Discover Immunity

RESEARCH Open Access

Deep meta-analysis of human pangenome data reveals triplication of the IGH constant gene locus

Pietro D'Addabbo[1], Rossella Cianci[2*] and Domenico Frezza[3]

*Correspondence:
Rossella Cianci
rossella.cianci@unicatt.it
[1]Department of Bioscience, Biotechnology and Environment, University of Bari "Aldo Moro", Via E. Orabona, 4, Bari 70125, Italy
[2]Department of Translational Medicine and Surgery, Catholic University of Rome, Fondazione Policlinico Universitario "A. Gemelli", IRCCS, Largo A. Gemelli, 8, Rome 00168, Italy
[3]Department of Biology-Enrico Calef, University of Rome Tor Vergata, Viale Della Ricerca Scientifica, Rome 00133, Italy

Abstract

Variability in immunogenetics is essential for immune system diversity, enhancing disease resistance and adaptation to pathogens. A key evolutionary mechanism that enhances immunoglobulin (IG) diversity involves the duplication of genes. Here, we describe a genomic variant of the human Immunoglobulin heavy chain (IGH) constant gene locus, characterized by a triplication affecting both the chromosomal segment encoding the α (IGHA), ε (IGHE), and γ (IGHG) isotypes, and the 3′ Regulatory Region (3′RR), i.e. a cluster of enhancers that modulate class switch recombination and gene expression in B-lineage cells. While most mammal genomes contain a single copy of the 3′RR, the Hominoidea species harbour two copies due to a lineage specific segmental duplication that also involves the IGHA, IGHE, and IGHG genes. The presence of a triplication at this locus further increases the number of IGH subclasses, suggesting expanded variability and modulation of immunoglobulin class switching and production in B lymphocytes, with potential consequences for plasma cell maturation. The triplication was detected in samples from unrelated individuals across geographically distant populations (including China, South America, Africa), suggesting that it represents a common haplotype of the human IGH constant region. Moreover, the triplicated region includes the hs1.2 enhancer, whose *2 allele has been associated with both autoimmune diseases and heightened immunological response to viral infection. These observations point out the relevance of this haplotype variant for future studies on autoimmunity and host–pathogen interactions.

Keywords Immunoglobulin heavy chain (IGH) constant gene locus, 3′ Regulatory region (3′RR), hs1.2 enhancer, Class switch recombination, Genomic variant

1 Introduction

The immunoglobulin genes are located on human chromosome 14, near the q telomere. Segmental duplications (SD) in this genomic region increase the complexity of the locus and contribute to the variability in immune response [30, 31]. The V, D, and J genes encode for the variable domain of antibodies, and their number results from multiple rounds of SD [30]. In modern humans, there are roughly 57 V, 23 D, and 6 J functional

genes together with almost the same number of pseudogenes [3], Ivana [25]. The locus encompassing both the Immunoglobulin heavy chain (IGH) gene cluster and the 3' Regulatory Region (3'RR) seems to have undergone fewer duplication events, therefore, in most mammals the IGH alpha gene and the 3'RR are present in a single copy. The genome of the Hominoidea had, instead, a lineage specific segmental duplication, i.e. two copies of a portion of the IGH constant-genes region [12, 33]. This duplication produced in humans two subclasses of IGHA, the IGHE gene and its pseudogene and expanded the IGHG subclasses to four, thereby slightly increasing the variability of the IGH constant region [15].

The 3'RR contains a cluster of enhancers with a fundamental role in the adaptive immune response, acting as modulators of the IGH class switch recombination in B-lineage cells on chromosome 14 [7, 13, 28]. Moreover, the polymorphic enhancer hs1.2 within the human 3'RR1 includes alleles with internal duplications that are significantly linked to both autoimmune diseases and differential responses to viral infections [10]. A study on longevity revealed a decreased homozygosity of hs1.2 for the *2 allele in older individuals, suggesting that the allele may confer some disadvantage in later life, potentially due to effects on the immune system and inflammation [26]. The hs1.2 *2 allele has also been linked to stronger immune responses in children under five, correlating with high serum levels of IGHG and IGHA, conferring improved immune protection against environmental antigens in European populations [32]. Women carrying the *2 allele exhibited milder COVID-19 symptoms and better vaccination responses, likely due to estrogen binding to the 3'RR1 hs1.2 region, enhancing IG transcription [10, 20]. However, in adulthood, the same allele is associated with an increased susceptibility to autoimmune diseases, such as rheumatoid arthritis, especially in women, reflecting its influence on immune function [1, 6, 11, 16]. Therefore, the hs1.2 alleles confer both immunological risks and benefits and appear to have reached a balance in their evolutionary history.

Detecting variability in the human genome is a major goal of public international consortia, which collect, curate and maintain large databases of sequenced human genomes. These repositories highlight common genetic variations and are widely used by the biomedical research community. Moreover, advances in sequencing technologies now enable fully reconstructed haplotypes, thus increasing sensitivity for characterizing structural variants even in regions with extensive SD [14, 29]. These technical advancements have given rise to the concept of a 'human pangenome', i.e. the collection of all normal genomic alternatives in the human species. Accordingly, the Human Pangenome Reference Consortium (HPRC) produced phased diploid assemblies from lymphoblastoid cells belonging to individuals from all over the world [23].

Carefully annotating the 464 available sequences from the HPRC, we detected a structural variant, i.e. a triplication of the genomic region containing IGH genes and 3'RR loci (Fig. 1). This region has largely been described till now as a duplication in the reference human genome (Fig. 2) and a triplication was mentioned only in a few anecdotical reports [2, 4, 17, 19]. This triplication is potentially relevant to both physiological immune responses and autoimmune disease in humans. Indeed, the triplicated region yields three IGHA subclasses, six IGHG subclasses and two functioning IGHE, with the pseudogene acting as the third element of the triplication (Fig. 1). Thus, an additional expansion of the variability of the IGH constant region with enhanced modulation of

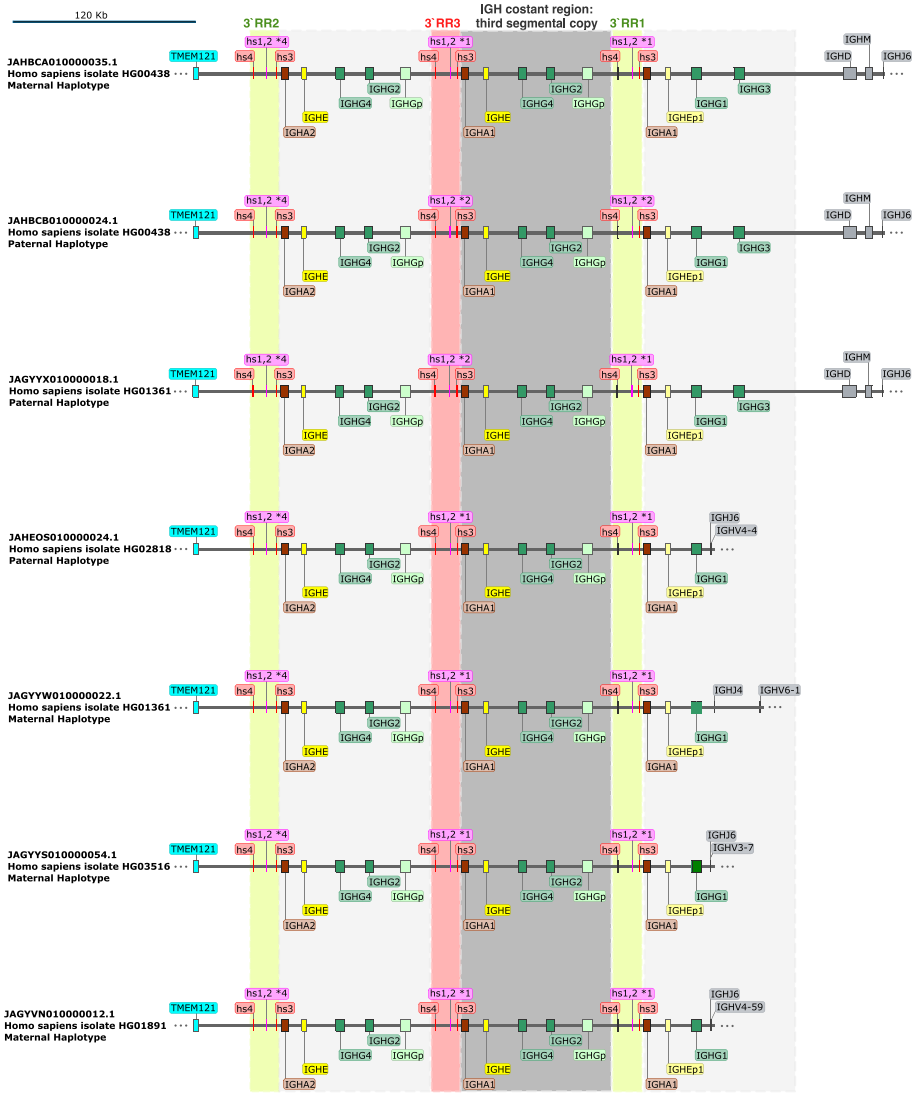

Fig. 1 Annotation of the H and 3'RR locus in 7 contigs belonging to different human haplotypes. The dark grey highlight the genomic segment harboring the third copy of the H locus, while the pale red highlights the third copy of the 3'RR. Color code is: brown – IGHA; yellow – IGHE; green – IGHG; red – hs4 and hs3; purple -hs1,2; pale colors indicate pseudogenes

immunoglobulin class switching and IG expression in plasma cell is possible in individuals carrying the triplication in their genome (Fig. 3).

2 Materials and methods

Sequence searches were performed using blast against the 464 HPRC haplotypes available in the UCSC genome browser (https://hgdownload.soe.ucsc.edu/hubs/HPRC/index.html) [8, 27], as well as using blast2seq to compare the identified contig with the T2T and hg38 releases of the human genome. We used a set of records from GenBank or Ensembl as a reference for each gene and enhancer mapping to the IGH constant locus: hs1.2 (allele *4 AY530201.1, allele *3 AJ544220.1, allele *2 AJ544219.1, and allele *1 AJ544218.1), hs4 (AF013725.1), hs3 (AF013718.1), IGHA2 (ENST00000497872.4), IGHE (ENST00000641420.1), IGHG4 (ENST00000641978.1), IGHG2 (ENST00000641095.1),

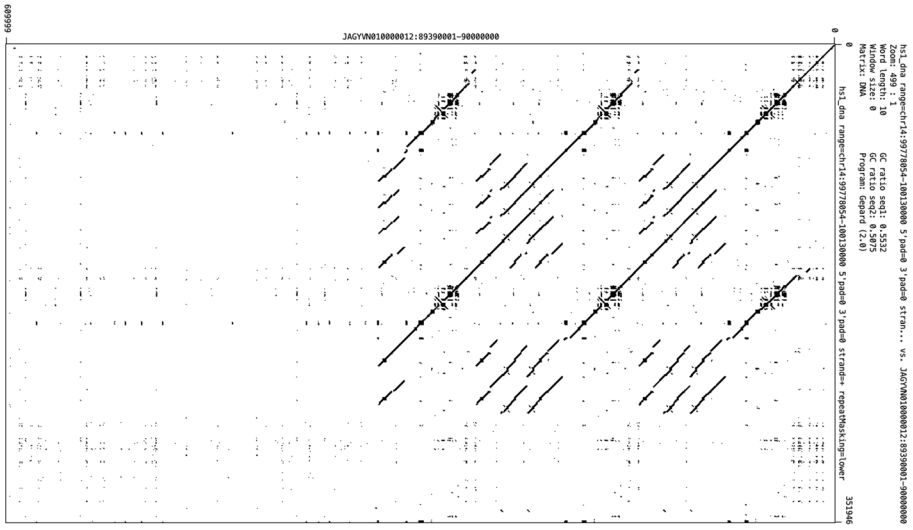

Fig. 2 Gepard dot plot comparison between two locus version: the duplication in the T2T assembly (hs1_dna, vertical) and the triplication in the JAGYVN010000012 contig (horizontal)

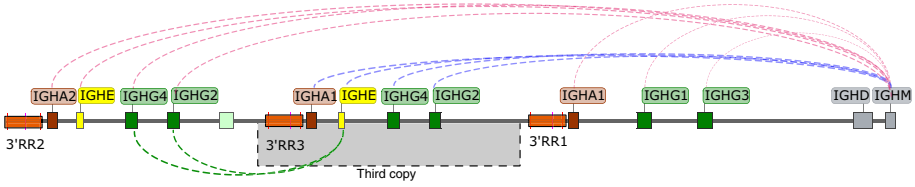

Fig. 3 Hypotheses of class switch recombination by somatic rearrangements in a haplotype with the triplication. The blue and green dashed lines are switches peculiar of this haplotype; the red ones are yet possible in haplotypes that host the duplication

IGHA1 (ENST00000641837.1), IGHG1 (ENST00000390548.6), IGHG3 (ENST00000641136.1), and IGHGP (ENST00000390555.3).

To detect the telomeric limits of the IGHC locus, the contigs were analysed to detect V, D, and J genes using IgBLAST, a tool that was developed at NCBI specifically for the identification of immunoglobulin genes [34]. Mapping of TMEM121 (NM_025268.4) was also used to define the centromeric boundaries of the IGHC locus. Gene and enhancer annotation were manually curated, and graphical representations of the region were generated with SnapGene software (https://www.snapgene.com/).

The IMGT database (https://www.imgt.org/) was queried to verify the annotation stat us of the haplotypes [21].

Dot plot diagrams were generated using Gepard (https://cube.univie.ac.at/research/software-databases/genome-pair-rapid-dotter-gepard/).

3 Results

We analysed 464 human phased haploid assemblies derived from lymphoblastoid cells [8, 23, 29] and identified contigs containing the genome segment spanning from *TMEM121* to the V, D, or J genes. We then annotated immunoglobulin genes and 3'RR enhancers in these selected contigs. Since currently available tools are primarily designed to analyse the VDJ region and often overlook the constant gene region and

downstream regulatory elements, manual curation of the annotations was necessary to ensure completeness and accuracy.

We expected to find 2 copies of the 3'RR in all human haploid assemblies, consistent with the human hg38 and T2T human genome references. A single copy of the region was also expected in samples derived from lymphoblastoid cells, which undergo to somatic rearrangements of the gene region, leading to V-domain diversity, loss of chromosomal segments, and IGH class switch [5]. Also, presence of copy number variation could explain the finding of a single copy IGHC region [22]. Unfortunately, fragmentation of assemblies frequently splits the IGHC region across multiple contigs, making it impossible to exclude the presence of undetected triplications in genomes lacking a contiguous triplicated assembly.

We identified 19 *Homo sapiens* isolates (HG00438, HG00621, HG00673, HG00733, HG00735, HG00741, HG01071, HG01109, HG01123, HG01243, HG01258, HG01361, HG01891, HG01928, HG02080, HG02145, HG02818, HG03492, HG03516), each containing at least one contig with more than one match to the hs1,2 sequence. We then selected all contigs harbouring the IGHC locus from these isolates. In total, 36 contigs were retained, including 23 contigs with two copies of the 3'RR, 6 contigs with a single copy, and 7 contigs with three copies of the region, hereafter named 3'RR1, 3'RR3 and 3'RR2 (Supplementary Table 1; Fig. 1; Table 1).

The 7 contigs with three copies of the 3'RR come from 5 unrelated individuals, each belonging to different populations (Table 1).

In each sample, all three 3'RR contained an hs1.2 enhancer (Supplementary Table 2). We found the *4 allele in all samples in the 3'RR1, while the *1 and *2 alleles are 6 and 1 times in the 3'RR2, and 5 and 2 times in the 3'RR3, respectively (Fig. 1; Table 1; Supplementary Table 2). It is noteworthy that the contig JAGYYX010000018.1 harboured three different alleles of the hs1.2 enhancer within the same DNA molecule (Fig. 1; Table 1; Supplementary Table 2).

Moreover, we found that in samples where somatic rearrangements did not involve the H constant locus, each of the 3'RRs was flanked by an IGHA gene, an IGHE gene or pseudogene, and a pair of IGHG genes, thus extending the limits of the triplicated region to include these genes (Fig. 1).

The International Immunogenetics Information System (IMGT) repository already recorded the NA19240.pri.mat.f1_v2 haplotype, to which the JAHEOL010000061.1 contig belongs, but the identification of the second copy of the IGHG4 gene, which is present in this region, is missing from this annotation. The IMGT database reports that a 185 kb region between IGHEP1 and IGHG4 contains genes that do not conform to the expected gene order and do not match the already annotated alleles (https://www.imgt.org/textes/IMGTrepertoire/index.php?section=LocusGenes&repertoire=locus&species=human&group=IGH&assembly=NA19240.pri.mat.f1_v2) [24]. Therefore, the database appears to treat this haplotype as an anomaly peculiar to the NA19240 individual genome, and not as a common variant of the human genome as we suggest, even though this statement is certainly waiting for additional confirmation before the validation as a novel variant.

To gather additional evidence for this genomic variant, we analysed the findings of a recent study by Gong et al. [18], which analysed a cohort of 945 Han Chinese genomes. This study reported a list of the genomic variants identified in that cohort, and the

Table 1 Human haplotypes with three copies of the 3′RR locus, their assemblies, allele composition, and population origin

Contig	Assembly	Contig count; N50	Haplotype origin	Locus completeness	hs1.2 allele in			Isolate	Population
					3`RR2	3`RR3	3`RR1		
JAHBCA010000035.1	GCA_018471515.1	250; 54.9 Mb	maternal	Yes	*4	*1	*1	HG00438	Han Chinese South
JAHBCB010000024.1	GCA_018472595.1	276; 48.1 Mb	paternal	Yes	*4	*2	*2		
JAGYYX010000018.1	GCA_018469705.1	288; 47.2 Mb	paternal	Yes	*4	*2	*1	HG01361	Colombian in Medellin, Colombia
JAGYYW010000022.1	GCA_018469685.1	295; 45.1 Mb	maternal	No	*4	*1	*1		
JAHEOS010000024.1	GCA_018503575.1	566, 18.2 Mb	paternal	No	*4	*1	*1	HG02818	Gambian in Western division of Gambia
JAGYYS010000054.1	GCA_018469415.1	369; 55.5 Mb	maternal	No	*4	*1	*1	HG03516	Esan in Nigeria
JAGYVN010000012.1	GCA_018467155.1	331; 81.1 Mb	maternal	No	*4	*1	*1	HG01891	African Caribbean in Barbados

variant DUP_59060 appears to correspond to the IGHC triplication described in the present work. The DUP_59060 is a 73 Kb variant contains a third copy of the H constant region and was detected in 4.92% of the samples. This observation may represent the minimum frequency for the triplication in the Han population, as its true prevalence may be underestimated due to both somatic recombination events that delete the supernumerary copy and limits of the assembling software while discriminating among different copies in SD regions.

4 Discussion

The Combination of long-read PacBio whole-genome sequencing and Strand-seq data has enabled the generation of fully phased diploid genome assemblies without relying on parent–child trio data, thus providing a more accurate representation of human genomic variation [14, 29]. By closely examining human haplotype sequences of the constant gene locus, we identified a tandem triplication of a region previously believed to exist predominantly as a duplication in the human genome. An insertion haplotype has already been reported in studies from 1993 and 1994 works by Bottaro et al., who described a 150 kb large sequence in an Italian family using in vitro methods, that differ substantially from current large-scale sequencing approaches. Recently, Gornitzka et al. detected the presence of a large (~ 120 kb) duplication in the IGHC region using Oxford Nanopore Technology (ONT) whole-genome sequencing [17]. Similarly, Jana et al. reported the same finding in two African-ancestry subjects (i.e. HG01891 and HG03516) carrying an additional copy of the IGHE—IGHA1 region, resulting in a tandem triplication of the constant genes [19]. These reports remain anectodical and have not been further investigated. Nevertheless, this finding is particularly relevant, as it suggests that structural variability in the constant region may have been underestimated, with potential implications for class switch recombination and, consequently, for the diversity and regulation of antibody responses. Both cited studies identified the triplication in African-ancestry haplotypes; our data indicate that the triplication variant is also shared by other populations.

Class switch recombination follows a sequential process through progressive deletion of constant genes [13]. It may be hypothesised that, in haplotypes with a duplication, IGHE can only switch to IGHA, because the alpha gene is the only one retained on chromosome 14 after the somatic rearrangement and class switching involving HE. By contrast, in haplotypes with a triplication, an IGHE produced from the central epsilon copy may also switch to an IGHG, or to a different IGHE allele (Fig. 3, green dashed lines). These alternative switching pathways should be considered when evaluating immune pathologies and therapies [9, 10], since they challenge current assumptions about the fixed order and availability of constant genes during class switch recombination, potentially broadening antibody functional diversity and reshaping our understanding of immune-related diseases and treatments.

Notably, we are pointing out that the triplication also introduces an additional 3'RR, which includes a copy of the hs1.2 enhancer. This structural context raises the possibility that not only the allele type, but also its position within the triplication, could modulate immune responses.

Indeed, it is also known that the *2 allele of the hs1.2 enhancer is associated with autoimmune diseases and heightened immune responses [11, 20]. However, to our

knowledge, existing protocols studying its function have not been specifically designed to distinguish the effects of hs1.2 polymorphisms at different positions within the triplication. Adopting long-read technology could help address this question in future studies.

The IMGT public repository has already recorded a single haplotype containing an additional 185 kb region between the IGHEP1 and IGHG4 with an unexpected gene order, corresponding to the triplication, despite the lack of the annotation of one IGHG4 copy. Our findings clarify the origin of this apparently unusual order, suggesting that the haplotype with the triplication should be reconsidered for a complete annotation of the identified genes, and included among the common variants of the human genome. Data from recent work on the Han Chinese population, which detected the same triplication in that cohort, further support its frequency in the human population.

Finally, our identification of a triplication variant in the H and 3'RR locus highlights that valuable information can indeed be uncovered among the pangenome data. Fully mining this genomic resource requires a critical re-examination of automated analyses, as available databases remain underexploited. Both bioinformatic characterization and the human inspection of the full spectrum of genetic variations are essential for the understanding human health and diseases.

5 Conclusions

In conclusion, we have identified an alternative haplotype of the human H constant gene locus that includes an additional copy of both the gene cluster and 3' regulatory region. This triplication expands the number of H subclasses, enhancing isotypic variability. Furthermore, the triplication of the hs1.2 enhancer, whose alleles are associated with both autoimmune diseases and heightened immune responses to viral infections, underscores the relevance of this finding for future immunological research. New in vitro studies targeting production in individuals carrying the triplicated H constant genes will help clarify its functional impact.

Abbreviations

IG	Immunoglobulin
IGH	Immunoglobulin heavy chain
3'RR	3' Regulatory Region
SD	Segmental duplication

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s44368-026-00018-x.

Supplementary file 1: List of the Homo sapiens isolates whose contigs contain the IGHC locus, and the number of 3'RR identified in each one.

Supplementary file 2: Mapping information of genes and enhancers in contigs harboring three copies of the 3'RR

Acknowledgements

We sincerely thank Mariano Rocchi for his valuable feedback and insightful suggestions during the review of this manuscript, and Sharon Cox for carefully revising the English language of the manuscript

Author contributions

Pietro D'Addabbo: bioinformatic formal analysis and investigation; writing—original draft preparation; writing—review and editing. Rossella Cianci: investigation; writing—original draft preparation; writing—review and editing. Domenico Frezza: investigation; writing—original draft preparation; writing—review and editing; supervision. All authors read and approved the final manuscript. The work has not been published previously and is not under consideration for publication elsewhere.

Funding

Pietro D'Addabbo work has been supported by the "National Centre for HPC, Big Data and Quantum Computing" research project [CUP: H93C22000450007].

Data availability

PRO-Seq data were deposited into the Gene Expression Omnibus database under accession number GSE85337 and are available at the following URL: [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE85337] (https:/www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE85337) . Example from: [https://www.nature.com/articles/s41559-017-0447-5] (https:/www.nature.com/articles/s41559-017-0447-5&__hstc=99596993.98ef68e75d901e41ea1d011dafd0de90.1744213197249.1762851559053.1767857926962.9&__hssc=99596993.1.1767857926962&__hsfp=1570198612) Metadata and sequences related to both *Homo sapiens* isolates and haplotypes were available in the Human Pangenome Reference Consortium assembly hub site at the following URL: [https://hgdownload.soe.ucsc.edu/hubs/HPRC/index.html] (https:/hgdownload.soe.ucsc.edu/hubs/HPRC/index.html) Gene sequence data were available in GenBank and Ensembl public repositories; we listed below the accession numbers and URL of each record used as reference to analyse the genome haplotypes: - AY530201.1 at [https://www.ncbi.nlm.nih.gov/nuccore/AY530201.1] (https:/www.ncbi.nlm.nih.gov/nuccore/AY530201.1) - AJ544220.1 at [https://www.ncbi.nlm.nih.gov/nuccore/AJ544220.1] (https:/www.ncbi.nlm.nih.gov/nuccore/AJ544220.1) - AJ544219.1 at [https://www.ncbi.nlm.nih.gov/nuccore/AJ544219.1] (https:/www.ncbi.nlm.nih.gov/nuccore/AJ544219.1) - AJ544218.1 at [https://www.ncbi.nlm.nih.gov/nuccore/AJ544218.1] (https:/www.ncbi.nlm.nih.gov/nuccore/AJ544218.1) - AF013725.1 at [https://www.ncbi.nlm.nih.gov/nuccore/AF013725.1] (https:/www.ncbi.nlm.nih.gov/nuccore/AF013725.1) - AF013718.1 at [https://www.ncbi.nlm.nih.gov/nuccore/AF013718.1] (https:/www.ncbi.nlm.nih.gov/nuccore/AF013718.1) - NM_025268.4 at [https://www.ncbi.nlm.nih.gov/nuccore/NM_025268.4] (https:/www.ncbi.nlm.nih.gov/nuccore/NM_025268.4) - ENST00000497872.4 at [https://grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000497872.4] (https:/grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000497872.4) - ENST00000641420.1 at [https://grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000641420.1] (https:/grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000641420.1) - ENST00000641978.1 at [https://grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000641978.1] (https:/grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000641978.1) - ENST00000641095.1 at [https://grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000641095.1] (https:/grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000641095.1) - ENST00000641837.1 at [https://grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000641837.1] (https:/grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000641837.1) - ENST00000390548.6 at [https://grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000390548.6] (https:/grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000390548.6) - ENST00000641136.1 at [https://grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000641136.1] (https:/grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000641136.1) - ENST00000390555.3 at [https://grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000390555.3] (https:/grch37.ensembl.org/Homo_sapiens/Transcript/Summary?t=ENST00000390555.3) IMGT data about the NA19240 isolates were available at [https://www.imgt.org/textes/IMGTrepertoire/index.php?section=LocusGenes&repertoire=locus&species=human&group=IGH&assembly=NA19240.pri.mat.f1_v2] (https:/www.imgt.org/textes/IMGTrepertoire/index.php?section=LocusGenes&repertoire=locus&species=human&group=IGH&assembly=NA19240.pri.mat.f1_v2) The VCF file containing all SV calls from the Gong et al. (2025) study was publicly available for download at [https://www.biosino.org/node/analysis/detail/OEZ007028] (https:/www.biosino.org/node/analysis/detail/OEZ007028)

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 21 May 2025 / Accepted: 12 January 2026

Published online: 17 January 2026

References:

1. Blanco E, Pérez-Andrés M, Arriba-Méndez S, Contreras-Sanfeliciano T, Criado I, Pelak O, et al. Age-associated distribution of normal B-cell and plasma cell subsets in peripheral blood. J Allergy Clin Immunol. 2018;141:2208-2219.e16. https://doi.org/10.1016/j.jaci.2018.02.017.
2. Bottaro A, Gallina R, Brusco A, Cariota U, Boccazzi C, Barilaro MR, et al. Familial clustering of IGHC deletions and duplications: functional and molecular analysis. Immunogenetics. 1993;37:356–63. https://doi.org/10.1007/BF00216800.
3. Briney BS, Willis JR, Hicar MD, Thomas JW, Crowe JE. Frequency and genetic characterization of V(DD)J recombinants in the human peripheral blood antibody repertoire. Immunology. 2012;137:56–64. https://doi.org/10.1111/j.1365-2567.2012.03605.x.
4. Brusco A, Cariota U, Bottaro A, Boccazzi C, Plebani A, Ugazio AG, et al. Structural and immunologic analysis of gene triplications in the Ig heavy chain constant region locus. J Immunol. 1994;152:129–35.
5. Chi X, Li Y, Qiu X. V(D)J recombination, somatic hypermutation and class switch recombination of immunoglobulins: mechanism and regulation. Immunology. 2020;160:233. https://doi.org/10.1111/IMM.13176.
6. Cianci R, Giambra V, Mattioli C, Esposito M, Cammarota G, Scibilia G, et al. Increased frequency of Ig heavy-chain HS1,2-A enhancer *2 allele in dermatitis herpetiformis, plaque psoriasis, and psoriatic arthritis. J Invest Dermatol. 2008;128:1920–4. https://doi.org/10.1038/jid.2008.40.
7. Cianci R, Mancino G, Galli E, Serone E, Massoud R, D'Addabbo P, et al. New insight of human-IgH 3'regulatory regions in immunoglobulins switch. Gene. 2023;862:147254. https://doi.org/10.1016/j.gene.2023.147254.

8. Clawson H, Lee BT, Raney BJ, Barber GP, Casper J, Diekhans M, et al. Genark: towards a million UCSC genome browsers. Genome Biol. 2023;24:217. https://doi.org/10.1186/s13059-023-03057-x.
9. Colucci M, De Santis E, Totti B, Miroballo M, Tamiro F, Rossi G, et al. Associations between allelic variants of the human IgH 3' regulatory region 1 and the immune response to BNT162b2 mRNA vaccine. Vaccines. 2021;9:1207. https://doi.org/10.3390/vaccines9101207.
10. Colucci M, Frezza D, Gambassi G, De Vito F, Iaquinta A, Massaro MG, et al. Functional associations between polymorphic regions of the human 3'IgH locus and COVID-19 disease. Gene. 2022;838:146698. https://doi.org/10.1016/j.gene.2022.146698.
11. D'Addabbo P, Frezza D, Sulentic CEW. Evolutive emergence and divergence of an Ig regulatory node: an environmental sensor getting cues from the aryl hydrocarbon receptor? Front Immunol. 2023;14:364. https://doi.org/10.3389/FIMMU.2023.996119.
12. D'Addabbo P, Scascitelli M, Giambra V, Rocchi M, Frezza D. Position and sequence conservation in Amniota of polymorphic enhancer HS1.2 within the palindrome of IgH 3'regulatory region. BMC Evol Biol. 2011;11:71. https://doi.org/10.1186/1471-2148-11-71.
13. Dauba A, Näser E, Andrieux D, Cogné M, Denizot Y, Khamlichi AA. The immunoglobulin heavy chain super enhancer controls class switch recombination in developing B cells. Sci Rep. 2024;14:7370. https://doi.org/10.1038/s41598-024-57576-z.
14. Duitama J. Phased genome assemblies. Methods Mol Biol. 2023;2590:273–86. https://doi.org/10.1007/978-1-0716-2819-5_16.
15. Flanagan JG, Rabbitts TH. Arrangement of human immunoglobulin heavy chain constant region genes implies evolutionary duplication of a segment containing γ, ε and α genes. Nature. 1982;300:709–13. https://doi.org/10.1038/300709a0.
16. Frezza D, Tolusso B, Giambra V, Gremese E, Marchini M, Nowik M, et al. Polymorphisms of the IgH enhancer HS1.2 and risk of systemic lupus erythematosus. Ann Rheum Dis. 2012;71:1309–15. https://doi.org/10.1136/ard.2010.147025.
17. Gornitzka MB, Røsjø E, Jana U, Ford EE, Tourancheau A, Lees WD, et al. Ultra-long sequencing for contiguous haplotype resolution of the human immunoglobulin heavy-chain locus. Genome Res. 2025. https://doi.org/10.1101/gr.280400.125.
18. Gong J, Sun H, Wang K, Zhao Y, Huang Y, Chen Q, et al. Long-read sequencing of 945 Han individuals identifies structural variants associated with phenotypic diversity and disease susceptibility. Nat Commun. 2025;16:1494. https://doi.org/10.1038/s41467-025-56661-9.
19. Jana U, Rodriguez OL, Lees W, Engelbrecht E, Vanwinkle Z, Peres A, et al. The human IG heavy chain constant gene locus is enriched for large structural variants and coding polymorphisms that vary among human populations. Cell Genom. 2025. https://doi.org/10.1016/j.xgen.2025.101058.
20. Jones BG, Penkert RR, Xu B, Fan Y, Neale G, Gearhart PJ, et al. Binding of estrogen receptors to switch sites and regulatory elements in the immunoglobulin heavy chain locus of activated B cells suggests a direct influence of estrogen on antibody expression. Mol Immunol. 2016;77:97–102. https://doi.org/10.1016/j.molimm.2016.07.015.
21. Lefranc M-P. IMGT, the International ImMunoGeneTics Information System. Cold Spring Harb Protoc. 2011;2011:pdb.top115. https://doi.org/10.1101/pdb.top115.
22. Lefranc M-P, Lefranc G. IMGT®homo sapiens IG and TR loci, gene order, CNV and haplotypes: new concepts as a paradigm for jawed vertebrates genome assemblies. Biomolecules. 2022;12(3):381. https://doi.org/10.3390/biom12030381.
23. Liao W-W, Asri M, Ebler J, Doerr D, Haukness M, Hickey G, et al. A draft human pangenome reference. Nature. 2023;617:312–24. https://doi.org/10.1038/s41586-023-05896-x.
24. Manso T, Folch G, Giudicelli V, Jabado-Michaloud J, Kushwaha A, Nguefack Ngoune V, et al. IMGT® databases, related tools and web resources through three main axes of research and development. Nucleic Acids Res. 2022;50(D1):D1262–72. https://doi.org/10.1093/nar/gkab1136.
25. Mikocziova I, Greiff V, Sollid LM. Immunoglobulin germline gene variation and its impact on human disease. Genes Immun. 2021;22:205–17. https://doi.org/10.1038/s41435-021-00145-5.
26. Napolioni V, Serone E, Iacoacci V, Carpi FM, Giambra V, Frezza D. The functional VNTR of IGH enhancer HS1.2 associates with human longevity and interacts with TNFA promoter diplotype in a population of Central Italy. Gene. 2014;551:201–5. https://doi.org/10.1016/j.gene.2014.08.057.
27. Perez G, Barber GP, Benet-Pages A, Casper J, Clawson H, Diekhans M, et al. The UCSC Genome Browser database: 2025 update. Nucleic Acids Res. 2025. https://doi.org/10.1093/NAR/GKAE974.
28. Pinaud E, Marquet M, Fiancette R, Péron S, Vincent-Fabert C, Denizot Y, et al. The IgH locus 3' regulatory region: pulling the strings from behind. Adv Immunol. 2011;110:27–70. https://doi.org/10.1016/B978-0-12-387663-8.00002-8.
29. Porubsky D, Ebert P, Audano PA, Vollger MR, Harvey WT, Marijon P, et al. Fully phased human genome assembly without parental data using single-cell strand sequencing and long reads. Nat Biotechnol. 2020;39(3):302–8. https://doi.org/10.1038/s41587-020-0719-5.
30. Pramanik S, Cui X, Wang H-Y, Chimge N-O, Hu G, Shen L, et al. Segmental duplication as one of the driving forces underlying the diversity of the human immunoglobulin heavy chain variable gene region. BMC Genomics. 2011;12:78. https://doi.org/10.1186/1471-2164-12-78.
31. Rodriguez OL, Safonova Y, Silver CA, Shields K, Gibson WS, Kos JT, et al. Genetic variation in the immunoglobulin heavy chain locus shapes the human antibody repertoire. Nat Commun. 2023;14:4419. https://doi.org/10.1038/s41467-023-40070-x.
32. Serone E, Daleno C, Principi N, Porretti L, Iacoacci V, Gargioli C, et al. The change in Ig regulation from children to adults disconnects the correlation with the 3'rr hs1.2 polymorphism. BMC Immunol. 2014;15:45. https://doi.org/10.1186/s12865-014-0045-0.
33. Woof JM, Kerr MA. IgA function – variations on a theme. Immunology. 2004;113:175–7. https://doi.org/10.1111/j.1365-2567.2004.01958.x.
34. Ye J, Ma N, Madden TL, Ostell JM. IgBLAST: an immunoglobulin variable domain sequence analysis tool. Nucleic Acids Res. 2013. https://doi.org/10.1093/NAR/GKT382.

Publisher's Note