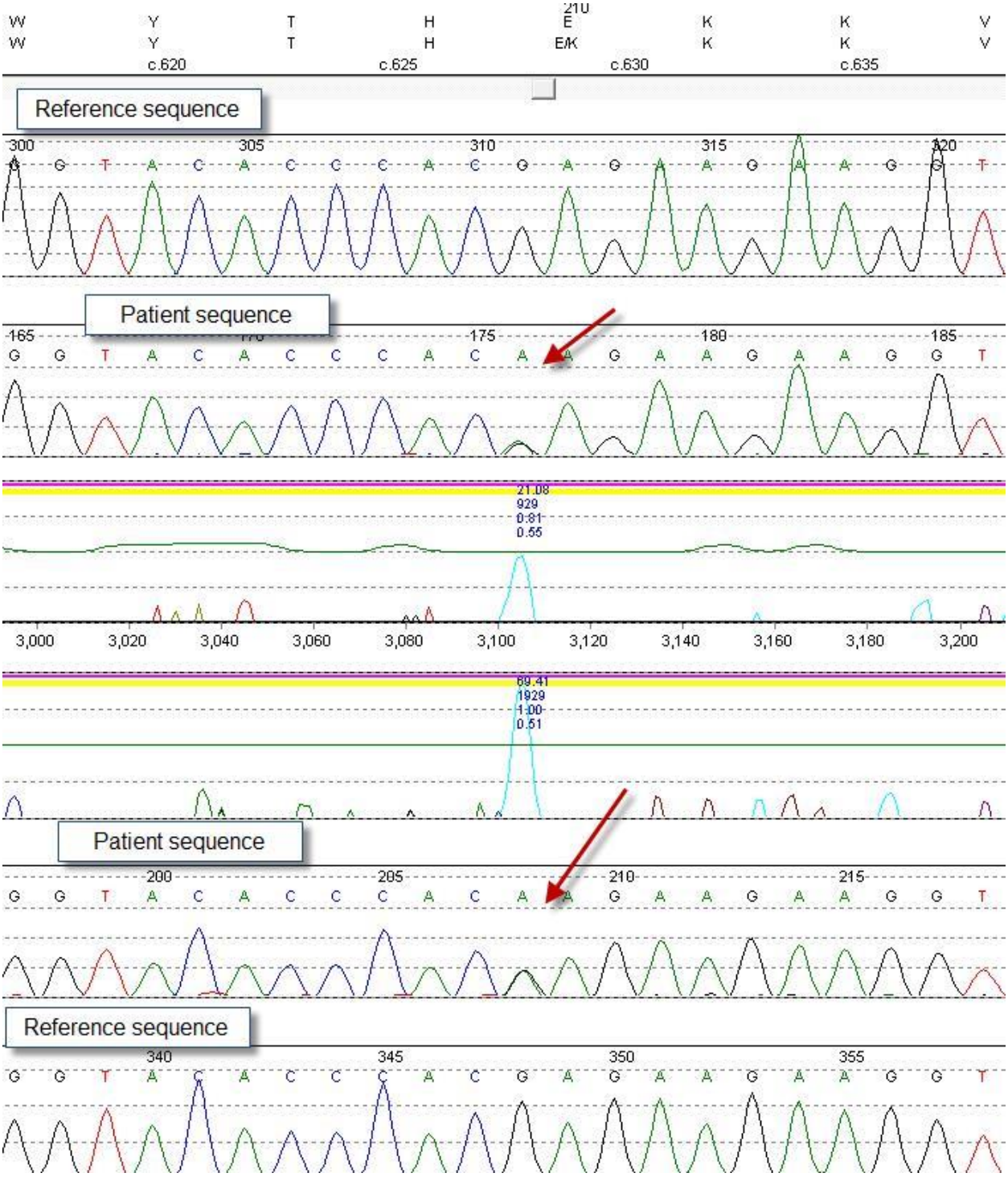
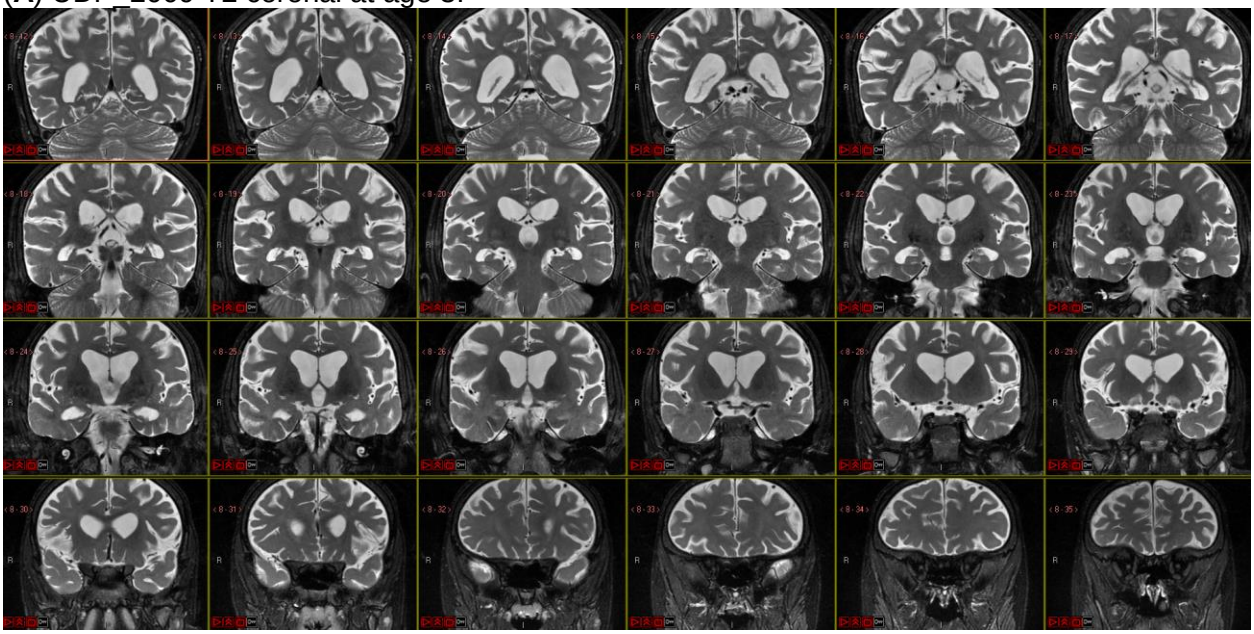
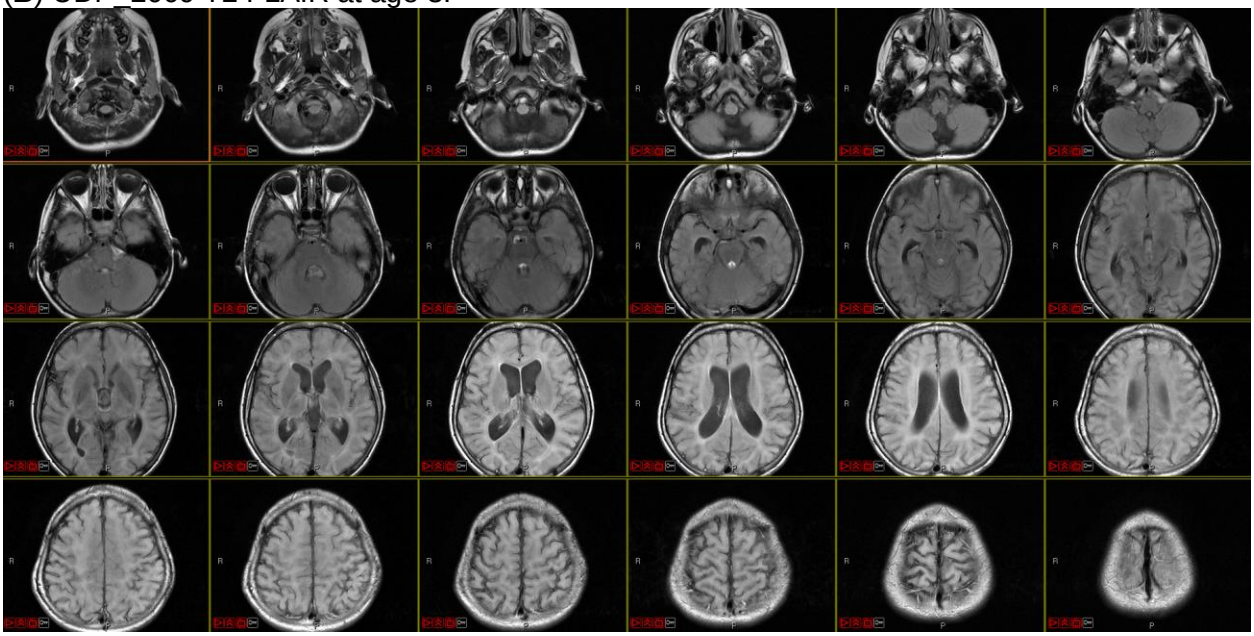


Figure S3. Confirmatory Sanger sequencing [NM_014233.2:c.628G>A]

(**A**) UDP_2009 T2 coronal at age 3.



(**B**) UDP_2009 T2 FLAIR at age 3.



(**C**) UDP_7621 T2 FLAIR at age 33

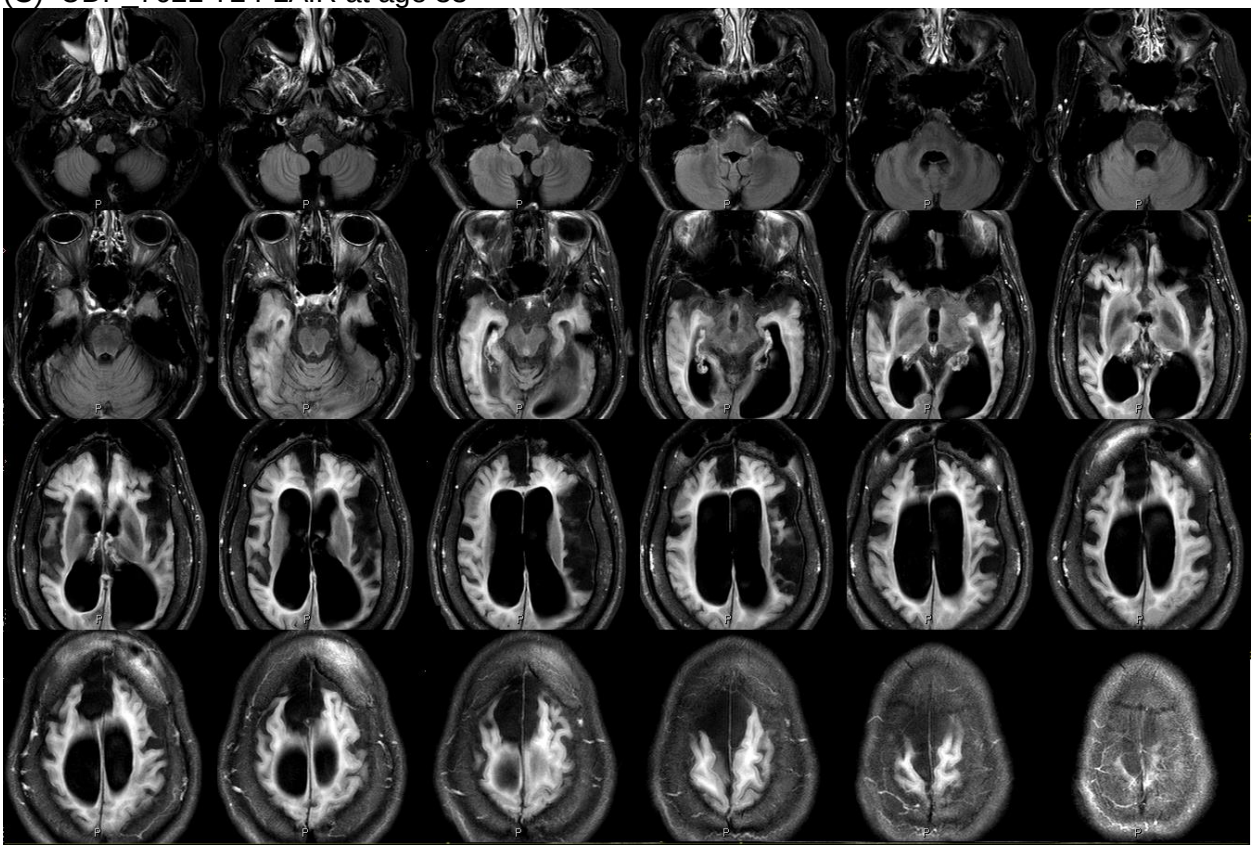


Figure S4. MRI images from UDP_2009 (**A** and **B**) and UDP_7621 (**C**).

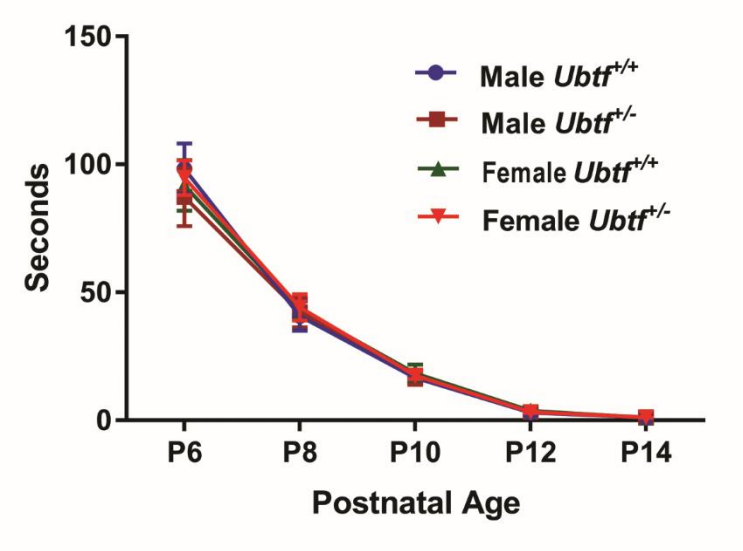


Figure S5. Righting reflex assays were performed prior to weaning in *Ubtf*$^{+/+}$ (N = 6 male and 5 female) and *Ubtf*$^{+/-}$ (N = 4 male and 5 female) mice.

Supplementary References

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and 1000 Genome Project Data Processing Subgroup. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics,* **25,** 2078-2079.

Quinlan, A.R. and Hall, I. M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics,* **26,** 841-842.

Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J. and Prins, P. (2015) Sambamba: fast processing of NGS alignment formats. *Bioinformatics*, **31**, 2032-2034.