

The American Journal of Human Genetics

Supplemental Data

**Allelic Mutations of *KITLG*, Encoding KIT Ligand,  
Cause Asymmetric and Unilateral Hearing Loss  
and Waardenburg Syndrome Type 2**

Celia Zazo Seco, Luciana Serrão de Castro, Josephine W. van Nierop, Matías Morín, Shalini Jhangiani, Eva J.J. Verver, Margit Schraders, Nadine Maiwald, Mieke Wesdorp, Hanka Venselaar, Liesbeth Spruijt, Jaap Oostrik, Jeroen Schoots, Baylor-Hopkins Center for Mendelian Genomics, Jeroen van Reeuwijk, Stefan H. Lelieveld, Patrick L.M. Huygen, María Insenser, Ronald J.C. Admiraal, Ronald J.E. Pennings, Lies H. Hoefsloot, Alejandro Arias-Vásquez, Joep de Ligt, Helger G. Yntema, Joop H. Jansen, Donna M. Muzny, Gerwin Huls, Michelle M. van Rossum, James R. Lupski, Miguel Angel Moreno-Pelayo, Henricus P. M. Kunst, and Hannie Kremer

## **Supplemental clinical information, figures, tables and legends**

### Supplemental Note

#### *Family W09-1628*

Individual V:7. Otoscopy revealed a fragile appearance of the malleus and an underdeveloped superior part of the auricle. These abnormalities were also seen in his brother (V:8) with normal hearing. CT imaging revealed no inner ear anomalies.

A dip at 4 kHz in the right ear was observed in pure tone audiometry. Such a dip is specifically associated with noise exposure although no excessive noise exposure was reported by the affected individual.

Individual III:6. The affection status of III:6 is indicated with a question mark in the pedigree (Figure 1) because his hearing loss (HL) might well be caused by other factors than the segregating genetic defect (Figure 1 and Figure S5). The HL in the left ear of individual III:6 had an onset between 45.56 and 46.46 years of age. Before the age of 45 years hearing was normal, as shown in Figure S3.

Individual V:10. This individual has a symmetric, mild mixed HL which can be explained by his recurrent otitis media.

Individual III:2. This individual, indicated to have balance problems while walking on uneven ground and when walking on a straight line. He is a spouse and does not carry the genetic defect segregating in the family.

Individual III:5. With regard to hair and eye color, individual III:5 was lighter than all other family members. She heterozygously carries the rs12821256 G allele that is associated with blond hair (Figure S5).<sup>1</sup>

#### *Family S1489*

Otoscopy did not reveal any abnormalities for the affected father and son (II:1 and III:1, respectively) and there was no evidence of non-genetic causes of the HL. MRI did neither demonstrate inner ear abnormalities nor cerebellopontine angle pathology. Propositus III:1 indicated to have suffered from atopic dermatitis in the past.

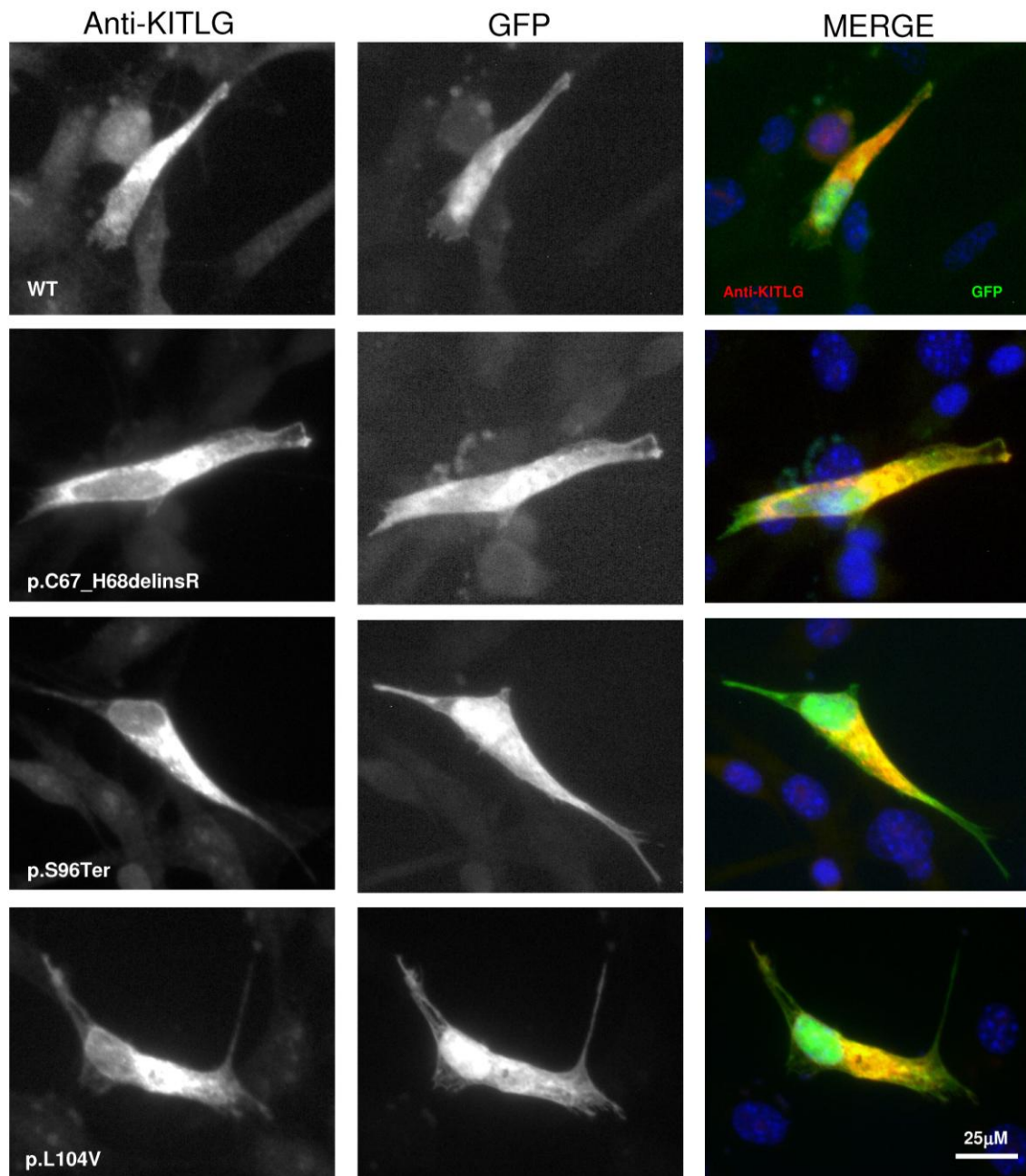


Figure S1. Expression of wild-type and aberrant soluble KITLG as detected by immunofluorescence in transfected NIH3T3 cells used in the ELISA assays.

Cells were permeabilized and incubated with anti-KITLG and Alexa488 goat anti mouse (left column), GFP signal (middle column) and merged (right column). Nuclei were stained with Hoechst (in blue). “WT” means wild-type.

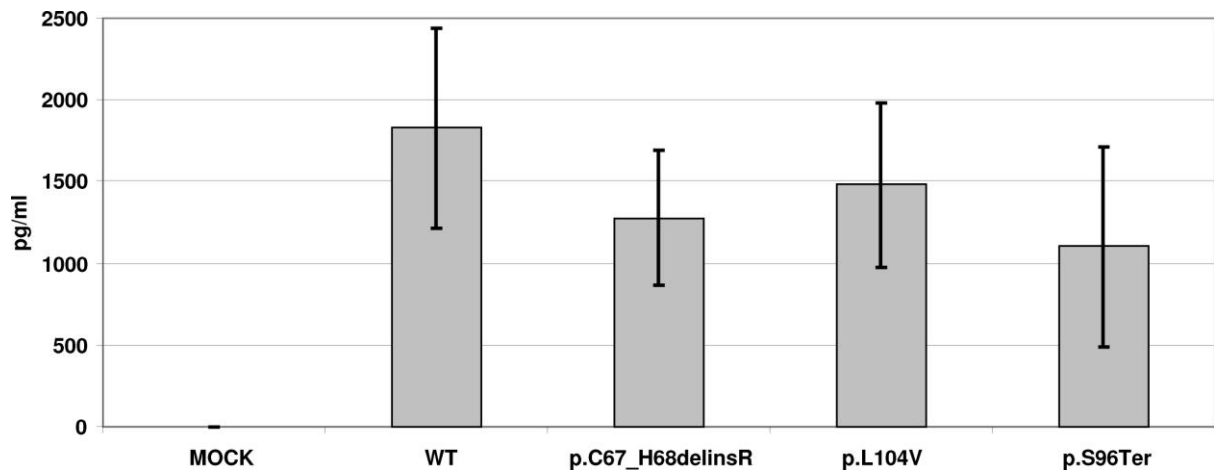
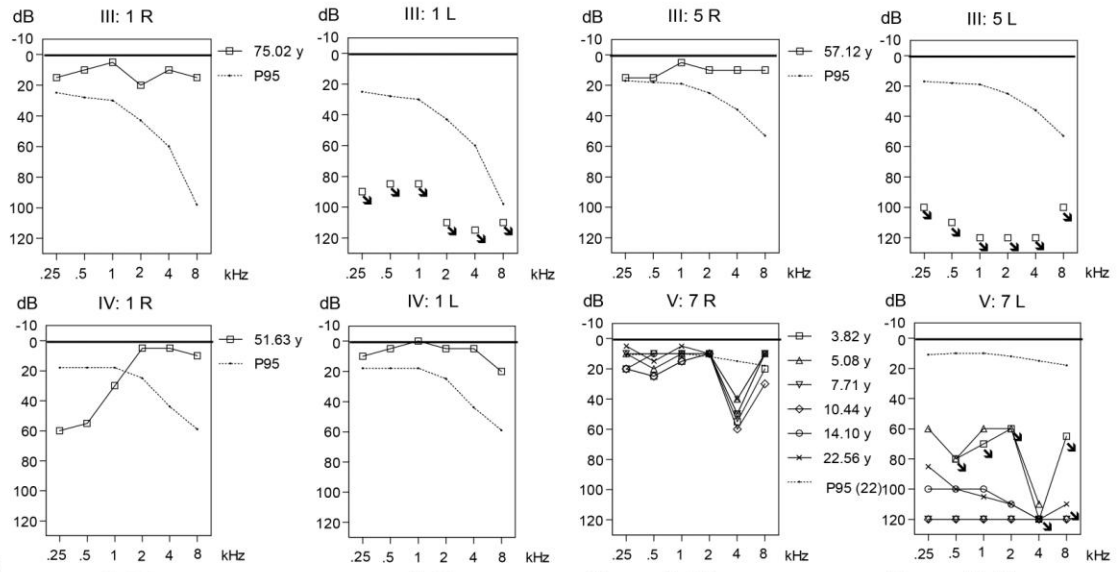
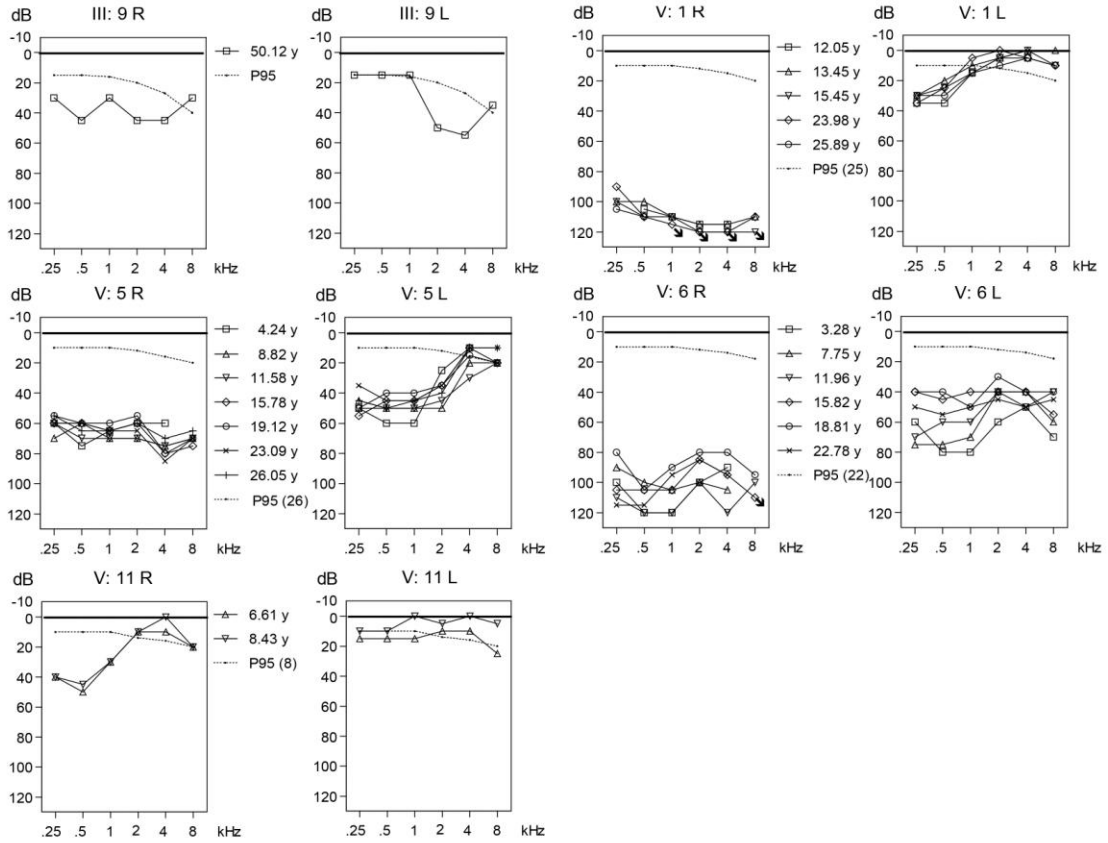


Figure S2: KITLG concentrations determined by ELISA. NIH3T3 cells transfected with a set of plasmids encoding FLAG-tagged wild-type and aberrant transmembrane KITLG were lysed and 10  $\mu$ l of cell extracts were analysed by ELISA. Expression of wild-type and aberrant KITLG was readily detected. Data are means  $\pm$ SD. “WT” means wild-type.

**A****B**

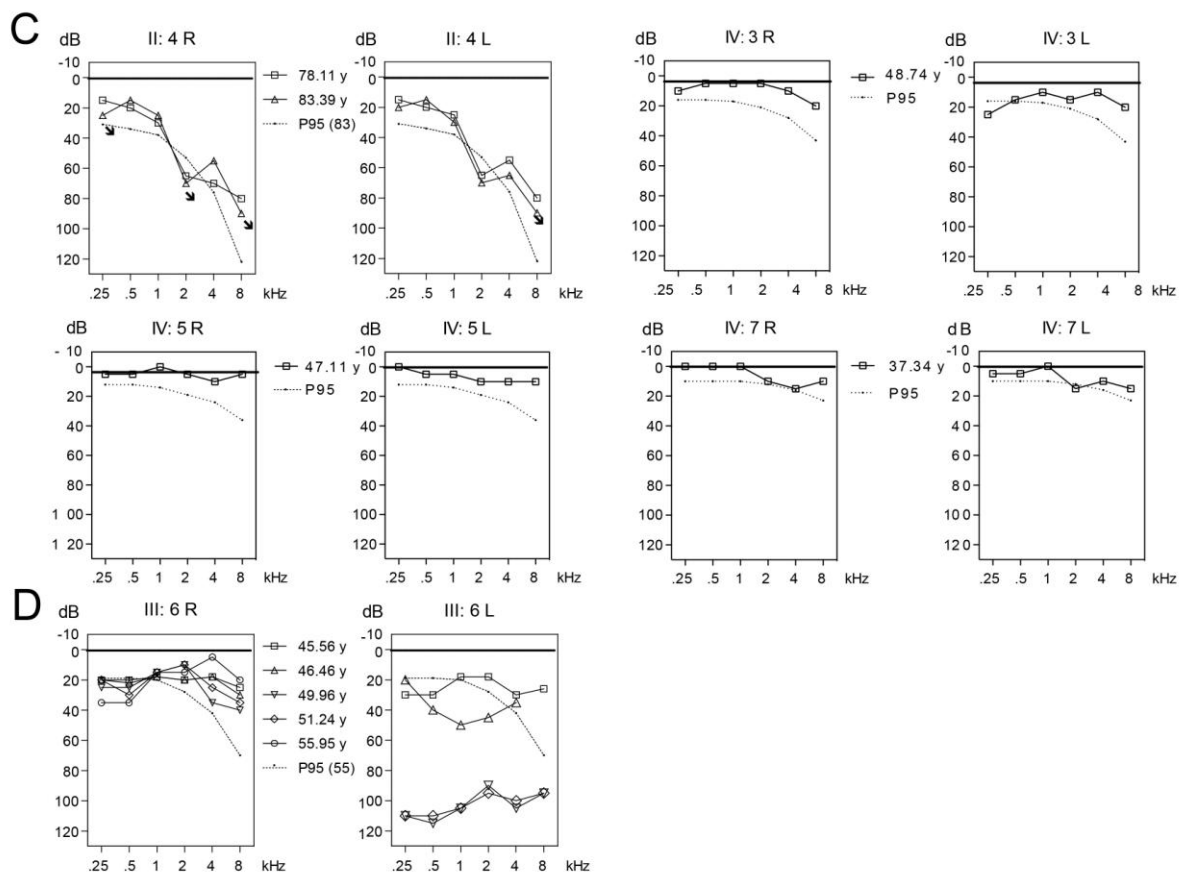


Figure S3. Pure tone air-conduction audiograms of individuals of family W09-1628 who have the *KITLG* c.286\_303delinsT variant. Thresholds of the right (R) and left (L) ears are shown separately. Age is indicated in years between the audiograms of R and L ears. When available, longitudinal data are shown. The dashed lines in the audiograms represent the P<sub>95</sub> thresholds related to the person's sex and age at the last measurement. Audiograms of individuals with UHL (A), AHL (B), normal hearing (C) and questionable affection status (D) are indicated.

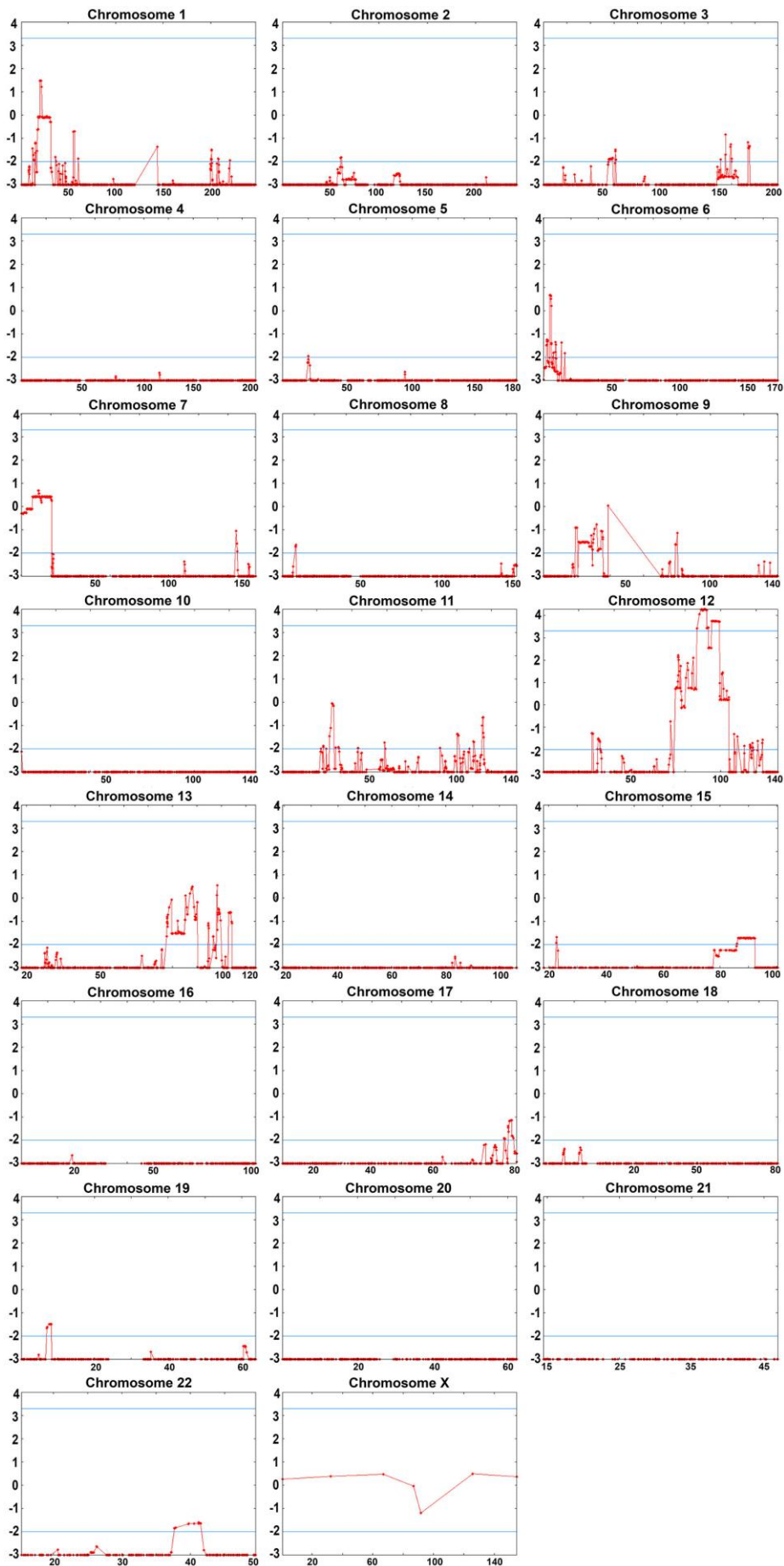




Figure S4. Genome wide LOD scores as calculated using SuperLink online SNP 1.1 software. The Y axis represents LOD scores and the X axis genetic distance in centimorgans (cM), per chromosome. The blue lines mark a LOD score of 3.3 indicating genome wide significance and a LOD score of -2.0 indicating exclusion of linkage. In the calculations, windows of 10 SNPs were used; penetrance was set at 70% and the disease allele frequency at 0.001. There is only one region in chromosome 12q21.32-q23.1, delimited by rs10459171 and rs35723, (chr12:87,808,426-100,960,087; GRCh37, hg19) with a significant maximum LOD score of 4.27 for rs7132875 and rs7309222.

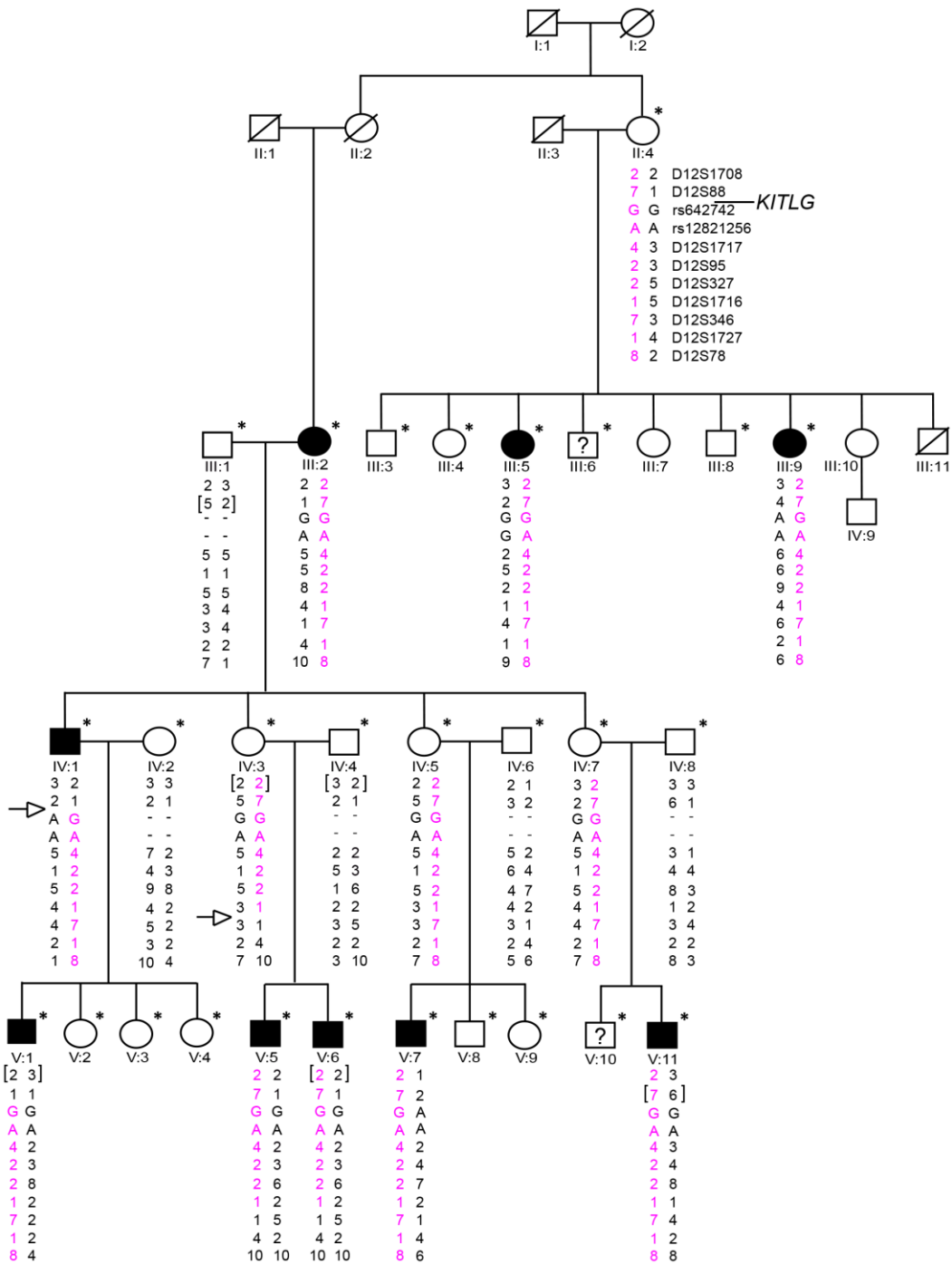


Figure S5. Genotypes of VNTR markers and SNPs in regulatory sequences of *KITLG*. VNTR marker order is according to the Marshfield map. In brackets, the genotypes are indicated which were inferred from the pedigree information. For privacy reasons, the genotypes for

the disease allele have not been indicated for unaffected individuals with the c.286\_303delinsT mutation except for those with affected offspring. Recombination events in individuals IV:1 and IV:3 define the disease allele (in pink) flanked by D12S88 to D12S346. Only one individual of unclear affection status (III:6) and three symptomatic individuals with the mutation (III:9, IV:1, V:7) were heterozygous for the A allele of SNP rs6427402 (allele frequency 92% in West Africans<sup>2</sup>). The G allele of rs12821256, which is associated with blond hair, is most prevalent in northern Europe.<sup>1</sup> Only individual III:5 has a G allele for rs12821256. Her hair and eye color are lighter blond and blue, respectively, when compared to the other family members (Table S4).

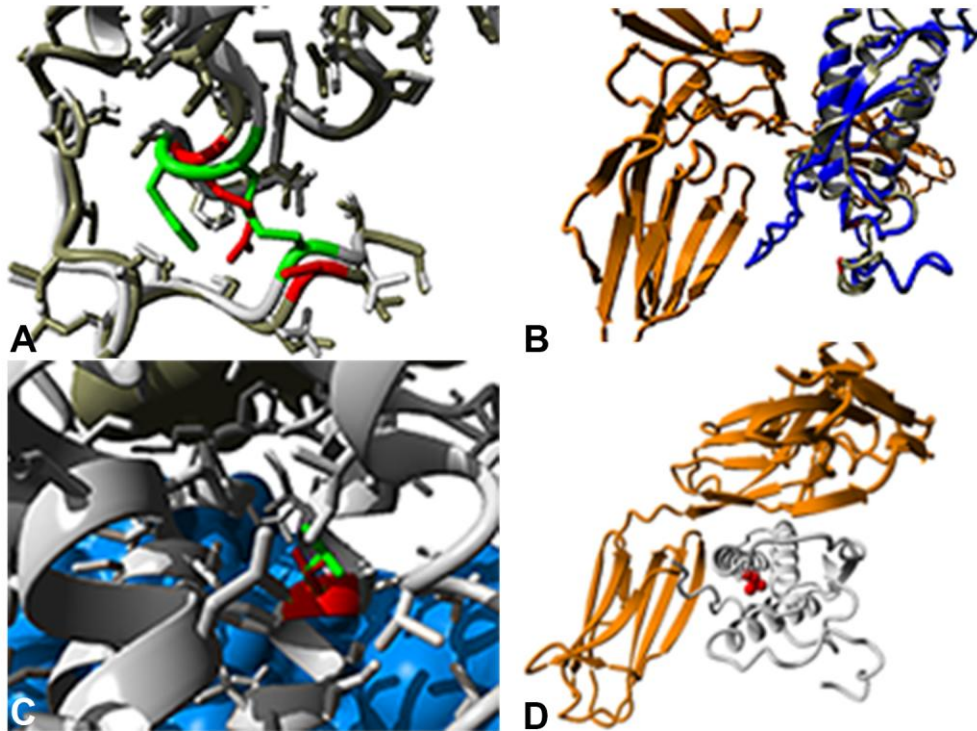


Figure S6. A molecular modeling of the KITLG variants p.His67\_Cys68delinsArg and p.Leu104Val using PDB files 1cfs and 2o26. The model for KITLG and KIT was constructed using the PDB files 1scf and 2o26 as templates.<sup>3</sup> The 1scf file contains the structure of human KITLG. In order to study the interaction of KITLG with the KIT receptor we constructed the Kit-Kitl complex using PDB-file 2o26. The 2o26 file contains the structures for KITLG and KIT from mouse, which are 77% and 67% identical to respectively KITLG and KIT in humans. The generated human KIT model was superposed on the mouse KIT-KITLG complex. The YASARA<sup>4</sup> and WHAT IF Twinset<sup>5</sup> were employed (with default parameters) for model building and subsequent analyses.<sup>6</sup> A) A close-up of the loop affected by the p.His67\_Cys68delinsArg variant. In green, the intramolecular disulfide bond formed by Cys68 and Cys163 is presented. The inserted Arg and Cys163 that is left without a binding partner are presented in red. Loss of the Cys68-Cys163 intramolecular bond is predicted to affect the local structure of the loop and the overall active KITLG conformation. The wild-type protein is in grey, the substituted in moss-green. B) Overview of the KIT receptor in

orange, the wild-type KITLG in blue, and the KITLG with p.His67\_Cys68delinsArg in grey. The amino acid residue affected by the substitution is in red. C) A close-up of the  $\alpha$ -helix affected by the p.Leu104Val variant. In green, the wild-type Leu104 is shown and in red the variant Val104. Valine is smaller than leucine and it prefers  $\beta$ -sheets rather than  $\alpha$ -helices. Therefore, p.Leu104Val might have a disruptive effect on the secondary structure of KITLG. D) Overview of the KIT receptor (orange) and the KITLG (grey) with the substituted Leu104, the side chain of this residue is shown as red balls.

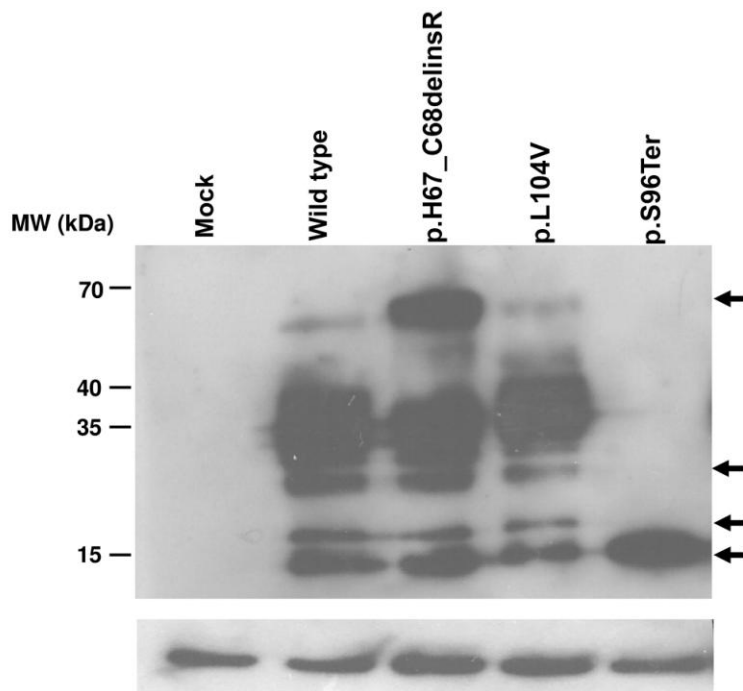


Figure S7. KITLG detection on Western blot of NIH3T3 cells transfected with constructs encoding transmembrane wild-type, p.His67\_Cys68delinsArg, p.Leu104Val and p.Ser96Ter KITLG. Transfected NIH3T3 cells were homogenized in lysis buffer (150 mM NaCl, 50 mM Tris HCl, pH 7.4, 5 mM EGTA, 5 mM EDTA, 1% Triton X-100, 25 lg/ml leupeptin, and 1 mM phenylmethylsulphonylfluoride) and kept on ice for 30 min. After centrifugation at 10,000g (4°C) for 5 min, 50 µg of protein were mixed 1:1 with SDS-PAGE sample buffer (50 mM Tris-Cl, pH 6.8, 50 mM DTT, 2% SDS, 10% glycerol, 5% βME). Proteins were separated in a SDS-PAGE gel and transferred electrophoretically to nitrocellulose membranes. After blocking with 5% skimmed dry milk in 1% TBS (pH 7.6), 0.1% Tween 20 for 1 hr at RT, blots were incubated with the mouse anti-FLAG antibody (1:1000) (SIGMA, F1804)) overnight at 4°C. After washing, membranes were incubated with sheep anti-mouse IgG (1:10000) conjugated with HRP (Amersham Biosciences) for 1 hr at RT. Bands were visualized by enhanced chemiluminiscence (Clarity Western ECL substrate; Biorad) and exposed to X-ray film at different time points. Beta-actin was used as loading control. Arrows

indicate bands of ~15, ~16, ~28 and ~68 kDa, respectively. Shown data is a representative of three independent experiments. Besides the ~28 kDa band corresponding to the expected size of FLAG-tagged KITLG, several other bands spanning a molecular weight range of about 28-35 kDa were detected, indicating extensive and heterogeneous glycosylation.<sup>7,8</sup> An additional band of 15 kDa was readily detected for the truncated p.Ser96Ter KITLG and was also present in the other samples, but with less intensity. This protein might represent the usage of an alternative ATG codon, downstream of the codon of Serine 96, and in frame with the three FLAG tags. A lighter second band of around 16 kDa was also detected in all samples which might be the result of proteolytic processing of the protein. Although the transmembrane isoform, expressed in this experiment, lacks the proteolytic cleavage site (encoded by exon 6), alternative proteolytic might have occurred as has been proposed for the murine Kitl.<sup>9</sup> Finally, a band of ~68 kDa was detected which is more intense for the p.His67\_Cys68delinsArg mutation than for the other mutations and control suggesting an increased tendency to aggregate. The bands of about 16 and 15 kDa were not detected when we used the anti-KITLG antibody that is directed against the N-terminal region of the human KITLG (data not shown).

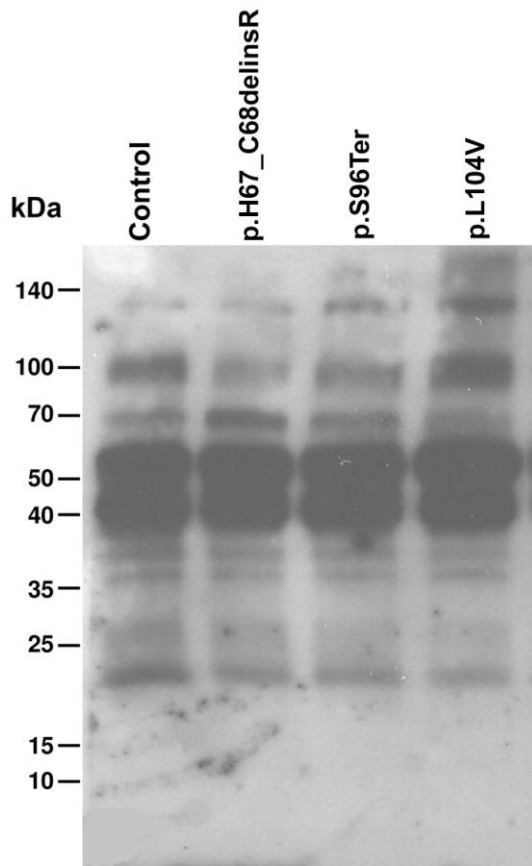
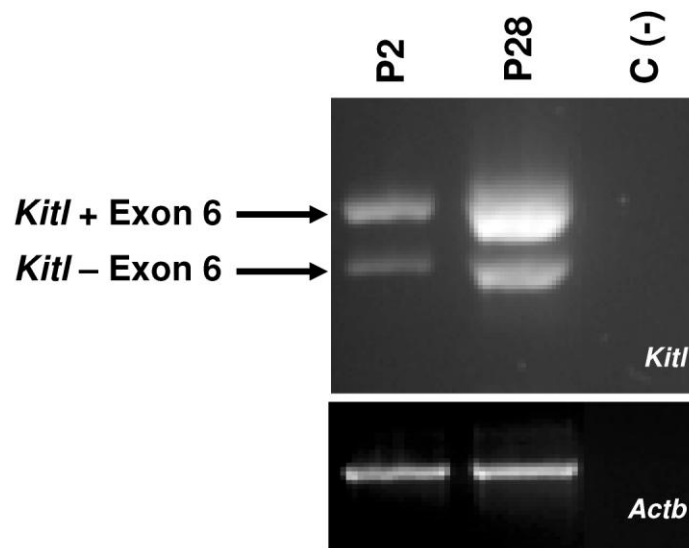


Figure S8. p.His67\_Cys68delinsArg, p.Ser96Ter and p.Leu104Val aberrations do not lead to clear alterations of KITLG levels in blood. To analyze KITLG expression in human blood, 1 ml of blood was collected from individuals II:1 of family 12-01744, V:1 of family W09-1628 and II:1 of family S1489. Blood samples were centrifuged for 15 min at 15000 rpm at 4 °C. One  $\mu$ l of supernatant was used for Western blot analysis with the anti-KITLG antibody (abcam, ab52603, anti-SCF) as primary antibody (1:1000) and the sheep anti-rabbit IgG (1:10000) conjugated with HRP (Amersham Biosciences) as secondary antibody. Western blot analysis was performed as described in the legend of Figure S7. Samples were analyzed on 12% SDS-PAGE followed by Western blotting with polyclonal anti-KITLG. No differences were observed between samples of control individuals and those of individuals with a KITLG mutation. There was no band visible that could expectedly represent the truncated p.Ser96Ter KITLG.



A



B

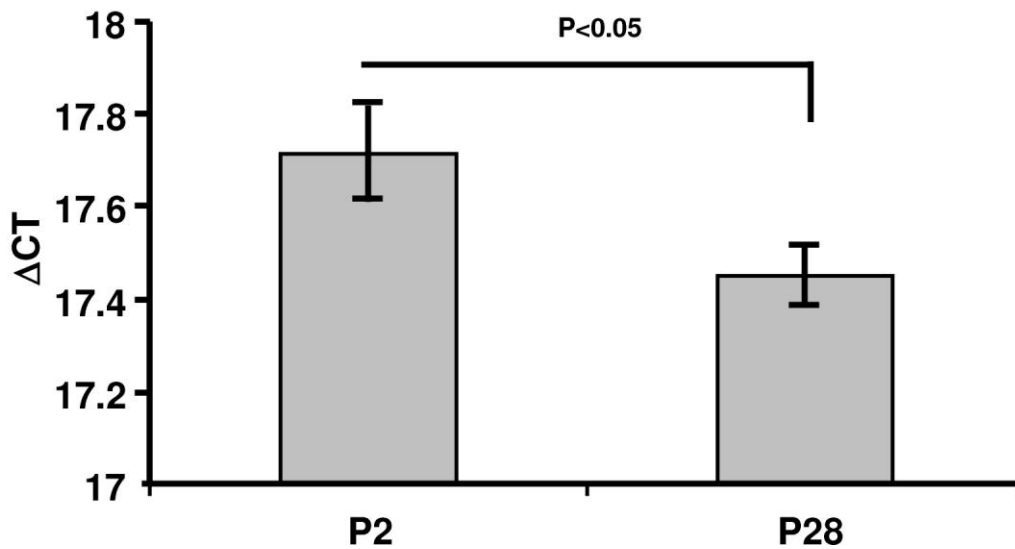


Figure S9. *Kitl* expression in mouse cochlea. Cochlea cDNA derived from cochlea RNA of postnatal (P) days P2 and P18 129sv wild-type mice was kindly provided by Prof. Guy Richardson (Sussex University, UK). PCR was performed with this cDNA as template and primers for *Actb* and *Kitl* and subsequently size-separated on 1% agarose gel according to standard protocols. To quantify *Kitl* expression, one  $\mu$ l of cDNA was used as template for quantitative PCR with SYBR Green (11066420, SYBR Green I Master, Roche Diagnostics). The gene for ribosomal 18S rRNA was employed as housekeeping gene for data normalization. The differences in expression were calculated by subtracting the obtained CT values of *Kitlg* and 18S rRNA, respectively, according to the following equation:  $\Delta CT = CT$

(*Kitlg*) – CT (18S rRNA). A) Transcripts of *Kitl* in P2 and P28 mouse cochlea. *Actb* was employed as control. B) Quantitative PCR revealed changes of relative *KITLG* transcript levels during functional maturation of the mouse cochlea. For statistical evaluation the Student's t-test was performed. Error bars represent standard error of means.

## Supplemental tables and legends:

Table S1. Primer sequences.

Target	Primer	Oligonucleotides
rs642742	Forward	TGGGGTTTAAAGCATAGACTG
	Reverse	TGTGGTAATGGGTGAGTCTG
rs12821256	Forward	AAGTATGCCCAAAGGATAAGG
	Reverse	GAGTGCTTTGTTCCAACCTAGTC
<i>KITLG</i> _mRNA_ NM_003994.5_BamHI_F	Forward	CGCTGCGGATCCTTATGAAGAAGACA
<i>KITLG</i> _mRNA_ NM_003994.5_XhoI_R1	Reverse	CGAAAGTAAACAGTGTGACTCGAGCCACAA
<i>KITLG</i> _mRNA_ NM_003994.5_XhoI_R2	Reverse	ACAAGCCACTCGAGCACTTCTTGAAACTCTC
<i>KITLG</i> _intron ex3-4_gateway_F	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGGAAGTTTGAGACAGCCTGG
<i>KITLG</i> _intron ex4-5_gateway_R	Reverse	GGGGACCACTTTGTACAAAGAAAGCTGGGTCGCAGGACTGATTTTGCATTG
pCI-Neo-Rho-Insert Fw	Forward	CGGAGGTCAACAACGAGTCT
pCI-Neo-Rho-Insert Rev	Reverse	AGGTGTAGGGGATGGGAGAC
<i>KITLG</i> _Exon1	Forward	GCTCCAGAACAGCTAAACGG
	Reverse	CAGCTGCAAGTCCCAGG
<i>KITLG</i> _Exon2	Forward	GGTGAGCATAGCTTGAATGC
	Reverse	GCAGTGGTGTGCAACTAGC
<i>KITLG</i> _Exon3	Forward	TGCTTTTCTCCAAAGCACTAC
	Reverse	CAATAAGCAAGCTCCTAAATAGC
<i>KITLG</i> _Exon4	Forward	TGCAGGCACTTGTAACTCTG
	Reverse	CTTGTTTTCTTTGCTCATCTG
<i>KITLG</i> _Exon5	Forward	CATACGTAAAACAGCCATCTAAAAG
	Reverse	TCTTGCTACAGCTTAACACAGAGG
<i>KITLG</i> _Exon6	Forward	AGACCTGGCAAGTAATCTGG
	Reverse	AGAAGACCATAAAAAGGAATAACAG
<i>KITLG</i> _Exon7	Forward	CTTGCAAAAAGCAATATTCAGC
	Reverse	TCAGCTAGAGGGATTGTCTCC
<i>KITLG</i> _Exon8	Forward	CAGCTTCTTTATTACTCTTTAGCCC
	Reverse	ATGGAGATGCAGCACTGAAG
<i>KITLG</i> _Exon9	Forward	AGAAGGTAGGCTCCCCATTC
	Reverse	ACAATGGAGCCTGTCATGG
<i>KITLG</i> _control testing_p.His67_Cys68delinsArg	Forward	6FAM CTCCGAGTTTATGGCACTTAC
	Reverse	CTTGTTTTCTTTGCTCATCTG
<i>KITLG</i> _control testing_p.Ser96Ter	Forward	GACAGCTTGACTGATCTTCTGG
	Reverse	GAGTTTTCTTTCACGCACTCC
<i>Kitl</i> _qPCR_cochlea	Forward	TATGATAACCCTCAACTATGTCGC
	Reverse	TTGTCCAGAAGAGTAGTCAAGC
18S rRNA	Forward	CGGCTACCACATCCAAGGA
	Reverse	AATTACCGCGGCTGCTG
<i>Actb</i>	Forward	GATGGTGGGCATGGGTCAGA
	Reverse	GGATGTCCACGTCACACTTCATGA
<i>Kitl</i>	Forward	CGGGAATCCTGTGACTGATAA
	Reverse	CCACTGTGCGAAGGTAACAA

For the primer design to amplify *KITLG*, NM\_003994.5 and NM\_000899.4 were used. For *Actb*, uc009ajk.1 was employed and for *Kitl*, uc007gxp.1.

Table S2. Overview of the results of audiometry, otoscopic examination, and vestibular function tests in family W09-1628.

Subject	Age at evaluation	Type of hearing	PTA (1,2,4 kHz)		Speech discrimination		Otosopic examination	Vestibular function test	
			R	L	R	L		Rotary chair	Caloric test
II:4 <sup>a</sup>	83	Normal	50	55	40% at 100 dB	30% at 100 dB	Normal	NT	NT
III:1	75	UHL	12	≥108	100% at 60 dB	0% at 80 dB	Normal	NT	NT
III:5	57	UHL	8	>120	100% at 60 dB	0% at 120 dB	Normal	Asymmetry Hyporeflexia	Asymmetry to the detriment of L
III:6 <sup>b</sup>	55	AHL	18	116	0% at 120 dB	100% at 60 dB	Normal	NT	NT
III:9	50	AHL	40	40	90% at 100 dB	88% at 80 dB	Normal	NT	NT
IV:1	51	UHL	13	3	100% at 80 dB	NT	Normal	Hyporeflexia	Asymmetry to the detriment of R and hyporeflexia R
IV:3 <sup>a</sup>	48	Normal	8	12	NT	NT	Normal	NT	NT
IV:5 <sup>a</sup>	47	Normal	5	5	NT	NT	Normal	NT	NT
IV:7 <sup>a</sup>	37	Normal	8	8	NT	NT	Normal	NT	NT
V:1	25	AHL	>120	8	0% at 120 dB	100% at 60 dB	Normal	Normal	Asymmetry to the detriment of R and hyporeflexia R
V:5	26	AHL	65	32	80% at 110 dB	100% at 80 dB	Normal	Normal	Normal
V:6	22	AHL	92	48	0% at 120 dB	100% at 90 dB	Normal	Normal	Normal
V:7	22	UHL	20	108	95% at 60 dB	0% at 120 dB	R+L: fragile malleus	Hyporeflexia	Bilateral marginal caloric hyporeflexia
V:11	8	AHL	17	12	97% at 80 dB	97% at 50 dB	Normal	NT	NT

Listed individuals are heterozygous for the *KITLG* c.286\_303delinsT mutation. Data for unaffected individuals with the mutation but without affected offspring are not indicated. <sup>a</sup> normal hearing; <sup>b</sup> other causes of HL cannot be excluded; NT, not tested. PTA of 1,2,4 kHz of the last audiogram is presented.

Table S3: Overview of the results of the hematological evaluations in family W09-1628.

Subject	Age at evaluation	Gender	p.Ser96Ter	HL	Anemia by history	Haemoglobin Mmol/L	Hematocrit	Erythrocytes 10 <sup>12</sup> /L	MCHC mmol/L	MCV fL	MCH fmol	RDW %	Trombocytes 10 <sup>9</sup> /L
III: 1	63	F	Yes	Yes	No	8.20	0.39	4.27	20.90	92.00	1.92	12.80	216.00
III:5	62	F	Yes	Yes	No	8.60	0.40	4.31	21.30	94.00	2.00	13.40	286.00
III:6	63	M	Yes	Yes	No	9.90	0.44	4.93	22.30	90.00	2.01	13.00	206.00
III:10	53	F	No	No	Yes	9.40	0.44	5.11	21.30	86.00	1.84	12.80	296.00
IV: 1	56	M	Yes	Yes	Yes	8.80	0.40	4.43	21.80	91.00	1.99	12.70	204.00
IV: 3	53	F	Yes	No	Yes	7.50	0.36	3.88	20.60	94.00	1.93	12.80	190.00
IV: 7	42	F	Yes	No	No	8.40	0.38	4.15	22.40	90.00	2.02	13.20	289.00
IV:9	25	F	No	No	Yes	9.10	0.41	4.55	22.00	91.00	2.00	12.80	309.00
V: 1	31	M	Yes	Yes	No	9.00	0.41	4.49	21.90	92.00	2.00	12.50	246.00
V: 2	29	F	No	No	No	8.30	0.41	4.52	20.40	90.00	1.84	13.10	252.00
V: 3	26	F	No	No	No	8.00	0.38	4.11	20.80	94.00	1.95	12.70	245.00
V: 5	31	M	Yes	Yes	No	8.70	0.41	4.32	21.30	95.00	2.01	12.30	250.00
V: 6	28	M	Yes	Yes	No	9.10	0.43	4.45	21.40	96.00	2.04	12.40	169.00
V: 11	13	M	Yes	Yes	No	8.50	0.37	4.41	22.70	85.00	1.93	12.50	314.00

Subject	Age at evaluation	Gender	p.Ser96Ter	HL	Leukocytes10 <sup>9</sup> /L	Reticulocytes permille	Reticulocytes absolute 10 <sup>9</sup> /L	Reticulocyte haemoglobin equivalent fmol	Auto diff. neutrophils 10 <sup>9</sup> /L	Auto diff. lymphocytes 10 <sup>9</sup> /L	Auto diff. monocytes 10 <sup>9</sup> /L	Auto diff. eosinophilic granulocytes 10 <sup>9</sup> /L	Auto diff. granulocytes 10 <sup>9</sup> /L
III: 1	63	F	Yes	Yes	7.30	10.00	41.00	2.12	3.15	2.92	0.54	0.60	0.09
III:5	62	F	Yes	Yes	7.20	10.00	41.80	2.31	4.02	2.45	0.62	0.12	0.03
III:6	63	M	Yes	Yes	5.00	8.00	40.40	2.29	2.28	2.06	0.48	0.13	0.02
III:10	53	F	No	No	7.20	8.00	41.40	2.10	3.43	2.90	0.72	0.15	0.04
IV: 1	56	M	Yes	Yes	7.40	7.00	30.10	2.27	5.42	1.38	0.40	0.13	0.03
IV: 3	53	F	Yes	No	5.60	10.00	37.20	2.26	3.09	1.93	0.42	0.13	0.02
IV: 7	42	F	Yes	No	6.40	10.00	39.40	2.27	3.74	2.20	0.30	0.12	0.03
IV:9	25	F	No	No	6.20	11.00	50.10	2.21	2.91	2.57	0.60	0.10	0.02
V: 1	31	M	Yes	Yes	4.20	6.00	28.70	2.14	2.14	1.64	0.32	0.04	0.02
V: 2	29	F	No	No	7.10	12.00	52.40	2.16	4.30	1.99	0.62	0.14	0.04
V: 3	26	F	No	No	6.60	10.00	39.00	2.10	3.78	2.13	0.41	0.25	0.02
V: 5	31	M	Yes	Yes	8.10	12.00	51.00	2.45	4.95	2.40	0.62	0.13	0.01
V: 6	28	M	Yes	Yes	4.50	13.00	58.70	2.42	2.10	1.72	0.45	0.24	0.01
V: 11	13	M	Yes	Yes	6.00	19.00	82.50	2.17	2.76	2.36	0.58	0.30	0.03

The normal values are 1) for males: haemoglobin 8.4-10.8 mmol/L; hematocrit 0.41-0.53; erythrocytes  $4.5-5.9 \times 10^{12}/L$ ; MCV 80-100 fL; trombocytes:  $150-400 \times 10^9/L$ ; leucocytes  $4-11 \times 10^9/L$ ; 2) for females: haemoglobin 7.4-9.9 mmol/L; hematocrit 0.36-0.46; erythrocytes  $4-5.2 \times 10^{12}/L$ ; MCV 80-100 fL; trombocytes:  $150-400 \times 10^9/L$ ; leucocytes  $4-11 \times 10^9/L$ .

MCHC, Mean corpuscular haemoglobin concentration; MCV, mean corpuscular volume; RDW, red blood cell distribution width.

Table S4. Overview of the results of the dermatological evaluations in family W09-1628.

Pedigree	Age at evaluation	Gender	p.Ser96Ter	HL	Skin type	Eye colour	Dystopia canthorum	Hair colour	Age at greying	Skin depigmentation patches	Skin hypopigmentation patches	Skin hyperpigmentation patches	Poliosis	Naevi
III: 1	63	F	Yes	Yes	x	blue	No	dark blond	60	No	x	x	x	x
III:5	62	F	Yes	Yes	I	light blue	No	white blond	50	Yes	Yes	No	Yes	<50
III:6	63	M	Yes	Yes	I	blue	No	blond	40	No	Yes	No	No	<50
III:10	53	F	No	No	II	blue	No	dark blond	no greying	Yes	Yes	No	No	50-100
IV: 1	56	M	Yes	Yes	I	blue	No	blond	50	No	Yes	No	No	<50
IV: 3	53	F	Yes	No	II	blue	No	blond	no greying	Yes	Yes	No	No	<50
IV: 7	42	F	Yes	No	II	blue	No	blond	no greying	Yes	Yes	No	No	50-100
IV:9	25	F	No	No	I	blue	No	dark blond	no greying	No	No	No	No	<50
V: 1	31	M	Yes	Yes	II	blue	No	light blond	no greying	No	No	No	No	<50
V: 2	29	F	No	No	II	blue	No	dark blond	no greying	No	No	No	No	<50
V: 3	26	F	No	No	I	blue	No	blond	no greying	No	Yes	No	Yes	<50
V: 5	31	M	Yes	Yes	II	blue	No	dark blond	no greying	No	Yes	No	No	<50
V: 6	28	M	Yes	Yes	II	blue	No	dark blond	no greying	Yes	No	No	No	<50

x; examinations could not be performed.

Table S5. Filter steps applied on total CNV data from WES of individuals III:9 and IV:1 of family W09-1628.

<b>Filter steps</b>	<b># Variants</b>
Total CNVs	26
CNVs $\geq$ 5Kb	24
Exonic CNV	22
< 70% overlapping with control population	4
Linkage region 12q21.31-q23.1	0

Indicated is the number of CNVs, compared to reference genome GRCh37 (hg19), after every filter step. Total CNVs indicates the number of CNVs called in individuals III:9 and IV:1. The control dataset consisted of 721 genotypes of control individuals with Cytoscan HD arrays (Affymetrix) and 1251 genotypes of control individuals with Affymetrix 6.0 arrays (Affymetrix).



## References

1. Guenther, C.A., Tasic, B., Luo, L., Bedell, M.A., and Kingsley, D.M. (2014). A molecular basis for classic blond hair color in Europeans. *Nat. Genet.* *46*, 748-752.
2. Miller, C.T., Beleza, S., Pollen, A.A., Schluter, D., Kittles, R.A., Shriver, M.D., and Kingsley, D.M. (2007). cis-Regulatory changes in Kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell* *131*, 1179-1189.
3. Yuzawa, S., Opatowsky, Y., Zhang, Z., Mandiyan, V., Lax, I., and Schlessinger, J. (2007). Structural basis for activation of the receptor tyrosine kinase KIT by stem cell factor. *Cell* *130*, 323-334.
4. Krieger, E., Koraimann, G., and Vriend, G. (2002). Increasing the precision of comparative models with YASARA NOVA--a self-parameterizing force field. *Proteins* *47*, 393-402.
5. Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* *8*, 52-56, 29.
6. Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., and Karplus, K. (2009). Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins* *77 Suppl 9*, 114-122.
7. Jiang, X., Gurel, O., Mendiaz, E.A., Stearns, G.W., Clogston, C.L., Lu, H.S., Osslund, T.D., Syed, R.S., Langley, K.E., and Hendrickson, W.A. (2000). Structure of the active core of human stem cell factor and analysis of binding to its receptor kit. *EMBO J.* *19*, 3192-3203.
8. Arakawa, T., Yphantis, D.A., Lary, J.W., Narhi, L.O., Lu, H.S., Prestrelski, S.J., Clogston, C.L., Zsebo, K.M., Mendiaz, E.A., Wypych, J., et al. (1991). Glycosylated and unglycosylated recombinant-derived human stem cell factors are dimeric and have extensive regular secondary structure. *J. Biol. Chem.* *266*, 18942-18948.
9. Majumdar, M.K., Feng, L., Medlock, E., Toksoz, D., and Williams, D.A. (1994). Identification and mutation of primary and secondary proteolytic cleavage sites in murine stem cell factor cDNA yields biologically active, cell-associated protein. *J. Biol. Chem.* *269*, 1237-1242.