Published in final edited form as:
Nature. 2017 August 31; 548(7669): 597–601. doi:10.1038/nature23670.

Public antibodies to malaria antigens generated by two *LAIR1* insertion modalities

Kathrin Pieper[#1], **Joshua Tan**[#1,2], **Luca Piccoli**[#1], **Mathilde Foglierini**[1,3], **Sonia Barbieri**[1], **Yiwei Chen**[1,4], **Chiara Silacci-Fregni**[1], **Tobias Wolf**[1,4], **David Jarrossay**[1], **Marica Anderle**[1], **Abdirahman Abdi**[5], **Francis M. Ndungu**[5], **Ogobara K. Doumbo**[6], **Boubacar Traore**[6], **Tuan M. Tran**[7], **Said Jongo**[8], **Isabelle Zenklusen**[9,10], **Peter D. Crompton**[11], **Claudia Daubenberger**[9,10], **Peter C. Bull**[12], **Federica Sallusto**[1,4], and **Antonio Lanzavecchia**[1,4]

[1]Institute for Research in Biomedicine, Università della Svizzera Italiana, Via Vincenzo Vela 6, 6500 Bellinzona, Switzerland [2]Nuffield Department of Clinical Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford, OX3 9DU, UK [3]Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland [4]Institute for Microbiology, ETH Zurich, Wolfgang-Pauli-Strasse 10, 8093 Zurich, Switzerland [5]KEMRI-Wellcome Trust Research Programme, CGMRC, PO Box 230, 80108 Kilifi, Kenya [6]Malaria Research and Training Centre, University of Sciences, Technique, and Technology of Bamako, Mali [7]Division of Infectious Diseases, Indiana University School of Medicine, Indianapolis, IN, USA [8]Ifakara Health Institute, Bagamoyo Clinical Trial Unit, Tanzania [9]Swiss Tropical and Public Health Institute, Clinical Immunology Unit, 4002 Basel, Switzerland [10]University of Basel, Petersplatz 1, 4003 Basel, Switzerland [11]Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD, USA [12]Department of Pathology, University of Cambridge, Cambridge, UK

[#] These authors contributed equally to this work.

Abstract

We previously described two donors in whom the extracellular domain of LAIR1, a collagen-binding inhibitory receptor encoded on chromosome 191, was inserted between the V and the DJ segments of an antibody. This insertion generated, through somatic mutations, broadly reactive antibodies against RIFINs, a type of variant antigen expressed on the surface of *Plasmodium falciparum*-infected erythrocytes (IEs)2. To investigate how frequently such antibodies are produced in response to malaria infection, we screened plasma from two large cohorts of

Correspondence and requests for materials should be addressed to A.L. (lanzavecchia@irb.usi.ch).

Author Contributions K.P. characterized genomic DNA, analyzed the data and wrote the manuscript; J.T. isolated new LAIR1-containing antibodies, analyzed the data and wrote the manuscript; L.P. produced mutant antibodies, analyzed the data and wrote the manuscript; M.F. performed bioinformatics analysis; S.B. sequenced and expressed antibodies; Y.C. helped in genomic sequencing; C.S.F. immortalized memory B cells; T.W. helped in MinION sequencing; D.J. performed cell sorting and analysis; M.A. performed protein analysis; A.A. performed *P. falciparum* culture; F.M.N., S.J., O.K.D., B.T., I.Z., C.D. and P.C.B. provided cohort samples; T.M.T. and P.D.C. provided cohort samples and analysed the relationship between LAIR1-containing antibodies and malaria risk; F.S. provided supervision; A.L. provided overall supervision, analyzed the data and wrote the manuscript.

The authors declare competing financial interests. A.L. is the scientific founder and shareholder of Humabs BioMed. F.S. is a shareholder of Humabs BioMed.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

individuals living in malaria-endemic regions. We report that 5-10% of malaria-exposed individuals, but none of the European blood donors tested, have high levels of LAIR1-containing antibodies that dominate the response to IEs without conferring enhanced protection against febrile malaria. By analyzing the antibody-producing B cell clones at the protein, cDNA and gDNA level, we characterized additional *LAIR1* insertions between the V and DJ segments and discovered a second insertion modality whereby the *LAIR1* exon encoding the extracellular domain and flanking intronic sequences are inserted into the switch region. By exon shuffling, this mechanism leads to the production of bispecific antibodies in which the LAIR1 domain is precisely positioned at the elbow between the VH and CH1 domains. Additionally, in one donor the gDNA encoding the VH and CH1 domains was deleted, leading to the production of a camel-like LAIR1-containing antibody. Sequencing of the switch regions of memory B cells from European blood donors revealed frequent templated inserts originating from transcribed genes that, in rare cases, comprised exons with orientation and frame compatible with expression. Collectively, these results reveal different modalities of *LAIR1* insertion that lead to public and dominant antibodies against IEs and suggest that insertion of templated DNA represents an additional mechanism of antibody diversification that can be selected in the immune response against pathogens and exploited for B cell engineering.

LAIR1-containing antibodies were initially isolated from two Kenyan donors who were selected from a large cohort of more than 500 individuals for their capacity to produce broadly reactive antibodies to *P. falciparum*-IEs2. To establish the prevalence of LAIR1-containing antibodies among malaria-exposed individuals, we screened plasma samples from two large cohorts in Tanzania3 and Mali4. To identify LAIR1-containing antibodies irrespective of their specificity for parasite isolates, we developed a two-determinant immunoassay using beads coated with anti-LAIR1 or a control antibody. Six out of 112 Tanzanian donors (5.4%) and 57 out of 656 Malian donors (8.7%) had detectable levels of LAIR1-containing IgG (Fig. 1a). In addition, 2-4% of African donors had LAIR1-containing IgM, with no or variable levels of LAIR1-containing IgG (Fig. 1b, c). In contrast, only 3 and 4 out of 1043 European blood donors showed a low positivity in the IgG and IgM LAIR1 assay, respectively. The presence of LAIR1-containing antibodies was confirmed by the isolation of 52 immortalized B cell clones from seven East and West African donors (Extended Data Table 1 and Supplementary Table 1), while we were not able to isolate LAIR1-containing monoclonal antibodies from four European donors that showed serum reactivity. The finding that 5-10% of individuals living in malaria-endemic regions produce LAIR1-containing antibodies is suggestive of a public antibody response.

To investigate the contribution of LAIR1-containing antibodies to the response to IEs, we dissected this response at the polyclonal and monoclonal level. Staining of IEs with plasma from selected individuals with LAIR1-containing antibodies revealed that the majority of IEs was recognized by the LAIR1 antibodies, while only a minority of IEs was recognized by conventional IgG (see examples in Fig. 1d). Furthermore, when the immortalized memory B cell clones from four donors were analyzed for reactivity to IEs and for the presence of LAIR1, all of the anti-IE monoclonal antibodies from three donors and most of the antibodies from the fourth donor contained the *LAIR1* insert (Fig. 1e). These findings suggest that in certain individuals circulating antibody and memory B cell responses are

dominated by LAIR1-containing antibodies, a finding that may be explained both by their breadth and by clonal expansion.

To investigate the nature of the *LAIR1* insertion, we sequenced cDNA and gDNA from B cell clones isolated from different individuals. As previously reported for the first two Kenyan donors2, the B cell clones isolated from four Malian and Tanzanian donors (E, F, O and Q) contained an insertion of the *LAIR1* exon with flanking intronic sequences between the V and DJ segments, positioning the LAIR1 domain in the CDR3 loop (Fig. 2a and Extended Data Fig. 1). The size of the insert and the partial splicing of the upstream intronic region differed between donors, but were identical in the sister clones isolated from each individual, indicating that in each donor the LAIR1-containing antibody response is monoclonal.

Strikingly, B cell clones from three additional donors showed a different insertion modality (Fig. 2b-d). The cDNA of clones isolated from donors M (Malian) and J (Tanzanian) contained only the *LAIR1* exon, which was precisely located between the JH and CH1. In both cases, gDNA analysis revealed that a DNA fragment comprising the *LAIR1* exon flanked by intronic sequences was inserted into the switch μ region (Extended Data Fig. 2) and, by alternative splicing, gave rise to two mRNA variants with or without the *LAIR1* insert. This was confirmed by the production of antibodies with or without LAIR1 in similar proportions by a single B cell clone (Fig. 2e). Another example of LAIR1 insertion into the switch region was observed in donor B (Kenyan), from whom we isolated a B cell clone (MGB47) producing a truncated LAIR1-containing IgG3 heavy chain without an attached light chain (Fig. 2f, g). In this clone, the gDNA carried multiple deletions that removed most of the VDJ and the entire CH1 region, leading to the production of a camel-like antibody5 (Fig. 2d). Taken together, the above findings highlight a new modality of exon insertion in the switch region that can add an extra domain to an antibody.

The two insertion modalities result in the production of antibodies with non-conventional structures in which an additional domain is inserted in the CDR3 or in the elbow between the VH and CH1 (Fig. 3a). To investigate the effect of the insert position on antibody specificity, we designed different constructs in which unmutated LAIR1, mutated LAIR1 or other Ig-like domains were inserted into the CDR3 or into the elbow region of an antibody of known specificity that was used as a scaffold (Fig. 3b). Antibody constructs carrying LAIR1 stained IEs and were recognized by an anti-LAIR1 antibody, independent of the LAIR1 position in the scaffold. While insertion of LAIR1 into the CDR3 of an antibody specific for the granulocyte-macrophage colony-stimulating factor (GM-CSF)6 abolished binding to GM-CSF, insertion of LAIR1, programmed cell death-1 (PD1), or signaling lymphocytic activation molecule family member (SLAM) Ig-like domains into the elbow region did not affect binding to GM-CSF. This indicates that the VH-CH1 elbow is permissive for insertions of different domains without affecting the original antibody specificity and may therefore be suitable for the generation of bispecific antibodies.

To analyze the role of somatic mutations we aligned and compared the LAIR1 sequences of 52 antibodies (Extended Data Fig. 3 and Extended Data Table 1). *LAIR1* inserts between the V and DJ segments carried several amino acid substitutions clustering at hot spots around

positions 67, 77 and 102 that determined distinct patterns of reactivity with parasite isolates, as well as loss of collagen binding (Fig. 3c). In contrast, LAIR1 inserted into the switch region carried only a few substitutions, which were however sufficient to abolish collagen binding. In particular, the camel-like antibody MGB47, which had the highest level of amino acid substitutions among those with inserts in the switch region, showed a considerable breadth since it stained 8 out of the 9 parasites tested. Interestingly, we found that unmutated LAIR1 bound to a few isolates when tested at 1 μg ml^{-1} and to all parasites when tested at a 100-fold higher concentration (Fig. 3c). We conclude that the unmutated LAIR1 domain binds with low affinity to most parasite isolates and that mutations can increase affinity and modify the spectrum of cross-reactivity. Furthermore, the finding that collagen binding is lost even in cases where the somatic mutation mechanism is less effective, as in the case of insertions into the switch region, suggests that there is strong pressure to “redeem” this BCR from autoreactivity7.

The insertion of an extra exon in the switch region represents a new modality of antibody diversification, analogous to exon shuffling8,9, that has the potential to generate a panoply of bispecific antibodies. To ask how generally and how frequently templated DNA sequences are inserted in the switch region, we isolated gDNA from switched memory B cells of European blood donors, amplified the switch regions and sequenced them using the Illumina platform (Extended Data Fig. 4). Using a bioinformatics pipeline, we identified templated inserts at a frequency of approximately one in $>10^3$ B cells, with the length of the inserts ranging from <100 to >1000 nucleotides (Fig. 4a, Extended Data Fig. 5 and Supplementary Tables 2–4). Switch region inserts could also be detected using long-read MinION sequencing of intact amplicons, which provided a suitable platform for insert identification, in spite of its high error rate10. Using MinION, we confirmed the identity of several inserts using biological replicates and estimated a higher frequency of templated inserts in the range of one in a few hundred switched memory B cells (Fig. 4a and Extended Data Fig. 6). In contrast, no insert was detected in the switch region of naïve B cells (Fig 4c). Interestingly, most of the inserts were derived from genic regions from all chromosomes and in particular from genes expressed in B cells, such as *PAX5* and *EBF1* (Fig. 4b,e,f and Extended Data Fig. 7). The genic inserts, which account for 75% of all inserts, were derived from introns, exons and, in some cases, comprised an entire exon with preserved splice sites (Fig. 4d). A fraction of the latter had the correct frame and orientation for the potential expression of an extra protein sequence in the Ig elbow (Extended Data Fig. 8). Taken together, the above findings indicate that templated inserts derived from transcribed genes are frequently found in the switch regions of memory B cells.

We have shown that LAIR1-containing antibodies are produced by two insertion modalities and have a prevalence of 5-10% of individuals exposed to malaria infection, suggesting that they may contribute to acquired immunity to blood-stage parasites. However, in spite of their breadth and opsonizing activity2, the presence of LAIR1-containing antibodies did not confer improved protection against febrile malaria (Extended Data Fig. 9), a finding that may be explained by the fact that the LAIR1-containing antibodies recognize only a fraction of cultured parasites and may allow the selection of escape mutants. A thorough investigation of the role of the LAIR1-containing antibodies *in vivo* will require the isolation of autologous parasites from the individuals who possess these antibodies.

It is unusual that in all cases observed so far, the LAIR1-containing antibodies are produced by a single expanded B cell clone that dominates the antibody response to IEs in these individuals. The finding that unmutated LAIR1 has the inherent ability to bind to IEs explains how the insertion of this domain results in the generation of public antibodies with a common specificity. Furthermore, the fact that this domain binds to all parasite isolates tested, albeit with low affinity, suggests a mechanism for the extraordinary clonal expansion and selection of mutated antibody variants with improved affinity and breadth by repeated infections with different *P. falciparum* parasites. These findings illustrate, in a biologically relevant system, the power of clonal selection driven by both antigen binding and loss of self-reactivity. Furthermore, the binding of IEs to LAIR1 suggests the possibility that the parasite might target this inhibitory receptor to modulate the host immune response.

The *LAIR1* insertion in the switch region resulting in the expression of a new domain in the elbow between the VH and CH1 domains represents a new and possibly general example of protein engineering by exon shuffling8,9, as suggested by the frequent occurrence of templated inserts. While insertions in the CDR3 loop result in monospecific antibodies, insertions in the switch region do not affect the specificity of the original antibody but rather add a second specificity which is found in approximately half of the antibodies produced due to alternative splicing. We are not aware of deliberate attempts to engineer the antibody elbow, and in this context, nature has shown considerable ingenuity by taking advantage of the exon shuffling principle.

It has been reported that in mice, chronic *P. chabaudi* infection promotes genomic instability leading to chromosomal translocations involving the switch region11. Although we cannot exclude that the LAIR1 insertions in the switch region observed in African donors may have been promoted by malaria infection, the frequent templated insertions found in European blood donors would be consistent with a general mechanism that does not necessarily rely on malaria infection. Nevertheless, the data suggest that it is the exposure to the malaria parasite that selects the rare B cells with a LAIR1 insertion in the VDJ or switch region. Finally, it remains to be established whether chromosomal translocations and templated insertions share a common mechanism.

The finding of templated inserts in the switch region of switched memory B cells, but not naïve B cells, suggests that the insertions occur in germinal centers as a consequence of unconventional repair of AID-induced double-strand DNA breaks. Similarly, repair of RAG-induced double-strand breaks during B cell development in the bone marrow may give rise to insertions in the CDR3. Our finding that switch region inserts are derived from transcribed genes suggests an involvement of nascent RNA as an insert template and recent publications have shown that nascent RNA12,13 as well as foreign RNA14,15 can be used to repair double-strand DNA breaks. Furthermore, multiple templated inserts from transcribed genes have been observed in engineered double-strand breaks in a human cell line and mouse pro-B cells14,16. However, we cannot exclude accessible DNA as a primary substrate. The possibility of inducing human naïve B cells to switch *in vitro* offers an opportunity to study the mechanism of templated insertions and to engineer B cells by manipulating the factors involved in DNA repair and by offering different substrates.

Materials and Methods

Serum and plasma samples

Kenyan plasma samples were obtained from adults living in a malaria-endemic region within Kilifi County. Tanzanian serum and plasma samples were obtained from healthy male volunteers, malaria negative at study enrolment and during PBMC collection, HIV and Hepatitis B and C negative, age 25.4 ± 2.8 (mean +/- SD)3. The Mali study was conducted in the rural village of Kalifabougou where intense *P. falciparum* transmission occurs from June through December each year. 610 individuals (310 males and 300 females) were enrolled, age ranging from 1 to 26 years (mean 9). The cohort has been described in detail elsewhere4. Smaller numbers of sera were also obtained from adults in the Fulani and Dogon ethnic groups in Mantéourou, Mali. 48 individuals (28 males and 20 females) were enrolled, age ranging from 21 to 57 years (mean 39.7).

Ethics approval

In all cases, written informed consent was obtained from the participants (or guardians of participating children) before inclusion in the study. The acquisition and use of the Kenyan plasma samples were approved by the Kenya Medical Research Institute Scientific and Ethics Review Unit (protocol number: KEMRI-SERU 3149), as well as the Oxford Tropical Research Ethics Committee. The Tanzanian samples were obtained with informed consent from the trial participants. The clinical trial was conducted according to Good Clinical Practices and with authorization from the Institutional Review Boards of the Ifakara Health Institute, the National Institute for Medical Research Tanzania, the Tanzanian Food and Drugs Authority and the Commission cantonale d'éthique de la recherche sur l'être humain du canton de Vaud, Switzerland. The trial is registered at ClinicalTrials.gov Identifier: NCT01949909. The Mali study was approved by The Ethics Committee of the Faculty of Medicine, Pharmacy and Dentistry at the University of Sciences, Technique and Technology of Bamako, and the Institutional Review Board of the National Institute of Allergy and Infectious Diseases, National Institutes of Health. Written informed consent was obtained from participants or parents or guardians of participating children prior to inclusion in the Mali study.

Parasite culture

P. falciparum parasites were initially obtained from children who were diagnosed with malaria in Kilifi County, Kenya. The parasites were adapted to *in vitro* culture and cultivated using standard protocols17. The *P. falciparum* laboratory line 3D7 was also cultured under the same conditions. $3D7^{D21}$ is derived by enrichment of 3D7 parasites reactive with the LAIR1-containing monoclonal antibody MGD21. Parasites were cryopreserved at the late trophozoite stage in small aliquots for subsequent use in assays.

Screening of sera or plasma with bead-based immunoassay

Sera or plasma were tested for the presence of LAIR1-containing antibodies using a two-determinant bead-based immunoassay. Anti-goat IgG microbeads (Spherotech) were coated with goat anti-human LAIR1 (R&D Systems, AF2664) and 40× SYBR Green I

(ThermoFisher Scientific) or with goat anti-human EGF (R&D Systems, AF-259-NA) without SYBR Green I for 20 min at room temperature. The beads were washed, mixed, and incubated with the sera at a 1/30 dilution for 1 h at room temperature under shaking conditions. Beads coated with anti-LAIR1 were differentiated from control beads coated with anti-EGF based on SYBR Green staining. Serum antibody binding was detected using 2.5 μg ml^{-1} Alexa Fluor 647-conjugated donkey anti-human IgG (Jackson ImmunoResearch, 709-606-098) or Alexa Fluor 647-conjugated donkey anti-human IgM (Jackson ImmunoResearch, 709-606-073). FACS Diva (version 6.2) was used for acquisition of samples and Flow-Jo (version 10.1) was used for FACS analysis. Delta MFI (median fluorescence intensity) was calculated by subtracting the MFI of the anti-EGF control beads from that of the anti-LAIR1 beads in the IgG or IgM channel.

B-cell immortalization and isolation of monoclonal antibodies

IgM or IgG memory B cells were isolated from frozen peripheral blood mononuclear cells (PBMCs) by magnetic cell sorting with anti-CD19-PECy7 antibodies (BD, 341113) and mouse anti-PE microbeads (Miltenyi Biotec, 130-048-081), followed by FACS sorting using goat Alexa Fluor 647-conjugated anti-human IgG (Jackson ImmunoResearch, 109-606-170) or anti-human IgM (Jackson ImmunoResearch, 109-606-129) and PE-labeled anti-human IgD (BD, 555779). As previously described18, sorted B cells were immortalized with Epstein-Barr virus (EBV) and plated in single cell cultures in the presence of CpG-DNA (2.5 μg ml^{-1}) and irradiated PBMC-feeder cells. Two weeks post-immortalization, the culture supernatants were tested (at a 2/3 dilution) for the presence of LAIR1-containing antibodies using the bead-based immunoassay described above. For several donors, the culture supernatants were also tested for the ability to bind to IEs from a mixture of four parasite isolates (3D7-$MGD21^{+}$, 9106, 9605 and 11019) by flow cytometry. Briefly, cryopreserved IEs were thawed, stained with 10× SYBR Green I for 30 min at room temperature, and incubated with the B-cell supernatants for 1 hour at 4°C. Detection of antibody binding was done with 2.5 μg ml^{-1} Alexa Fluor 647-conjugated goat anti-human IgG.

Sequence analysis of antibody cDNA

cDNA was synthesized from selected B-cell cultures and both heavy chain and light chain variable regions (VH and VL) were sequenced as previously described19. The usage of VH and VL genes and the number of somatic mutations were determined by analyzing the homology of VH and VL sequences of monoclonal antibodies to known human V, D and J genes in the IMGT database20. Antibody-encoding sequences were amplified and sequenced with primers specific for the V and J regions of the given antibody. Sequences were aligned with Clustal Omega21.

Mutation analysis of *LAIR1* inserts

The number of somatic mutations in the *LAIR1* inserts was obtained by analyzing the homology of the inserts to the original *LAIR1* genomic sequence (sequence from Ensembl genome database: ENSG00000167613). Amino acid sequences of *LAIR1* inserts of all the antibodies discovered were grouped for each donor and aligned to the original unmutated LAIR1 sequence using Clustal Omega21. The replacement to silent mutation ratio (R/S)

values were calculated at each codon for each donor. Mean (R/S) values above 2.9, indicative of positive selection, were used to highlight hot spots in the LAIR1 extracellular domain (PDB 3RP1) using the BioLuminate software (Schrödinger, LLC, New York, NY, 2016 v2.4).

Production of recombinant antibodies, antibody variants and fusion proteins

Antibody heavy and light chains were cloned into human IgG1, Igκ and Igλ expression vectors and expressed by transient transfection of Expi293F Cells (ThermoFisher Scientific) using polyethylenimine (PEI). Cell lines were routinely tested for mycoplasma contamination. The antibodies were affinity purified by protein A chromatography (GE Healthcare). Variants of the GCE536, MGD21, MGM1 and MGJ5 antibodies were produced by inserting or exchanging the mutated or unmutated *LAIR1* sequences or by substituting these sequences with the Ig-like domains of PD1 and SLAM genes (sequences from Ensembl genome database: ENSG00000188389 and ENSG00000117090, respectively). The antibody constructs were tested for staining of 9215 IEs and binding values (%) at 1 μg ml^{-1} antibody concentration were calculated by interpolation of binding curves fitted to a sigmoidal curve model (Graphpad Prism 7).

Amplification of antibody gDNA

Genomic DNA was isolated from B-cell clones with a commercial kit (QIAGEN). For analysis of the camel-like antibody MGB47, 3'RACE22 with the CH2-γ-REV1 primer (gagaccttgcacttgtactccttgcc) was used for amplification of truncated heavy chain mRNAs. Heavy-chain variable to constant gDNA of MGB47 was amplified using an upstream 5' VH3-21 primer (gggtccatattgtgatcctgagtctggg) and CH2-γ-REV1 as the reverse primer. After PCR amplification with LongAmp Taq Polymerase (New England Biolabs), the 6000bp amplicon was cloned into a vector using the TOPO XL PCR cloning kit (ThermoFisher) and sequenced by plasmid-NGS-sequencing (Microsynth, Switzerland). All other LAIR1 switch and V-DJ inserts were analyzed by PCR amplification of gDNA and Sanger Sequencing using donor-specific forward and universal reverse primers: donE_FW (cctggagggtcttctgcttgctggc), donF_FW (cctcctgctggtggcagctccc), donJ/M_FW1 (atggagtttgggctgagctgggttttcc), donJ_FW2 (gtgagtgaacacgagtgagagaaacagtgg), donM_FW2 (gagtgaacatgagtgagaaaaactggatttgtgtgg), donO/Q_FW (atgaaacatctgtggttctt), 3'J6_REV (ggcatcggaaaatccacagaggctcc), IgG_CH1_REV1 (tcttgtccaccttggtgttgct), IgG_CH1_REV2 (gtagtccttgaccaggcagc), IgM_CH2_REV1 (ggacacctgaatctgccggggactgaaaccc), and IgM_CH2_REV2 (ctggtcaccttgtaggtcgtgggcccag).

Switch region PCR and Illumina sequencing

Genomic DNA (gDNA) was isolated from FACS-sorted human naïve ($CD19^+$ $CD27^-$ IgM^+) or memory B cells ($CD19^+$ $CD27^+$ IgG^+/IgA^+) using a commercial kit (QIAGEN). Switch region PCRs on memory B cell gDNA were performed using LongAmp Taq Polymerase (New England Biolabs) in 50 μl reaction volumes with incubation for 3 min at 95°C, followed by 30 cycles of 95°C for 40 s, 60°C for 30 s, 65°C for 3 min and a final extension for 10 min at 65°C. The upstream switch-μ forward primer S-μ-FW (cacccttgaaagtagcccatgccttcc) was combined with different reverse primers. IgG-switched B cell DNA was amplified using S-γ-REV (cctgcctcccagtgtcctgcattacttctg)23 , which binds 3'

of switch-γ-regions, or CH2-γ-REV1, which binds in the IgG-CH2 constant region, to allow amplification of alleles carrying a CH1 deletion. DNA deriving from IgA^+ sorted cells was amplified with primer S-α-REV (ctcagtccaacacccaccactcc). All reverse primers mentioned were designed to allow amplification of various IgG and IgA subclasses. The switch-μ region of naïve B cell gDNA was amplified combining the S-μ-FW primer with S-μ-REV (ggaacgcagtgtagactcagctgagg). The PCR reaction was performed using Herculase II Fusion DNA Polymerases (Agilent) with 1 M betaine and 3% DMSO in a 50 μl volume at 98°C for 4 min followed by 30 cycles of 98°C for 40 s, 58°C for 30 sec and 72°C for 4 min, with a final extension for 10 min at 72°C. Size-selected, purified switch amplicons were sent to GATC Biotech (Germany) for library preparation, barcoding, and Illumina MiSeq sequencing.

MinION sequencing

Oxford Nanopore Technology (ONT) was used to generate biological and technical replicates of Illumina MiSeq sequencing runs. For biological replicates, barcodes (BC) were introduced by the addition of recommended BC-sequences to S-μ and S-γ primers and PCR amplification. The sequencing library was prepared using the Nanopore 2D sequencing kit SQK-LSK207, followed by loading onto Nanopore flow cells FLO-MIN106 and sequencing with the MinION Mk1B sequencer for up to 20 h.

IE binding assays

Recombinant monoclonal antibodies were produced in 293 Expi cells and tested for the ability to stain 3D7-$MGD21^+$ and eight Kenyan parasite isolates2 by flow cytometry. Cryopreserved IEs were thawed, stained with 10× SYBR Green I for 30 min at room temperature, and incubated with serial dilutions of the recombinant antibodies for 20 min at room temperature. Antibody binding was detected with 2.5 μg ml^{-1} goat Alexa Fluor 647-conjugated anti-human IgG.

Selected sera were screened for the presence of LAIR1-containing antibodies that could bind to IEs. A mixture of four parasite isolates (3D7-$MGD21^+$, 9106, 9605 and 11019) was stained with 10 μg ml^{-1} DAPI for 30 min at room temperature and incubated with test sera at a 1/30 dilution for 20 min at room temperature. The IEs were then incubated with goat anti-human LAIR1 (R&D Systems, AF2664) for 20 min at room temperature. The binding of LAIR1-containing antibodies to the IE surface was detected by the simultaneous addition of Alexa Fluor 488-conjugated donkey anti-goat IgG (Jackson ImmunoResearch, 705-546-147) and Alexa Fluor 647-conjugated donkey anti-human IgG (Jackson ImmunoResearch, 709-606-098).

ELISA

Total IgGs were quantified using 96-well MaxiSorp plates (Nunc) coated with goat anti-human IgG (SouthernBiotech, 2040-01) using Certified Reference Material 470 (ERMs-DA470, Sigma-Aldrich) as a standard. To test specific binding of antibody constructs, ELISA plates were coated with 2 μg ml^{-1} of type I recombinant human collagen (Millipore, CC050), 2 μg ml^{-1} of an anti-human LAIR1 antibody (clone DX26, BD Biosciences 550810), 1 μg ml^{-1} of recombinant human GM-CSF (Gentaur), 2 μg ml^{-1} of an anti-PD1 or

an anti-SLAM antibody (R&D Systems, AF1086 and AF164). Plates were blocked with 1% bovine serum albumin (BSA) and incubated with titrated antibodies, followed by AP-conjugated goat anti-human IgG, Fcγ fragment specific (Jackson Immuno Research, 109-056-098). Plates were then washed, substrate (p-NPP, Sigma) was added and plates were read at 405 nm.

Western blots

B cell supernatants containing secreted antibodies were diluted in H_2O, 4x sample loading buffer (Life Technologies) and 10x sample reducing agent (Life Technologies) and loaded onto precast gels with a 4-12% acrylamide gradient (Invitrogen). The iBlot2 apparatus (Life Technologies) was used for protein transfer to PVDF membranes followed by blocking for 1 h at room temperature with 3% BSA in TBS. The membrane was incubated with different combinations of primary and secondary antibodies diluted in TBS/1% BSA for 1 h at room temperature with 2 sequential TBS incubations to wash the membrane between incubations. For detection of IgG, anti-human IgG-biotinylated antibody (Southern Biotech, 2040-08) was used at 1 μg ml^{-1}, followed by 25 ng ml^{-1} streptavidin-horseradish peroxidase (HRP) (Jackson ImmunoResearch, 016-030-084). IgM isotypes were stained with 10 μg ml^{-1} unlabeled goat anti-human IgM (Southern Biotech, 2020-01) and 8 ng ml^{-1} donkey anti-goat HRP (Jackson ImmunoResearch, 705-036-147). To detect LAIR1-containing antibodies, a polyclonal goat anti-human LAIR1 antibody (R&D) at 2 μg ml^{-1} was combined with secondary donkey anti-goat HRP. Membranes were developed with ECL-substrate on a Las4000 imager (General Electric Company).

Surface staining of B cell lines

EBV immortalized B cells were stained with different fluorescently labeled antibodies to detect surface immunoglobulin expression and LAIR1 co-staining (Alexa Fluor 647-conjugated anti-human IgG, Jackson ImmunoResearch, 109-606-170; FITC-conjugated anti-human kappa, DAKO, F0434, PE-conjugated anti-human lambda, DAKO, R0437; PE-conjugated anti-human LAIR1 clone DX26, BD Bioscience, 550811). Cells were analyzed by flow cytometry and FlowJo software. Dead cells were excluded from the analysis by 4',6-Diamidine-2'-phenylindole dihydrochloride (DAPI) staining. The time point of the analysis was selected for optimal BCR expression levels and for downregulation of the original LAIR1 receptor, because EBV cell lines downregulate the inhibitory receptor LAIR1 at an early time point after immortalization but may also decrease surface-Ig levels after certain passages.

Genomic insertion analysis after Illumina sequencing

We generated a computational pipeline to analyze targeted amplicons of 300 bp paired-end (PE) reads obtained by MiSeq Illumina sequencing methodology (Extended Data Fig. 5). Raw sequences reads were trimmed to remove adapter contamination and poor-quality base calls using Trim Galore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/, v0.4.2, parameters --Illumina --paired -q 20 --length 99). In addition, assessment of the PE reads was performed by average quality score per base position, using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Trimmed reads were aligned in paired-end mode to the GRCh37 human genome assembly using Burrows-Wheeler

Alignment tool (v0.7.12, parameters: bwa mem)24 . The switch region of the IgH locus was defined on GRCh37, with the following coordinates: chr14:106050000-106337000. To find insertions in the IgH switch region locus, we selected genomic ranges when sequence coverage was above 2 reads and/or above 40 reads, thus, generating two respective and separated workflows, with a sequence length comprises between 50 bp and 2000 bp. Genome coverage was processed by Bedtools (http://bedtools.readthedocs.io/en/latest/index.html, v2.26.0) and a dedicated python script using pysam (https://github.com/pysam-developers/pysam) was written to identify potential insert. Briefly, in both workflows (2 and 40 reads), potential inserts were assigned if they fulfilled the following criteria: (i) one mapping read in 5' end is chimeric with the switch region (ii) one mapping read in 3' end is chimeric with the switch region (iii) two discordant reads have their mates read mapped in the switch region. Afterwards, inserts coming from 2 and 40 minimum reads coverage were merged according to the following rule: if the difference between two inserts that overlap was equal or below 10 bp, we kept the shortest one, otherwise the longest one is kept. The list of potential inserts was annotated using GENCODE v1925 and BEDOPS tools26 and individually validated by a *de novo* assembly of the contig sequence. Non-chimeric, but properly paired reads mapping the insert coordinates and chimeric reads corresponding to the encompassing mate pairs and spanning mate pairs, were extracted using SAMtools27. Reads mapping to the switch region were extracted only if they were spanning mate pairs related to the insert coordinates. Selected reads were uniquely mapped to the region of interest (no XA tag), with a minimum mapping quality of 5. Original reads sequences were retrieved, pull together and used as input files to perform a *de novo* assembly using the Trinity software28. Finally, to validate a contig sequence for each insert, we used BLAST29 (command-line version, v2.5.0+). The consensus insert sequences were "blasted" against the switch region and we removed the inserts that had an alignment with at least half of their sequence length at a minimum 80 % identity (parameters: -task megablast -dust no -perc_identity 80). Then, we blasted each contig sequence against the switch region sequence and the consensus insert sequence (parameters: -task blastn -dust no -perc_identity 70) to confirm if the contig was made of the complete insert sequence and if the contig contained two flanking sequences of at least 50bp that matched the switch region. The shortest contig that fulfilled the criteria mentioned above was selected for each insert.

Bioinformatic analysis of MinION sequencing

To analyze targeted amplicon 2-4 kb reads obtained by ONT sequencing methodology, we developed a pipeline (Extended Data Fig. 6). Raw sequences reads were quality-filtered using Metrichor basecaller (https://metrichor.com/s/). 2D reads with a sequence length above 1000bp (or 2000bp depending on the primers used for the amplification) were aligned against GRCh37 human genome assembly with LAST software30 (parameters: last-train and lastal -p ONT_fasta_sequences.par last-split -m1e-6). Then we parsed LAST output and selected reads that contain an insert (minimum 50 bp length), two flanking regions of minimum 100 bp mapping to the switch region (switch locus defined as chr14:106050000-106337000) and allowing a gap of 100 bp maximum between the insert and the switch region. Finally, we merged the insert coordinates of the overlapping inserts with bedtools and annotated the inserts list with GENCODE v1925 using BEDOPS tool26. Scripts are available at MinION and Illumina inserts coordinates were merged with Bedtools

(merge command with default parameter). The circular genomic representation of the inserts has been generated using the Circos software31. Switch inserts with genic (intron or exon-intron) origin were subjected to an enrichment analysis using EnrichR and the Human Gene Atlas as the gene-set library32,33.

Statistical analysis

The number n described in the figure legends refers to the number of independent experiments. The analysis of the relationship between the presence of LAIR1-containing antibodies and protection from malaria was performed in the R statistical environment (v3.2.5). Two-tailed Fisher's exact tests were performed to investigate the association between LAIR1-containing antibodies and protection from febrile malaria.

Code availability

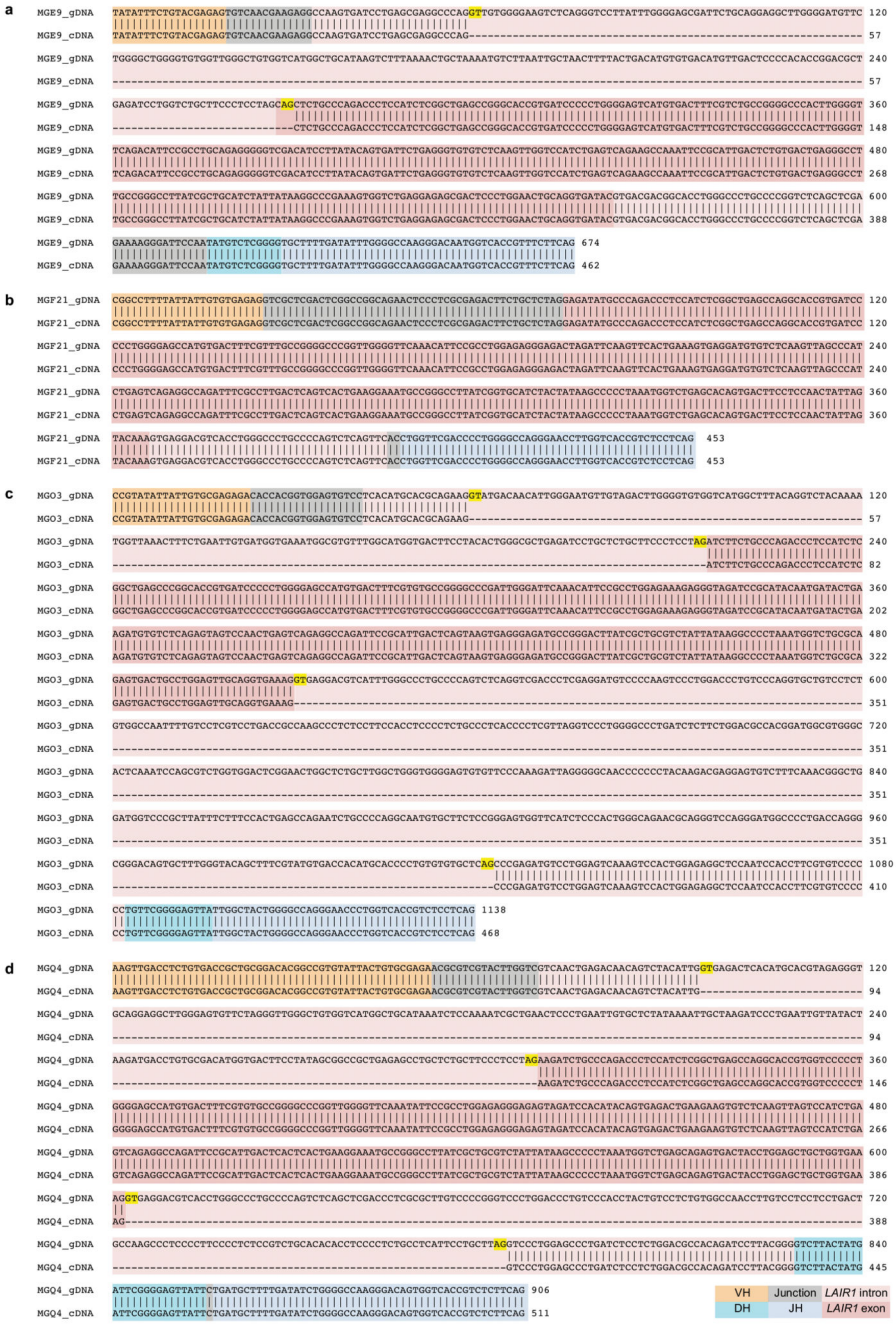
Scripts for Illumina and MinION sequence analysis are available at https://bitbucket.org/mathildefog/switchillumina and https://bitbucket.org/mathildefog/switchminion, respectively.

Data availability

Sequence data of the monoclonal antibodies isolated in this study have been deposited in GenBank (https://www.ncbi.nlm.nih.gov/genbank/) with the accession codes indicated in Supplementary Table 1. The NGS data for switch region sequencing are deposited in NCBI Sequence Read Archive (SRA) with the accession code PRJNA382214 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA382214). The analysis of these sequences are provided in Supplementary Tables 2, 3 and 4.

Europe PMC Funders Author Manuscripts

Extended Data



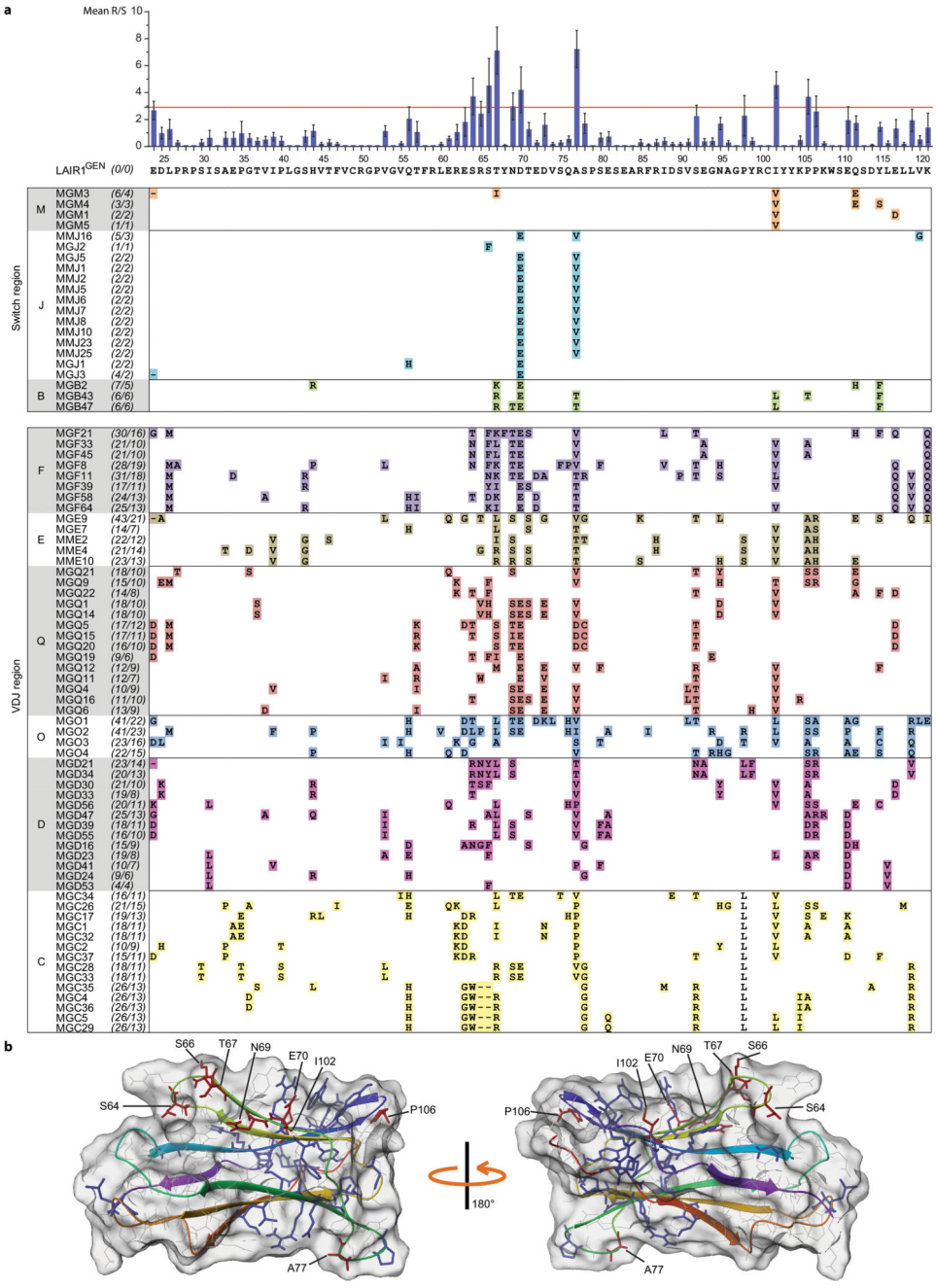
Extended Data Figure 1. Alignment of gDNA and cDNA sequences of LAIR1-containing antibodies.

Shown is one representative antibody from each donor. **a**, MGE9 (donor E), **b**, MGF21 (donor F), **c**, MGO3 (donor O), **d**, MGQ4 (donor Q).

Extended Data Figure 2. Genomic sequences of switch regions containing *LAIR1* inserts.

Shown is one representative antibody for each donor; **a**, MMJ5 (donor J), **b**, MGM5 (donor M), **c**, MGB47 (donor B). The chromosome coordinates of the insertion sites are indicated in blue and green.

Europe PMC Funders Author Manuscripts

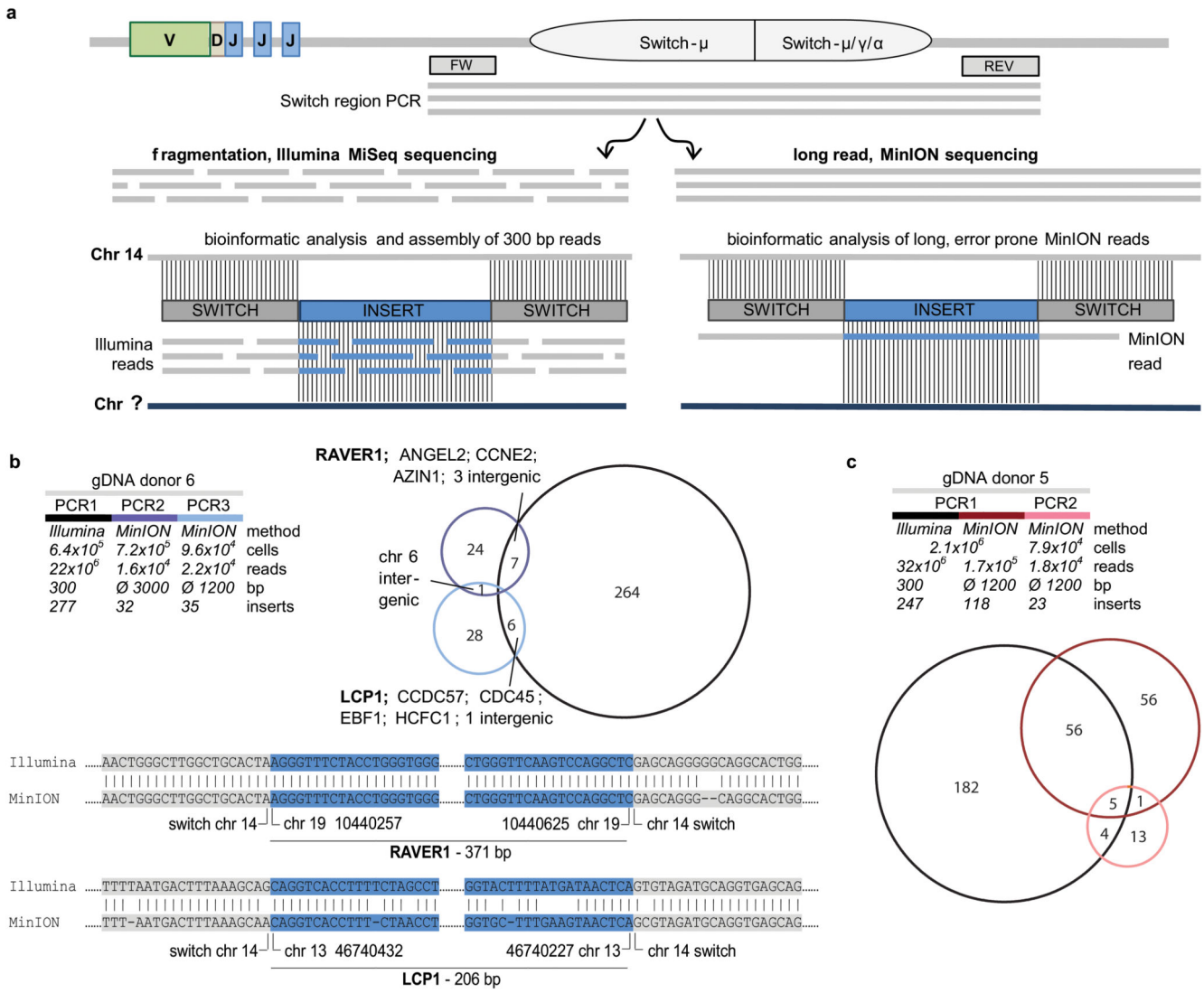


Extended Data Figure 3. Somatically mutated and conserved regions in the LAIR1 domains inserted in the VDJ or in the switch region.

a, Amino acid substitutions in antibodies isolated from different donors and mean R/S ratios at each residue. The mutational analysis takes into consideration the germline LAIR1 alleles found in each donor. In donor C, the P98L substitution is uncolored because it may arise from polymorphism, since the donor is heterozygous at this position. The number of nucleotide mutations and amino acid substitutions are reported in brackets next to the

Europe PMC Funders Author Manuscripts

antibody names. **b,** Graphic representation of mutational hot spots (red) and of most conserved regions (blue).

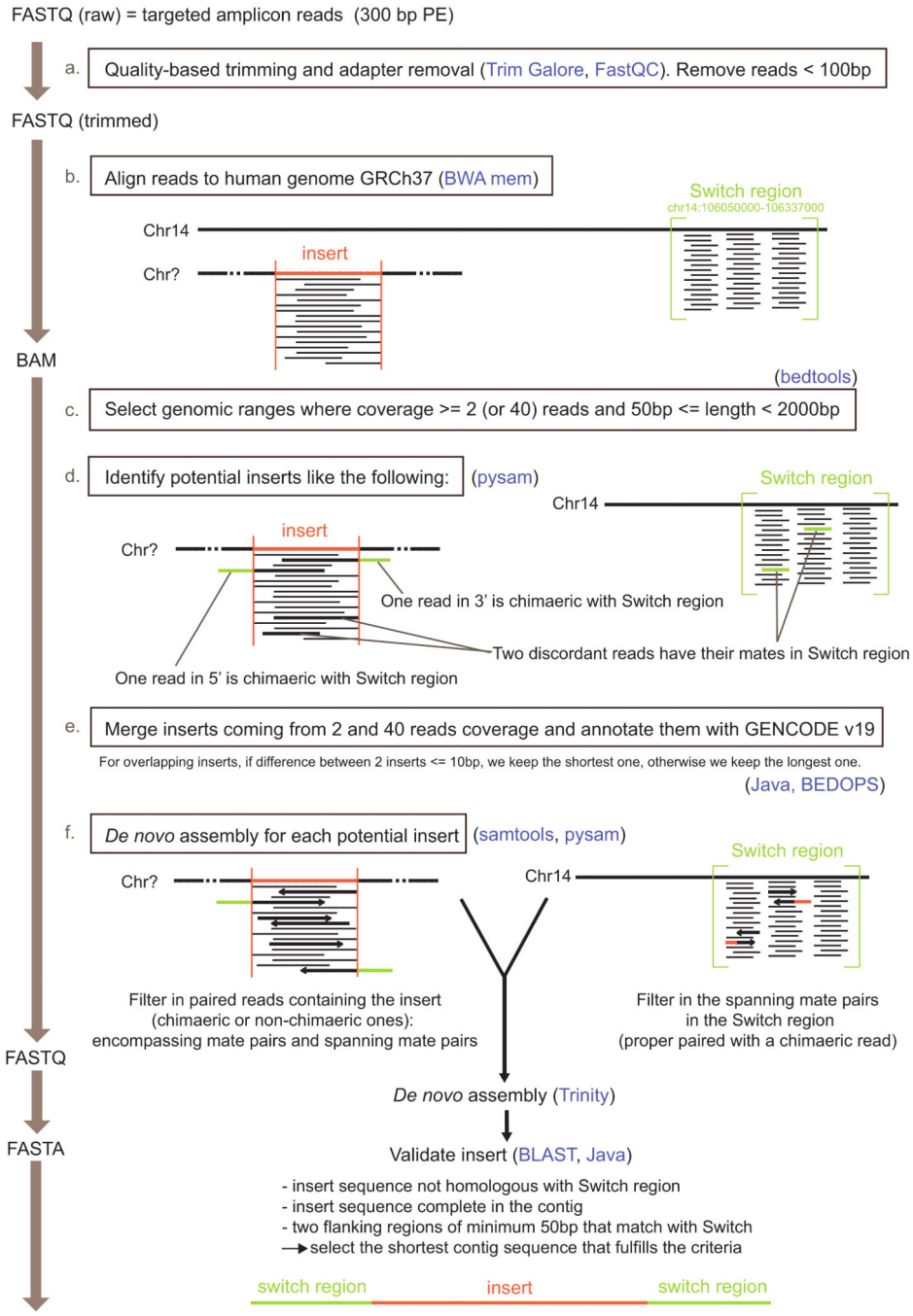


Extended Data Figure 4. Validation of switch region inserts combining Illumina and MinION technologies.

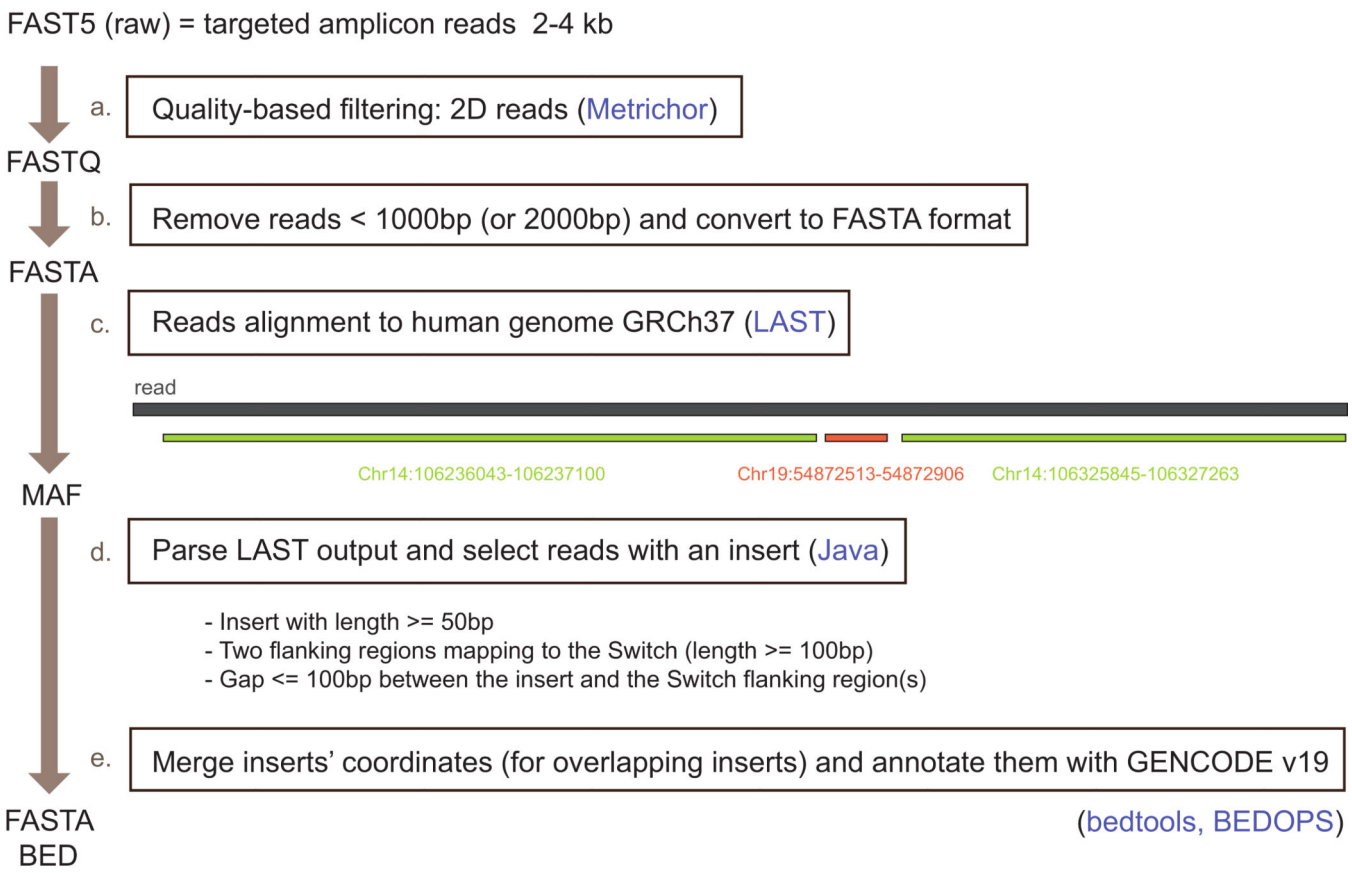
a, Illumina and MinION workflows. Switch regions of polyclonal naïve or IgG/IgA switched B cells were amplified by PCR. For Illumina sequencing, PCR amplicons were fragmented, re-amplified during library preparation and sequenced using the 2x300 bp MiSeq system. The bioinformatic analysis included the assembly of contiguous, chimeric reads. For insert confirmation, independently generated PCR-barcoded primary products were sequenced with MinION technology and analyzed with a different bioinformatic approach for long, error-prone MinION reads. **b**, Multiple identical switch inserts for donor 6 were confirmed in biological replicate experiments with independent technical and analytical setups. Shown are the experimental designs, shared and unique reads in a Venn diagram and an alignment of Illumina and MinION sequences covering the switch insertion

Europe PMC Funders Author Manuscripts

sites for two examples (*LCP1, RAVER1*). **c**, Shared and unique switch inserts in technical and biological replicate experiments of donor 5.

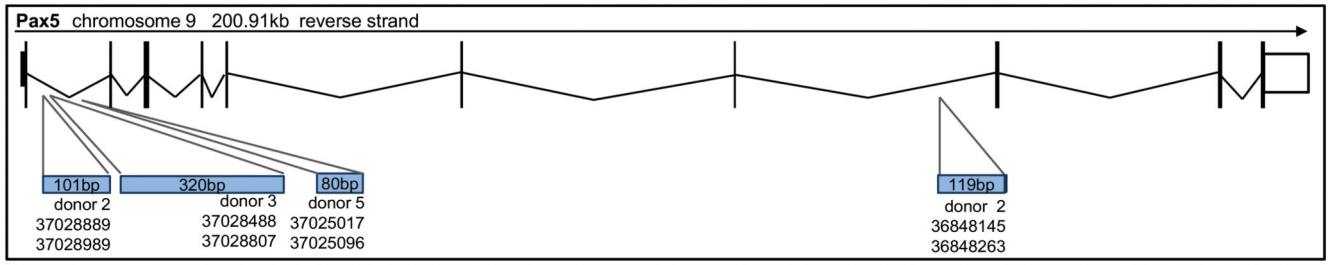


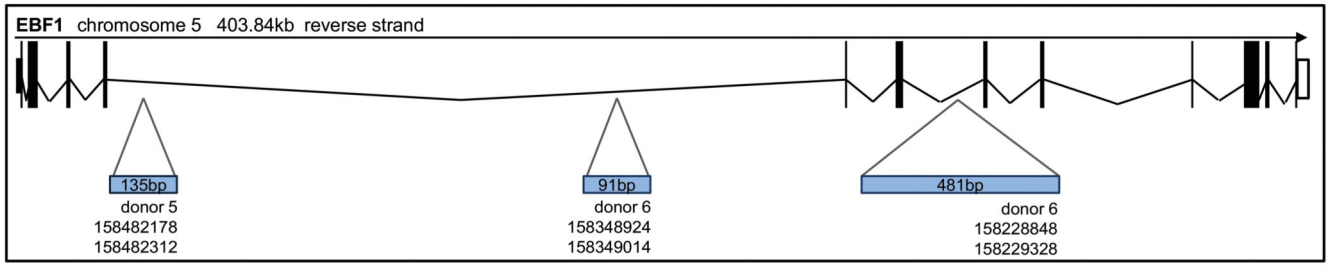
Extended Data Figure 5. Pipeline for data analysis using the Illumina platform.
Shown is the scheme of the bioinformatics workflow used for the analysis of Illumina sequences.



Extended Data Figure 6. Pipeline for data analysis using the MinION technology.
Shown is the scheme of the bioinformatics workflow used for the analysis.

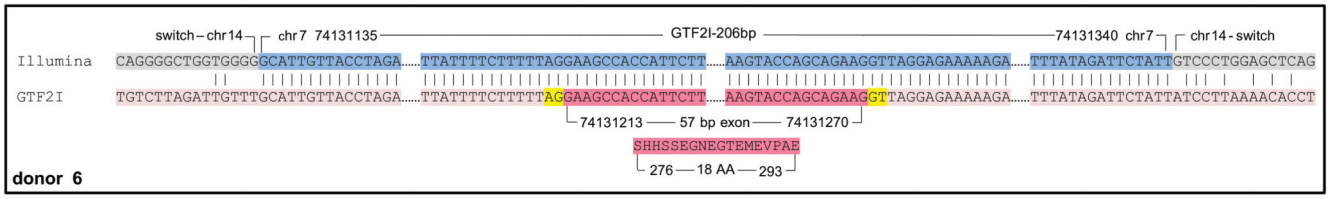
Europe PMC Funders Author Manuscripts

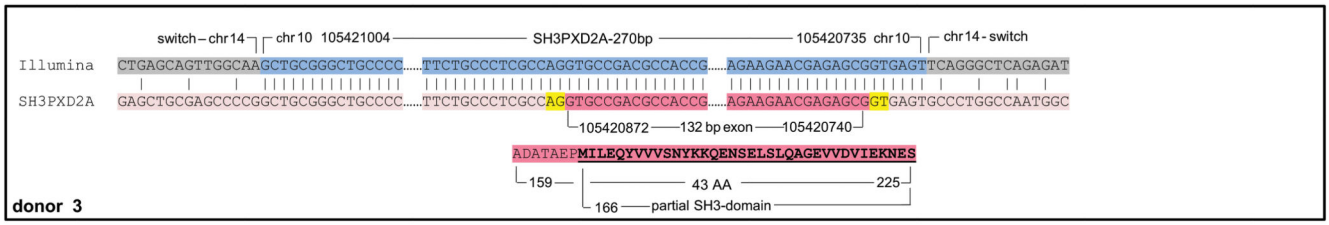




multiple insert-donor genes	
6 inserts	*SNHG1*
5 inserts	*MALAT1*
4 inserts	*CUX1, PAX5, LMNA, IRF8, AFF3, ACTB, AC010976.2*
3 inserts	*TXNDC11, SLC20A1, SETBP1, EBF1, SNHG5, SND1, SETX, RPL28, PACS1, NEAT1, ENO1, CSNK1E, BCL11A, AC096579.13*
2 inserts	*PRKD2, LCP, IKZF3, GAS2L1, FBXO11, ATG16L2, ACTG1, DGKH, HECTD4, HNRNPA1, SLAMF1, EIF3B, KDM2B, MBP, PPP2R5C, SCAPER, SNAPC4, ZNF423, ZZEF1, TCF4, RUFY3, RPTOR, RP11-63A11.1, MGRN1, INPP5D, HMGB1, ETS1, CDT1, FAM53B, HUWE1, RCC1, ST3GAL3, FANCA, L3MBTL2, MKNK2, RAB26, RPL36, ARHGEF2, CASKIN1, CSK, EEFSEC, EIF4A2, ESR2, IKZF1, MYC, NRXN2, NUP214, PKIG, RCC2, ZC3H7A, WDR66, AC096579.7, AFTPH, MAML3, PHF12, WIPF2*

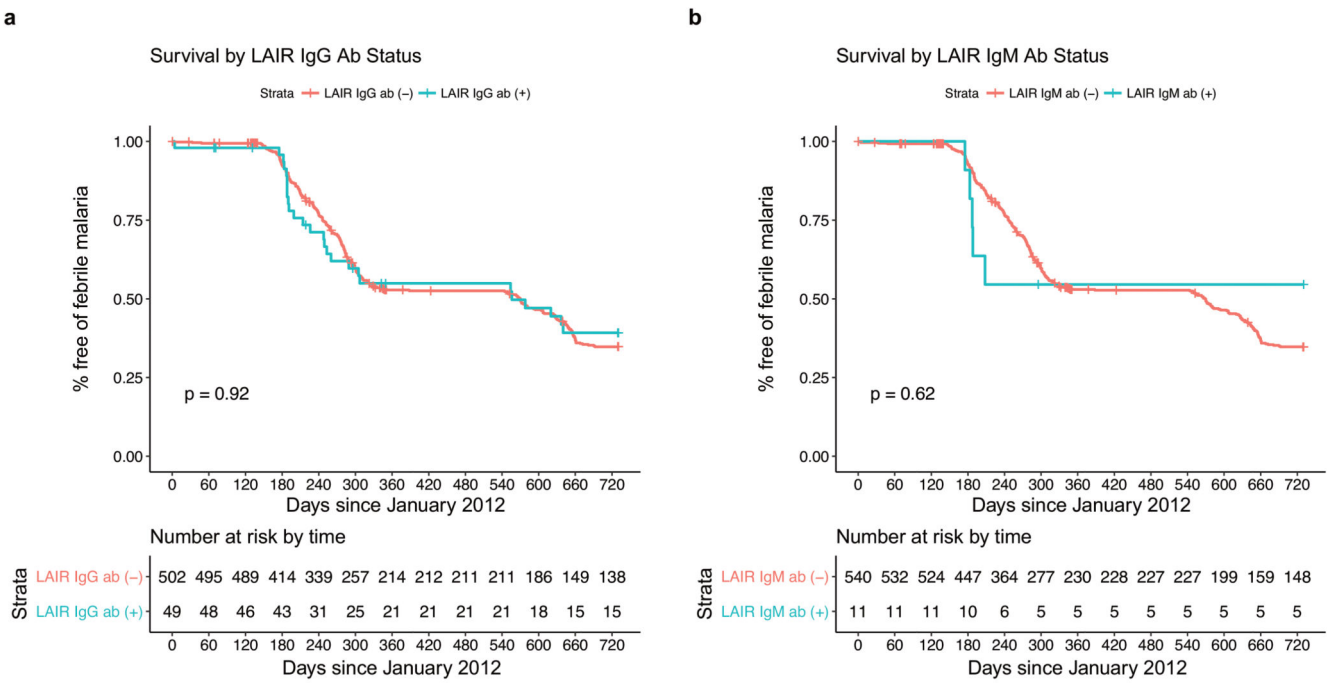
Extended Data Figure 7. Examples of genes that donate multiple inserts.
Shown is the original position of the inserts donated by *PAX5* and *EBF1* as well as a list of genes that donated two or more inserts.





Extended Data Figure 8. Examples of potentially functional inserts.
Shown is the alignment of the contig sequence and the genomic region from which the insert was derived, as well as the potential amino acid sequence inserted between the VH and CH1.

Europe PMC Funders Author Manuscripts



Extended Data Figure 9. Relationship between LAIR1-IgG or LAIR1-IgM status with protection from febrile malaria.

Shown is the clinical status of 551 members of the Malian cohort, stratified by LAIR1-containing IgG (**a**) or LAIR1-containing IgM (**b**) status, over the years 2012 and 2013. Febrile malaria is defined as parasite density ≥2500 asexual parasites per μl of blood and an axillary temperature of ≥37.5°C.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Extended Data Table 1

V gene and insert usage of LAIR1-containing antibodies.

Isotype and V(D)J gene usage of heavy chain and light chain of mAbs containing LAIR1 in the switch or in the VDJ region. D genes of the mAbs containing a V(D)J insert cannot always be properly predicted by IMGT. Mutations of the *LAIR1* insert are shown as % of identity to genomic unmutated *LAIR1* exon. GL = germline; nd = not determined.

	Donor	mAb	Isotype		Heavy chain VDJ genes (% identity to GL)					Light chain VJ genes (% identity to GL)				LAIR1 mutations (% identity to GL)
Switch region	M (Malian)	MGM1	IgG1	λ	*VH3-30*	(79.5)	*D3-16*	*JH4*	(89.4)	*VL3-10*	(88.2)	*JL1*	(81.1)	(99.3)
		MGM3	IgG1	λ	*VH3-30*	(86.8)	*D3-16*	*JH4*	(87.2)	*VL3-10*	(88.5)	*JL1*	(78.4)	(98.0)
		MGM4	IgG1	λ	*VH3-30*	(93.4)	*D3-16*	*JH4*	(91.5)	*VL3-10*	(97.1)	*JL1*	(83.8)	(99.0)
		MGM5	IgG1	λ	*VH3-30*	(88.5)	*D3-16*	*JH4*	(83.0)	*VL3-10*	(91.0)	*JL1*	(81.1)	(99.7)
	J (Tanzanian)	MGJ1	IgG1	κ	*VH3-20*	(83.0)	*D6-25*	*JH3*	(87.8)	*VK1-5*	(89.3)	*JK3*	(91.9)	(99.3)
		MGJ2	IgG3	κ	*VH3-20*	(90.6)	*D6-19*	*JH3*	(91.8)	*VK1-5*	(96.1)	*JK3*	(91.9)	(99.7)
		MGJ3	IgG1	κ	*VH3-20*	(85.4)	*D2-21*	*JH3*	(89.8)	*VK1-5*	(88.5)	*JK3*	(89.2)	(98.6)
		MGJ5	IgG1	κ	*VH3-20*	(87.5)	*D6-19*	*JH3*	(89.8)	*VK1-5*	(98.9)	*JK3*	(97.4)	(99.3)
		MMJ1	IgM	κ	*VH3-20*	(88.2)	*D6-19*	*JH3*	(85.7)	*VK1-5*	(90.0)	*JK3*	(86.5)	(99.3)
		MMJ2	IgM	κ	*VH3-20*	(90.6)	*D6-19*	*JH3*	(81.6)	*VK1-5*	(93.2)	*JK3*	(86.5)	(99.3)
		MMJ5	IgM	κ	*VH3-20*	(91.7)	*D6-19*	*JH3*	(83.7)	*VK1-5*	(92.5)	*JK3*	(86.5)	(99.3)
		MMJ6	IgM	κ	*VH3-20*	(91.0)	*D6-19*	*JH3*	(79.6)	*VK1-5*	(91.4)	*JK3*	(86.5)	(99.3)
		MMJ7	IgM	κ	*VH3-20*	(90.6)	*D6-19*	*JH3*	(81.6)	*VK1-5*	(92.8)	*JK3*	(86.5)	(99.3)
		MMJ8	IgM	κ	*VH3-20*	(90.6)	*D6-19*	*JH3*	(81.6)	*VK1-5*	(93.2)	*JK3*	(86.5)	(99.3)
		MMJ10	IgM	κ	*VH3-20*	(90.6)	*D6-19*	*JH3*	(81.6)	*VK1-5*	(93.2)	*JK3*	(86.8)	(99.3)
		MMJ16	IgM	κ	*VH3-20*	(90.6)	*D6-19*	*JH3*	(81.6)	*VK1-5*	(93.2)	*JK3*	(86.5)	(98.3)
		MMJ23	IgM	κ	*VH3-20*	(91.7)	*D6-19*	*JH3*	(83.7)	*VK1-5*	(92.5)	*JK3*	(86.5)	(99.3)
		MMJ25	IgM	κ	*VH3-20*	(90.6)	*D6-19*	*JH3*	(83.7)	*VK1-5*	(92.1)	*JK3*	(86.5)	(99.3)
	B (Kenyan)	MGB2	IgG3	nd	*nd*	(nd)	*nd*	*nd*	(nd)	*nd*	(nd)	*nd*	(nd)	(97.6)
		MGB43	IgG3	nd	*nd*	(nd)	*nd*	*nd*	(nd)	*nd*	(nd)	*nd*	(nd)	(98.0)
		MGB47	IgG3	nd	*nd*	(nd)	*nd*	*nd*	(nd)	*nd*	(nd)	*nd*	(nd)	(98.0)
VDJ region	E (Tanzanian)	MGE7	IgG1	λ	*VH1-46*	(89.9)	*D2-15*	*JH3*	(85.7)	*VL2-14*	(91.3)	*JL7*	(86.5)	(95.2)
		MGE9	IgG1	λ	*VH1-46*	(88.5)	*D3-10*	*JH3*	(91.8)	*VL2-14*	(93.4)	*JL7*	(91.9)	(85.4)

Europe PMC Funders Author Manuscripts

Donor	mAb	Isotype		Heavy chain VDJ genes (% identity to GL)					Light chain VJ genes (% identity to GL)				LAIR1 mutations (% identity to GL)
	MME2	IgM	λ	*VH1-46*	(86.1)	*D1-26*	*JH3*	(87.8)	*VL2-14*	(94.4)	*JL7*	(91.7)	(92.5)
	MME4	IgM	λ	*VH1-46*	(86.8)	*D2-21*	*JH3*	(91.8)	*VL2-14*	(93.4)	*JL7*	(85.7)	(92.9)
	MME10	IgM	λ	*VH1-46*	(86.8)	*D1-26*	*JH3*	(91.8)	*VL2-14*	(92.4)	*JL7*	(88.6)	(92.2)
F (Tanzanian)	MGF8	IgG2	κ	*VH4-38-2*	(93.8)	*D5-18*	*JH5*	(90.0)	*VK3-15*	(94.3)	*JK2*	(92.1)	(90.5)
	MGF11	IgG1	κ	*VH4-38-2*	(92.4)	*D5-18*	*JH5*	(96.0)	*VK3-15*	(92.5)	*JK2*	(94.7)	(89.5)
	MGF21	IgG1	κ	*VH4-38-2*	(88.5)	*D7-27*	*JH5*	(92.0)	*VK3-15*	(92.8)	*JK2*	(92.1)	(89.8)
	MGF33	IgG1	κ	*VH4-38-2*	(91.3)	*D3-3*	*JH5*	(92.0)	*VK3-15*	(92.1)	*JK2*	(97.4)	(92.9)
	MGF39	IgG1	κ	*VH4-38-2*	(94.1)	*D5-18*	*JH5*	(96.0)	*VK3-15*	(95.0)	*JK2*	(92.3)	(94.2)
	MGF45	IgG1	κ	*VH4-38-2*	(91.3)	*D3-3*	*JH5*	(92.0)	*VK3-15*	(91.8)	*JK2*	(97.4)	(92.9)
	MGF58	IgG1	κ	*VH4-38-2*	(92.4)	*D5-18*	*JH5*	(92.0)	VK3-20	(97.2)	*JK2*	(97.4)	(91.8)
	MGF64	IgG1	κ	*VH4-38-2*	(91.3)	*D5-18*	*JH5*	(96.0)	*VK3-15*	(91.8)	*JK2*	(97.4)	(91.5)
O (Malian)	MGO1	IgG1	κ	*VH4-59*	(87.4)	*D3-10*	*JH4*	(87.2)	*VK3D-20*	(90.4)	*JK5*	(100.0)	(86.1)
	MGO2	IgG1	κ	*VH4-59*	(89.8)	*D2-2*	*JH4*	(85.1)	*VK3D-20*	(94.0)	*JK5*	(97.3)	(86.1)
	MGO3	IgG1	κ	*VH4-59*	(89.1)	*D3-10*	*JH4*	(85.1)	*VK3D-20*	(90.8)	*JK5*	(92.1)	(92.2)
	MGO4	IgG1	κ	*VH4-59*	(91.6)	*D3-10*	*JH4*	(80.9)	*VK3D-20*	(94.3)	*JK5*	(97.3)	(92.5)
Q (Malian)	MGQ1	IgG1	κ	*VH4-59*	(90.9)	*D3-10*	*JH3*	(98.0)	*VK2-30*	(95.9)	*JK1*	(94.1)	(93.9)
	MGQ4	IgG1	κ	*VH4-59*	(94.0)	*D3-10*	*JH3*	(98.0)	*VK2-30*	(94.2)	*JK1*	(91.2)	(96.6)
	MGQ5	IgG2	κ	*VH4-59*	(89.9)	*D3-10*	*JH3*	(100.0)	*VK2-30*	(91.8)	*JK1*	(85.3)	(94.2)
	MGQ6	IgG1	κ	*VH4-59*	(94.0)	*D3-10*	*JH3*	(98.0)	*VK2-30*	(95.6)	*JK1*	(94.1)	(95.6)
	MGQ9	IgG1	κ	*VH4-59*	(92.3)	*D3-10*	*JH3*	(100.0)	*VK2-30*	(95.6)	*JK1*	(97.1)	(94.9)
	MGQ11	IgG1	κ	*VH4-59*	(93.3)	*D3-10*	*JH3*	(95.9)	*VK2-30*	(95.9)	*JK1*	(100.0)	(95.9)
	MGQ12	IgG1	κ	*VH4-59*	(92.7)	*D3-10*	*JH3*	(100.0)	*VK2-30*	(94.2)	*JK1*	(94.1)	(95.9)
	MGQ14	IgG1	κ	*VH4-59*	(90.2)	*D3-10*	*JH3*	(98.0)	*VK2-30*	(95.6)	*JK1*	(94.1)	(93.9)
	MGQ15	IgG2	κ	*VH4-59*	(91.2)	*D3-10*	*JH3*	(100.0)	*VK2-30*	(92.5)	*JK1*	(85.3)	(94.2)
	MGQ16	IgG1	κ	*VH4-59*	(94.4)	*D3-10*	*JH3*	(98.0)	*VK2-30*	(96.3)	*JK1*	(97.1)	(96.3)
	MGQ19	IgG1	κ	*VH4-59*	(90.5)	*D3-10*	*JH3*	(95.4)	*VK2-30*	(97.3)	*JK1*	(97.1)	(96.9)
	MGQ20	IgG1	κ	*VH4-59*	(92.6)	*D3-10*	*JH3*	(100.0)	*VK2-30*	(95.9)	*JK1*	(91.2)	(94.6)
	MGQ21	IgG1	κ	*VH4-59*	(88.1)	*D3-10*	*JH3*	(100.0)	*VK2-30*	(97.6)	*JK1*	(100.0)	(93.9)
	MGQ22	IgG1	κ	*VH4-59*	(88.1)	*D3-10*	*JH3*	(100.0)	*VK2-30*	(93.8)	*JK1*	(100.0)	(95.2)

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank M. Nussenzweig and H. Wardemann for providing reagents for antibody cloning and expression. This work was supported in part by the Swiss Vaccine Research Institute, by the European Research Council (grant no. 670955 BROADimmune) and by the Fondazione Aldo e Cele Daccò. The Mali study was funded by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health. A.L. and F.S. are supported by the Helmut Horten Foundation. This paper is published with the permission of the Director of KEMRI.

References

1. Meyaard L. The inhibitory collagen receptor LAIR-1 (CD305). J Leukoc Biol. 2008; 83:799–803. [PubMed: 18063695]
2. Tan J, et al. A LAIR1 insertion generates broadly reactive antibodies against malaria variant antigens. Nature. 2016; 529:105–109. [PubMed: 26700814]
3. Mwakasungula S, et al. Red blood cell indices and prevalence of hemoglobinopathies and glucose 6 phosphate dehydrogenase deficiencies in male Tanzanian residents of Dar es Salaam. Int J Mol Epidemiol Gen. 2014; 5:185–194.
4. Tran TM, et al. An intensive longitudinal cohort study of Malian children and adults reveals no evidence of acquired immunity to *Plasmodium falciparum* infection. Clin Infect Dis. 2013; 57:40–47. [PubMed: 23487390]
5. Muyldermans S. Nanobodies: natural single-domain antibodies. Annu Rev Biochem. 2013; 82:775–797. [PubMed: 23495938]
6. Piccoli L, et al. Neutralization and clearance of GM-CSF by autoantibodies in pulmonary alveolar proteinosis. Nat Commun. 2015; 6:7375. [PubMed: 26077231]
7. Reed JH, Jackson J, Christ D, Goodnow CC. Clonal redemption of autoantibodies by somatic hypermutation away from self-reactivity during human immunization. J Exp Med. 2016; 213:1255–1265. [PubMed: 27298445]
8. Long M, Betrán E, Thornton K, Wang W. The origin of new genes: glimpses from the young and old. Nat Rev Genet. 2003; 4:865–875. [PubMed: 14634634]
9. Kolkman JA, Stemmer Willem PC. Directed evolution of proteins by exon shuffling. Nat Biotechnol. 2001; 19:423–428. [PubMed: 11329010]
10. Jain M, et al. Improved data analysis for the MinION nanopore sequencer. Nature Methods. 2015; 12:351–356. [PubMed: 25686389]
11. Robbiani DF, et al. *Plasmodium* infection promotes genomic instability and AID-dependent B cell lymphoma. Cell. 2015; 162:727–737. [PubMed: 26276629]
12. Keskin H, et al. Transcript-RNA-templated DNA recombination and repair. Nature. 2014; 515:436–439. [PubMed: 25186730]
13. Chakraborty A, et al. Classical non-homologous end-joining pathway utilizes nascent RNA for error-free double-strand break repair of transcribed genes. Nat Commun. 2016; 7:13049. [PubMed: 27703167]
14. Onozawa M, et al. Repair of DNA double-strand breaks by templated nucleotide sequence insertions derived from distant regions of the genome. Proc Natl Acad Sci USA. 2014; 111:7729–7734. [PubMed: 24821809]
15. Onozawa M, Goldberg L, Aplan PD. Landscape of insertion polymorphisms in the human genome. Genome Biol Evol. 2015; 7:960–968. [PubMed: 25745018]
16. Rommel PC, Oliveira TY, Nussenzweig MC, Robbiani DF. RAG1/2 induces genomic insertions by mobilizing DNA into RAG1/2-independent breaks. J Exp Med. 2017; 214:815–831. [PubMed: 28179379]

17. Trager W, Jensen J. Human malaria parasites in continuous culture. Science. 1976; 193:673–675. [PubMed: 781840]
18. Traggiai E, et al. An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus. Nat Med. 2004; 10:871–875. [PubMed: 15247913]
19. Tiller T, et al. Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. J Immunol Methods. 2008; 329:112–124. [PubMed: 17996249]
20. Lefranc M-P, et al. IMGT, the international ImMunoGeneTics information system. Nucleic Acids Res. 2009; 37:D1006–12. [PubMed: 18978023]
21. Goujon M, et al. A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res. 2010; 38:W695–W699. [PubMed: 20439314]
22. Picelli S, et al. Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc. 2014; 9:171–181. [PubMed: 24385147]
23. Mills F, Mitchell M, Harindranath N, Max E. Human Ig S gamma regions and their participation in sequential switching to IgE. J Immunol. 1995; 155:3021–3036. [PubMed: 7673720]
24. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv.org q-bio.GN. 2013
25. Harrow J, et al. GENCODE: The reference human genome annotation for The ENCODE Project. Genome Research. 2012; 22:1760–1774. [PubMed: 22955987]
26. Neph S, et al. BEDOPS: high-performance genomic feature operations. Bioinformatics. 2012; 28:1919–1920. [PubMed: 22576172]
27. Li H, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–2079. [PubMed: 19505943]
28. Haas BJ, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013; 8:1494–1512. [PubMed: 23845962]
29. Camacho C, et al. BLAST : architecture and applications. BMC Bioinformatics. 2009; 10:421. [PubMed: 20003500]
30. Kielbasa SM, Wan R, Sato K, Horton P, Frith MC. Adaptive seeds tame genomic sequence comparison. Genome Research. 2011; 21:487–493. [PubMed: 21209072]
31. Krzywinski M, et al. Circos: An information aesthetic for comparative genomics. Genome Research. 2009; 19:1639–1645. [PubMed: 19541911]
32. Chen EY, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 2013; 14:128. [PubMed: 23586463]
33. Kuleshov MV, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016; 44:W90–W97. [PubMed: 27141961]

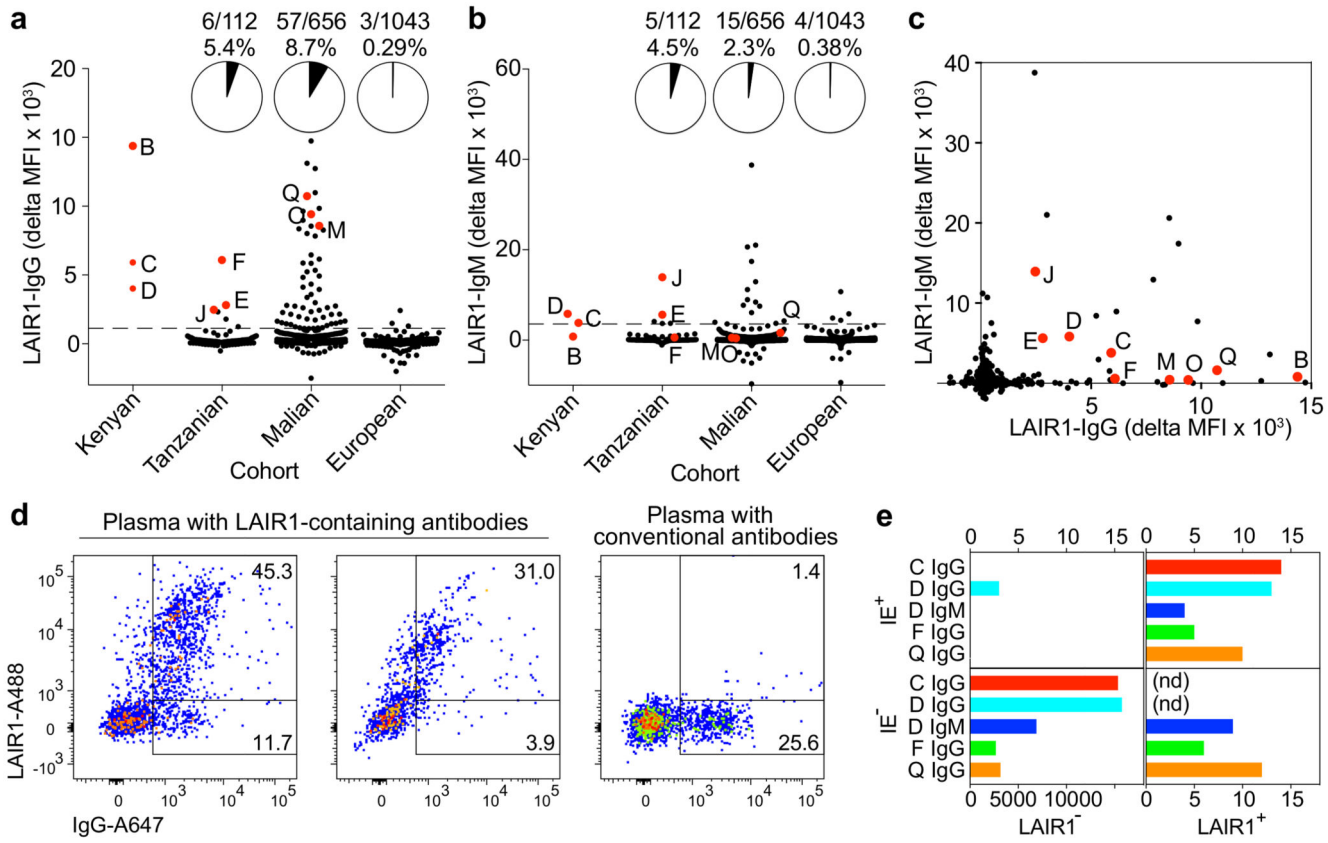


Figure 1. Prevalence and dominance of LAIR1-containing antibodies in malaria-endemic regions.

a, b, Prevalence of LAIR1-containing IgG and IgM in African individuals living in malaria-endemic regions and in European blood donors. Donors from whom LAIR1-containing antibodies were isolated are named and highlighted in red. **c**, Comparison between LAIR1-containing IgG and IgM values. MFI, median fluorescence intensity. Data points with a delta MFI value below -2000 are not shown. **d**, Staining of IEs by LAIR1-containing IgG and conventional IgG from three representative donors. **e**, Dominance of LAIR1-containing B cell clones among memory B cells specific for IEs. Monoclonal antibodies isolated from immortalized memory B cells were classified based on their ability to bind to IEs and the presence of a *LAIR1* insert. Bars show number of IgG or IgM monoclonal antibodies isolated from each donor. For gating strategy, see Supplementary Figure 1. nd = not determined.

Europe PMC Funders Author Manuscripts

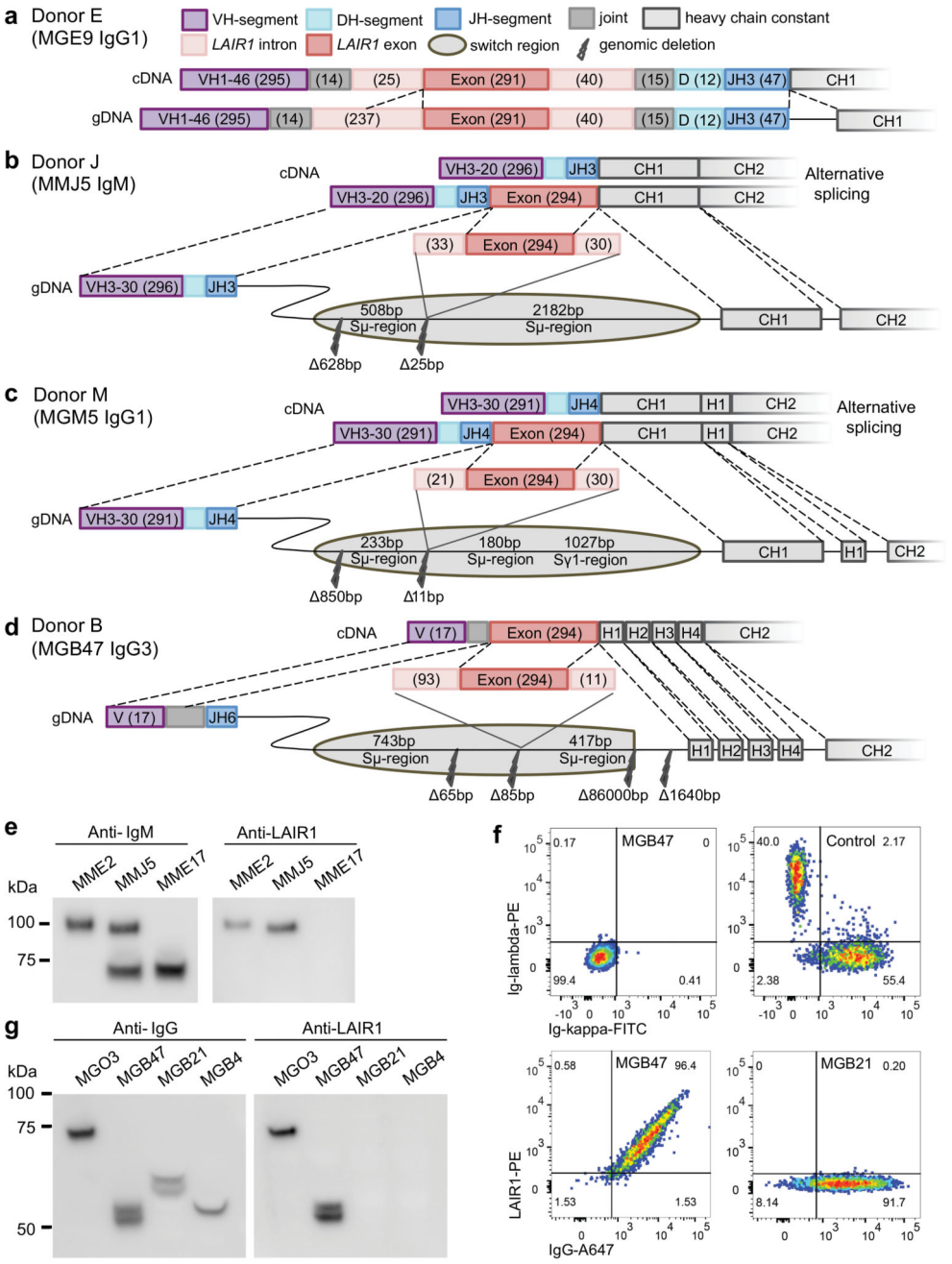


Figure 2. LAIR1-containing antibodies produced by two insertion modalities.

a-d, cDNA and gDNA organization in representative B cell clones of different donors. Donors E and J are Tanzanian, donor M is Malian and donor B is Kenyan. **e**, Western blot analysis of culture supernatants of B cell clones with *LAIR1* insertion in the VDJ (MME2, donor E) or in the switch region (MMJ5, donor J) or a conventional antibody (MME17), (n=4). **f**, Surface staining of clone MGB47 showing LAIR1/IgG co-expression and lack of light chain (n=3). A positive control and a LAIR1-negative clone (MGB21) are shown for comparison. For gating strategy, see Supplementary Figure 1. **g,** Western blot analysis of

culture supernatant of the camel-like clone MGB47 (n=4). Also shown are supernatants from clones MGO3 (LAIR1 insertion in VDJ), MGB21 (IgG3 control) and MGB4 (IgG1 control). For gel source data, see Supplementary Fig. 2.

Europe PMC Funders Author Manuscripts

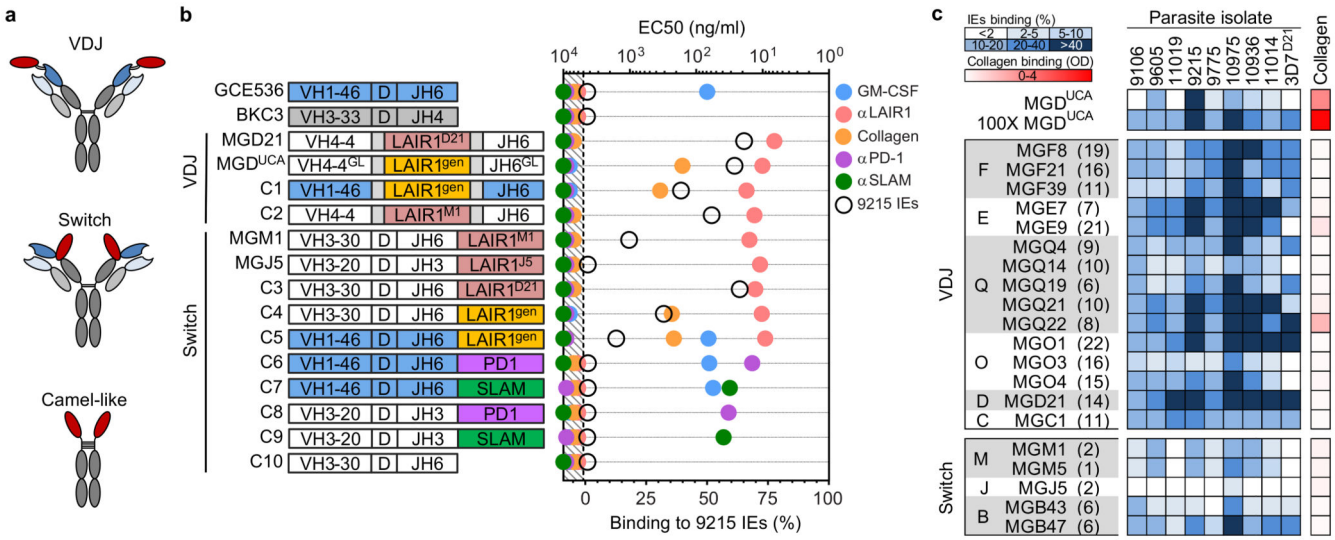


Figure 3. The influence of insert position and somatic mutations on antibody specificity.

a, Schematic representation of LAIR1-containing antibodies produced by different insertion modalities. **b**, Scheme of the constructs (C1-C10) containing LAIR1 or other Ig-like domains in different positions which were tested for binding to a set of antigens or anti-domain antibodies by ELISA or FACS. The construct domains and their binding values to the cognate ligands are color coded as depicted in the figure. The V and J segments that do not contribute to any binding are not colored. BKC3, in grey, is a negative control. $LAIR1^{D21}$, $LAIR1^{M1}$ and $LAIR1^{J5}$ are the exons from MGD21, MGM1 and MGJ5 antibodies. $LAIR1^{gen}$ is the unmutated genomic sequence of LAIR1 encoded on chromosome 19. MGD^{UCA} = unmutated common ancestor of donor D antibodies; GL = germline. Data are from one experiment out of two. **c**, Binding of LAIR1-containing antibodies to erythrocytes infected with nine parasite isolates and to human collagen (n=1). Values refer to binding at a concentration of 1 μg ml^{-1}. The MGD^{UCA} was also tested at 100 μg ml^{-1}. The number of amino acid substitutions is reported in brackets next to the antibody names.

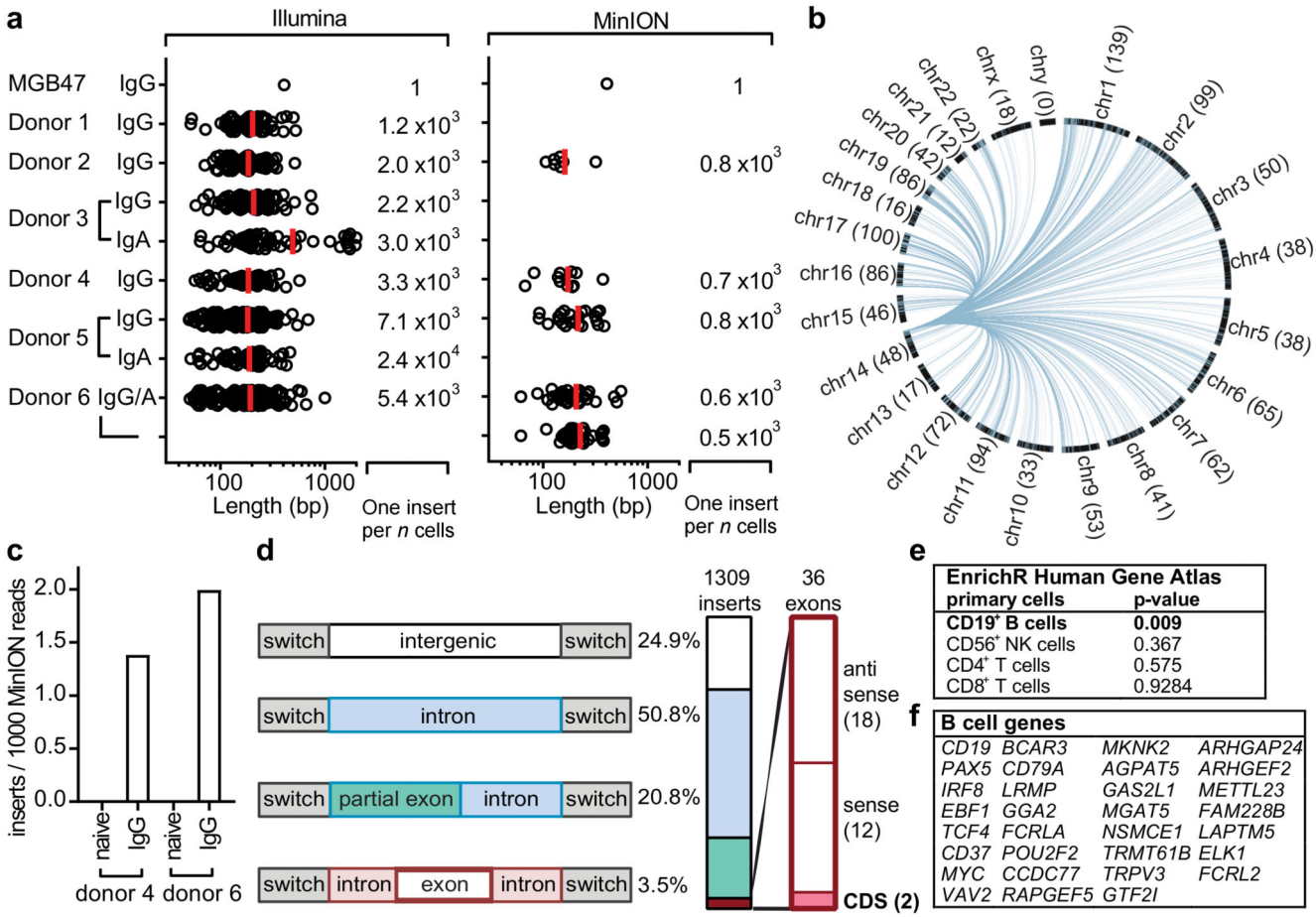


Figure 4. Frequent occurrence of templated inserts in the switch region.

a, Inserts in the switch region were detected by amplification and Illumina MiSeq sequencing (left) of 6 polyclonal samples of primary B cells from European donors and of the MGB47 monoclonal cell line as a control. Biological replicates were analyzed with MinION technology (right). Shown are the size distribution of all detected inserts and the estimated number of B cells that need to be analysed to detect one insert. For each technology, two independent experiments with 2-3 donors were performed. Red lines show mean values. **b**, Circos plot showing the origin of the inserts from different chromosomes. **c**, Insert frequencies in switch-μ regions of naïve B cells and in μ-γ joint regions of IgG memory B cells. **d**, Frequency of intergenic and different types of genic inserts. The red bar shows the number of exon-containing inserts with preserved splice sites in anti-sense and sense orientation. Inserts with the correct orientation and frame of the coding sequence (CDS) are highlighted in pink. **e**, EnrichR analysis on the Human Gene Atlas reveals significant enrichment for B cell genes. **f**, List of B cell-specific genes donating inserts.

Europe PMC Funders Author Manuscripts