New Insights into Hallux Valgus by Whole Exome Sequencing Study

Jun Jia [1,#], Junyi Li [2,#], Huiqi Qu [3,#], Mengyu Li [2], Sipeng Zhang [2], Jun Hao [4], Xinyi Gao [2], Xinyi Meng [2], Yan Sun [2], Hakon Hakonarson [3,5,6], Xiantie Zeng [1,*], Qianghua Xia [2,*], Jin Li [2,*]

Supplementary Materials:

Figure S1: the pipeline of WES data analyses.

Figure S2: the appearance of the feet of family members of family 1.

Figure S3: the appearance of the feet of family members of family 2.

Figure S4: the appearance of the feet of family members of family 3.

Figure S5: digital radiography of proband (II-1) in family 1.

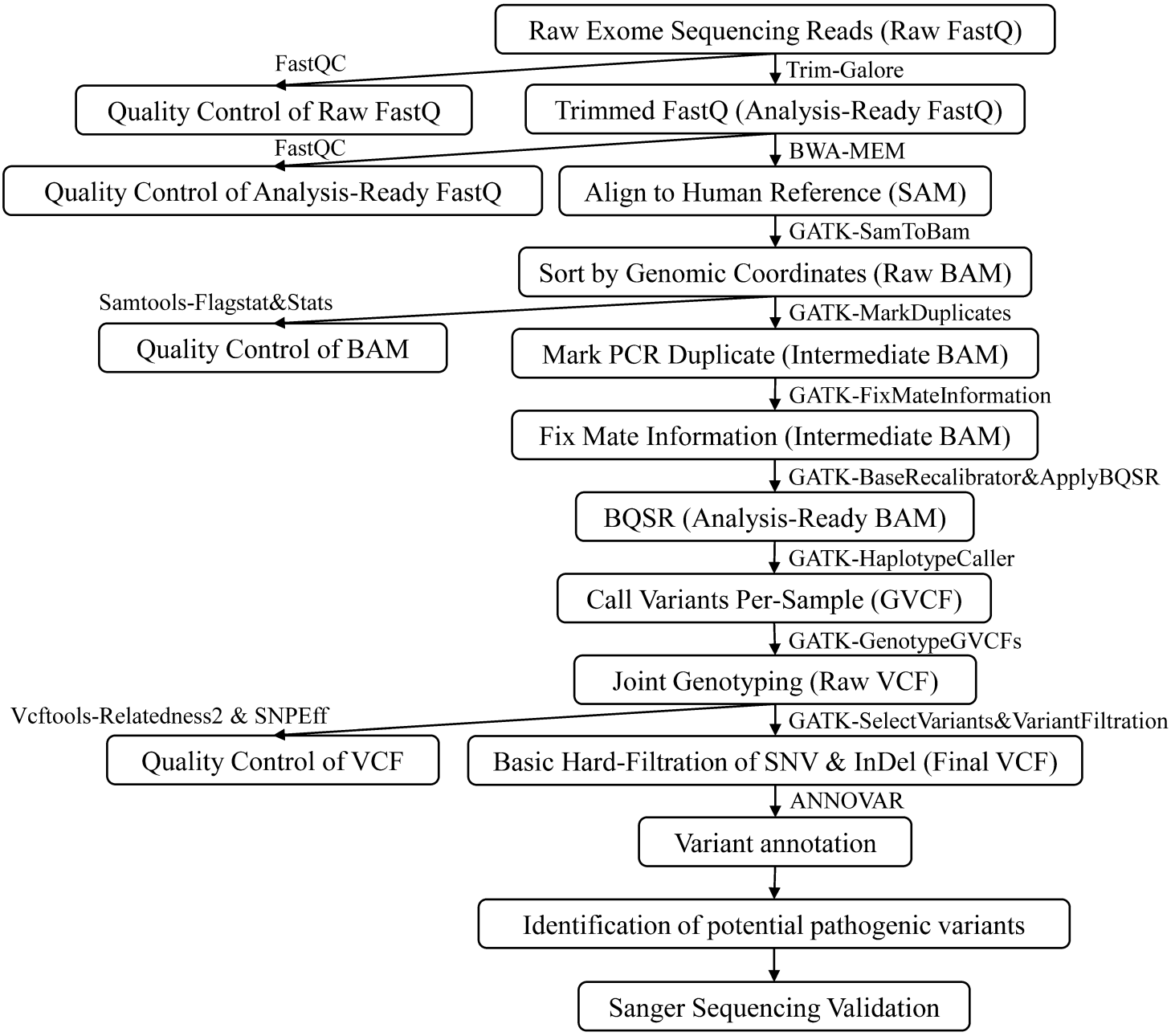
Figure S6: digital radiography of proband (I-2) in family 2.

Figure S7: digital radiography of proband (II-1) in family 3.

Table S1: PCR primers for Sanger sequencing.

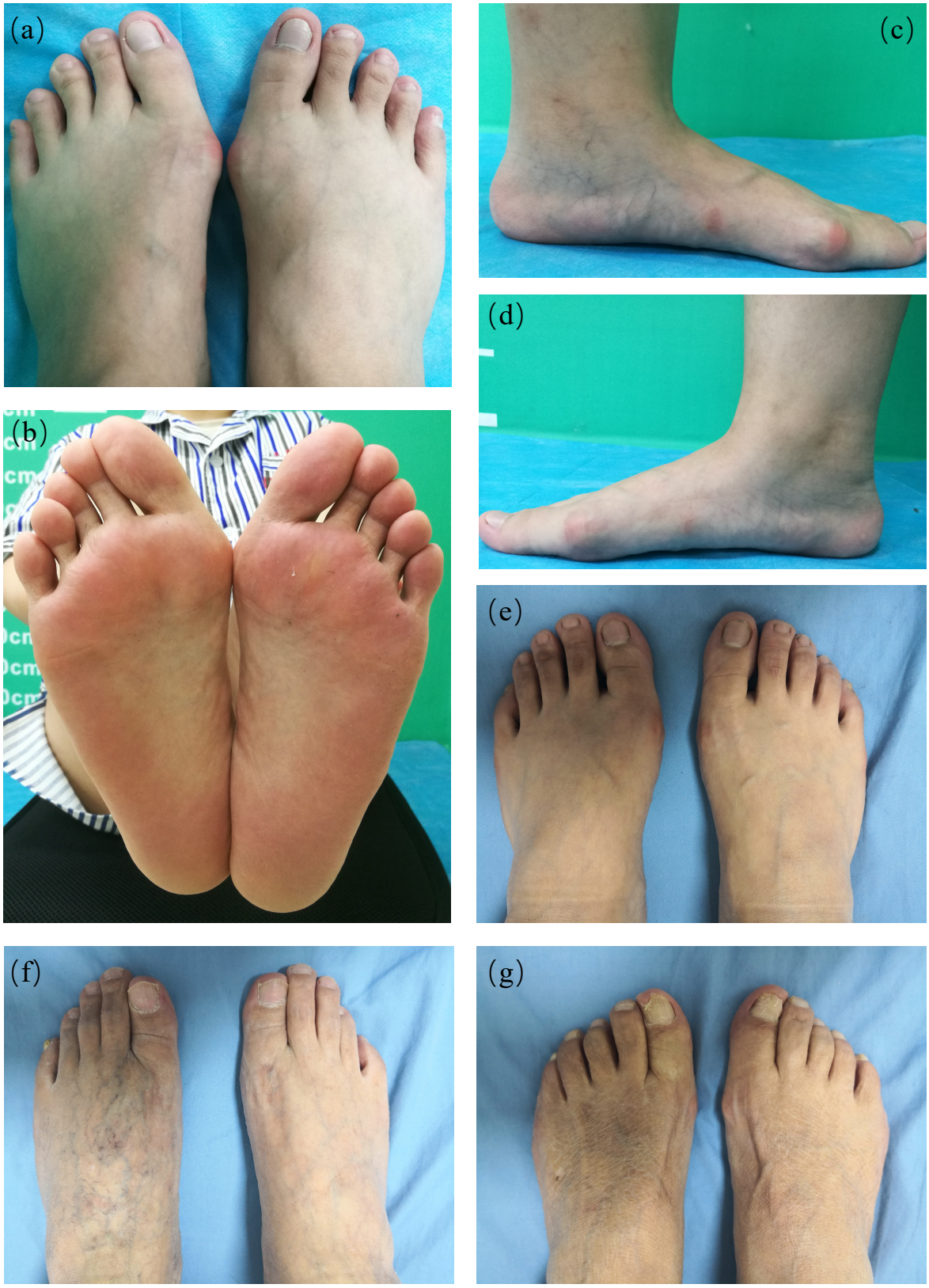
Table S2: the sequencing coverage quality parameters of each sample.

Supplementary Figure 1. The pipeline of whole exome sequencing data analyses.



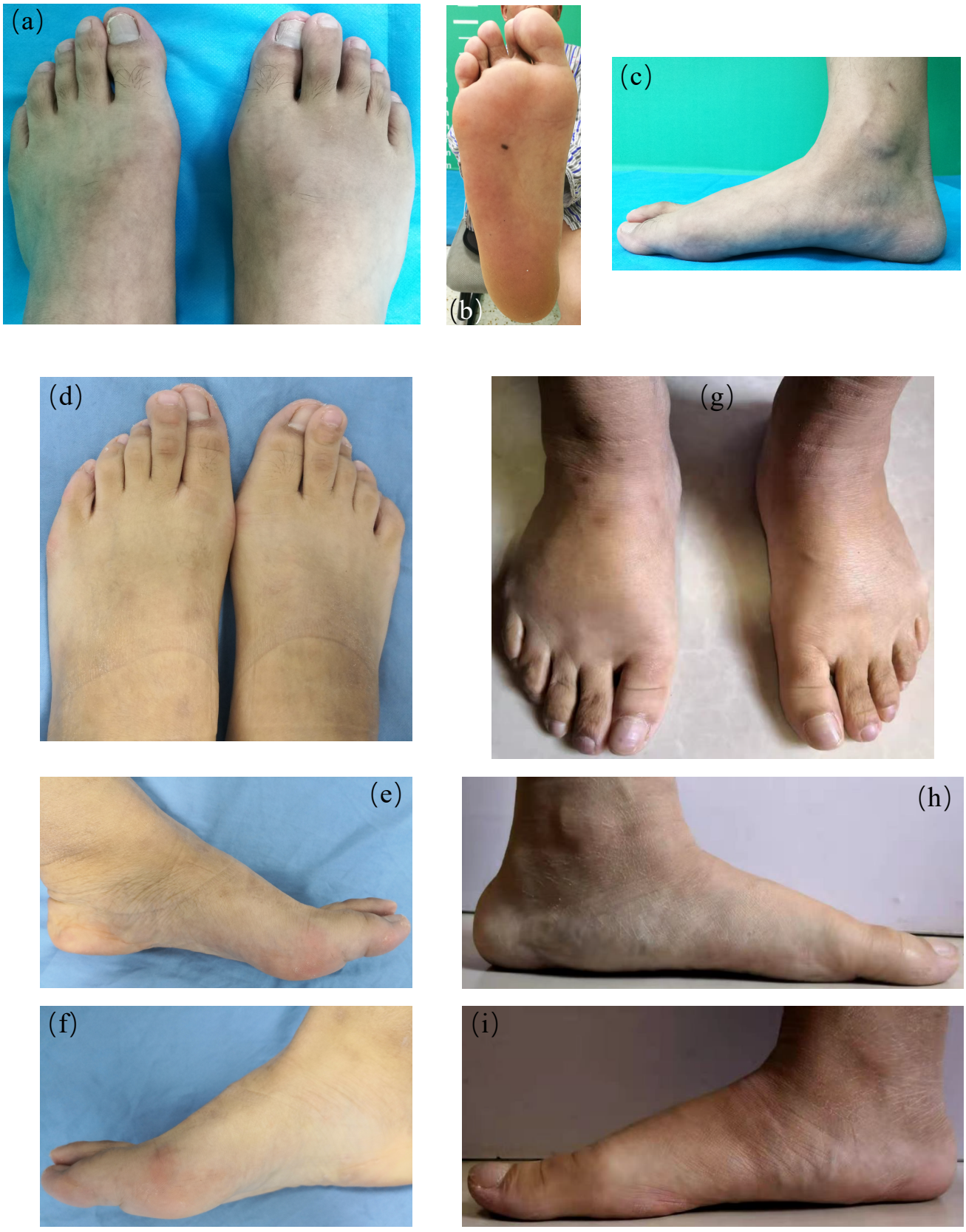
This diagram shows the general processing flow of the whole exome sequencing data of hallux valgus families. The main purpose of each step is described in the rounded rectangle, and the format of the data generated in the corresponding step is listed in the parenthesis. The software or tool used in each step is marked next to the single-line arrow.

Supplementary Figure 2. The appearance of the feet of family members of family 1.



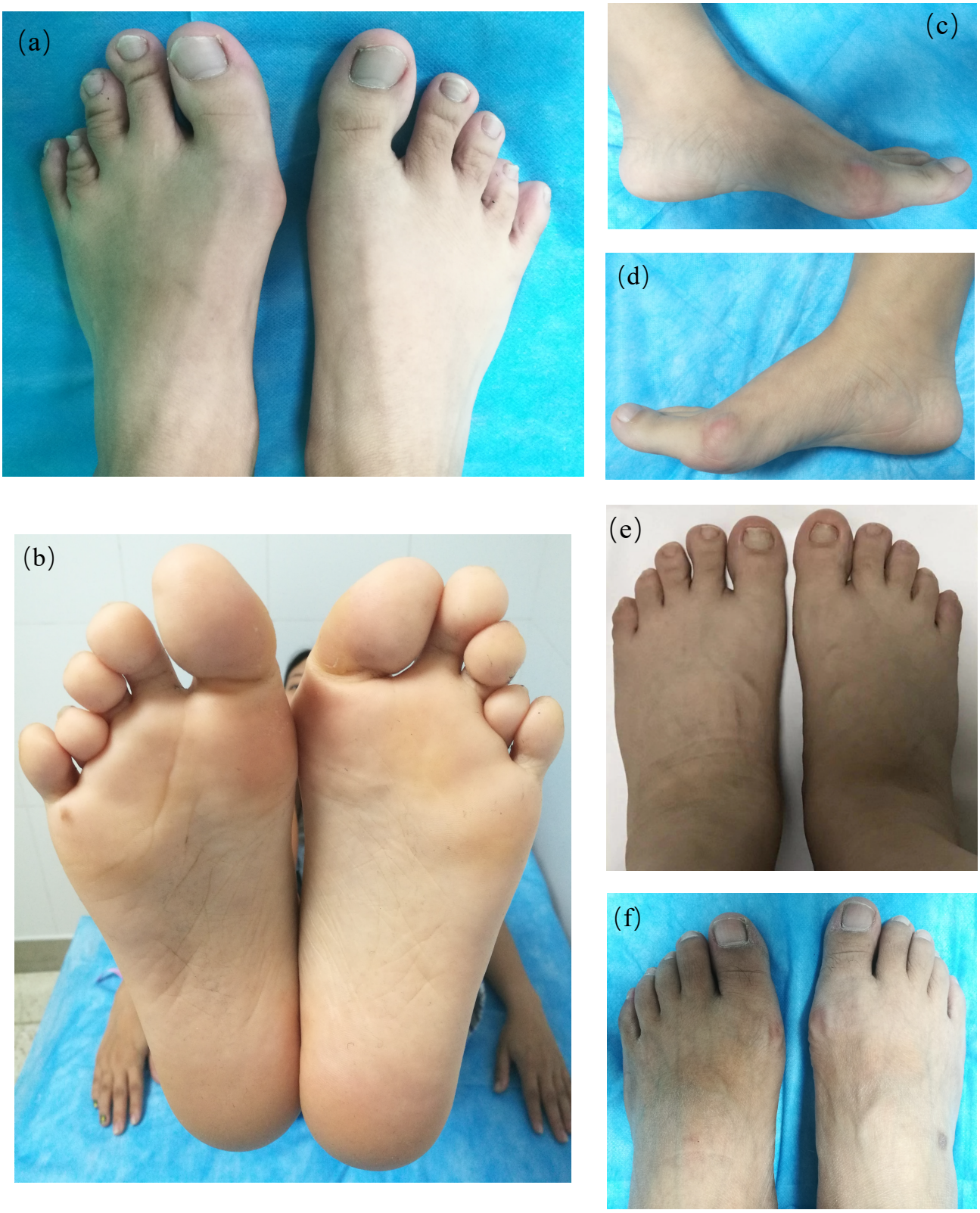
(a-d) II-1 (proband); **(e)** II-2 (proband's sister); **(f)** I-1 (proband's father); **(g)** I-2 (proband's mother).

Supplementary Figure 3. The appearance of the feet of family members of family 2.



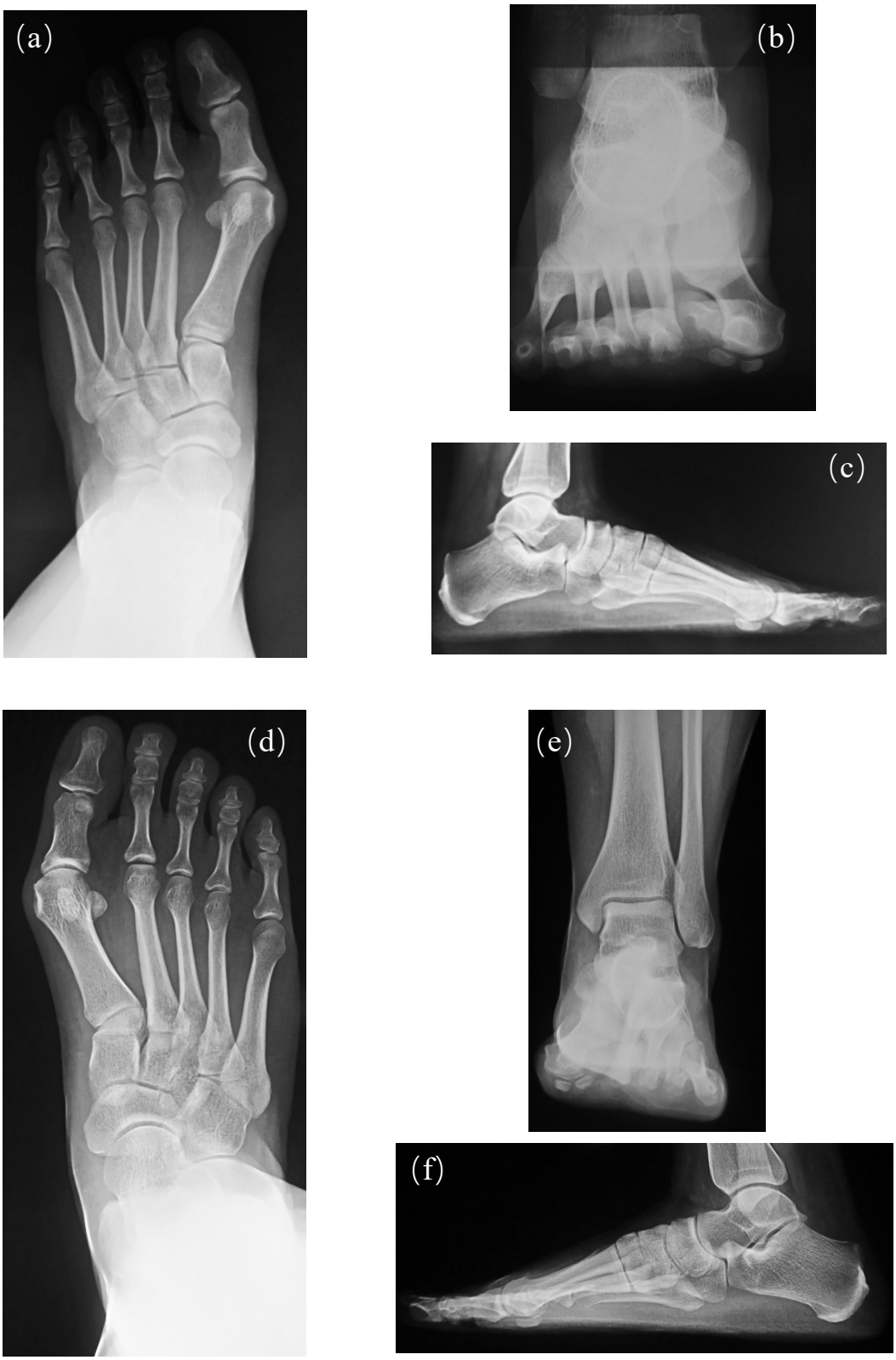
(a-c) I-2 (proband); **(d-f)** II-1 (proband's daughter); **(g-i)** I-1 (proband's wife).

Supplementary Figure 4. The appearance of the feet of family members of family 3.



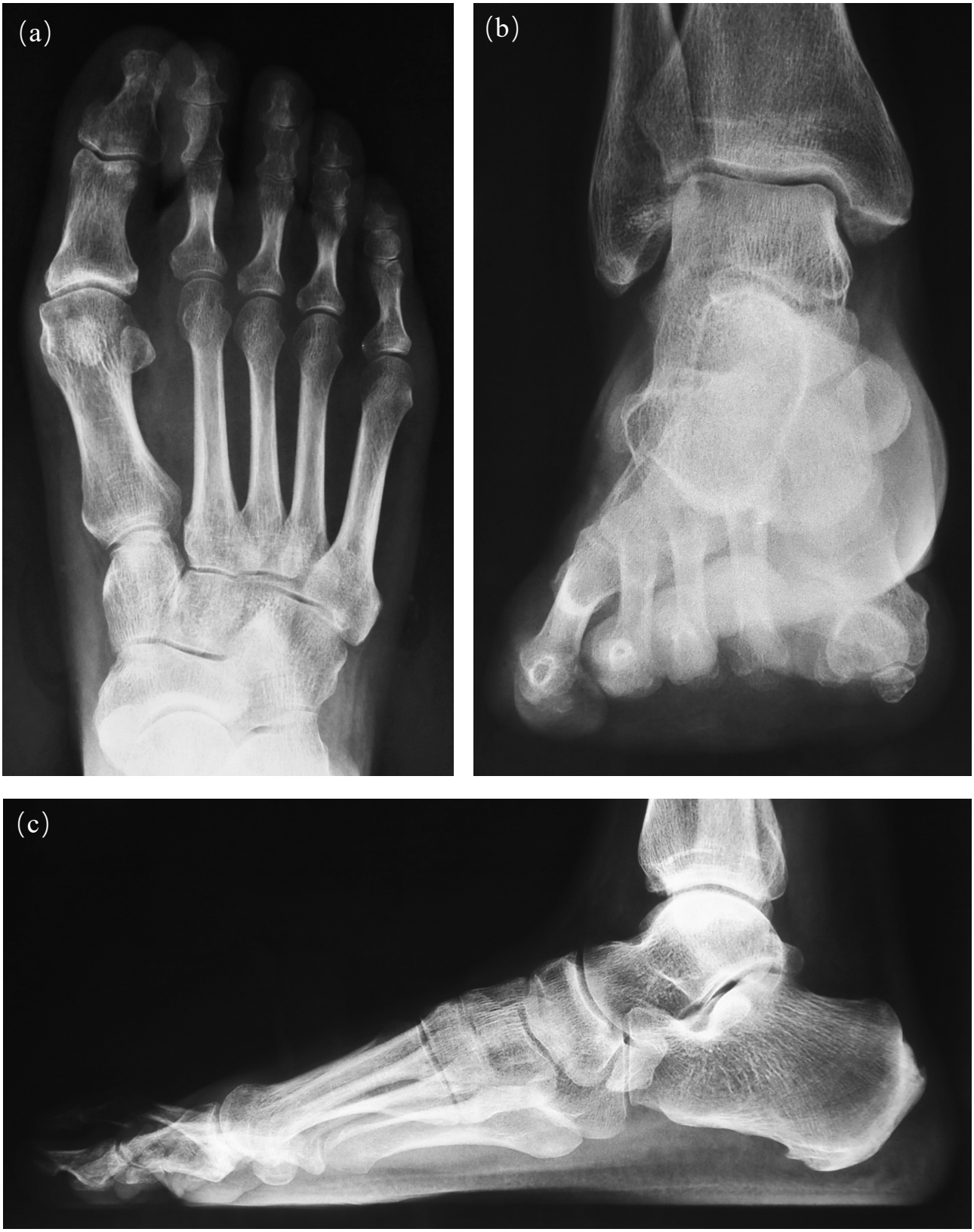
(a-d) II-1 (proband); **(e)** I-1 (proband's mother); **(f)** I-2 (proband's father).

Supplementary Figure 5. Digital radiography of proband (II-1) in family 1.



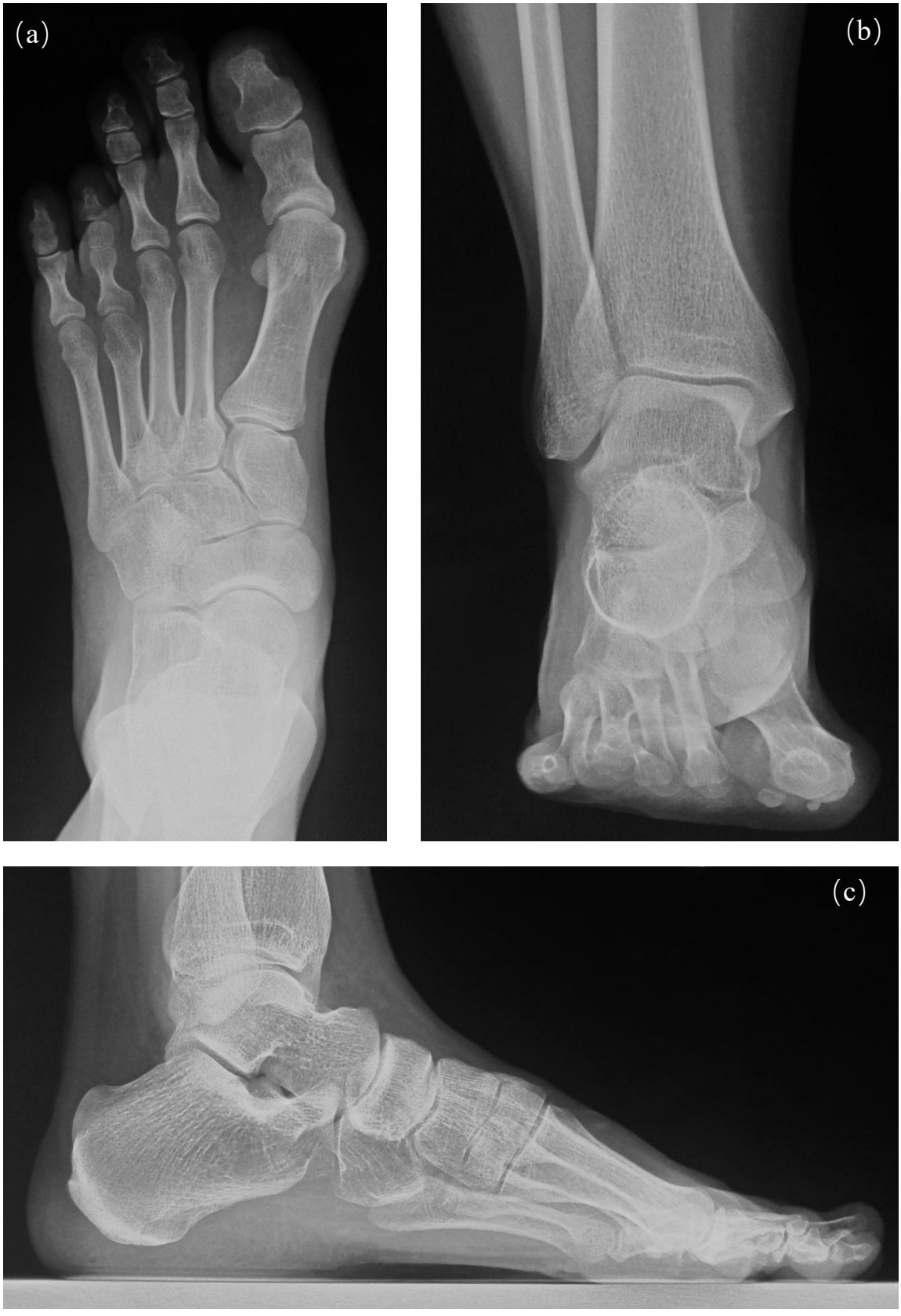
(a-c) frontal, axial, and lateral films of the left foot;
(d-f) frontal, axial, and lateral films of the right foot.

Supplementary Figure 6. Digital radiography of proband (I-2) in family 2.



(a-c) frontal, axial, and lateral films of the right foot.

Supplementary Figure 7. Digital radiography of proband (II-1) in family 3.



(a-c) frontal, axial, and lateral films of the left foot.

Supplementary Table 1. PCR primers for Sanger sequencing.

Genomic coordinates	Product length	Forward primer 5'→3'	Reverse primer 5'→3'
Chr5:40958328	389bp	ACCTGAAATGGGCTCTTGAA	GCACAGTGGGGCTGTTACTT
Chr2:179482520	397bp	AGGCACAGTCTGGGTTTCTAA	TCTCACTGACATGGGAACCA
Chr2:238258801	386bp	ACCTCTCTACCCCTGCCAAT	TTCCACCGAAATCTCAAAGG
Chr5:145499963	359bp	GCCACAGAGTCCCAAAAAGA	TTCCTCCTCCCTCAGCAGTA
Chr6:110062680	463bp	GCTGCACTGCAATTCTGAAG	TGGATTTTGGATGGGAAGAG
Chr21:44486410	464bp	GGAAGCTAGGTTGGGACACA	GTGTAGGGGAGGTGAGTGGA

Supplementary Table 2. The sequencing coverage quality parameters of each sample.

Sample Name	% GC	Ins.Size	>= 1X	>= 5X	>= 10X	>= 30X	>= 50X	Median cov	Mean cov	% Aligned	M Aligned	M Total reads	% Error
F1-I-1	49%	170	99.8%	99.7%	99.4%	97.0%	91.8%	138X	42.9X	99.9%	109.7	109.9	0.26%
F1-I-2	49%	173	99.6%	99.3%	98.6%	91.4%	77.2%	82X	25.6X	99.9%	66.3	66.4	0.26%
F1-II-1	49%	180	99.6%	99.5%	99.4%	98.1%	95.2%	171X	53.9X	99.9%	133.1	133.3	0.26%
F1-II-2	49%	176	99.6%	99.5%	99.4%	98.3%	96.1%	190X	59.9X	99.9%	150.7	150.9	0.25%
F2-I-1	49%	185	99.6%	99.4%	99.1%	95.8%	88.3%	113X	35.4X	99.9%	88.2	88.4	0.25%
F2-I-2	50%	192	99.9%	99.7%	99.5%	97.7%	93.6%	150X	47.2X	99.9%	118.0	118.1	0.27%
F2-II-1	49%	186	99.7%	99.5%	99.2%	96.9%	91.7%	134X	42.0X	99.9%	107.8	107.9	0.24%
F3-I-1	50%	174	99.6%	99.5%	99.3%	97.6%	94.0%	160X	50.1X	99.8%	126.3	126.5	0.26%
F3-I-2	45%	171	99.8%	99.5%	98.5%	86.4%	66.6%	67X	20.4X	100.0%	69.4	69.4	0.30%
F3-II-1	50%	190	99.6%	99.5%	99.4%	98.4%	96.1%	189X	59.9X	99.9%	145.8	146.0	0.27%
Average	**48.9%**	**179.7**	**99.68%**	**99.51%**	**99.18%**	**95.76%**	**89.06%**	**139.4X**	**43.73X**	**99.9%**	**111.53**	**111.68**	**0.26%**

%GC = mean GC content; Ins.Size = median insert size; >=nX = fraction of genome with at least nX coverage; Median cov = median coverage; Mean cov = mean coverage; %Aligned = % mapped reads; M Aligned = number of mapped reads (millions); M Total reads = number of reads (millions); Error rate = alignment error rate (i.e. total edit distance over the number of mapped bases).