Figure 1A, left panel: One-way ANOVA: Femur Cortical Thickness = Genotype : Means, SD, n, paired differences and p-values

For all results in Figure 1, a One-Way ANOVA model was fitted to the 4 genotypes. Normality and Equal Variance assumptions of the residuals was tested and passed in each case. Pairwise comparisons were adjusted by Tukey procedure and the 'Difference of Differences' contrast was also estimated.

For femur cortical thickness in Figure 1A, left panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$. Tnfrsf11b$^{-/-}$ has significantly lower values on average than CD19-Cre(p=0.0003) & CD19-Cre;Tnfrsf11b$^{f/f}$(p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p<0.0001).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0150*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.7644	.	.
Test	Equal variance test	.	.	.		.	0.2453	.	.
Mean	A)Tnfrsf11b$^{+/+}$	6	0.13	0.01		.	.	.	.
Mean	B)Tnfrsf11b$^{-/-}$	7	0.09	0.01		.	.	.	.
Mean	C)CD19-Cre	8	0.12	0.01		.	.	.	.
Mean	D)CD19-Cre;Tnfrsf11b$^{f/f}$	8	0.12	0.01		.	.	.	.
Diff	A)Tnfrsf11b$^{+/+}$	.	.	.	B)Tnfrsf11b$^{-/-}$	0.03	<.0001*	<.0001*	0.0354*
Diff	A)Tnfrsf11b$^{+/+}$	.	.	.	C)CD19-Cre	0.01	0.1566	0.4753	0.4449
Diff	A)Tnfrsf11b$^{+/+}$	.	.	.	D)CD19-Cre;Tnfrsf11b$^{f/f}$	0.00	0.4009	0.8278	0.8285
Diff	B)Tnfrsf11b$^{-/-}$	.	.	.	C)CD19-Cre	-0.02	<.0001*	0.0003*	0.0354*
Diff	B)Tnfrsf11b$^{-/-}$	.	.	.	D)CD19-Cre;Tnfrsf11b$^{f/f}$	-0.03	<.0001*	<.0001*	0.0354*
Diff	C)CD19-Cre	.	.	.	D)CD19-Cre;Tnfrsf11b$^{f/f}$	-0.00	0.5188	0.9130	0.8804
Contrast	(1-2)-(3-4)	.	0.04	0.01		.	<.0001*	.	.

Figure 1A, right panel: One-way ANOVA: L4 Cortical Thickness = Genotype : Means, SD, n, paired differences and p-values

For L4 cortical thickness in Figure 1A, right panel, by the Tukey p-values, $Tnfrsf11b^{+/+}$ has significantly higher values on average than $Tnfrsf11b^{-/-}$(p<0.0001). $Tnfrsf11b^{-/-}$ has significantly lower values on average than CD19-Cre(p<0.0001) & CD19-Cre;$Tnfrsf11b^{f/f}$(p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p<0.0001).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0032*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.5427	.	.
Test	Equal variance test	.	.	.		.	0.5244	.	.
Mean	A)$Tnfrsf11b^{+/+}$	6	0.06	0.00		.	.	.	.
Mean	B)$Tnfrsf11b^{-/-}$	7	0.04	0.01		.	.	.	.
Mean	C)CD19-Cre	8	0.05	0.00		.	.	.	.
Mean	D)CD19-Cre;$Tnfrsf11b^{f/f}$	8	0.06	0.00		.	.	.	.
Diff	A)$Tnfrsf11b^{+/+}$	.	.	.	B)$Tnfrsf11b^{-/-}$	0.02	<.0001*	<.0001*	0.0393*
Diff	A)$Tnfrsf11b^{+/+}$	.	.	.	C)CD19-Cre	0.00	0.2021	0.5654	0.5425
Diff	A)$Tnfrsf11b^{+/+}$	.	.	.	D)CD19-Cre;$Tnfrsf11b^{f/f}$	0.00	0.7558	0.9890	0.9974
Diff	B)$Tnfrsf11b^{-/-}$	.	.	.	C)CD19-Cre	-0.01	<.0001*	<.0001*	0.0393*
Diff	B)$Tnfrsf11b^{-/-}$	.	.	.	D)CD19-Cre;$Tnfrsf11b^{f/f}$	-0.01	<.0001*	<.0001*	0.0393*
Diff	C)CD19-Cre	.	.	.	D)CD19-Cre;$Tnfrsf11b^{f/f}$	-0.00	0.2925	0.7074	0.8988
Contrast	(1-2)-(3-4)	.	0.02	0.00		.	<.0001*	.	.

Figure 1B, left panel: One-way ANOVA: Femur Cancellous BVTV = Genotype : Means, SD, n, paired differences and p-values

For femur cancellous BV/TV in Figure 1B, left panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$(p<0.0001). Tnfrsf11b$^{-/-}$ has significantly lower values on average than CD19-Cre(p<0.0001) & CD19-Cre;Tnfrsf11b$^{f/f}$(p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p<0.0001).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0785	.	.
Test	S-W Normality, residuals	.	.	.		.	0.2793	.	.
Test	Equal variance test	.	.	.		.	0.1724	.	.
Mean	A)Tnfrsf11b$^{+/+}$	6	0.04	0.01		.	.	.	.
Mean	B)Tnfrsf11b$^{-/-}$	7	0.00	0.00		.	.	.	.
Mean	C)CD19-Cre	8	0.03	0.01		.	.	.	.
Mean	D)CD19-Cre;Tnfrsf11b$^{f/f}$	8	0.04	0.01		.	.	.	.
Diff	A)Tnfrsf11b$^{+/+}$	.	.	.	B)Tnfrsf11b$^{-/-}$	0.04	<.0001*	<.0001*	0.0287*
Diff	A)Tnfrsf11b$^{+/+}$	.	.	.	C)CD19-Cre	0.01	0.0453*	0.1783	0.5850
Diff	A)Tnfrsf11b$^{+/+}$	.	.	.	D)CD19-Cre;Tnfrsf11b$^{f/f}$	0.00	0.5562	0.9322	0.8660
Diff	B)Tnfrsf11b$^{-/-}$	.	.	.	C)CD19-Cre	-0.03	<.0001*	0.0001*	0.0197*
Diff	B)Tnfrsf11b$^{-/-}$	.	.	.	D)CD19-Cre;Tnfrsf11b$^{f/f}$	-0.03	<.0001*	<.0001*	0.0197*
Diff	C)CD19-Cre	.	.	.	D)CD19-Cre;Tnfrsf11b$^{f/f}$	-0.01	0.1152	0.3797	0.5850
Contrast	(1-2)-(3-4)	.	0.05	0.01		.	<.0001*	.	.

Figure 1B, right panel: One-way ANOVA: L4 Cancellous BV/TV = Genotype : Means, SD, n, paired differences and p-values

For L4 cancellous BV/TV in Figure 1B, right panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$(p<0.0001), CD19-Cre(p-0.0124), & CD19-Cre;Tnfrsf11b$^{f/f}$(p=0.0465). Tnfrsf11b$^{-/-}$ has significantly lower values on average than CD19-Cre(p<0.0001) & CD19-Cre;Tnfrsf11b$^{f/f}$(p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p<0.0001).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0100*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.0712	.	.
Test	Equal variance test	.	.	.		.	0.1173	.	.
Mean	A)Tnfrsf11b$^{+/+}$	6	0.17	0.02		.	.	.	.
Mean	B)Tnfrsf11b$^{-/-}$	7	0.06	0.01		.	.	.	.
Mean	C)CD19-Cre	8	0.14	0.01		.	.	.	.
Mean	D)CD19-Cre;Tnfrsf11b$^{f/f}$	8	0.15	0.02		.	.	.	.
Diff	A)Tnfrsf11b$^{+/+}$	.	.	.	B)Tnfrsf11b$^{-/-}$	0.11	<.0001*	<.0001*	0.0287*
Diff	A)Tnfrsf11b$^{+/+}$	.	.	.	C)CD19-Cre	0.03	0.0025*	0.0124*	0.2467
Diff	A)Tnfrsf11b$^{+/+}$	.	.	.	D)CD19-Cre;Tnfrsf11b$^{f/f}$	0.03	0.0101*	0.0465*	0.3237
Diff	B)Tnfrsf11b$^{-/-}$	.	.	.	C)CD19-Cre	-0.08	<.0001*	<.0001*	0.0197*
Diff	B)Tnfrsf11b$^{-/-}$	.	.	.	D)CD19-Cre;Tnfrsf11b$^{f/f}$	-0.09	<.0001*	<.0001*	0.0197*
Diff	C)CD19-Cre	.	.	.	D)CD19-Cre;Tnfrsf11b$^{f/f}$	-0.01	0.5369	0.9227	0.8353
Contrast	(1-2)-(3-4)	.	0.12	0.01		.	<.0001*	.	.

Figure 2A, left panel: One-way ANOVA: Femur Cortical Thickness= Genotype : Means, SD, n, paired differences and p-values

For all results in Figure 2, a One-Way ANOVA model was fitted to the 4 genotypes. Normality and Equal Variance assumptions of the residuals was tested. Transformations were used when those assumptions failed. Pairwise comparisons were adjusted by Tukey procedure and the 'Difference of Differences' contrast was also estimated.

For Rank-Transformed femur cortical thickness in Figure 2A, left panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0151). Tnfrsf11b$^{-/-}$ has significantly lower values on average than Tnfrsf11b$^{f/f}$(p=0.0002). Tnfrsf11b$^{f/f}$ has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0263).

The (1-2)-(3-4) difference of differences contrast is not statistically significant(p=0.3527).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.5094	.	.
Test	S-W Normality, residuals	.	.	.		.	0.8187	.	.
Test	Equal variance test	.	.	.		.	0.5051	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.13	0.01		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	7	0.09	0.01		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	5	0.12	0.01		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	6	0.10	0.01		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.03	0.0001*	0.0006*	0.0731
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	0.00	0.9219	0.9996	1.0000
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.02	0.0031*	0.0151*	0.0731
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.03	0.0002*	0.0013*	0.0731
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-0.01	0.2089	0.5745	0.8993
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.02	0.0056*	0.0263*	0.1217
Contrast	(1-2)-(3-4)	.	0.01	0.01		.	0.3527	.	.

Figure 2A, right panel: One-way ANOVA: L4 Cortical Thickness = Genotype : Means, SD, n, paired differences and p-values

For L4 cortical thickness in Figure 2A, right panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$(p<0.0001) and Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0006). Tnfrsf11b$^{-/-}$ has significantly lower values on average than Tnfrsf11b$^{f/f}$(p<0.0001). Tnfrsf11b$^{f/f}$ has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0025).

The (1-2)-(3-4) difference of differences contrast is not statistically significant(p=0.3242).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0214*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.5163	.	.
Test	Equal variance test	.	.	.		.	0.3742	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.06	0.00		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	7	0.04	0.01		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	5	0.06	0.00		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	6	0.05	0.01		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.02	<.0001*	<.0001*	0.0546
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	0.00	0.7207	0.9832	0.9433
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.01	0.0001*	0.0006*	0.0546
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.01	<.0001*	0.0002*	0.0548
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-0.00	0.2847	0.6940	0.9433
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.01	0.0005*	0.0025*	0.0548
Contrast	(1-2)-(3-4)	.	0.00	0.00		.	0.3242	.	.

Figure 2B, left panel: One-way ANOVA: Square-root-Transformed Femur Cancellous BV/TV = Genotype : Means, SD, n, paired differences and p-values

For Square-root-Transformed femur cancellous BV/TV in Figure 2B, left panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$(p<0.0001) and Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0043). Tnfrsf11b$^{-/-}$ has significantly lower values on average than Tnfrsf11b$^{f/f}$(p<0.0001) & Dmp1-Cre;Tnfrsf11b$^{f/f}$(p<0.0001). Tnfrsf11b$^{f/f}$ has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0027).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p=0.0006).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0951	.	.
Test	S-W Normality, residuals	.	.	.		.	0.8382	.	.
Test	Equal variance test	.	.	.		.	0.2617	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.20	0.04		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	7	0.04	0.02		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	5	0.20	0.03		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	6	0.13	0.02		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.16	<.0001*	<.0001*	0.0431*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.01	0.6992	0.9790	0.9472
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.06	0.0008*	0.0043*	0.0920
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.17	<.0001*	<.0001*	0.0465*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-0.10	<.0001*	<.0001*	0.0431*
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.07	0.0005*	0.0027*	0.0470*
Contrast	(1-2)-(3-4)	.	0.09	0.02		.	0.0006*	.	.

Figure 2B, right panel: One-way ANOVA: L4 Cancellous BV/TV = Genotype : Means, SD, n, paired differences and p-values

For L4 cancellous BV/TV in Figure 2B, right panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$(p<0.0001) and Dmp1-Cre;Tnfrsf11b$^{f/f}$(p<0.0001). Tnfrsf11b$^{-/-}$ has significantly lower values on average than Tnfrsf11b$^{f/f}$(p<0.0001). Tnfrsf11b$^{f/f}$ has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p<0.0001).

The (1-2)-(3-4) difference of differences contrast is not statistically significant(p=0.0982).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0050*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.4319	.	.
Test	Equal variance test	.	.	.		.	0.1958	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.17	0.02		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	7	0.06	0.01		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	5	0.17	0.02		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	9	0.08	0.01		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.11	<.0001*	<.0001*	0.0287*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	0.00	0.9277	0.9997	0.9834
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.09	<.0001*	<.0001*	0.0287*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.11	<.0001*	<.0001*	0.0349*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-0.02	0.0166*	0.0731	0.0742
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.09	<.0001*	<.0001*	0.0287*
Contrast	(1-2)-(3-4)	.	0.02	0.01		.	0.0982	.	.

Figure 2E: One-way ANOVA: Square-root-Transformed Cortical Porosity = Genotype : Means, SD, n, paired differences and p-values

For Square-root-Transformed cortical porosity in Figure 2E, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly lower values on average than Tnfrsf11b$^{-/-}$(p<0.0001). Tnfrsf11b$^{-/-}$ has significantly higher values on average than Tnfrsf11b$^{f/f}$(p<0.0001) & Dmp1-Cre;Tnfrsf11b$^{f/f}$(p<0.0001). Tnfrsf11b$^{f/f}$ has significantly lower values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$ (p=0.0251).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p<0.0001).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0006*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.3277	.	.
Test	Equal variance test	.	.	.		.	0.2502	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.47	0.44		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	7	4.14	0.72		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	5	0.27	0.25		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	6	1.22	0.41		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	-3.66	<.0001*	<.0001*	0.0431*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	0.21	0.5043	0.9034	0.8850
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-0.75	0.0184*	0.0795	0.0920
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	3.87	<.0001*	<.0001*	0.0465*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	2.91	<.0001*	<.0001*	0.0431*
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-0.96	0.0053*	0.0251*	0.0470*
Contrast	(1-2)-(3-4)	.	-2.71	0.42		.	<.0001*	.	.

Figure 2F, left panel: One-way ANOVA: Rank-Transformed Serum OPG = Genotype : Means, SD, n, paired differences and p-values

For Rank-Transformed serum OPG in Figure 2F, left panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$(p<0.0001). Tnfrsf11b$^{-/-}$ has significantly lower values on average than Tnfrsf11b$^{f/f}$(p=0.0002) and Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0008).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p=0.0097).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0293*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.2990	.	.
Test	Equal variance test	.	.	.		.	0.0453*	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	19.83	7.33		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	7	4.00	0.00		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	5	20.00	3.67		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	10	15.90	6.35		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	15.83	<.0001*	0.0001*	0.0166*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.17	0.9594	0.9999	0.9978
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	3.93	0.1670	0.4967	0.8387
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-16.00	<.0001*	0.0002*	0.0166*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-11.90	0.0001*	0.0008*	0.0136*
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	4.10	0.1742	0.5113	0.7991
Contrast	(1-2)-(3-4)	.	11.73	4.17		.	0.0097*	.	.

Figure 2F, right panel: One-way ANOVA: Square-root-Transformed Serum RANKL = Genotype : Means, SD, n, paired differences and p-values

For Square-root-Transformed serum RANKL in Figure 2F, right panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly lower values on average than Tnfrsf11b$^{-/-}$ (p<0.0001) and Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0317). Tnfrsf11b$^{-/-}$ has significantly higher values on average than Tnfrsf11b$^{f/f}$(p<0.0001) & Dmp1-Cre;Tnfrsf11b$^{f/f}$(p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p<0.0001).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0001*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.1181	.	.
Test	Equal variance test	.	.	.		.	0.2137	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	2.94	4.60		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	7	48.14	2.58		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	5	9.40	1.10		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	10	9.07	5.10		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	-45.20	<.0001*	<.0001*	0.0377*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-6.46	0.0135*	0.0609	0.2015
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-6.14	0.0067*	0.0317*	0.1341
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	38.74	<.0001*	<.0001*	0.0465*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	39.06	<.0001*	<.0001*	0.0212*
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.32	0.8834	0.9988	0.6875
Contrast	(1-2)-(3-4)	.	-45.53	3.13		.	<.0001*	.	.

Figure 2G, panel 1: One-way ANOVA: Rank-Transformed Tnfrsf11b Gene Expression = Genotype : Means, SD, n, paired differences and p-values

For all results in Figure 2G, a One-Way ANOVA model was fitted to the 4 genotypes. Normality and Equal Variance assumptions of the residuals was tested. Transformations were used when those assumptions failed. Pairwise comparisons were adjusted by Tukey procedure and the 'Difference of Differences' contrast was also estimated.

For Rank-Transformed *Tnfrsf11b* gene expression in Figure 2G, panel 1, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$ (p<0.0001) and Dmp1-Cre;Tnfrsf11b$^{f/f}$ (p=0.0006). Tnfrsf11b$^{-/-}$ has significantly lower values on average than Tnfrsf11b$^{f/f}$(p<0.0001) & Dmp1-Cre;Tnfrsf11b$^{f/f}$ (p=0.0001). Tnfrsf11b$^{f/f}$ has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$ (p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p=0.0230).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.2492	.	.
Test	S-W Normality, residuals	.	.	.		.	0.7008	.	.
Test	Equal variance test	.	.	.		.	0.2859	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	19.67	3.39		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	6	3.50	1.76		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	4	21.75	2.22		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	9	11.00	2.74		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	16.17	<.0001*	<.0001*	0.0388*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-2.08	0.2369	0.6230	0.7103
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	8.67	<.0001*	<.0001*	0.0239*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-18.25	<.0001*	<.0001*	0.0567
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-7.50	<.0001*	0.0001*	0.0239*
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	10.75	<.0001*	<.0001*	0.0421*
Contrast	(1-2)-(3-4)	.	5.42	2.21		.	0.0230*	.	.

Figure 2G, panel 2: One-way ANOVA: Tnfsf11 Gene Expression = Genotype : Means, SD, n, paired differences and p-values

For *Tnfsf11* gene expression in Figure 2G, panel 2, by the Tukey p-values, Tnfrsf11b$^{f/f}$ has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$ (p=0.0273).

The (1-2)-(3-4) difference of differences contrast is not statistically significant(p=0.2810).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0320*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.0730	.	.
Test	Equal variance test	.	.	.		.	0.1136	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	1.37	0.49		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	6	1.04	0.26		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	4	0.95	0.11		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	9	0.89	0.18		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.33	0.0693	0.2525	0.6273
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	0.42	0.0371*	0.1488	0.6273
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.48	0.0058*	0.0273*	0.6273
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	0.10	0.6127	0.9549	0.8290
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.15	0.3399	0.7641	0.6273
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.05	0.7638	0.9899	0.8290
Contrast	(1-2)-(3-4)	.	0.27	0.25		.	0.2810	.	.

Figure 2G, panel 3: One-way ANOVA: TRAP Gene Expression = Genotype : Means, SD, n, paired differences and p-values

For *TRAP* gene expression in Figure 2G, panel 3, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly lower values on average than Tnfrsf11b$^{-/-}$(p<0.0001) and Dmp1-Cre;Tnfrsf11b$^{f/f}$ (p=0.0006). Tnfrsf11b$^{-/-}$ has significantly higher values on average than Tnfrsf11b$^{f/f}$(p<0.0001) & Dmp1-Cre;Tnfrsf11b$^{f/f}$ (p<0.0001). Tnfrsf11b$^{f/f}$ has significantly lower values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$ (p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p=0.0016).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0413*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.4045	.	.
Test	Equal variance test	.	.	.		.	0.5304	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	129.67	23.39		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	6	326.76	31.64		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	4	102.00	13.11		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	9	206.50	37.75		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	-197.09	<.0001*	<.0001*	0.0412*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	27.67	0.1759	0.5126	0.3207
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-76.84	0.0001*	0.0006*	0.0352*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	224.76	<.0001*	<.0001*	0.0617
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	120.25	<.0001*	<.0001*	0.0352*
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-104.51	<.0001*	<.0001*	0.0421*
Contrast	(1-2)-(3-4)	.	-92.59	25.50		.	0.0016*	.	.

Figure 2G, panel 4: One-way ANOVA: CTSK Gene Expression = Genotype : Means, SD, n, paired differences and p-values

For *CTSK* gene expression in Figure 2G, panel 4, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly lower values on average than Tnfrsf11b$^{-/-}$(p<0.0001) and Dmp1-Cre;Tnfrsf11b$^{f/f}$ (p=0.0006). Tnfrsf11b$^{-/-}$ has significantly higher values on average than Tnfrsf11b$^{f/f}$(p<0.0001) & Dmp1-Cre;Tnfrsf11b$^{f/f}$ (p<0.0001). Tnfrsf11b$^{f/f}$ has significantly lower values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$ (p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p=0.0002).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0440*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.5932	.	.
Test	Equal variance test	.	.	.		.	0.0811	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	122.88	30.31		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	6	335.79	38.85		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	4	104.04	8.31		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	9	208.44	23.93		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	-212.91	<.0001*	<.0001*	0.0412*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	18.83	0.3158	0.7354	0.8290
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-85.56	<.0001*	<.0001*	0.0239*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	231.75	<.0001*	<.0001*	0.0617
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	127.35	<.0001*	<.0001*	0.0239*
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-104.40	<.0001*	<.0001*	0.0421*
Contrast	(1-2)-(3-4)	.	-108.52	23.66		.	0.0002*	.	.

Figure 3A, left panel: One-way ANOVA: Femur Cortical Thickness = Genotype : Means, SD, n, paired differences and p-values

For all results in Figure 3, a One-Way ANOVA model was fitted to the 4 genotypes. Normality and Equal Variance assumptions of the residuals was tested. Transformations were used when those assumptions failed. Pairwise comparisons were adjusted by Tukey procedure and the 'Difference of Differences' contrast was also estimated.

For femur cortical thickness in Figure 3A, left panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$(p<0.0001). Tnfrsf11b$^{-/-}$ has significantly lower values on average than Tnfrsf11b$^{f/f}$(p<0.0001) & SOST-Cre;Tnfrsf11b$^{f/f}$ (p=0.0017).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p=0.0179).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0374*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.1819	.	.
Test	Equal variance test	.	.	.		.	0.1673	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.13	0.01		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	7	0.09	0.01		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	6	0.13	0.00		.	.	.	.
Mean	4)Sost-Cre;Tnfrsf11b$^{f/f}$	6	0.12	0.01		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.03	<.0001*	<.0001*	0.0531
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.00	0.5745	0.9398	0.8141
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	0.01	0.1018	0.3432	0.2577
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.04	<.0001*	<.0001*	0.0531
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	-0.02	0.0003*	0.0017*	0.0883
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	0.01	0.0331*	0.1346	0.0801
Contrast	(1-2)-(3-4)	.	0.02	0.01		.	0.0179*	.	.

Figure 3A, right panel: One-way ANOVA: L4 Cortical Thickness = Genotype : Means, SD, n, paired differences and p-values

For L4 cortical thickness in Figure 3A, right panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$(p<0.0001). Tnfrsf11b$^{-/-}$ has significantly lower values on average than Tnfrsf11b$^{f/f}$(p<0.0001) & SOST-Cre;Tnfrsf11b$^{f/f}$ (p=0.0002).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p=0.0061).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0268*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.0657	.	.
Test	Equal variance test	.	.	.		.	0.5629	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.06	0.00		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	7	0.04	0.01		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	6	0.06	0.00		.	.	.	.
Mean	4)Sost-Cre;Tnfrsf11b$^{f/f}$	6	0.06	0.01		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.02	<.0001*	<.0001*	0.0498*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.00	0.4185	0.8419	0.8863
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	0.00	0.4547	0.8706	0.7281
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.02	<.0001*	<.0001*	0.0498*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	-0.01	<.0001*	0.0002*	0.0637
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	0.00	0.1275	0.4070	0.7281
Contrast	(1-2)-(3-4)	.	0.01	0.00		.	0.0061*	.	.

Figure 3B, left panel: One-way ANOVA: Square Root-Transformed Femur Cancellous BV/TV = Genotype : Means, SD, n, paired differences and p-values

For Square Root-Transformed femur cancellous BV/TV in Figure 3B, left panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$(p<0.0001). Tnfrsf11b$^{-/-}$ has significantly lower values on average than Tnfrsf11b$^{f/f}$(p<0.0001) & SOST-Cre;Tnfrsf11b$^{f/f}$ (p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p<0.0001).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0004*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.7609	.	.
Test	Equal variance test	.	.	.		.	0.1448	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.20	0.04		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	7	0.04	0.02		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	6	0.21	0.03		.	.	.	.
Mean	4)Sost-Cre;Tnfrsf11b$^{f/f}$	6	0.20	0.01		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.16	<.0001*	<.0001*	0.0287*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.01	0.5140	0.9094	0.9886
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	-0.00	0.7624	0.9897	0.9886
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.17	<.0001*	<.0001*	0.0287*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	-0.17	<.0001*	<.0001*	0.0287*
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	0.01	0.7242	0.9839	0.9886
Contrast	(1-2)-(3-4)	.	0.15	0.02		.	<.0001*	.	.

Figure 3B, right panel: One-way ANOVA: L4 Cancellous BV/TV = Genotype : Means, SD, n, paired differences and p-values

For L4 cancellous BV/TV in Figure 3B, right panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$(p<0.0001) and SOST-Cre;Tnfrsf11b$^{f/f}$ (p=0.0050). Tnfrsf11b$^{-/-}$ has significantly lower values on average than Tnfrsf11b$^{f/f}$(p<0.0001) & SOST-Cre;Tnfrsf11b$^{f/f}$ (p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p<0.0001).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0118*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.3803	.	.
Test	Equal variance test	.	.	.		.	0.1129	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.17	0.02		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	7	0.06	0.01		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	6	0.16	0.01		.	.	.	.
Mean	4)Sost-Cre;Tnfrsf11b$^{f/f}$	6	0.14	0.01		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.11	<.0001*	<.0001*	0.0287*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	0.02	0.0651	0.2397	0.5224
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	0.04	0.0010*	0.0050*	0.1682
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.10	<.0001*	<.0001*	0.0287*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	-0.08	<.0001*	<.0001*	0.0287*
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	0.02	0.0739	0.2660	0.1909
Contrast	(1-2)-(3-4)	.	0.10	0.01		.	<.0001*	.	.

Figure 3E: One-way ANOVA: Square Root-Transformed Cortical Porosity = Genotype : Means, SD, n, paired differences and p-values

For Square Root-Transformed Cortical Porosity in Figure 3E, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly lower values on average than Tnfrsf11b$^{-/-}$($p<0.0001$). Tnfrsf11b$^{-/-}$ has significantly higher values on average than Tnfrsf11b$^{f/f}$($p<0.0001$) & SOST-Cre;Tnfrsf11b$^{f/f}$($p<0.0001$).

The (1-2)-(3-4) difference of differences contrast is also statistically significant($p<0.0001$).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0001*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.0779	.	.
Test	Equal variance test	.	.	.		.	0.1194	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.47	0.44		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	7	4.14	0.72		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	6	0.28	0.25		.	.	.	.
Mean	4)Sost-Cre;Tnfrsf11b$^{f/f}$	6	0.25	0.12		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	-3.66	<.0001*	<.0001*	0.0287*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	0.20	0.4624	0.8762	0.9886
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	0.22	0.4110	0.8354	0.9886
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	3.86	<.0001*	<.0001*	0.0287*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	3.89	<.0001*	<.0001*	0.0287*
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	0.02	0.9289	0.9997	0.9886
Contrast	(1-2)-(3-4)	.	-3.69	0.37		.	<.0001*	.	.

Figure 3F, panel 1: One-way ANOVA: Rank-transformed Tnfrsf11b Gene Expression = Genotype : Means, SD, n, paired differences and p-values

For Rank-transformed *Tnfrsf11b* gene expression in Figure 3F, panel 1, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$ (p<0.0001) and SOST-Cre;Tnfrsf11b$^{f/f}$ (p=0.0029). Tnfrsf11b$^{-/-}$ has significantly lower values on average than Tnfrsf11b$^{f/f}$(p<0.0001) & SOST-Cre;Tnfrsf11b$^{f/f}$ (p=0.0058).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p=0.044).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.2354	.	.
Test	S-W Normality, residuals	.	.	.		.	0.2037	.	.
Test	Equal variance test	.	.	.		.	0.4363	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	20.67	5.89		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	6	3.50	1.76		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	6	18.67	4.27		.	.	.	.
Mean	4)Sost-Cre;Tnfrsf11b$^{f/f}$	8	11.75	3.54		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	17.17	<.0001*	<.0001*	0.0388*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	2.00	0.4059	0.8313	0.6766
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	8.92	0.0005*	0.0029*	0.1112
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-15.17	<.0001*	<.0001*	0.0388*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	-8.25	0.0011*	0.0058*	0.0388*
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	6.92	0.0048*	0.0232*	0.0725
Contrast	(1-2)-(3-4)	.	10.25	3.23		.	0.0044*	.	.

Figure 3F, panel 2: One-way ANOVA: Tnfsf11 Gene Expression = Genotype : Means, SD, n, paired differences and p-values

For *Tnfsf11* gene expression in Figure 3F, panel 2, there are no statistically significant pairwise comparisons.

The (1-2)-(3-4) difference of differences contrast is not statistically significant(p=0.5579).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.7549	.	.
Test	S-W Normality, residuals	.	.	.		.	0.7996	.	.
Test	Equal variance test	.	.	.		.	0.1950	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	1.37	0.49		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	6	1.04	0.26		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	6	1.23	0.27		.	.	.	.
Mean	4)Sost-Cre;Tnfrsf11b$^{f/f}$	8	1.06	0.25		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.33	0.0965	0.3298	0.9940
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	0.13	0.4808	0.8892	0.9940
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	0.31	0.0938	0.3223	0.9940
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.19	0.3193	0.7403	0.9940
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	-0.02	0.9179	0.9996	0.9940
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	0.17	0.3354	0.7594	0.9940
Contrast	(1-2)-(3-4)	.	0.15	0.26		.	0.5579	.	.

Figure 3F, panel 3: One-way ANOVA: TRAP Gene Expression = Genotype : Means, SD, n, paired differences and p-values

For *TRAP* gene expression in Figure 3F, panel 3, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly lower values on average than Tnfrsf11b$^{-/-}$($p<0.0001$). Tnfrsf11b$^{-/-}$ has significantly higher values on average than Tnfrsf11b$^{f/f}$($p<0.0001$) & SOST-Cre;Tnfrsf11b$^{f/f}$ ($p<0.0001$).

The (1-2)-(3-4) difference of differences contrast is also statistically significant($p<0.0001$).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0001*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.6406	.	.
Test	Equal variance test	.	.	.		.	0.5028	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	129.67	23.39		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	6	326.76	31.64		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	6	102.47	18.28		.	.	.	.
Mean	4)Sost-Cre;Tnfrsf11b$^{f/f}$	8	121.23	22.63		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	-197.09	<.0001*	<.0001*	0.0412*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	27.20	0.0656	0.2417	0.2386
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	8.44	0.5270	0.9169	0.8660
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	224.29	<.0001*	<.0001*	0.0412*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	205.53	<.0001*	<.0001*	0.0412*
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	-18.76	0.1670	0.4954	0.4022
Contrast	(1-2)-(3-4)	.	-178.33	19.22		.	<.0001*	.	.

Figure 3F, panel 4: One-way ANOVA: CTSK Gene Expression = Genotype : Means, SD, n, paired differences and p-values

For *CTSK* gene expression in Figure 3F, panel 4, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly lower values on average than Tnfrsf11b$^{-/-}$(p<0.0001). Tnfrsf11b$^{-/-}$ has significantly higher values on average than Tnfrsf11b$^{f/f}$(p<0.0001) & SOST-Cre;Tnfrsf11b$^{f/f}$ (p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p<0.0001).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0001*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.7345	.	.
Test	Equal variance test	.	.	.		.	0.0584	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	122.88	30.31		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	6	335.79	38.85		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	6	101.03	13.58		.	.	.	.
Mean	4)Sost-Cre;Tnfrsf11b$^{f/f}$	8	126.14	22.48		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	-212.91	<.0001*	<.0001*	0.0412*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	21.85	0.1822	0.5259	0.8119
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	-3.26	0.8282	0.9961	0.9803
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)Tnfrsf11b$^{f/f}$	234.76	<.0001*	<.0001*	0.0412*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	209.66	<.0001*	<.0001*	0.0412*
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Sost-Cre;Tnfrsf11b$^{f/f}$	-25.11	0.1047	0.3513	0.3191
Contrast	(1-2)-(3-4)	.	-187.81	21.72		.	<.0001*	.	.

Figure 4A, left panel: One-way ANOVA: Femur Cortical Thickness = Genotype : Means, SD, n, paired differences and p-values

For all results in Figure 4, a One-Way ANOVA model was fitted to the 4 genotypes. Normality and Equal Variance assumptions of the residuals were tested and appropriate transformations used where needed. Three Pairwise comparisons and Three 'Difference of Differences' contrasts was also estimated with each group of 3 being adjusted by the False Discovery Rate Procedure.

For femur cortical thickness in Figure 4A, left panel, by the FDR adjusted p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$(p<0.0001). Tnfrsf11b$^{f/f}$ (Dmp1) has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0366). Tnfrsf11b$^{f/f}$ (Sost) has significantly higher values on average than Sost-Cre;Tnfrsf11b$^{f/f}$(p=0.0103).

The (1-2)-(3-4) and (1-2)-(5-6) difference of differences contrasts are also statistically significant(p=0.0046, 0.0085, adjusted p-values respectively).

Dep. Var.	*Type*	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Adjusted p-value*
femur_ctth	Test	S-W Normality, obser	.	.	.		.	0.1355	.
	Test	S-W Normality, resid	.	.	.		.	0.7669	.
	Test	Equal variance test	.	.	.		.	0.8227	.
	Mean	1)Tnfrsf11b$^{+/+}$	6	0.22	0.01		.	.	.
	Mean	2)Tnfrsf11b$^{-/-}$	6	0.16	0.02		.	.	.
	Mean	3)Tnfrsf11b$^{f/f}$ (Dmp1)	6	0.21	0.01		.	.	.
	Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	6	0.20	0.01		.	.	.
	Mean	5)Tnfrsf11b$^{f/f}$ (Sost)	6	0.21	0.01		.	.	.
	Mean	6)Sost-Cre;Tnfrsf11b$^{f/f}$	6	0.19	0.01		.	.	.
	Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.05	<.0001*	<.0001*
	Diff	3)Tnfrsf11b$^{f/f}$ (Dmp1)	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.02	0.0366*	0.0366*
	Diff	5)Tnfrsf11b$^{f/f}$ (Sost)	.	.	.	6)Sost-Cre;Tnfrsf11b$^{f/f}$	0.02	0.0069*	0.0103*
	Contrast	(1-2)-(3-4)	.	0.04	0.01		.	0.0015*	0.0046*
	Contrast	(1-2)-(5-6)	.	0.03	0.01		.	0.0057*	0.0085*
	Contrast	(3-4)-(5-6)	.	-0.01	0.01		.	0.6174	0.6174

Figure 4A, right panel: One-way ANOVA: L4 Cortical Thickness = Genotype : Means, SD, n, paired differences and p-values

For L4 cortical thickness in Figure 4A, right panel, by the FDR adjusted p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$($p<0.0001$). Tnfrsf11b$^{f/f}$ (Dmp1) has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$($p<0.0001$). Tnfrsf11b$^{f/f}$ (Sost) has significantly higher values on average than Sost-Cre;Tnfrsf11b$^{f/f}$($p=0.0025$).

The (1-2)-(5-6) and (3-4)-(5-6) difference of differences contrasts are also statistically significant($p<0.0001$, $p=0.0001$, adjusted p-values respectively).

Dep. Var.	*Type*	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Adjusted p-value*
spine_ctth	Test	S-W Normality, obser	.	.	.		.	0.0132*	.
	Test	S-W Normality, resid	.	.	.		.	0.7250	.
	Test	Equal variance test	.	.	.		.	0.2705	.
	Mean	1)Tnfrsf11b$^{+/+}$	6	0.09	0.01		.	.	.
	Mean	2)Tnfrsf11b$^{-/-}$	6	0.06	0.01		.	.	.
	Mean	3)Tnfrsf11b$^{f/f}$ (Dmp1)	6	0.09	0.01		.	.	.
	Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	6	0.06	0.01		.	.	.
	Mean	5)Tnfrsf11b$^{f/f}$ (Sost)	6	0.09	0.00		.	.	.
	Mean	6)Sost-Cre;Tnfrsf11b$^{f/f}$	6	0.08	0.00		.	.	.
	Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.04	<.0001*	<.0001*
	Diff	3)Tnfrsf11b$^{f/f}$ (Dmp1)	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.03	<.0001*	<.0001*
	Diff	5)Tnfrsf11b$^{f/f}$ (Sost)	.	.	.	6)Sost-Cre;Tnfrsf11b$^{f/f}$	0.01	0.0025*	0.0025*
	Contrast	(1-2)-(3-4)	.	0.00	0.00		.	0.7557	0.7557
	Contrast	(1-2)-(5-6)	.	0.03	0.00		.	<.0001*	<.0001*
	Contrast	(3-4)-(5-6)	.	0.02	0.00		.	<.0001*	<.0001*

Figure 4B, left panel: One-way ANOVA: Rank transformed Femur Cancellous BV/TV = Genotype : Means, SD, n, paired differences and p-values

For Rank transformed femur cancellous BV/TV in Figure 4B, left panel, by the FDR adjusted p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$($p<0.0001$). Tnfrsf11b$^{f/f}$ (Dmp1) has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$($p<0.0001$).

The (1-2)-(5-6) and (3-4)-(5-6) difference of differences contrasts are also statistically significant($p=0.0179$, 0.0179, adjusted p-values respectively).

Dep. Var.	*Type*	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Adjusted p-value*
r_femur_bvtv	Test	S-W Normality, obser	.	.	.		.	0.1578	.
	Test	S-W Normality, resid	.	.	.		.	0.1428	.
	Test	Equal variance test	.	.	.		.	0.0345*	.
	Mean	1)Tnfrsf11b$^{+/+}$	6	23.33	6.89		.	.	.
	Mean	2)Tnfrsf11b$^{-/-}$	6	4.17	2.86		.	.	.
	Mean	3)Tnfrsf11b$^{f/f}$ (Dmp1)	6	27.17	6.55		.	.	.
	Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	6	8.83	2.64		.	.	.
	Mean	5)Tnfrsf11b$^{f/f}$ (Sost)	6	26.33	6.35		.	.	.
	Mean	6)Sost-Cre;Tnfrsf11b$^{f/f}$	6	21.17	8.47		.	.	.
	Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	19.17	<.0001*	<.0001*
	Diff	3)Tnfrsf11b$^{f/f}$ (Dmp1)	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	18.33	<.0001*	<.0001*
	Diff	5)Tnfrsf11b$^{f/f}$ (Sost)	.	.	.	6)Sost-Cre;Tnfrsf11b$^{f/f}$	5.17	0.1477	0.1477
	Contrast	(1-2)-(3-4)	.	0.83	4.92		.	0.8666	0.8666
	Contrast	(1-2)-(5-6)	.	14.00	4.92		.	0.0079*	0.0179*
	Contrast	(3-4)-(5-6)	.	13.17	4.92		.	0.0119*	0.0179*

Figure 4B, right panel: One-way ANOVA: Square Root transformed L4 Cancellous BV/TV = Genotype : Means, SD, n, paired differences and p-values

For Square Root transformed L4 cancellous BV/TV in Figure 4B, right panel, by the FDR adjusted p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$(p<0.0001). Tnfrsf11b$^{f/f}$ (Dmp1) has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p<0.0001). Tnfrsf11b$^{f/f}$ (Sost) has significantly higher values on average than Sost-Cre;Tnfrsf11b$^{f/f}$(p=0.0447).

The (1-2)-(3-4), (1-2)-(5-6), (3-4)-(5-6) difference of differences contrasts are also statistically significant(p=0.0460, <0.0001, <0.0001, adjusted p-values respectively).

Dep. Var.	*Type*	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Adjusted p-value*
sqrt_spine_bvtv	Test	S-W Normality, obser	.	.	.		.	0.0002*	.
	Test	S-W Normality, resid	.	.	.		.	0.4572	.
	Test	Equal variance test	.	.	.		.	0.5197	.
	Mean	1)Tnfrsf11b$^{+/+}$	6	0.42	0.02		.	.	.
	Mean	2)Tnfrsf11b$^{-/-}$	6	0.12	0.02		.	.	.
	Mean	3)Tnfrsf11b$^{f/f}$ (Dmp1)	6	0.48	0.04		.	.	.
	Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	6	0.23	0.02		.	.	.
	Mean	5)Tnfrsf11b$^{f/f}$ (Sost)	6	0.44	0.02		.	.	.
	Mean	6)Sost-Cre;Tnfrsf11b$^{f/f}$	6	0.41	0.03		.	.	.
	Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.30	<.0001*	<.0001*
	Diff	3)Tnfrsf11b$^{f/f}$ (Dmp1)	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.25	<.0001*	<.0001*
	Diff	5)Tnfrsf11b$^{f/f}$ (Sost)	.	.	.	6)Sost-Cre;Tnfrsf11b$^{f/f}$	0.03	0.0447*	0.0447*
	Contrast	(1-2)-(3-4)	.	0.05	0.02		.	0.0460*	0.0460*
	Contrast	(1-2)-(5-6)	.	0.27	0.02		.	<.0001*	<.0001*
	Contrast	(3-4)-(5-6)	.	0.22	0.02		.	<.0001*	<.0001*

Figure 4E: One-way ANOVA: Rank transformed Cortical Porosity = Genotype : Means, SD, n, paired differences and p-values

For Rank transformed cortical porosity in Figure 4E, by the FDR adjusted p-values, Tnfrsf11b$^{+/+}$ has significantly lower values on average than Tnfrsf11b$^{-/-}$($p<0.0001$). Tnfrsf11b$^{f/f}$ (Dmp1) has significantly lower values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$($p<0.0001$). Tnfrsf11b$^{f/f}$ (Sost) has significantly lower values on average than Sost-Cre;Tnfrsf11b$^{f/f}$($p=0.0018$).

The (1-2)-(5-6) difference of differences contrast is also statistically significant($p=0.0009$).

Dep. Var.	*Type*	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Adjusted p-value*
r_porosity_percent	Test	S-W Normality, obser	.	.	.		.	0.0001*	.
	Test	S-W Normality, resid	.	.	.		.	0.0333*	.
	Test	Equal variance test	.	.	.		.	0.4908	.
	Mean	1)Tnfrsf11b$^{+/+}$	6	9.33	3.27		.	.	.
	Mean	2)Tnfrsf11b$^{-/-}$	6	33.00	2.83		.	.	.
	Mean	3)Tnfrsf11b$^{f/f}$ (Dmp1)	6	11.83	6.15		.	.	.
	Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	6	28.00	2.37		.	.	.
	Mean	5)Tnfrsf11b$^{f/f}$ (Sost)	6	10.00	4.90		.	.	.
	Mean	6)Sost-Cre;Tnfrsf11b$^{f/f}$	6	18.83	5.78		.	.	.
	Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	-23.67	<.0001*	<.0001*
	Diff	3)Tnfrsf11b$^{f/f}$ (Dmp1)	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-16.17	<.0001*	<.0001*
	Diff	5)Tnfrsf11b$^{f/f}$ (Sost)	.	.	.	6)Sost-Cre;Tnfrsf11b$^{f/f}$	-8.83	0.0018*	0.0018*
	Contrast	(1-2)-(3-4)	.	-7.50	3.64		.	0.0483*	0.0531
	Contrast	(1-2)-(5-6)	.	-14.83	3.64		.	0.0003*	0.0009*
	Contrast	(3-4)-(5-6)	.	-7.33	3.64		.	0.0531	0.0531

Figure 4F, left panel: One-way ANOVA: Serum OPG= Genotype : Means, SD, n, paired differences and p-values

For serum OPG in Figure 4F, left panel, by the FDR adjusted p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$(p<0.0001). Tnfrsf11b$^{f/f}$ (Sost) has significantly higher values on average than Sost-Cre;Tnfrsf11b$^{f/f}$(p=0.0410).

The (1-2)-(3-4) and (1-2)-(5-6) difference of differences contrasts are also statistically significant(p<0.0001, <0.0001, adjusted p-values respectively).

Dep. Var.	*Type*	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Adjusted p-value*
opg_elisa	Test	S-W Normality, obser	.	.	.		.	0.0001*	.
	Test	S-W Normality, resid	.	.	.		.	0.5460	.
	Test	Equal variance test	.	.	.		.	0.0640	.
	Mean	1)Tnfrsf11b$^{+/+}$	6	558.04	127.74		.	.	.
	Mean	2)Tnfrsf11b$^{-/-}$	6	0.00	0.00		.	.	.
	Mean	3)Tnfrsf11b$^{f/f}$ (Dmp1)	6	526.46	71.03		.	.	.
	Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	11	472.80	93.91		.	.	.
	Mean	5)Tnfrsf11b$^{f/f}$ (Sost)	6	573.75	44.59		.	.	.
	Mean	6)Sost-Cre;Tnfrsf11b$^{f/f}$	9	477.45	67.01		.	.	.
	Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	558.04	<.0001*	<.0001*
	Diff	3)Tnfrsf11b$^{f/f}$ (Dmp1)	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	53.66	0.1921	0.1921
	Diff	5) Tnfrsf11b$^{f/f}$ (Sost)	.	.	.	6)Sost-Cre;Tnfrsf11b$^{f/f}$	96.30	0.0273*	0.0410*
	Contrast	(1-2)-(3-4)	.	504.38	61.20		.	<.0001*	<.0001*
	Contrast	(1-2)-(5-6)	.	461.74	62.24		.	<.0001*	<.0001*
	Contrast	(3-4)-(5-6)	.	-42.64	58.25		.	0.4687	0.4687

Figure 4F, right panel: One-way ANOVA: Square Root transformed Serum RANKL = Genotype : Means, SD, n, paired differences and p-values

For Square Root transformed serum RANKL in Figure 4F, right panel, by the FDR adjusted p-values, Tnfrsf11b$^{+/+}$ has significantly lower values on average than Tnfrsf11b$^{-/-}$(p<0.0001). Tnfrsf11b$^{f/f}$ (Dmp1) has significantly lower values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p<=0.0072).

The (1-2)-(3-4) and (1-2)-(5-6) difference of differences contrasts are also statistically significant(p<0.0001, <0.0001, adjusted p-values respectively).

Dep. Var.	*Type*	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Adjusted p-value*
sqrt_rankl_elisa	Test	S-W Normality, obser	.	.	.		.	0.0001*	.
	Test	S-W Normality, resid	.	.	.		.	0.8537	.
	Test	Equal variance test	.	.	.		.	0.3601	.
	Mean	1)Tnfrsf11b$^{+/+}$	6	4.24	3.75		.	.	.
	Mean	2)Tnfrsf11b$^{-/-}$	5	42.95	1.80		.	.	.
	Mean	3)Tnfrsf11b$^{f/f}$ (Dmp1)	5	2.38	3.84		.	.	.
	Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	9	7.15	3.26		.	.	.
	Mean	5)Tnfrsf11b$^{f/f}$ (Sost)	6	7.73	2.46		.	.	.
	Mean	6)Sost-Cre;Tnfrsf11b$^{f/f}$	9	9.88	1.39		.	.	.
	Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	-38.71	<.0001*	<.0001*
	Diff	3)Tnfrsf11b$^{f/f}$ (Dmp1)	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	-4.77	0.0048*	0.0072*
	Diff	5)Tnfrsf11b$^{f/f}$ (Sost)	.	.	.	6)Sost-Cre;Tnfrsf11b$^{f/f}$	-2.15	0.1591	0.1591
	Contrast	(1-2)-(3-4)	.	-33.94	2.33		.	<.0001*	<.0001*
	Contrast	(1-2)-(5-6)	.	-36.56	2.28		.	<.0001*	<.0001*
	Contrast	(3-4)-(5-6)	.	-2.62	2.17		.	0.2372	0.2372

Supplementary Figure 1: Two-sample t-tests between genotypes

By two-sample t-tests, or Wilcoxon Rank Sum tests where data were not normal, Tnfrsf11b$^{+/+}$ genotype animals showed statistically significantly higher on average values of female femur cortical thickness($p<0.001$), female L4 cortical thickness($p<0.001$), female femur cancellous BV/TV($p=0.0012$),female L4 cancellous BV/TV($p=0.0034$), male femur cortical thickness($p=0.0057$), male L4 cortical thickness($p<0.001$), male femur cancellous BV/TV($p<0.001$), and male L4 cancellous BV/TV($p<0.001$) than Tnfrsf11b$^{-/-}$ genotype animals.

Analysis Variable	*Genotype*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Female Femur Cortical Thickness.	Tnfrsf11b$^{-/-}$	7	0.09	0.01	0.09600	0.0670	0.1140	t-test, equal variances	<0.001*
	Tnfrsf11b$^{+/+}$	6	0.13	0.01	0.12400	0.1140	0.1380		
	Difference	.	-0.03	0.01	.	.	.		

Analysis Variable	*Genotype*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Female L4 Cortical Thickness	Tnfrsf11b$^{-/-}$	7	0.04	0.01	0.04100	0.0370	0.0540	t-test, equal variances	<0.001*
	Tnfrsf11b$^{+/+}$	6	0.06	0.00	0.05900	0.0540	0.0600		
	Difference	.	-0.02	0.00	.	.	.		

Analysis Variable	*Genotype*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Female Femur Cancellous BV/TV	Tnfrsf11b$^{-/-}$	7	0.00	0.00	0.00100	0.0001	0.0044	t-test, unequal variances	0.0012*
	Tnfrsf11b$^{+/+}$	6	0.04	0.01	0.04085	0.0195	0.0595		
	Difference	.	-0.04	0.01	.	.	.		

Analysis Variable	*Genotype*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Female L4 Cancellous BV/TV	Tnfrsf11b$^{+/+}$	6	0.17	0.02	0.18735	0.1373	0.1912	Rank Sum test	0.0034*
	Tnfrsf11b$^{-/-}$	7	0.06	0.01	0.06440	0.0394	0.0775		
	Difference	.	0.11	0.02	.	.	.		

Supplementary Figure 1: Two-sample t-tests between genotypes

Analysis Variable	*Genotype*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Male Femur Cortical Thickness.	Tnfrsf11b$^{+/+}$	6	0.13	0.01	0.12950	0.1180	0.1390	t-test, equal variances	0.0057*
	Tnfrsf11b$^{-/-}$	4	0.10	0.02	0.09550	0.0880	0.1220		
	Difference	.	0.03	0.01	.	.	.		

Analysis Variable	*Genotype*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Male L4 Cortical Thickness	Tnfrsf11b$^{+/+}$	7	0.06	0.00	0.05900	0.0570	0.0650	t-test, equal variances	<0.001*
	Tnfrsf11b$^{-/-}$	4	0.04	0.00	0.03800	0.0350	0.0400		
	Difference	.	0.02	0.00	.	.	.		

Analysis Variable	*Genotype*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Male Femur Cancellous BV/TV	Tnfrsf11b$^{+/+}$	6	0.05	0.01	0.05360	0.0394	0.0650	t-test, unequal variances	<0.001*
	Tnfrsf11b$^{-/-}$	4	0.00	0.00	0.00190	0.0001	0.0044		
	Difference	.	0.05	0.01	.	.	.		

Analysis Variable	*Genotype*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Male L4 Cancellous BV/TV	Tnfrsf11b$^{+/+}$	7	0.19	0.01	0.18940	0.1645	0.2063	t-test, equal variances	<0.001*
	Tnfrsf11b$^{-/-}$	4	0.06	0.01	0.05615	0.0531	0.0779		
	Difference	.	0.13	0.01	.	.	.		

Supplementary Figure 2A, left panel: One-way ANOVA: Femur Cortical Thickness = Genotype : Means, SD, n, paired differences and p-values

For all results in Supplementary Figure 2, a One-Way ANOVA model was fitted to the 4 genotypes. Normality and Equal Variance assumptions of the residuals was test and passed in each case. Pairwise comparisons were adjusted by Tukey procedure and the 'Difference of Differences' contrast was also estimated.

For femur cortical thickness in Supplementary Figure 2A, left panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$ (p<0.0001). Tnfrsf11b$^{-/-}$ has significantly lower values on average than CD19-Cre(p<0.0001) & CD19-Cre;Tnfrsf11b$^{f/f}$(p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p<0.0001).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0144*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.5791	.	.
Test	Equal variance test	.	.	.		.	0.5431	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.22	0.01		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	6	0.16	0.02		.	.	.	.
Mean	3)CD19-Cre	10	0.20	0.01		.	.	.	.
Mean	4)CD19-Cre;Tnfrsf11b$^{f/f}$	8	0.20	0.01		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.05	<.0001*	<.0001*	0.0406*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)CD19-Cre	0.01	0.1155	0.3809	0.4010
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)CD19-Cre;Tnfrsf11b$^{f/f}$	0.01	0.0880	0.3086	0.4010
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)CD19-Cre	-0.04	<.0001*	<.0001*	0.0315*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)CD19-Cre;Tnfrsf11b$^{f/f}$	-0.04	<.0001*	<.0001*	0.0315*
Diff	3)CD19-Cre	.	.	.	4)CD19-Cre;Tnfrsf11b$^{f/f}$	0.00	0.8078	0.9947	0.9507
Contrast	(1-2)-(3-4)	.	0.05	0.01		.	<.0001*	.	.

Supplementary Figure 2A, right panel: One-way ANOVA: L4 Cortical Thickness = Genotype : Means, SD, n, paired differences and p-values

For L4 cortical thickness in Supplementary Figure 2A, right panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$ (p<0.0001) and CD19-Cre(p=0.0045). Tnfrsf11b$^{-/-}$ has significantly lower values on average than CD19-Cre(p<0.0001) & CD19-Cre;Tnfrsf11b$^{f/f}$(p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p<0.0001).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0052*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.3458	.	.
Test	Equal variance test	.	.	.		.	0.7702	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.09	0.01		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	6	0.06	0.01		.	.	.	.
Mean	3)CD19-Cre	10	0.08	0.01		.	.	.	.
Mean	4)CD19-Cre;Tnfrsf11b$^{f/f}$	8	0.08	0.01		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.04	<.0001*	<.0001*	0.0400*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)CD19-Cre	0.01	0.0045*	0.0219*	0.1012
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)CD19-Cre;Tnfrsf11b$^{f/f}$	0.01	0.0451*	0.1781	0.3203
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)CD19-Cre	-0.03	<.0001*	<.0001*	0.0305*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)CD19-Cre;Tnfrsf11b$^{f/f}$	-0.03	<.0001*	<.0001*	0.0305*
Diff	3)CD19-Cre	.	.	.	4)CD19-Cre;Tnfrsf11b$^{f/f}$	-0.00	0.3310	0.7560	0.8299
Contrast	(1-2)-(3-4)	.	0.04	0.00		.	<.0001*	.	.

Supplementary Figure 2B, left panel: One-way ANOVA: Femur Cancellous BV/TV = Genotype : Means, SD, n, paired differences and p-values

For femur cancellous BV/TV in Supplementary Figure 2B, left panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$ (p<0.0001). Tnfrsf11b$^{-/-}$ has significantly lower values on average than CD19-Cre(p<0.0001) & CD19-Cre;Tnfrsf11b$^{f/f}$(p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p<0.0001).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0016*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.2197	.	.
Test	Equal variance test	.	.	.		.	0.3279	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.04	0.01		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	6	0.00	0.00		.	.	.	.
Mean	3)CD19-Cre	10	0.05	0.01		.	.	.	.
Mean	4)CD19-Cre;Tnfrsf11b$^{f/f}$	8	0.05	0.01		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.04	<.0001*	<.0001*	0.0406*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)CD19-Cre	-0.01	0.1048	0.3536	0.5450
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)CD19-Cre;Tnfrsf11b$^{f/f}$	-0.00	0.4175	0.8427	0.7811
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)CD19-Cre	-0.05	<.0001*	<.0001*	0.0311*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)CD19-Cre;Tnfrsf11b$^{f/f}$	-0.04	<.0001*	<.0001*	0.0311*
Diff	3)CD19-Cre	.	.	.	4)CD19-Cre;Tnfrsf11b$^{f/f}$	0.00	0.3806	0.8090	0.9388
Contrast	(1-2)-(3-4)	.	0.04	0.01		.	<.0001*	.	.

Supplementary Figure 2B, right panel: One-way ANOVA: L4 Cancellous BV/TV = Genotype : Means, SD, n, paired differences and p-values

For L4 cancellous BV/TV in Supplementary Figure 2B, right panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Tnfrsf11b$^{-/-}$ (p<0.0001), but has significantly lower values on average than CD19-Cre(p=0.0003) & CD19-Cre;Tnfrsf11b$^{f/f}$(p<0.0001). Tnfrsf11b$^{-/-}$ has significantly lower values on average than CD19-Cre(p=0.0003) & CD19-Cre;Tnfrsf11b$^{f/f}$(p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p<0.0001).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0001*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.1235	.	.
Test	Equal variance test	.	.	.		.	0.2161	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.18	0.02		.	.	.	.
Mean	2)Tnfrsf11b$^{-/-}$	6	0.02	0.01		.	.	.	.
Mean	3)CD19-Cre	10	0.25	0.02		.	.	.	.
Mean	4)CD19-Cre;Tnfrsf11b$^{f/f}$	8	0.25	0.03		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Tnfrsf11b$^{-/-}$	0.16	<.0001*	<.0001*	0.0247*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)CD19-Cre	-0.07	<.0001*	<.0001*	0.0237*
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)CD19-Cre;Tnfrsf11b$^{f/f}$	-0.06	<.0001*	<.0001*	0.0237*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	3)CD19-Cre	-0.23	<.0001*	<.0001*	0.0237*
Diff	2)Tnfrsf11b$^{-/-}$	.	.	.	4)CD19-Cre;Tnfrsf11b$^{f/f}$	-0.23	<.0001*	<.0001*	0.0237*
Diff	3)CD19-Cre	.	.	.	4)CD19-Cre;Tnfrsf11b$^{f/f}$	0.00	0.8299	0.9963	0.9510
Contrast	(1-2)-(3-4)	.	0.16	0.02		.	<.0001*	.	.

Supplementary Figure 2C: Two-sample t-tests on serum OPG and RANKL, tested by genotypes, within each age.

`No significant differences between genotypes were detected for either variable or age.`

Analysis Variable	*Age group*	*Genotype*	*n*	*Mean*	*Std. Dev.*	*Test*	*p-value*
OPG_ELISA	*A)5 weeks*	A) CD19-Cre;Tnfrsf11b$^{f/f}$	8	3222.37	371.70	t-test, equal variances	0.6506
		B) CD19-Cre	5	3323.18	393.58		
		Difference	.	-100.80	379.80		

Analysis Variable	*Age group*	*Genotype*	*n*	*Mean*	*Std. Dev.*	*Test*	*p-value*
OPG_ELISA	*B)4 months*	A) CD19-Cre;Tnfrsf11b$^{f/f}$	7	2556.74	597.49	t-test, equal variances	0.2233
		B) CD19-Cre	8	2941.20	566.07		
		Difference	.	-384.45	580.78		

Analysis Variable	*Age group*	*Genotype*	*n*	*Mean*	*Std. Dev.*	*Test*	*p-value*
RANKL_ELISA	*A)5 weeks*	A) CD19-Cre;Tnfrsf11b$^{f/f}$	8	92.49	52.03	t-test, equal variances	0.5358
		B) CD19-Cre	5	74.78	41.94		
		Difference	.	17.71	48.61		

Analysis Variable	*Age group*	*Genotype*	*n*	*Mean*	*Std. Dev.*	*Test*	*p-value*
RANKL_ELISA	*B)4 months*	A) CD19-Cre;Tnfrsf11b$^{f/f}$	7	119.74	65.83	t-test, equal variances	0.1778
		B) CD19-Cre	7	80.82	29.00		
		Difference	.	38.92	50.87		

Supplementary Figure 3A: Two-sample tests within tissue types

By a two-sample t-test, the CD19-Cre genotype showed statistically significantly higher(mean of 1.02 vs. 0.15) on average *Tnfrsf11b* exon 2 than the CD19-Cre;Tnfrsf11b$^{f/f}$ genotype within +CD19 Cells(p<0.001).

By a two-sample Wilcoxon Rank Sum test(used in place of t-test due to non-normal data), the CD19-Cre and CD19-Cre;Tnfrsf11b$^{f/f}$ genotypes showed no statistically significant difference in *Tnfrsf11b* exon 2 within -CD19 Cells(p=0.817).

Analysis Variable	*CD19 + or - (only for b cells)*	*geno*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Tnfrsf11b exon 2	*+CD19 Cells*	A)CD19-Cre	2	1.02	0.08	1.02385	0.96569	1.08202	t-test, equal variances	<0.001*
		B)CD19-Cre;Tnfrsf11b	4	0.15	0.04	0.14835	0.11406	0.20778		
		Difference	.	0.87	0.05	.	.	.		

Analysis Variable	*CD19 + or - (only for b cells)*	*geno*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Tnfrsf11b exon 2	*-CD19 Cells*	A)CD19-Cre	2	0.98	0.11	0.97615	0.90112	1.05117	Rank Sum test	0.817
		B)CD19-Cre;Tnfrsf11b	4	0.87	0.12	0.92558	0.69292	0.93260		
		Difference	.	0.11	0.11	.	.	.		

Supplementary Figure 3A: Two-sample tests within tissue types

By two-sample t-tests, the CD19-Cre and CD19-Cre;Tnfrsf11b$^{f/f}$ genotypes showed no statistically significant differences in *Tnfrsf11b* exon 2 in Cortical Bone(p=0.4927), Liver(p=0.9639), or Muscle(p=0.9752) tissues.

Analysis Variable	*Tissue*	*geno*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Tnfrsf11b exon 2	*Cortical Bone*	A)CD19-Cre	6	1.00	0.21	1.01251	0.66708	1.25941	t-test, equal variances	0.4927
		B)CD19-Cre;Tnfrsf11b$^{f/f}$	6	0.93	0.12	0.93169	0.75074	1.11653		
		Difference	.	0.07	0.17	.	.	.		

Analysis Variable	*Tissue*	*geno*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Tnfrsf11b exon 2	*Muscle*	A)CD19-Cre	6	1.00	0.09	1.02754	0.87233	1.07537	t-test, equal variances	0.9752
		B)CD19-Cre;Tnfrsf11b$^{f/f}$	5	1.00	0.10	0.97724	0.91252	1.17129		
		Difference	.	0.00	0.09	.	.	.		

Analysis Variable	*Tissue*	*geno*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Tnfrsf11b exon 2	*Liver*	A)CD19-Cre	6	1.00	0.07	1.00959	0.90246	1.08939	t-test, equal variances	0.9639
		B)CD19-Cre;Tnfrsf11b$^{f/f}$	6	1.00	0.07	1.01593	0.92020	1.11368		
		Difference	.	-0.00	0.07	.	.	.		

Supplementary Figure 3B: Two-sample tests within tissue types

By two-sample t-tests, the Tnfrsf11b$^{f/f}$ and Dmp1-Cre;Tnfrsf11b$^{f/f}$ genotypes showed statistically significant differences in *Tnfrsf11b* exon 2 in Cortical Bone(p<0.001), and Muscle(p<0.001) tissues, but not in Liver(p=0.8254)

Analysis Variable	*Tissue*	*geno*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Tnfrsf11b exon 2	*Cortical Bone*	A)Tnfrsf11b$^{f/f}$	4	1.00	0.11	0.99811	0.89824	1.10554	t-test, equal variances	<0.001*
		B)Dmp1-Cre;Tnfrsf11b$^{f/f}$	9	0.56	0.08	0.53837	0.47111	0.67429		
		Difference	.	0.44	0.09	.	.	.		

Analysis Variable	*Tissue*	*geno*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Tnfrsf11b exon 2	*Muscle*	A)Tnfrsf11b$^{f/f}$	5	1.00	0.04	0.99362	0.94165	1.06412	t-test, equal variances	<0.001*
		B)Dmp1-Cre;Tnfrsf11b$^{f/f}$	4	0.63	0.04	0.63847	0.57283	0.66570		
		Difference	.	0.37	0.04	.	.	.		

Analysis Variable	*Tissue*	*geno*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Tnfrsf11b exon 2	*Liver*	A)Tnfrsf11b$^{f/f}$	5	1.00	0.09	0.94463	0.90916	1.11537	t-test, equal variances	0.8254
		B)Dmp1-Cre;Tnfrsf11b$^{f/f}$	10	0.99	0.10	0.97570	0.84636	1.19746		
		Difference	.	0.01	0.10	.	.	.		

Supplementary Figure 3C: Two-sample tests within tissue types

By two-sample t-tests, the Tnfrsf11b$^{f/f}$ genotypes shows statistically significantly higher values of *Tnfrsf11b* exon 2 on average than the Sost-Cre;Tnfrsf11b$^{f/f}$ genotype in Cortical Bone(p=<0.001), Liver(p=0.0207), and Muscle(p=0.0419) tissues.

Analysis Variable	*Tissue*	*geno*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Tnfrsf11b exon 2	*Cortical Bone*	A)Tnfrsf11b$^{f/f}$	6	1.00	0.08	0.99521	0.91012	1.12094	t-test, equal variances	<0.001*
		B)Sost-Cre;Tnfrsf11b$^{f/f}$	5	0.59	0.06	0.57029	0.52240	0.67102		
		Difference	.	0.41	0.07	.	.	.		

Analysis Variable	*Tissue*	*geno*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Tnfrsf11b exon 2	*Muscle*	A)Tnfrsf11b$^{f/f}$	6	1.00	0.11	0.98398	0.87690	1.17968	t-test, equal variances	0.0419*
		B)Sost-Cre;Tnfrsf11b$^{f/f}$	8	0.84	0.14	0.87229	0.62817	1.02662		
		Difference	.	0.16	0.13	.	.	.		

Analysis Variable	*Tissue*	*geno*	*n*	*Mean*	*Std. Dev.*	*Median*	*Minimum*	*Maximum*	*Test*	*p-value*
Tnfrsf11b exon 2	*Liver*	A)Tnfrsf11b$^{f/f}$	6	1.00	0.07	1.00495	0.90107	1.12299	t-test, equal variances	0.0207*
		B)Sost-Cre;Tnfrsf11b$^{f/f}$	8	0.89	0.08	0.92926	0.73968	0.96115		
		Difference	.	0.11	0.08	.	.	.		

Supplementary Figure 4A: One-way ANOVA: Rank-Transformed Female Femur Cortical Thickness = Genotype : Means, SD, n, paired differences and p-values

For all results in Supplementary Figure 4, a One-Way ANOVA model was fitted to the 4 genotypes. Normality and Equal Variance assumptions of the residuals was tested. Transformations were used when those assumptions failed. Pairwise comparisons were adjusted by Tukey procedure and the 'Difference of Differences' contrast was also estimated.

For Rank-Transformed female femur cortical thickness in Supplementary Figure 4A, by the Tukey p-values, Dmp1-Cre has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0033). Tnfrsf11b$^{f/f}$ has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0287)

The (1-2)-(3-4) difference of differences contrast is also statistically significant (p=0.0021).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.4151	.	.
Test	S-W Normality, residuals	.	.	.		.	0.0661	.	.
Test	Equal variance test	.	.	.		.	0.4170	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	11.50	5.88		.	.	.	.
Mean	2)Dmp1-Cre	6	17.25	5.69		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	5	14.70	5.35		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	6	5.00	3.58		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Dmp1-Cre	-5.75	0.0705	0.2546	0.5644
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-3.20	0.3220	0.7418	0.8425
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	6.50	0.0432*	0.1687	0.4350
Diff	2)Dmp1-Cre	.	.	.	3)Tnfrsf11b$^{f/f}$	2.55	0.4278	0.8488	0.8425
Diff	2)Dmp1-Cre	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	12.25	0.0006*	0.0033*	0.1972
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	9.70	0.0061*	0.0287*	0.2434
Contrast	(1-2)-(3-4)	.	-15.45	4.35		.	0.0021*	.	.

Supplementary Figure 4B, left panel: One-way ANOVA: Square-root-Transformed Female Femur Cancellous BV/TV = Genotype : Means, SD, n, paired differences and p-values

For Square-root-Transformed female femur cancellous BV/TV in Supplementary Figure 4B, left panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0006). Dmp1-Cre has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p<0.0001). Tnfrsf11b$^{f/f}$ has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0015).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p=0.0007).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.6742	.	.
Test	S-W Normality, residuals	.	.	.		.	0.7150	.	.
Test	Equal variance test	.	.	.		.	0.1379	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.20	0.02		.	.	.	.
Mean	2)Dmp1-Cre	6	0.22	0.03		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	5	0.20	0.03		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	6	0.13	0.02		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Dmp1-Cre	-0.02	0.2311	0.6119	0.9769
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	0.00	0.8874	0.9989	1.0000
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.07	0.0001*	0.0006*	0.0618
Diff	2)Dmp1-Cre	.	.	.	3)Tnfrsf11b$^{f/f}$	0.02	0.2015	0.5602	0.9769
Diff	2)Dmp1-Cre	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.09	<.0001*	<.0001*	0.0618
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.07	0.0003*	0.0015*	0.0627
Contrast	(1-2)-(3-4)	.	-0.09	0.02		.	0.0007*	.	.

Supplementary Figure 4B, right panel: One-way ANOVA: Female L4 Cancellous BV/TV = Genotype : Means, SD, n, paired differences and p-values

For female L4 cancellous BV/TV in Supplementary Figure 4B, right panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p<0.0001). Dmp1-Cre has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p<0.0001). Tnfrsf11bf/f has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p<0.0001).

The (1-2)-(3-4) difference of differences contrast is not statistically significant(p=0.0982).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0107*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.3366	.	.
Test	Equal variance test	.	.	.		.	0.3243	.	.
Mean	1)Tnfrsf11b+/+	6	0.16	0.01		.	.	.	.
Mean	2)Dmp1-Cre	6	0.18	0.02		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	5	0.17	0.02		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	9	0.08	0.01		.	.	.	.
Diff	1)Tnfrsf11b+/+	.	.	.	2)Dmp1-Cre	-0.01	0.2021	0.5636	0.9978
Diff	1)Tnfrsf11b+/+	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.01	0.3379	0.7622	0.9978
Diff	1)Tnfrsf11b+/+	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.08	<.0001*	<.0001*	0.0239*
Diff	2)Dmp1-Cre	.	.	.	3)Tnfrsf11bf/f	0.00	0.7867	0.9926	0.9978
Diff	2)Dmp1-Cre	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.09	<.0001*	<.0001*	0.0239*
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.09	<.0001*	<.0001*	0.0287*
Contrast	(1-2)-(3-4)	.	-0.10	0.01		.	<.0001*	.	.

Supplementary Figure 4C: One-way ANOVA: Rank-Transformed Male Femur Ct. Th. = Genotype : Means, SD, n, paired differences and p-values

`For Rank-Transformed male femur cortical thickness in Supplementary Figure 4C, by the Tukey p-values, there are no statistically significant pairwise comparisons.`

`The (1-2)-(3-4) difference of differences contrast is not statistically significant(p=0.0762).`

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.6294	.	.
Test	S-W Normality, residuals	.	.	.		.	0.1007	.	.
Test	Equal variance test	.	.	.		.	0.3243	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	11.17	4.07		.	.	.	.
Mean	2)Dmp1-Cre	6	9.83	5.71		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	2	14.50	0.71		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	4	4.00	4.08		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Dmp1-Cre	1.33	0.6234	0.9572	0.9096
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-3.33	0.3897	0.8113	0.8124
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	7.17	0.0301*	0.1199	0.4985
Diff	2)Dmp1-Cre	.	.	.	3)Tnfrsf11b$^{f/f}$	-4.67	0.2344	0.6116	0.9096
Diff	2)Dmp1-Cre	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	5.83	0.0696	0.2470	0.4985
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	10.50	0.0196*	0.0815	0.4985
Contrast	(1-2)-(3-4)	.	-9.17	4.79		.	0.0762	.	.

Supplementary Figure 4D, left panel: One-way ANOVA: Male Femur Cancellous BV/TV = Genotype : Means, SD, n, paired differences and p-values

For male femur cancellous BV/TV in Supplementary Figure 4D, left panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0045). Dmp1-Cre has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p=0.0384).

The (1-2)-(3-4) difference of differences contrast is not statistically significant(p=0.0534).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.4739	.	.
Test	S-W Normality, residuals	.	.	.		.	0.2195	.	.
Test	Equal variance test	.	.	.		.	0.3442	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.05	0.01		.	.	.	.
Mean	2)Dmp1-Cre	6	0.05	0.01		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	2	0.05	0.00		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	4	0.03	0.01		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Dmp1-Cre	0.00	0.7166	0.9819	0.9951
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.00	0.7848	0.9921	0.9951
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.02	0.0045*	0.0206*	0.1543
Diff	2)Dmp1-Cre	.	.	.	3)Tnfrsf11b$^{f/f}$	-0.00	0.5975	0.9476	0.9951
Diff	2)Dmp1-Cre	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.02	0.0087*	0.0384*	0.1543
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.02	0.0147*	0.0627	0.4985
Contrast	(1-2)-(3-4)	.	-0.02	0.01		.	0.0534	.	.

Supplementary Figure 4D, right panel: One-way ANOVA: Male L4 Cancellous BV/TV = Genotype : Means, SD, n, paired differences and p-values

For male L4 cancellous BV/TV in Supplementary Figure 4D, right panel, by the Tukey p-values, Tnfrsf11b$^{+/+}$ has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p<0.0001). Dmp1-Cre has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p<0.0001). Tnfrsf11b$^{+/+}$ has significantly higher values on average than Dmp1-Cre;Tnfrsf11b$^{f/f}$(p<0.0001).

The (1-2)-(3-4) difference of differences contrast is also statistically significant(p<0.0001).

Type	*Test/Level of geno*	*n*	*Mean*	*Std Dev*	*2nd level of geno*	*Difference of means*	*p-value*	*Tukey p-value*	*Adjusted DSCF p-value (nonparam.)*
Test	S-W Normality, observed	.	.	.		.	0.0074*	.	.
Test	S-W Normality, residuals	.	.	.		.	0.5219	.	.
Test	Equal variance test	.	.	.		.	0.3809	.	.
Mean	1)Tnfrsf11b$^{+/+}$	6	0.18	0.02		.	.	.	.
Mean	2)Dmp1-Cre	6	0.20	0.03		.	.	.	.
Mean	3)Tnfrsf11b$^{f/f}$	3	0.18	0.02		.	.	.	.
Mean	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	4	0.08	0.00		.	.	.	.
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	2)Dmp1-Cre	-0.02	0.1603	0.4745	0.8762
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	3)Tnfrsf11b$^{f/f}$	0.00	0.7571	0.9887	1.0000
Diff	1)Tnfrsf11b$^{+/+}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.10	<.0001*	<.0001*	0.1543
Diff	2)Dmp1-Cre	.	.	.	3)Tnfrsf11b$^{f/f}$	0.02	0.1491	0.4501	0.8762
Diff	2)Dmp1-Cre	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.12	<.0001*	<.0001*	0.1543
Diff	3)Tnfrsf11b$^{f/f}$	.	.	.	4)Dmp1-Cre;Tnfrsf11b$^{f/f}$	0.10	<.0001*	<.0001*	0.2925
Contrast	(1-2)-(3-4)	.	-0.11	0.02		.	<.0001*	.	.

Supplementary Figure 4F: Two-sample tests within tissue types

A paired t-test on samples measured from the same 5 animals shows bone tissue has significantly higher *Tnfrsf11b*$^{+/+}$ gene expression than muscle tissue (mean difference of 8.75, p=0.0025)

Analysis Variable	*Difference*	*n*	*Mean*	*unadjusted p-value*
Tnfrsf11b$^{+/+}$ Gene Expression	Bone - Muscle	5	8.75	0.0025