High-throughput sequencing reveals similar molecular signatures for class switch recombination junctions for the γ and α isotypes

Hussein Issaoui, Nour Ghazzaui, Alexis Saintamand, Yves Denizot, François Boyer

High throughput sequencing reveals similar molecular signatures for class switch recombination junctions for the γ and α isotypes.

Hussein Issaoui[1,2], Nour Ghazzaui[1,2], Alexis Saintamand[4], Yves Denizot[1,3], François Boyer[1]

[1]UMR CNRS 7276, INSERM U1262, Limoges University, CBRS, rue Pr. Descottes, 87025 Limoges, France

Running title: γ and α CSR

[2]N.G. and H.I. contributed equally to this work.

[3]To whom correspondence should be addressed. Email: yves.denizot@unilim.fr

[4]present address: INSERM U1236, Rennes 1 University, France

Acknowledgments: This work was supported by ANR (projet EpiSwitch-3'RR 2016). N.G. was supported by a grant from Association de Spécialisation et d'Orientation Scientifique (Lebanon) and the municipality of Khiam (Lebanon). H.I. is supported by CORC (FJA/NP 2015-109) and University of Limoges. The authors are "Equipe Labellisée LIGUE 2018". We thank the GénoLim platform and the Cytometry platform of the University of Limoges for sequencing and cell sorting. F.B. is supported by Fondation Partenariale de l'Université de Limoges and ALURAD.

The authors declare no conflict of interest.

Author contributions: Y.D designed research. N.G. and H.I. performed research. A.S., F.B and Y.D. analysed data. N.G., H.I., F.B and Y.D. wrote the paper.

After encountering antigen, B-cells undergo class switch recombination (CSR) that substitutes the C_μ gene with C_γ, C_ε or C_α, thereby generating IgG, IgE and IgA antibodies with same antigenic specificity but new effector functions (1). The DNA-editing enzyme activation-induced deaminase (AID) is essential for CSR by targeting switch (S) regions preceding C_μ (namely the S_μ donor region) and C_γ, C_ε and C_α genes (namely the $S_{\gamma,\varepsilon,\alpha}$ acceptor regions) (1). *Cis*-and *trans*-controlled DNA double strand beaks are generated during this process (2-6). The recruitment of DNA repair factors that facilitate the end-joining process is a crucial step of class switch recombination. Two pathways are implicated in this end joining. The classical non-homogenous end joining (c-NHEJ) pathway ligates DNA ends with no or little homology. In contrast, the alternative end joining (A-EJ) pathways is used to ligate DNA ends with microhomology (3-6). Previous reports have suggested that IgG and IgA CSR might be differently regulated with a preferential use of c-NHEJ for γ CSR and A-EJ for α CSR (7, 8). We recently reported a computational tool (CSReport) for automatic analysis of CSR junctions sequenced by high-throughput sequencing (9) and used it to analyze the rare S_μ-σ_δ junctions form during IgD CSR (10, 11). We thus used CSReport and high-throughput sequencing to analyze the molecular signature of S_μ-$S_{\gamma 3}$, S_μ-$S_{\gamma 1}$ and S_μ-S_α junctions in *wt* mice more in depth.

Our research has been approved by our local ethics committee review board (Comité Régional d'Ethique sur l'Expérimentation Animale du Limousin, Limoges, France) and carried according the European guidelines for animal experimentation. Single-cell suspensions of spleen cells from wild-type (*wt*) 129 mice were cultured 4 days at $1x10^6$ cells/ml in RPMI 1640 with 10% fetal calf serum (FCS) and 5 μg/ml LPS, with or without addition of 20 ng/ml IL-4 or 2 ng/ml TGFβ (PeproTech, Rocky Hill, NJ). Splenocyte DNA was then extracted for investigation of S_μ-$S_{\gamma 3}$, S_μ-$S_{\gamma 1}$ and S_μ-S_α junctions. As previously described in detail (9), junctions were PCR amplified. Libraries of 200bp were prepared from the 1-2kb PCR

products of S_μ-$S_{\gamma1}$, S_μ-$S_{\gamma3}$ and S_μ-S_α amplification for Ion Proton sequencing ("GénoLim platform" of the Limoges University, France). Sequenced reads were then mapped to S_μ and acceptor $S_{\gamma1}$, $S_{\gamma3}$ and S_α regions using BLAST algorithm. The computational tool developed for experiments performs junction assembly, identifies breakpoints in S_μ, $S_{\gamma1}$, $S_{\gamma3}$ S_α, identifies junction structure (blunt, micro-homology or junction with insertions) and outputs a statistical summarization of identified junctions.

LPS, LPS+IL4 and LPS+TGFβ stimulated B-cell CSR to IgG3, IgG1 and IgA, respectively (2, 5, 12). We detected 4140, 3798 and 1955 S_μ-$S_{\gamma1}$, S_μ-$S_{\gamma3}$ and S_μ-S_α junctions, respectively. The structural profiles of all these junctions (blunt, micro-homology or junction with insertions) are reported in Fig. 1A. The positions of IgG_1, IgG_3 and IgA junctions in terms of distance from the forward PCR primer in S_μ are reported in Fig. 1B. Localizations of S_μ breakpoints within AID hotspots (AGCT, WRCY, RGYW) and other motifs are shown in Fig. 1C (both displayed along S_μ region and expressed in % of junctions). Analysis of 5000 synthetic junctions simulated from the random association of 100-bp S_μ segments with 100-bp segments of $S_{\gamma1}$, $S_{\gamma3}$ or S_α revealed a similar pattern of blunt and micro-homology junctions (junction with insertions are not produced with this numerical approach) (Fig. 1D) compared with true junctions (Fig. 1A). As shown in Fig1E, the frequency of S_μ-$S_{\gamma1}$, S_μ-$S_{\gamma3}$ and S_μ-S_α junctions with large (>5 bp) micro-homology is found higher in S_μ-S_α compared to S_μ-$S_{\gamma1}$ and S_μ-$S_{\gamma3}$, both for sequenced junctions and for randomly simulated junctions. Fig. 1F shows dotplots for S_μ vs $S_{\gamma1}$, S_μ vs $S_{\gamma3}$ and S_μ vs S_α sequence comparisons.

Confirming the validity of our technical approach, the structural profile of our S_μ-$S_{\gamma1}$ junctions was similar to that reported few weeks ago using a high-throughput translocation sequencing method (5). We also confirm the previously reported slight increase of small insertions for S_μ-S_γ compared to S_μ-S_α (8). Finally we demonstrated that the slight increase of

S_μ-S_α junctions with large micro-homology can be numerically reproduced using randomly generated synthetic junctions. As this simulation mimics a pure NHEJ process (linking two free DNA ends without any resection), it evidences that it is not necessary to invoke another molecular mechanism (such A-EJ) to explain the observed structure alteration of S_μ-S_α junctions. Those micro-homologies arise solely by chance and are favored by the repetitive structure of the S_α region and its high degree of similarity to S_μ region. In conclusion the structural profiles of S_μ-$S_{\gamma 1}$, S_μ-$S_{\gamma 3}$ and S_μ-S_α junctions are similar indicating same CSR process and partners whatever cytokine stimulations, length of the S acceptor region and its distance with the S_μ donor region. Analysis of the molecular signature of CSR junctions does not argue in favour of a preferential use of c-NHEJ and A-EJ for γ CSR and α CSR, respectively.

References

1. Pinaud E, Marquet M, Fiancette R, Péron S, Vincent-Fabert C, Denizot Y, Cogné M. The IgH locus 3' regulatory region: pulling the strings from behind. *Adv Immunol* 2011; **110:** 27-70.

2. Saintamand A, Rouaud P, Saad F, Rios G, Cogné M, Denizot Y. Elucidation of IgH 3' region regulatory role during class switch recombination via germline deletion. *Nature Commun* 2015; **6:** 7084.

3. Methot SP, Di Noia JM. Molecular mechanisms of somatic hypermutation and class switch recombination. *Adv Immunol* 2017; **133:** 37-87.

4. Chaudhuri J, Alt FW. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nat Rev Immunol* 2004; **4:** 541-552.

5. Panchakshari RA, Zhang X, Kumar V, Du Z, Wei PC, Kao J et al. DNA double-strand break response factors influence end-joining features of IgH class switch and general translocation junctions. *Proc Natl Acad Sci USA* 2018; **115:** 762-767.

6. Matthews A, Zheng S, DiMenna LJ, Chaudhuri J. Regulation of immunoglobulin class-switch recombination: choregraphy of noncoding transcription, targeted DNA deamination, and long-range DNA repair. *Adv Immunol* 2014; **122:** 1-57.

7. Pan-Hammarström Q, Jones AM, Lähdesmäki A, Zhou W, Gatti RA, Hammarström L, Gennery AR, Ehrenstein MR. Impact of DNA ligase IV on nonhomologous end joining pathways during class switch recombination in human cells. *J Exp Med* 2005; **201:** 189-194.

8. Björkman A, Du L, Felgentreff K, Rosner C, Kamdar RP, Kokaraki G, Matsumoto Y, Davies EG, van der Burg M, Notarangelo LD, Hammarström L, Pan-Hammarström Q. DNA-PKcs is involved in Ig class switch recombination in human B cells. *J Immunol* 2015; **195:** 5608-5615.

9. Boyer F, Boutouil H, Dalloul I, Dalloul Z, Cook-Moreau J, Aldigier JC, Carrion C, Herve B, Scaon E, Cogné M, Péron S. CSReport: a new computational tool designed for automatic analysis of class switch recombination junctions sequenced by high-throughput sequencing. *J Immunol* 2017; **198:** 4148-4155.

10. Ghazzaui N, Issaoui H, Saintamand A, Boyer F, Denizot Y. Analysis of IgD CSR junctions by high-throughput sequencing. *Immunol Lett* 2017; **188:** 86-88.

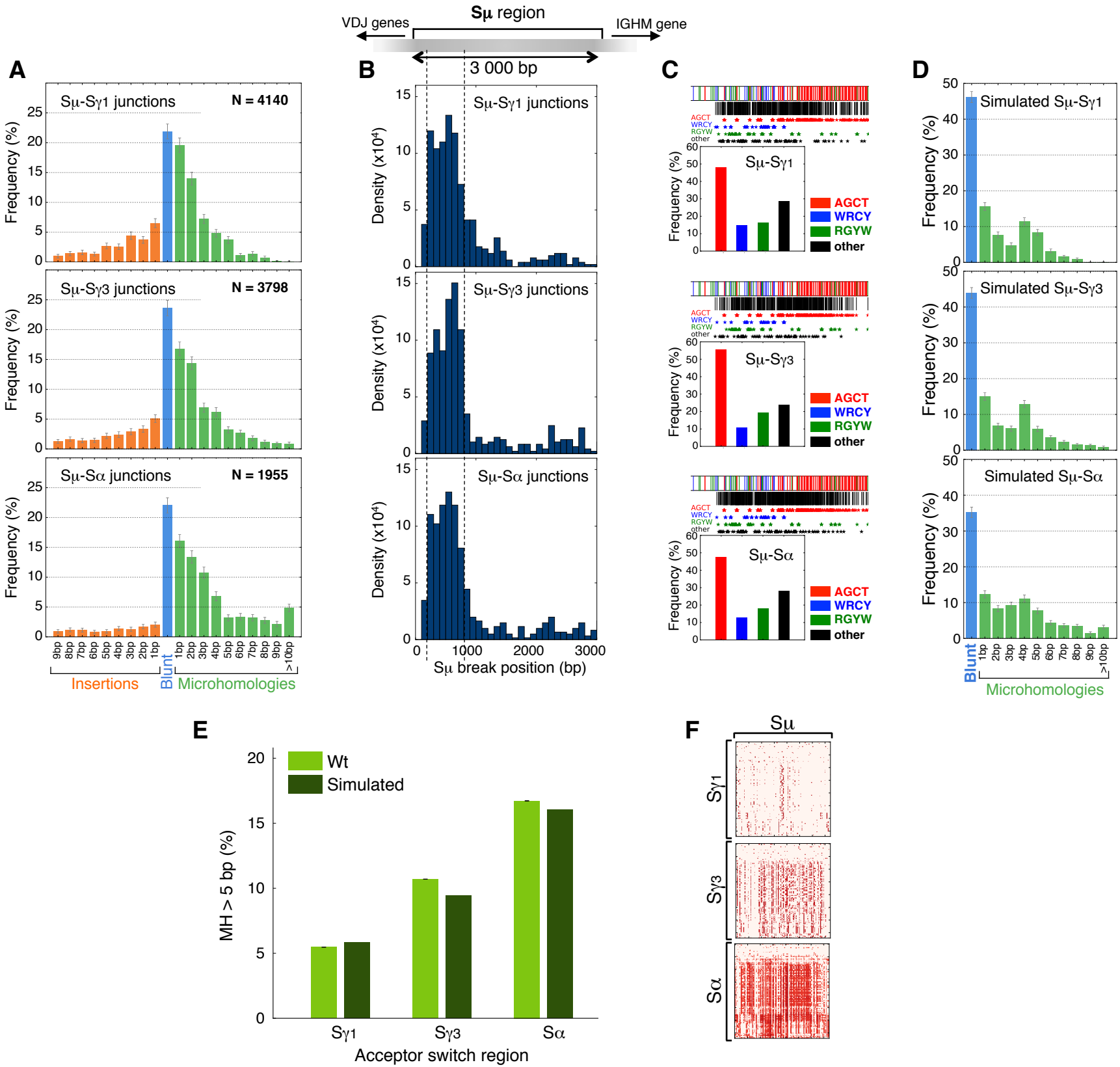
11. Issaoui H, Ghazzaui N, Saintamand A, Denizot Y, Boyer F. IgD class switch recombination is not controlled through the immunoglobulin heavy chain 3’ regulatory region super-enhancer. *Cell Mol Immunol* 2017;**14:** 871–874.

12. Issaoui H, Ghazzaui N, Saintamand A, Carrion C, Oblet C, Denizot Y. The IgH 3' regulatory region super-enhancer does not control IgA class switch recombination in the B1 lineage. *Cell Mol Immunol* 2018. in press. doi: 10.1038/cmi.2017.103. [Epub ahead of print]

Legend to Figure 1.

CSR in *wt* mice.

A: Structure profiles of S_μ-$S_{\gamma1}$, S_μ-$S_{\gamma3}$ and S_μ-S_α junctions. Junctions are classified in terms of junction types (junction with insertions, blunt junction or junction with micro-homology). B: Breakpoint localizations in S_μ for S_μ-$S_{\gamma1}$, S_μ-$S_{\gamma3}$ and S_μ-S_α junctions. C: (top) Location of breakpoints in respect of AID hotspots AGCT, WRCY, RGYW and other motifs along the first 1kb in S_μ. Identified breaks are shown as a black line, colocation with a sequence motif is indicated with a colored asterisk. (bottom) Frequency of hotspot/break colocation events. D: Structure profiles of synthetic S_μ-$S_{\gamma1}$, S_μ-$S_{\gamma3}$ and S_μ-S_α junctions. Three sets of 5000 synthetic junctions were simulated from the random association of 100-bp S_μ segments with 100-bp segments of $S_{\gamma1}$, $S_{\gamma3}$ or S_α and analyzed with the same computational tool (CSReport) as sequencing reads. E: Frequency of junctions with long micro-homology (>5bp) for S_μ-$S_{\gamma1}$, S_μ-$S_{\gamma3}$ and S_μ-S_α junctions. Simulated datasets are compared to the junctions identified from high-throughput sequencing. F: Sequence similarity dotplots of S_μ *vs* $S_{\gamma1}$, S_μ *vs* $S_{\gamma3}$ and S_μ *vs* S_α . Similarity was evaluated with a 20-bp window and reported when greater than 60%.


VDJ genes
Sμ region
IGHM gene
3 000 bp
A
Sμ-Sγ1 junctions
N = 4140
Sμ-Sγ3 junctions
N = 3798
Sμ-Sα junctions
N = 1955
Frequency (%)
9bp 8bp 7bp 6bp 5bp 4bp 3bp 2bp 1bp Blunt 1bp 2bp 3bp 4bp 5bp 6bp 7bp 8bp 9bp >10bp
Insertions
Microhomologies
B
Density (x10^4)
Sμ break position (bp)
C
AGCT
WRCY
RGYW
other
D
Simulated Sμ-Sγ1
Simulated Sμ-Sγ3
Simulated Sμ-Sα
E
Wt
Simulated
MH > 5 bp (%)
Sγ1
Sγ3
Sα
Acceptor switch region
F
Sμ