PNAS BRIEF REPORT | IMMUNOLOGY AND INFLAMMATION OPEN ACCESS

The immunoglobulin J chain is an evolutionarily co-opted chemokine

Kazuhiko Kawasaki[a], Yuko Ohta[b], Caitlin D. Castro[c], and Martin F. Flajnik[b,1]

Edited by Philippa Marrack, National Jewish Health, Denver, CO; received October 30, 2023; accepted December 12, 2023

The joining (J) chain regulates polymerization of multimeric Immunoglobulin(Ig)M and IgA, forming a disulfide bond to the C termini of their Ig heavy chains, and it controls IgM/IgA transport across mucosal epithelia. Like Ig itself and human-like adaptive immunity, J chain emerged in jawed vertebrates (gnathostomes), but its origin has remained mysterious since its discovery over 50 y ago. Here, we show unexpectedly that J chain is a member of the CXCL chemokine family. The J chain gene (*JCHAIN*) is linked to clustered *CXCL* chemokine loci in all gnathostomes except actinopterygians that lost *JCHAIN*. *JCHAIN* and most *CXCL* genes have four exons with the same intron phases, including the same cleavage site for the signal peptide/mature protein. The second exon of both genes encodes a CXC motif at the same position, and the lengths of exons 1 to 3 are similar. No other gene in the human secretome shares all of these characteristics. In contrast, intrachain disulfide bonds of the two proteins are completely different, likely due to modifications in J chain to direct Ig polymerization and mucosal transport. Crystal structures of CXCL8 and J chain share a conserved beta-strand core but diverge otherwise due to different intrachain disulfide bonds and extension of the J chain C terminus. Identification of this ancestral affiliation between J chain and CXCL chemokines addresses an age-old problem in immunology.

immunoglobulin | evolution | gene family | vertebrate | synteny

The adaptive immune system arose 500 Mya in vertebrates (1). Immunoglobulins (Igs) are crucial for human and all gnathostome (jawed vertebrate) immunity. IgM and IgD are primordial Ig isotypes, with IgM being best studied. Upon B cell activation, alternative splicing of IgHeavy(H) chain mRNA causes IgM secretion from plasma cells as multimers, mostly pentamers (2). The joining (J) chain was discovered in 1970 as a component of IgA (3), and IgM (4). It is expressed by most plasma cells, and IgM pentamerization/IgA dimerization is regulated by incorporation of J chain via disulfide bonding with a cysteine (Cys) in the IgH secretory tail (5, 6). J chain is required for transport of Igs across mucosal epithelia, and effector humoral immunity is dysregulated in knockout mice.

J chain is present in all gnathostomes except actinopterygians (ray-fins) where it was lost. J chain sequences have never revealed a relationship to any recognized gene family. J chain crystal structure was finally solved in 2020, in association with IgM and IgA (7, 8), which provided an appreciation of its function in polymerization and mucosal transport; however, no insight into J chain origins was uncovered.

Herein, we investigated evolutionary emergence of *JCHAIN* based on conserved synteny and surprising kinship with a major immune superfamily, the CXCL chemokines, which arose in jawless vertebrates (9). We found a strong relationship with these chemokines, likely unmasking J chain's derivation. In addition to informing the membership of J chain into this chemokine family, we speculate on how a chemokine can be recruited for a new function.

Results

Common Exon–Intron Structures of Linked *JCHAIN* and *CXCL* Genes. Since no homology-based programs uncovered a J chain connection with any gene family, we instead examined synteny of *JCHAIN* with other genes in gnathostomes and found that clusters of *CXCL* genes are invariably closely linked. Surprisingly, we detected ancestral characteristics in the exon–intron structure (Fig. 1 and Datasets S1 and S2). Most *CXCL* and all *JCHAIN* genes consist of four exons with 1-2-2 intron phases (intron 1 sits between the first and second nucleotides in a codon, and introns 2 and 3 between the second and third nucleotides). Exons 1 to 3 are similar in size, and the CXC motif (defining the CXCL family) is encoded at an equivalent position. Furthermore, their signal peptide/mature protein boundary coincides between exons 1 and 2. These common characteristics, and the ancient array of linked *JCHAIN* and *CXCL* genes (Fig. 2), indicate that *JCHAIN* arose from an adjacent *CXCL* gene by duplication in the gnathostome common ancestor.

Author affiliations: [a]Department of Anthropology, Pennsylvania State University, University Park, PA 16802; [b]Department of Microbiology and Immunology, University of Maryland, Baltimore, MD 21201; and [c]Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637

Author contributions: K.K. and M.F.F. designed research; K.K. and M.F.F. performed research; K.K., Y.O., and C.D.C. analyzed data; and K.K., Y.O., C.D.C., and M.F.F. wrote the paper.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: mflajnik@som.umaryland.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2318995121/-/DCSupplemental.

Published January 12, 2024.

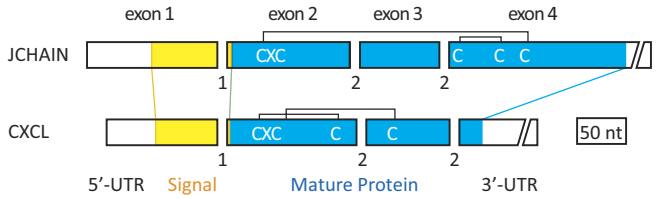

Fig. 1. Conserved gene organization of gnathostome *JCHAIN* and *CXCL* genes. Average length of four exons, 5′-untranslated region (5'-UTR), signal peptide (Signal; yellow)-coding region, and mature protein-coding region (blue) of *JCHAIN* and *CXCL* genes. The phase of introns is indicated between exons. "C" represents Cys residues forming the CXC motif or an intramolecular disulfide bond. (Scale bar: 50 nucleotides.)

Immunogenetics and Evolution of the *JCHAIN*/*CXCL* Family. Human *JCHAIN* links to 12 *CXCL* genes, forming cluster A and cluster B, and an isolated *CXCL13* on chromosome 4 (Fig. 2*A*). Many *CXCL* genes were also identified in three locations on the same chromosome in caecilians (amphibians). The arrangement of adjacent genes, like *SDAD1*, allowed us to assign three distinct regions in caecilians: cluster A, cluster B, and *CXCL13*-corresponding region. Notably, one or two *JCHAIN* genes were identified within or near cluster B (Datasets S1 and S2). Like human *CXCL8*, *CXCL8*-like genes encode an ELR motif, Glu residue, and WV motif at specific positions (10), and these genes were identified in caecilian cluster A (Fig. 2*A*). Both cluster A (containing *CXCL8*-like genes) and cluster B (no *CXCL8*-like genes) were also identified in coelacanth (sarcopterygians and lobe-fins) and bichir (actinopterygians). *JCHAIN* was found near cluster B in coelacanth but not detected (lost) in bichir or any other actinopterygians.

Chondrichthyans (cartilaginous fish) are the most basal vertebrates with a human-like adaptive immune system, including *jchain*. In whale shark, and six other sharks and rays (Elasmobranchii), two *cxcl* clusters are on the same chromosome, and one or more *jchain* genes reside within or near one cluster (Fig. 2*A* and Dataset S1). Unlike sarcopterygians, *jchain* is downstream of *cxcl* genes and clusters with *CXCL8*-like genes. A similar arrangement of *cxcl* genes and *jchain* was confirmed in elephant shark (Holocephali). The arrangement of these *cxcl* genes infers that *jchain* is adjacent to cluster A that contains *CXCL8*-like genes in chondrichthyans. Whereas the phases of the three introns are 1-2-2 in most osteichthyan (bony-vertebrate) *CXCL* genes (Fig. 2*A*), those in some chondrichthyan *cxcl* genes are 1-2-1, all residing in cluster B (Fig. 2*A*). *CXCL* genes with phase 1-2-1 introns are in cluster B of sarcopterygian coelacanth and lungfish (Datasets S1 and S2), reinforcing our conclusion that *JCHAIN* (always 1-2-2) links to cluster B in sarcopterygians but cluster A in chondrichthyans (Fig. 2*B*).

These data infer an ancestral arrangement of *jchain*, cluster A, and cluster B (Fig. 2*B*). In elephant shark and rays, *rufy3* is adjacent

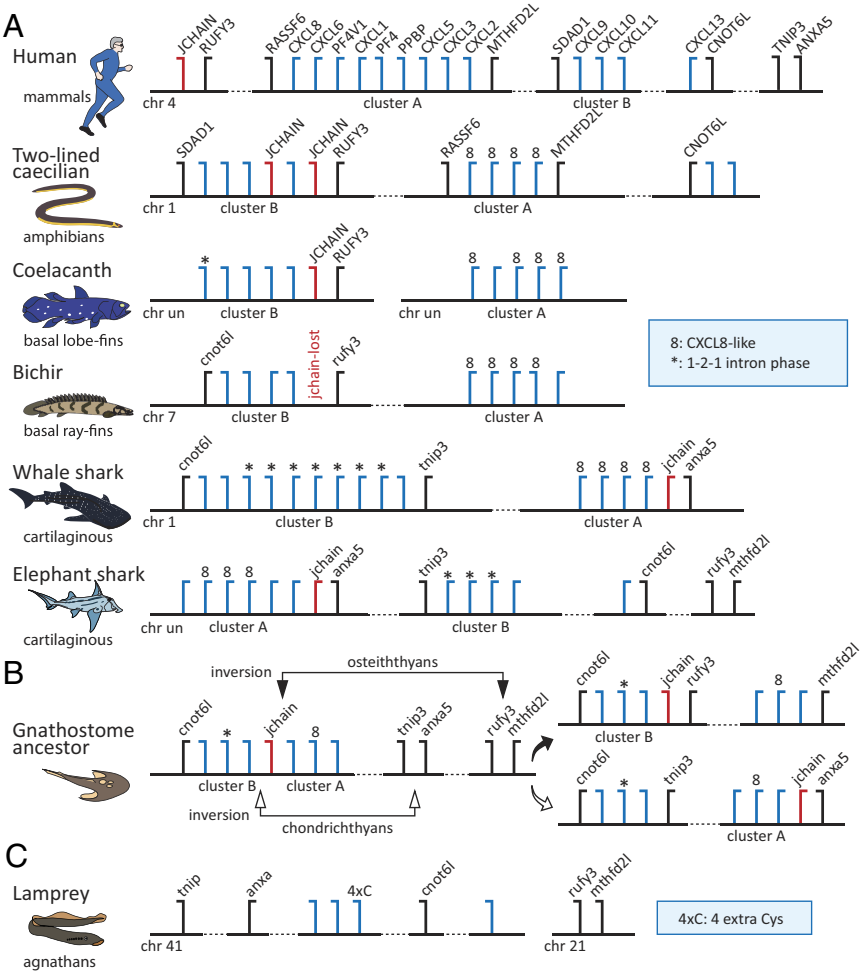

Fig. 2. *JCHAIN* and *CXCL* synteny conservation in jawed vertebrates. Transcriptional orientation and order of *JCHAIN* (red), *CXCL* (blue), and other genes (black). (*A*) Distribution of *JCHAIN*, *CXCL* clusters A and B, and adjacent genes in humans (chr 4), two-lined caecilian (chr 1), coelacanth [two contigs; chr unknown (un)], bichir (chr 7), whale shark (chr 1), and elephant shark. No *CCL* or other chemokine genes were found adjacent to *JCHAIN* (Dataset S4). (*B*) Putative arrangement of these genes in the gnathostome common ancestor and lineage-specific inversions in the common ancestor of osteichthyans (closed arrow) and chondrichthyans (open arrow). (*C*) Lamprey *CXCL* gene cluster (no J chain in Agnathans).

to *mthfd2l*, and *TNIP3* resides near *ANXA5* in humans (Fig. 2*A*). Similar arrangements are found in lamprey (agnathan) that lacks *jchain* (Fig. 2*C* and Datasets S1 and S2), implying their ancestral linkage. The primordial array of these genes suggests lineage-specific inversions: While *jchain* and cluster B were juxtaposed to *rufy3* in the osteichthyan common ancestor, *jchain* and cluster A were placed near *anxa5* in the chondrichthyan common ancestor (Fig. 2*B*).

***JCHAIN* and *CXCL* Genes and Encoded Proteins Are Uniquely Similar.** As described, the linked *JCHAIN* and *CXCL* genes have four exons with phase 1-2-2 introns and a CXC motif encoded in exon 2. We searched the entire human secretome (1,021/~21,000 protein-coding genes, https://www.proteinatlas.org/humanprotome/tissue/secretome) and found 163 genes with four exons with 16 genes having phase 1-2-2 introns that included *JCHAIN*, 10 *CXCL* genes, two *CCL* chemokine genes, and three others (Dataset S3). Among these four exon genes, CXC-encoded sequences were found in 50 genes, 23 of which were in exon 2. Importantly, these two characteristics, phase 1-2-2 introns and CXC-encoded sequence in exon 2 were found only in *JCHAIN* and *CXCL* genes among the entire secretome.

Chemokines are basic proteins (11), while tetrapod J chains are highly acidic (12), which calls into question the common origin. However, long ago, we found that nurse shark J chain is slightly basic (13), and *JCHAIN* encodes neutral or basic proteins in all nine cartilaginous fish and basal sarcopterygians (Datasets S1 and S2. These data suggest that the J chain was neutral/basic at its origins, unlike tetrapod J chains, further solidifying the J chain/chemokine relationship.

Discussion

Since its discovery, the origin of the J chain has remained mysterious (3). We posit *JCHAIN*'s genesis by conserved synteny and nearly identical gene structure with *CXCL* genes. This relationship has been overlooked for decades since J chain function is so distinct from that of chemokines, and placement of disulfide bonds (Fig. 1) has resulted in major modifications of the primordial CXCL structure (14). Thus, attempts to structurally align J chain and CXCL8 provided no convincing similarity, although a region of three core beta strands aligned for the proteins in most models.

While tetrapod J chain structure has been conserved, the J chain C-terminal region in basal Elasmobranchs is unique. The modification of tetrapod J chain was in concert with luminal transport, while J chain's function in Elasmobranchs is only for IgM polymerization (15). Unique among vertebrates, sharks have monomeric IgM regulated by J chain expression in plasma cells (2). The structure of shark J chain might be revelatory regarding its relationship with chemokines.

While J chain's function is well known, *JCHAIN* has other interesting transcriptional patterns besides expression in plasma cells (12). It is expressed in developing mouse dendritic cells (16), shark muscle (13), and lungfish intestinal epithelia (17). J chain also associates with IL12p40 in dendritic cells (18). Thus, there may be some primordial functions related to its chemokine ancestry. Note that synteny also explains *JCHAIN*'s major expression in plasma cells due to its recruitment of the *Igκ* light-chain gene promoter (19), which is syntenic to *JCHAIN*/*CXCL* genes in some amphibians.

Is there something special about J chain's chemokine ancestry besides this "one-off" evolutionary example? Note that one lamprey *cxcl* locus encodes a protein with four other Cys's, reminiscent of J chain but likely encoding a different type of molecule (Fig. 2*C*) (9). Perhaps there will be other examples in phylogeny of chemokine "transitioning," especially in the immune system.

Data, Materials, and Software Availability. All study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. We thank Louis Du Pasquier and Hanover Matz for editing. The work was supported by NIH grants 5R01AI170844 and 5R01AI140326.

1. M. F. Flajnik, A cold-blooded view of adaptive immunity. *Nat. Rev. Immunol.* **18**, 438–453 (2018).
2. C. D. Castro, Y. Ohta, H. Dooley, M. F. Flajnik, Noncoordinate expression of J-chain and Blimp-1 define nurse shark plasma cell populations during ontogeny. *Eur. J. Immunol.* **43**, 3061–3075 (2013).
3. M. S. Halpern, M. E. Koshland, Noval subunit in secretory IgA. *Nature* **228**, 1276–1278 (1970).
4. S. L. Morrison, M. E. Koshland, Characterization of the J chain from polymeric immunoglobulins (IgA-IgM-immunological specificity-primary structure). *Proc. Natl. Acad. Sci. U.S.A.* **69**, 124–128 (1972).
5. E. M. Yoo *et al.*, Structural requirements for polymeric immunoglobulin assembly and association with J chain. *J. Biol. Chem.* **274**, 33771–33777 (1999).
6. N. Lycke, L. Erlandsson, L. Ekman, K. Schon, T. Leanderson, Lack of J chain inhibits the transport of gut IgA and abrogates the development of intestinal antitoxic protection. *J. Immunol.* **163**, 913–919 (1999).
7. N. Kumar, C. P. Arthur, C. Ciferri, M. L. Matsumoto, Structure of the secretory immunoglobulin A core. *Science* **367**, 1008–1014 (2020).
8. Y. Li *et al.*, Structural insights into immunoglobulin M. *Science* **367**, 1014–1017 (2020).
9. Z. Sun *et al.*, The evolution and functional characterization of CXC chemokines and receptors in lamprey. *Dev. Comp. Immunol.* **116**, 103905 (2021).
10. K. Gangele, M. Jamsandekar, A. Mishra, K. M. Poluri, Unraveling the evolutionary origin of ELR motif using fish CXC chemokine CXCL8. *Fish Shellfish Immunol.* **93**, 17–27 (2019).
11. P. M. Murphy, "Chemokines and Chemokine Receptors" in *Fundamental Immunology*, M. F. Flajnik, N. J. Singh, S. Holland, Eds. (Wolters Kluwer, ed. 8, Philadelphia, 2022), pp. 344–385.
12. C. D. Castro, M. F. Flajnik, Putting J chain back on the map: How might its expression define plasma cell development? *J. Immunol.* **193**, 3248–3255 (2014).
13. V. S. Hohman *et al.*, J chain in the nurse shark: Implications for function in a lower vertebrate. *J. Immunol.* **170**, 6016–6023 (2003).
14. J. S. Boyles *et al.*, Discovery and characterization of a neutralizing pan-ELR+CXC chemokine monoclonal antibody. *MAbs* **12**, 1831880 (2020).
15. E. M. Flowers *et al.*, Identification of the Fc-alpha/mu receptor in Xenopus provides insight into the emergence of the poly-Ig receptor (pIgR) and mucosal Ig transport. *Eur. J. Immunol.* **51**, 2590–2606 (2021), 10.1002/eji.202149383.
16. E. Kallberg, T. Leanderson, A subset of dendritic cells express joining chain (J-chain) protein. *Immunology* **123**, 590–599 (2008).
17. L. Tacchi, E. Larragoite, I. Salinas, Discovery of J chain in African lungfish (*Protopterus dolloi*, Sarcopterygii) using high throughput transcriptome sequencing: Implications in mucosal immunity. *PLoS One* **8**, e70650 (2013).
18. K. Abdi *et al.*, Free IL-12p40 monomer is a polyfunctional adaptor for generating novel IL-12-like heterodimers extracellularly. *J. Immunol.* **192**, 6028–6036 (2014).
19. M. Sigvardsson, L. Olsson, E. Hogbom, T. Leanderson, Characterization of the joining chain (J-chain) promoter. *Scand. J. Immunol.* **38**, 411–416 (1993).