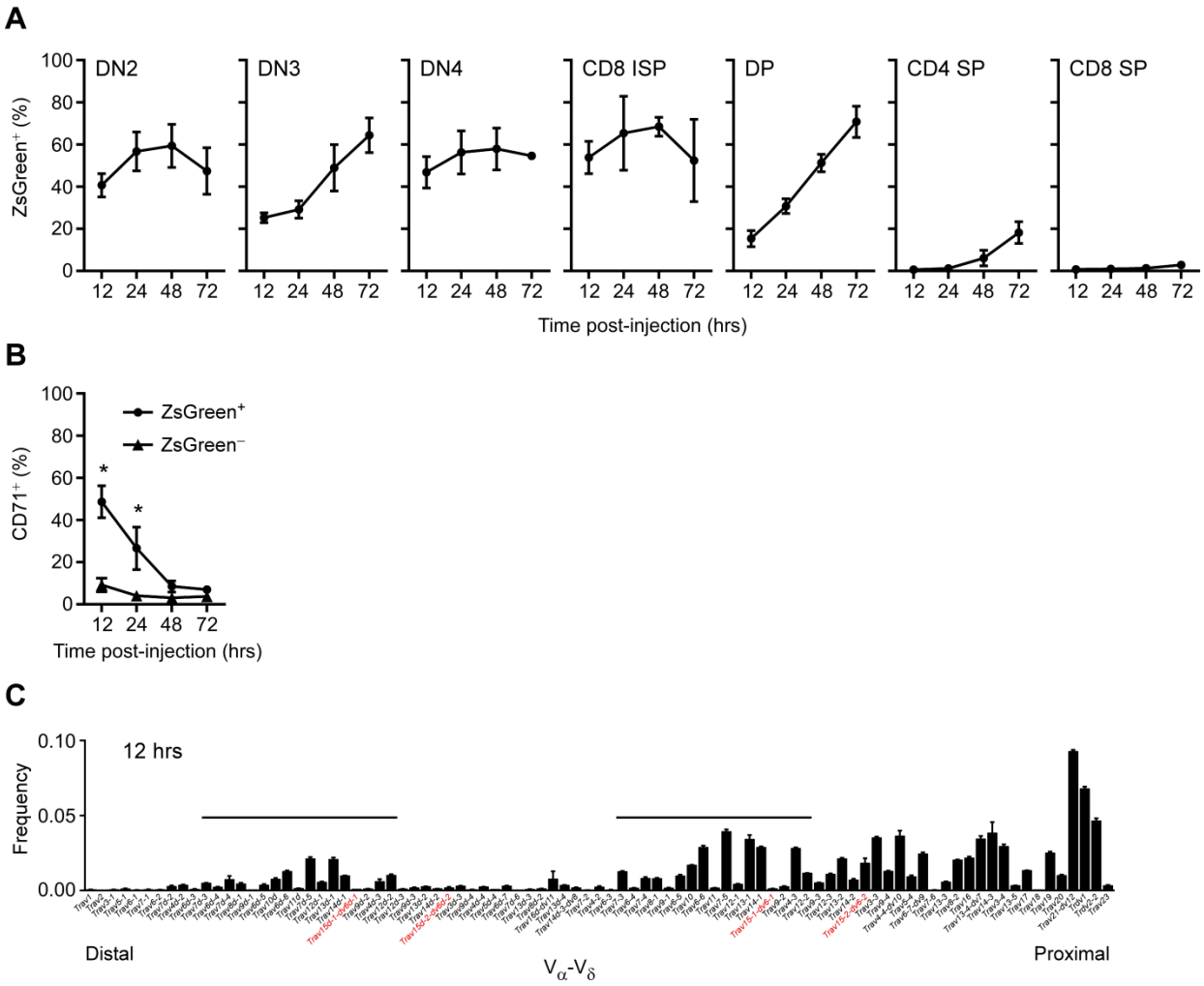

Figure S1. Inducible ZsGreen labeling of thymocytes in *Tcrd*^{CreER} *Rosa26*$^{fl\text{-}STOP\text{-}fl\text{-}ZsGreen}$ mice. Related to Figure 2. (A) Percentage ZsGreen expression in given thymocyte compartments as assessed by flow cytometry. Compartments were defined as follows: DN2, Lin$^-$CD4$^-$CD8$^-$CD3ε$^-$CD44$^+$CD25$^+$; DN3, Lin$^-$CD4$^-$CD8$^-$CD3ε$^-$CD44$^-$CD25$^+$; DN4, Lin$^-$CD4$^-$CD8$^-$CD3ε$^-$CD44$^-$CD25$^-$; CD8ISP, Lin$^-$CD4$^-$CD8$^+$CD3ε$^-$; DP, Lin$^-$CD4$^+$CD8$^+$CD3ε$^-$; CD4SP, Lin$^-$CD4$^+$CD8$^-$CD3ε$^+$; CD8SP, Lin$^-$CD4$^-$CD8$^+$CD3ε$^+$. Data represent the mean and s.d. of 3-7 mice per timepoint. (B) CD71 expression in ZsGreen$^+$ and ZsGreen$^-$ DP thymocyte populations as assessed by flow cytometry. CD71$^+$ and CD71$^-$ gates were determined using fluorescence-minus-one staining strategy. Data are presented as mean and s.d. of 3 mice per timepoint analyzed in two independent experiments. *$P \leq 0.0001$ by repeated-measures two-way ANOVA with Sidak's multiple comparisons test. (**C**) V_α segment usage in ZsGreen$^+$ DP thymocytes of *Tcrd*^{CreER} *Rosa26*$^{fl\text{-}STOP\text{-}fl\text{-}ZsGreen}$ mice isolated 12 hrs after tamoxifen injection. Data are identical to Figure 2B, top left panel, with V_α and V_δ segments identified and *Trav15d-dv6* family members highlighted in red. Lines in the V_α-V_δ plot identify homologous sets of central and central duplication V segments targeted for early rearrangement.

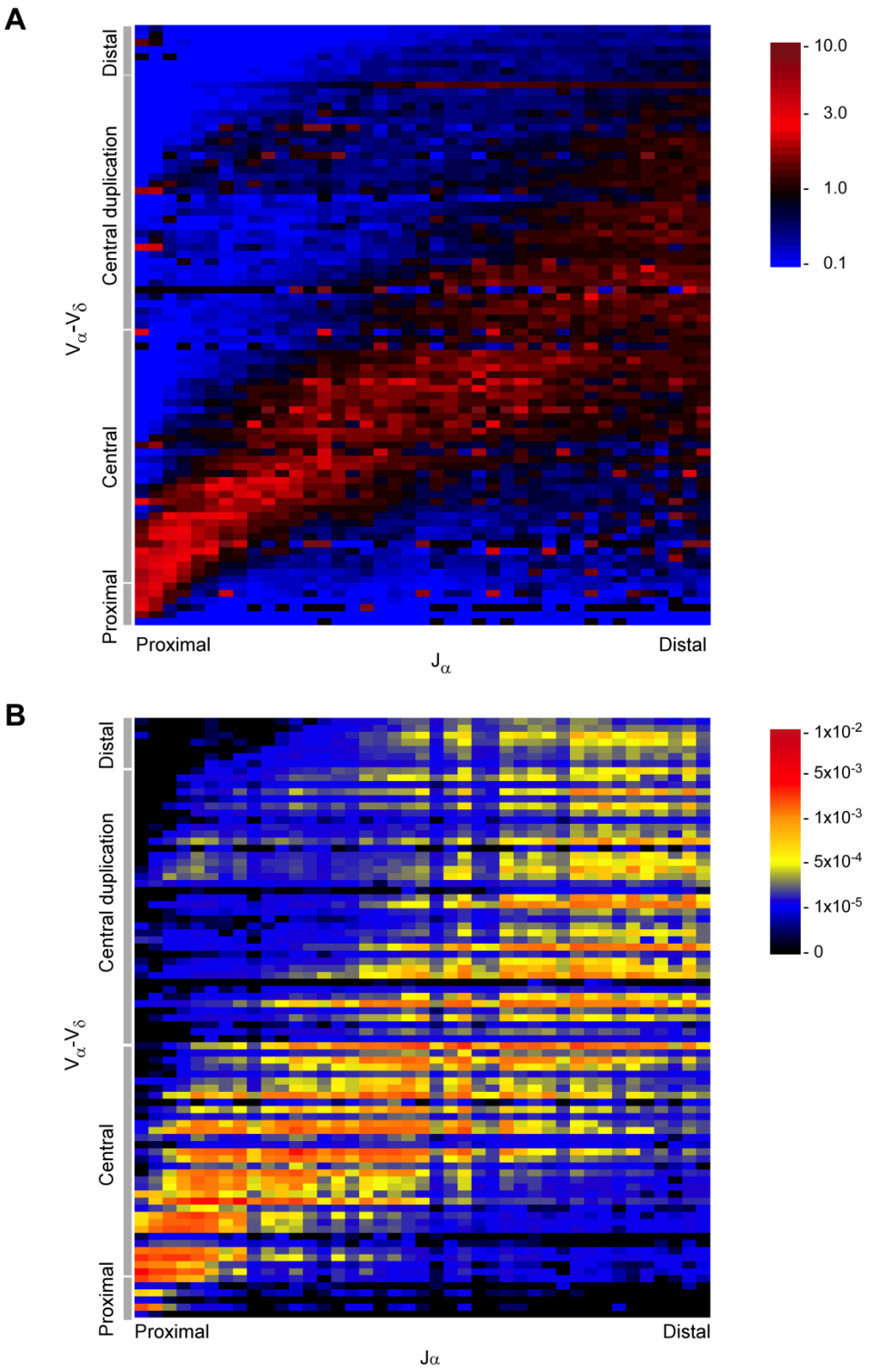

Figure S2. HYα *Tcra* repertoire in thymus and periphery. Related to Figure 3. (A) Difference map depicting ratio of average V_α-J_α rearrangement frequencies in HYα DP thymocytes to those in wild-type, using data in Figures 1B and 3B. Ratios greater than or equal to ten are binned into the same color. (B) V_α-J_α rearrangement frequencies in HYα $CD3^+CD8^+CD44^-CD62L^+$ splenocytes as determined by HTS of *Tcra* transcripts amplified using 5' RACE. Data are presented as the mean of two mice analyzed in one experiment.

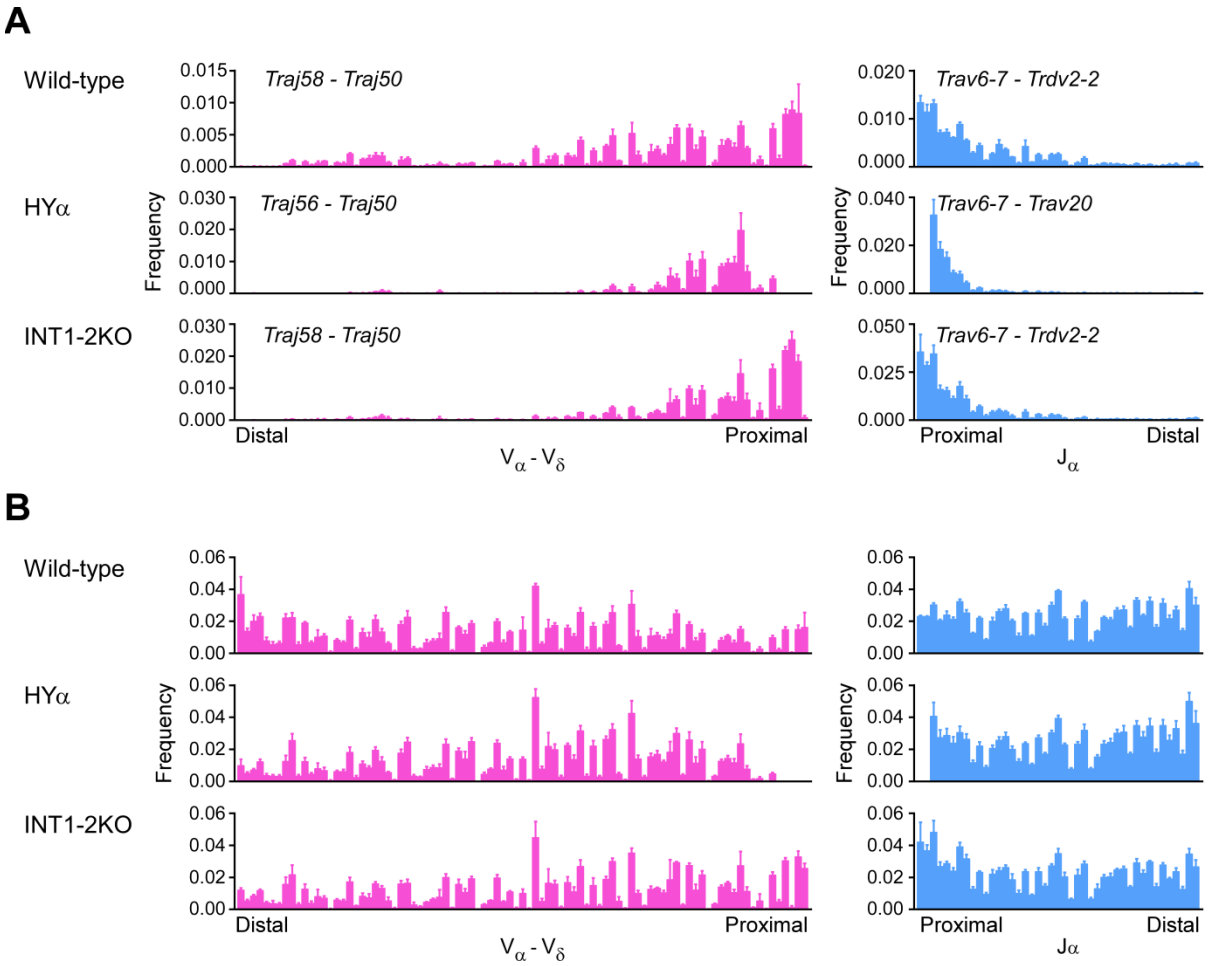

Figure S3. Comparison of V_α-J_α rearrangements in wild-type, HYα, and INT1-2-deficient thymocytes. Related to Figure 5. (A) Comparison of V_α usage with relatively proximal J_α segments (left), or of J_α segments used with relatively proximal V_α segments (right), in wild-type, HYα, and INT1-2-deficient DP thymocytes. Data are from Figures 1C, 3C, and 5C. (B) Comparison of total V_α usage (left) and total J_α usage (right) in wild-type, HYα, and INT1-2-deficient thymocytes. Data are from Figures 1B, 3B, and 5B and are presented as mean and s.d. of 3-4 mice for each genotype, analyzed in two independent experiments.

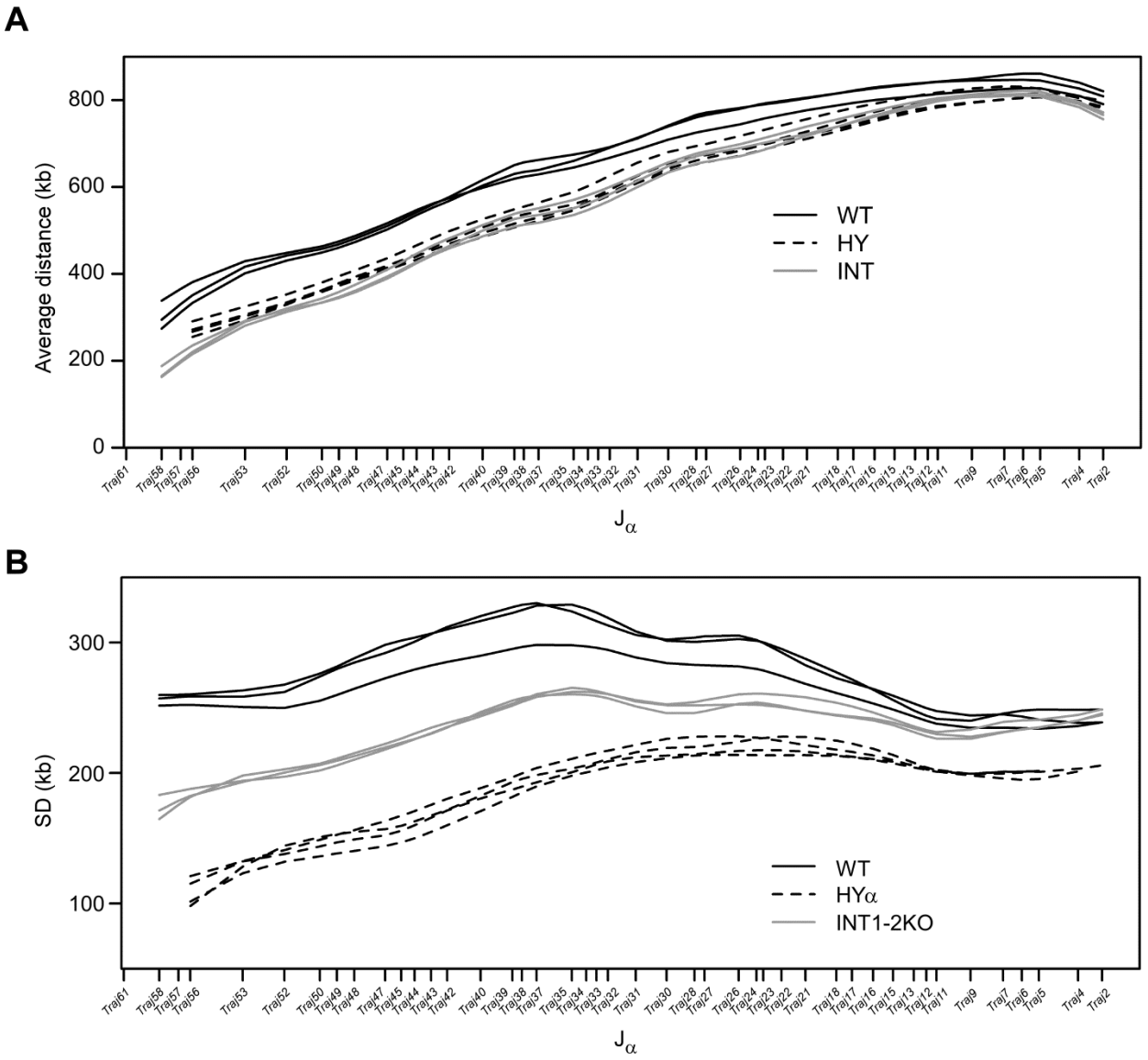

Figure S4. Quantitative analysis of V_α-J_α rearrangements in wild-type, HYα, and INT1-2-deficient thymocytes. Related to Figure 5. (A) Frequency-weighted means of the chromosomal coordinates of the V_α segments used with each J_α segment, with V_α chromosomal coordinates expressed as distance (in kb) from *Traj58*, and J_α segments plotted according to their chromosomal coordinates. (B) Frequency-weighted standard deviations of V_α segment usage (in kb) were plotted as a function of J_α segment. Each line was generated by fitting a smoothed spline function to the data from a single mouse. Statistical tests are presented in Table S2.

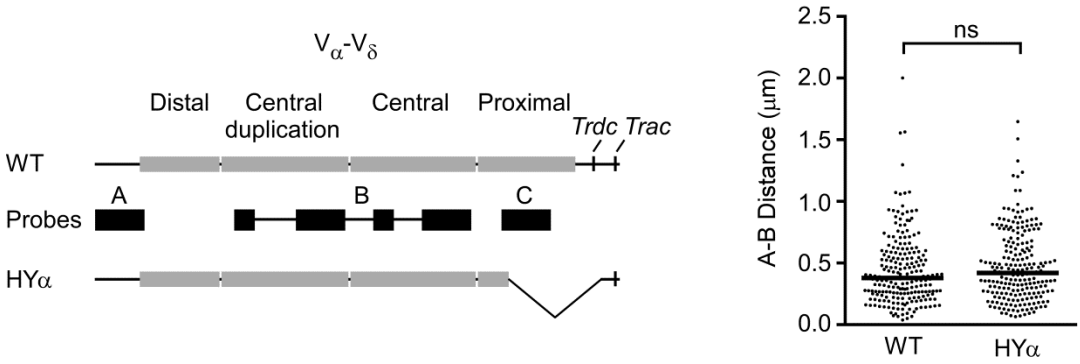

Figure S5. Extended V_α-V_δ conformation on wild-type and HYα *Tcra* alleles. Related to Figure 6. (A) Schematic of hybridization of DNA-FISH probes A, B, and C to wild-type and HYα *Tcra* alleles. (B) Distance between centers of probe A and B hybridization on individual alleles in DP thymocytes generated from *Rag2*$^{-/-}$ heterozygous for the HYα allele. The horizontal line denotes median distance. Measurements were compiled from 216 nuclei examined from slides prepared from three mice analyzed in two independent experiments. ns, not significant (P=0.16) by Mann-Whitney U test.

Table S1. Quality of repertoire analyses. Related to Figure 1.

		No. of sequences passing MiXCR QC				No. of unique clones			
Sample		Rep. 1	Rep. 2	Rep. 3	Rep.4	Rep. 1	Rep. 2	Rep. 3	Rep.4
wild-type DP		618,929	605,294	959,253	-	67,098	118,262	140,982	-
HYα DP		219,828	730,265	1,623,523	969,911	62,124	92,711	182,959	118,118
INT1-2-KO DP		885,547	558,082	960,521	-	107,483	83,785	125,284	-
HYα spleen		1,335,088	1,686,050	-	-	139,923	111,370	-	-
$ZsGreen^{+}$	12hr	958,297	1,115,107	-	-	25,758	31,480	-	-
	24hr	623,464	1,068,044	-	-	86,477	11,549	-	-
	48hr	923,259	1,270,032	-	-	146,222	34,606	-	-
	72hr	1,264,011	877053	-	-	187,444	14,753	-	-
$ZsGreen^{-}$	12hr	499,955	-	-	-	11,992	-	-	-
	24hr	733,689	-	-	-	14,409	-	-	-
	48hr	946,922	1,042,691	-	-	134,768	134,811	-	-

Table S2. V_α usage in different genotypes. Related to Figure 4.

	Average V_α position				V_α position variance			
Genotype	Initial V_α (SEM)[1]	p-value vs. WT[2]	Slope (SEM)[3]	p-value vs. WT[2]	Initial SD (SEM)[1]	p-value vs. WT[2]	Slope (SEM)[3]	p-value vs. WT[2]
Wild-type	331,660 ($1.56x10^4$)	–	14.41 (0.89)	–	240,876 ($6.6x10^3$)	–	3.66 (0.25)	–
HYα	259,777 ($2.06x10^4$)	0.01	12.76 (1.18)	0.165	107,898 ($8.7x10^3$)	<0.001	3.89 (0.33)	0.478
INT1-2KO	208,042 ($2.20x10^4$)	0.001	15.29 (1.26)	0.488	168,685 ($9.3x10^3$)	<0.001	4.11 (0.35)	0.2

[1]Distance in base pairs (bp) from the most V_α-proximal J_α segment.
[2]As determined by two-factor ANOVA.
[3]Expressed as V_α bp/J_α bp, determined between *Traj52* and *Traj37*