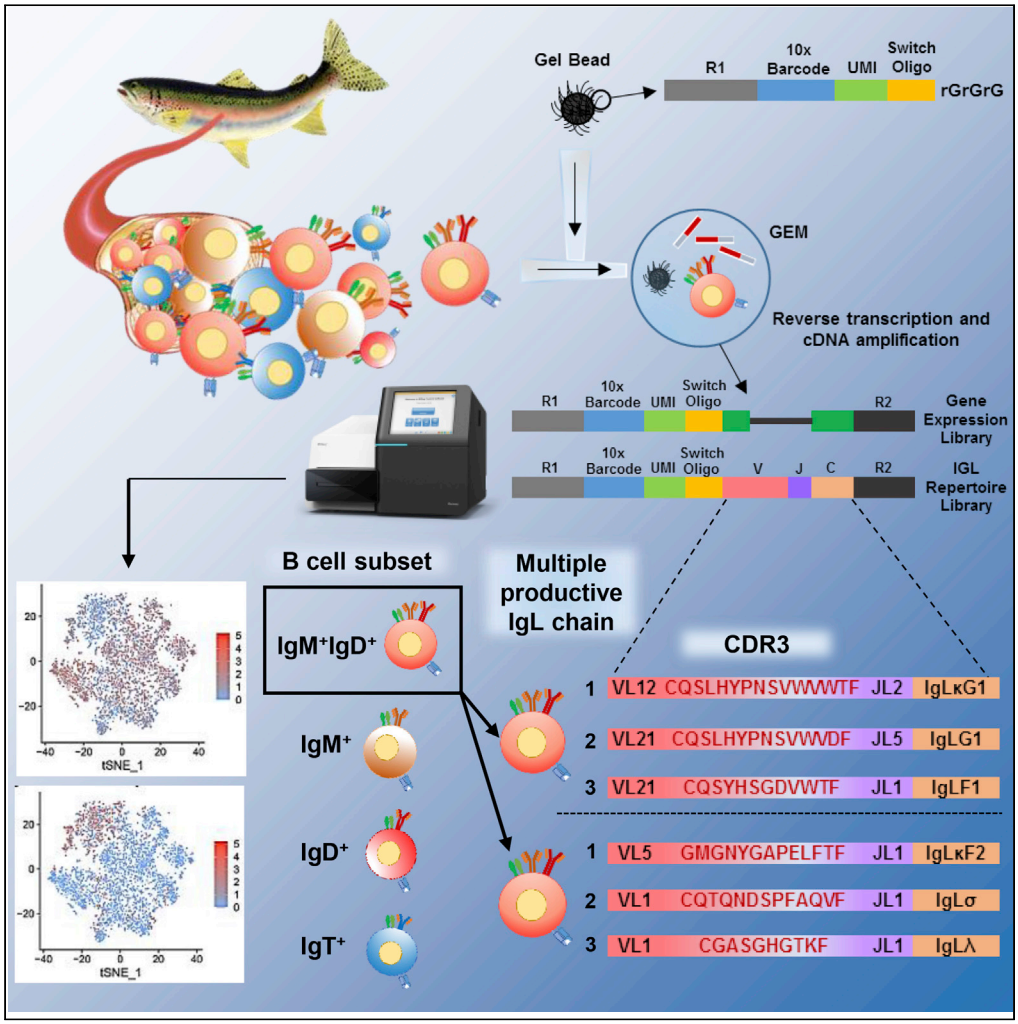
iScience

CellPress
OPEN ACCESS

Article

Individual B cells transcribe multiple rearranged immunoglobulin light chains in teleost fish



Pedro Perdiguero, Esther Morel, Patricia Díaz-Rosales, Carolina Tafalla

tafalla@inia.es

Highlights

Single teleost B cells transcribe different IgL chain sub-isotypes

In many cells, more than one transcript is productively rearranged as $V_LJ_LC_L$

These differently rearranged $V_LJ_LC_L$ genes render IgL chains with variable CDR3

This suggests that single fish B cells can produce Igs of different specificities

Perdiguero et al., iScience *24*, 102615
June 25, 2021 © 2021 The Author(s).
https://doi.org/10.1016/j.isci.2021.102615

iScience

CellPress
OPEN ACCESS

Article

Individual B cells transcribe multiple rearranged immunoglobulin light chains in teleost fish

Pedro Perdiguero,[1] Esther Morel,[1] Patricia Díaz-Rosales,[1] and Carolina Tafalla[1,2,*]

SUMMARY

B cells express a unique antibody protein which comprises two pairs of immunoglobulin (Ig) heavy (H) and light (L) chains. In addition to an invariable constant (C) region, IgH and IgL chains encompass a variable (V) region mediating antigen binding. This unique region stems from Ig V(D)J gene recombination, which generates diversity by assembling these gene segments into V_HDJ_H and V_LJ_L genes. To ensure that one B cell only expresses one antibody, V_HDJ_H rearrangement occurs only in one IgH locus (allelic exclusion), whereas V_LJ_L rearrangement only in either the κ or λ locus (isotype exclusion). However, teleosts express multiple IgLs encoded by distinct C_L genes. Using single-cell transcriptomics, we have demonstrated the transcription of distinct rearranged $V_LJ_LC_L$ genes in single rainbow trout B cells. Our results highlight the laxity of isotype exclusion in teleosts and strongly suggest that fish B cells can produce antibodies of different specificities.

INTRODUCTION

Antibodies, also termed immunoglobulins (Igs), are tetrameric molecules that include two identical heavy (H) chains and two identical light (L) chains. While IgH chains comprise one variable (V) domain and two to four constant (C) domains, IgL chains comprise one V domain and one C domain. V domains from paired IgH and IgL chains mediate antigen binding and account for the specificity of a given antibody. A recombined V_HDJ_H gene encodes each VH domain and results from the rearrangements of V, diversity (D) and joining (J) gene segments located in the IgH locus. In contrast, a recombined V_LJ_L gene encodes each VL domain and emerges from the rearrangement of V and J gene segments placed in the IgL locus. These V(D)J recombination events sequentially involve random selection of individual gene segments, generation of double-strand breaks in each gene segment by rearrangement activation gene 1 (RAG1) and RAG2 endonucleases, deletion of the intervening DNA, and ligation of the remaining gene segments (Rodgers, 2017). This complex process occurs in the bone marrow and generates antibody recognition diversity in an antigen-independent manner (Schroeder and Cavacini, 2010).

According to the nature of their C domain, encoded by C_H genes from the IgH locus, teleost express three distinct isotypes named IgM, IgD, and IgT, this last being specific to teleosts. When expressed as surface signal-transducing B cell receptors (BCR), specific IgM, IgD, or IgT combinations define discrete B cell subsets. Like in mammals, IgM^+IgD^+ cells constitute the main B cell subset in systemic immune tissues (Fillatreau et al., 2013; Simon et al., 2019). In these cells, IgM and IgD receptors are produced by alternative splicing of a long mRNA that includes the V_HDJ_H segment in addition to Cμ and Cδ and therefore are thought to express the same variable region (Geisberger et al., 2006). Upon activation by antigen, IgM^+IgD^+ B cells transcriptionally down-regulate surface IgD expression to become IgM^+IgD^- B cells, also defined as IgM^+ B cells, which include systemic IgM-secreting plasmablasts/plasma cells (Granja and Tafalla, 2019). Additionally, some IgM^+IgD^+ B cells lose surface IgM through a yet not well-defined class switch recombination event, generating IgM^-IgD^+ B cells, including IgD-secreting plasmablasts/plasma cells (Edholm et al., 2010; Perdiguero et al., 2019). Finally, IgT^+ B cells, including IgT-secreting plasmablasts/plasma cells, preferentially inhabit mucosal surfaces (Zhang et al., 2010). Consequently, these cells represent the main responders to mucosal antigens such as those from commensal bacteria, although systemic IgT responses have also been reported (Abos et al., 2018b; Castro et al., 2013). At a genomic level, three IgT genes (IgT1, IgT2 and IgT3) have been identified in teleost species such as rainbow trout (Zhang et al., 2017), but whether each IgT^+ B cell express only one or more of these IgTs is currently unknown.

[1]Animal Health Research Center (CISA-INIA), Valdeolmos, Madrid 28130, Spain

[2]Lead contact

*Correspondence: tafalla@inia.es

https://doi.org/10.1016/j.isci.2021.102615

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

As for IgL chains, most mammals studied to date express two isotypes located in two separate gene loci, which are known as Igκ and Igλ (Das et al., 2008). In mammals, these IgL gene loci have a translocon arrangement which comprised multiple V_L and J_L segments positioned upstream of a single C_L gene segment. In teleosts, IgL gene loci have a multi-clustered orientation, with small clusters containing V_L, J_L, and C_L scattered over multiple chromosomes (Daggfeldt et al., 1993; Ghaffari and Lobb, 1993; Zimmerman et al., 2008). In teleost species such as rainbow trout (*Oncorhynchus mykiss*), recent studies have defined four groups of IgL genes with sufficiently different sequences to be considered different isotypes (Daggfeldt et al., 1993; Edholm et al., 2009; Partula et al., 1996). In these IgL groups, three IgL1, two IgL2, two IgL3, and a single IgL4 sub-isotype have been identified (Zhang et al., 2016). IgL1 and IgL3 correspond to κ-like chains, currently named κG and κF, respectively. IgL2 corresponds to IgLσ and has been identified from cartilaginous fish to coelacanth and frogs (Guselnikov et al., 2018). Finally, the IgL4 isotype corresponds to a λ-like gene which up to date has been identified in cod, catfish and rainbow trout (Edholm et al., 2009; Guselnikov et al., 2018; Zhang et al., 2016). Although this isotype has been proposed by some authors as the ortholog of the λ isotype from tetrapods, based on specific gene features, the IgL4 isotype has been recently cataloged as a member of a new λ group (λ-2), as it was recently found that IgL4 emerged prior to the radiation of ray-finned fish (Guselnikov et al., 2018). Thus, in particular, in rainbow trout, eight IgL sub-isotypes are located in different clusters along the genome. IgL2 and IgL4 are located in chromosomes 13 and 17 respectively, whereas genes encoding the different IgL1 and IgL3 isotypes have been identified in different clusters throughout chromosomes 7, 15, 18, and 21 (Rego et al., 2020b). In total, 48 constant (C_L), 87 variable (V_L), and 59 joining (J_L) productive genes have been identified along the rainbow trout Swanson line genome with some specific differences in other lines, highlighting the complexity of the rainbow trout IgL repertoire (Rego et al., 2020b). In general, the expansion of IgL genes from teleost fish may be considered as part of the evolutionary genome duplication events that led to a similar expansion of distinct immune gene families (Glasauer and Neuhauss, 2014).

In mammals, V_HDJ_H rearrangements occur only in one IgH locus-targeted allele, which determines the so-called "one cell-one antibody" rule. This process, commonly defined as allelic exclusion, causes productively rearranged material from one allele to functionally silence any transcript derived from the second allele (Vettermann and Schlissel, 2010). Similarly, the V_LJ_L rearrangement occurs only in one Igκ locus- or Igλ locus-targeted allele through a process defined as isotype exclusion (Frippiat et al., 1995). However, recent evidence from mammals has begun to challenge this paradigm and points to the possibility that a significant percentage of B cells could express antibodies of at least two specificities (Shi et al., 2019). This probability might be even higher in species such as teleost fish, in which more IgL genes are present throughout the genome, which significantly increases the chances for one B cell to express more than one IgL chain.

To verify this hypothesis, we interrogated single B cells from rainbow trout by single cell transcriptomics. We found that individual cells from each major B cell subset commonly transcribed more than one IgL chain. In a high percentage of these cells, more than one of these transcribed IgL chains was productively rearranged as $V_LJ_LC_L$, demonstrating that rainbow trout individual B cells have the capacity to transcribe antibodies of different specificities. These results strongly suggest that individual B cells from teleosts can express antibodies that recognize different antigens, thereby challenging the "one cell-one antibody" paradigm.

RESULTS

Isolation and quality analysis of single B cells from rainbow trout

To acquire a transcriptional profile of B cells from the peripheral blood of rainbow trout at single-cell resolution, leukocyte isolation followed by cell sorting was carried out using blood from three healthy animals (Figure 1A). Given the paucity of B cell antibodies in this type of fish and to avoid the purification of B cells by using an anti-IgM that may trigger BCR-mediated activating signals, we sorted small MHC IIβ$^+$ lymphoid cells with low intracellular complexity. In peripheral blood from rainbow trout, these cells are almost exclusively comprised of B cells. Cell viability was more than 98% after sorting. Around 2,500 cells per sample (700 cells/μL) were used in a 10 × Chromium platform for single-cell RNA-sequencing using Chip A, which is specific for 5′ gene expression analysis (Figure 1A). Single cell cDNA libraries were sequenced using Illumina HiSeq 150PE and a total number of 73,934,335, 83,626,235 and 76,465,952 raw reads was obtained in each fish. After quality control and mapping using Cell Ranger software, the transcriptomes of 1,078, 2,189,

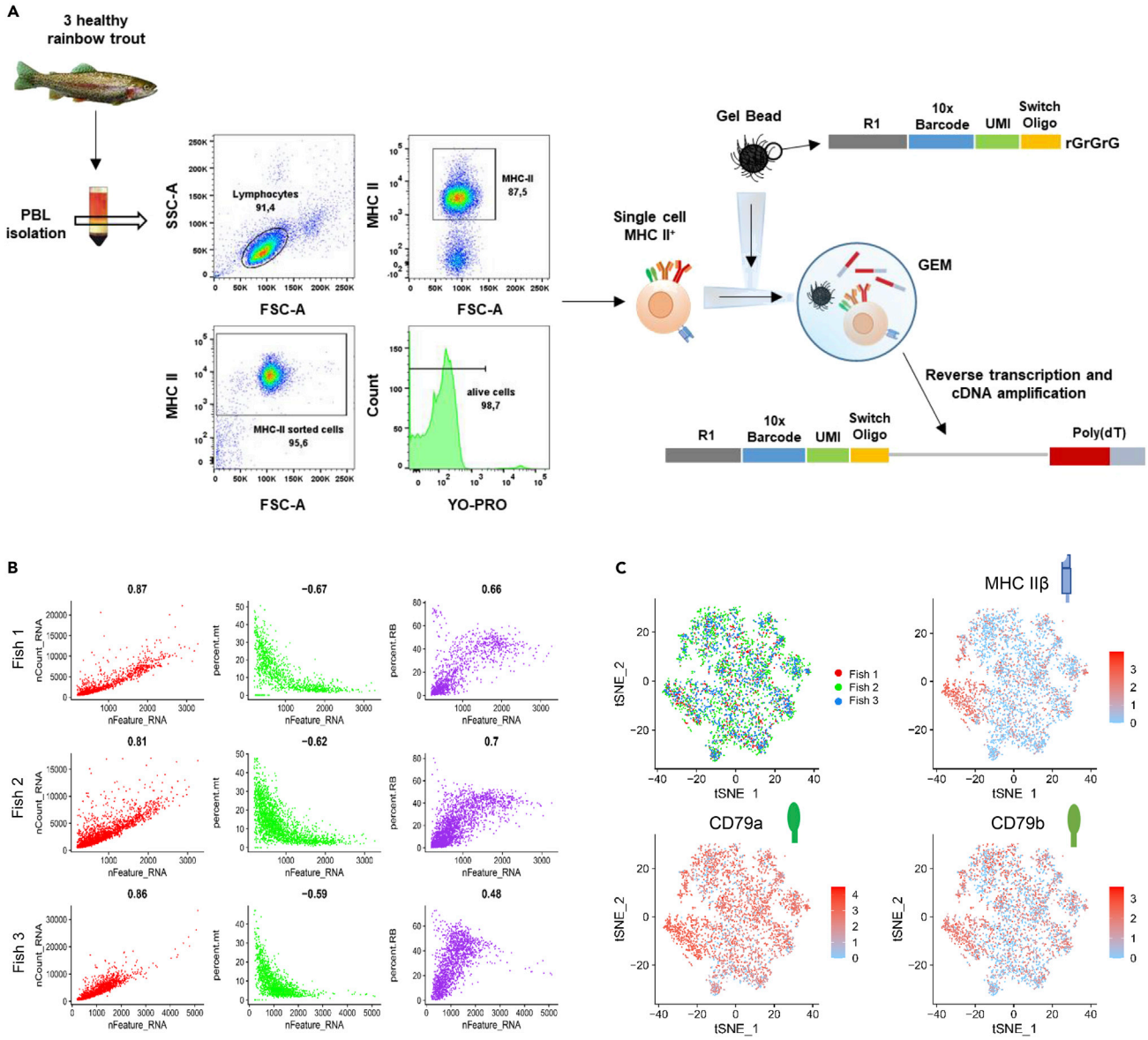


Figure 1. Isolation and quality analysis of MHC-IIβ$^+$ lymphoid cells from the peripheral blood of rainbow trout

(A) Schematics of workflow. Peripheral blood leukocytes (PBLs) were isolated from three healthy fish using Percoll gradients. Small MHC IIβ$^+$ lymphocytes with low intracellular complexity were then FACS sorted. Cell viability and purity were established after sorting. A total of 2,500 cells were used to conduct cell individualization in Gel Bead-In Emulsions (GEMs) using the 10 × Genomics A chip. In this process, single cells were included in GEMs together with specific gel beads and a reverse transcription mix. The global transcriptome from each cell was then amplified and sequenced for 5′ expression analysis.

(B) Feature scatterplot for quality control of cells isolated from fish 1, 2, and 3. Graphs show the total number of read counts, the percentage of reads mapping to mitochondrial genes, as well as the percentage of reads mapping ribosomal proteins in comparison to the total number of genes detected in each cell.

(C) t-distributed stochastic neighbor embedding (t-SNE) projection of cells after integration of data from each fish. Projection of cells showing MHC IIβ, CD79a and CD79b transcription. Color bar scales show log2 SCTransform (SCT) expression values.

and 1,489 single cells from each fish were acquired, detecting a median of 759, 688, and 1,025 genes per cell, respectively.

The analysis included a second step of cell quality control, which was performed using the Seurat software (v3.0). Initially, a filter was applied to each data set according to which only genes expressed in at least three cells from each fish were included in subsequent analysis. At this stage, cells with an identified number of

genes below 200 were discarded. Single cell sequencing data from each fish was analyzed to establish the distribution read count, the percentage of reads mapping to mitochondrial genes as well as the percentage of reads mapping ribosomal genes in relation to the total number of genes detected in each cell (Figure 1B). The distribution of genes and reads was found similar in each fish, making up for approximately 2,500 genes per cell or 15,000 counts per cell (Figure 1B). In fish 3, some cells with higher numbers of genes or reads were considered abnormal or potential doublets and thus excluded from further analysis. Of note, a wide distribution of mitochondrial reads was observed in each fish, reaching values of 50% (Figure 1B). Commonly, a high percentage of mitochondrial reads is associated with abnormal cells or cells that have been damaged during isolation. For this reason, cells showing a percentage of mitochondrial reads higher than 25% were considered abnormal and excluded from subsequent analysis. Similarly, a reduced number of cells with no mitochondrial reads were filtered out at this step. The percentage of mitochondrial reads and ribosomal reads showed reciprocally opposite trends when plotted versus the number of genes (Figure 1B). Both variables may introduce non-desired data variability and for this reason were taken into account during data normalization. After filtering, a total of 843, 1,814, and 1,327 cells from each fish were retained for subsequent analysis.

Expression matrices from cells isolated from each fish were integrated using the SCT transform method implemented in the Seurat software, which allowed a comparable distribution of cells from each fish while avoiding a batch effect (Figure 1C). In order to verify the correct isolation of B cells, multiple known markers of these lymphocytes were analyzed and visualized along a t-SNE (t-distributed stochastic neighbor embedding) cell projection. Thus, cell expressing transcripts for MHC IIβ (LOC100500791), CD79a (LOC110537828) and CD79b (LOC110491709) appeared widely distributed along the t-SNE projection used for data visualization (Figure 1C). CD79a, a recently defined B cell marker in salmonid fish (Penaranda et al., 2019), showed higher expression than MHC IIβ or CD79b, reaching 85.8%, 81.47%, and 92.16% of analyzed cells in fish 1, 2, and 3, respectively (Figure 1C).

IgH chain expression in single B cells from rainbow trout

IgH chains expressed by single B cells were analyzed by exploring reads mapping the constant regions of IgM, IgD, and IgT isotypes, including IgT1, IgT2, and IgT3 subtypes. IgM was expressed in 69.5%, 58.8%, and 63.9% of cells from fish 1, 2, and 3, respectively (Figure 2A). Similarly, IgD showed transcriptional activity in 63.9%, 49%, and 55.5% of cells from fish 1, 2, and 3, respectively (Figure 2A). The percentage of cells transcribing IgT1 was 6%, 10.6%, and 13.4% for fish 1, 2, and 3, respectively, while IgT3 was comparatively expressed in a lower fraction of cells from each fish (Figure 2A). Of note, IgT2 was identified in 18.5 and 24.9% of cells from fish 2 and fish 3, respectively, whereas no IgT2 was detected in fish 1 (Figure 2A). Similar IgM, IgD, and IgT frequencies were also observed in t-SNE projections. Both IgM and IgD showed a wide distribution, with IgM having higher expression values than IgD (Figure 2B). Remarkably, IgTs clustered together in the top region of t-SNE projections, with higher numbers of IgT2 transcripts per cell compared to IgT1 and IgT3 transcripts (Figure 2B).

When the transcriptional activity of different IgH chains was analyzed at the same time at a single cell level, cells transcribing different Ig combinations were evident in each fish at variable percentages. As previously established by flow cytometry (Abos et al., 2018a; Simon et al., 2019), the most common pattern of IgH chain co-expression was a combination of IgM and IgD, as $IgM^{+}IgD^{+}$ B cells accounted for approximately 52%, 33%, and 44% of cells in fish 1, 2, and 3, respectively (Figure 2C). The second and third most frequent patterns of IgH chain expression reflected cells that only transcribed IgM or IgD. Indeed, IgM^{+} B cells corresponded to 15%, 16% and 12% in fish 1, 2 and 3, respectively, whereas IgD^{+} B cells corresponded to 10%, 11%, and 7.2% in fish 1, 2, and 3, respectively (Figure 2C).

Cells expressing multiple IgTs were also identified and referred to hereafter as "multiple IgT cells". The combination IgT1 and IgT3 was identified in 0.2%–0.8% of cells, depending on the fish, whereas different combinations involving IgT2 were detected in fish 2 and 3 (Figure 2C). For instance, $IgT1^{+}IgT2^{+}$ cells accounted for 2.1% and 5% of cells in fish 2 and 3, respectively, whereas $IgT2^{+}IgT3^{+}$ cells constituted 0.5% and 1.1% of cells from fish 2 and 3, respectively (Figure 2C). Finally, $IgT1^{+}IgT2^{+}IgT3^{+}$ cells corresponded to 0.4 and 0.8% of cells from fish 2 and fish 3, respectively (Figure 2C).

Remarkably, some single cells showed hitherto undescribed transcriptional IgH expression patterns. These patterns included virtually any potential IgH combination and the percentage of cells expressing these

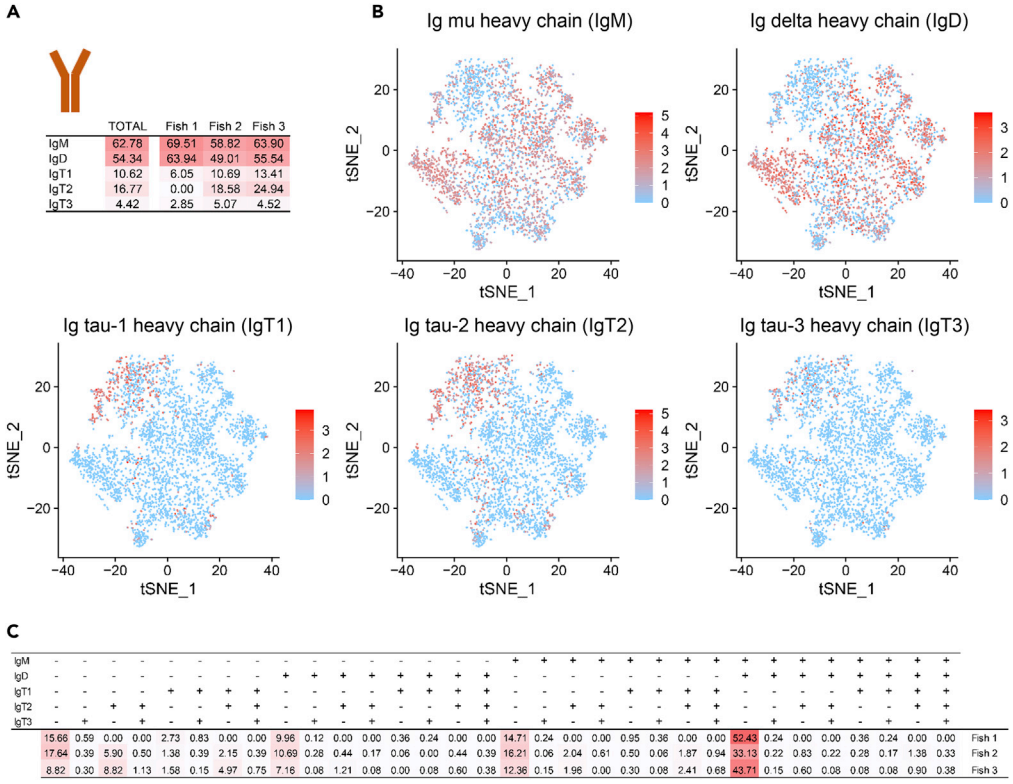


Figure 2. IgH chain expression in individual B cells from the peripheral blood of rainbow trout

(A) Heatmap showing percentages of cells expressing each IgH isotype from the total number of cells analyzed in each fish.

(B) t-SNE plots showing cells transcribing each IgH isotype. Color bar scales show log2 of SCTransform expression values.

(C) Heatmap showing the percentage of cells expressing each combination of IgH isotypes in each fish. + and – indicate expression of the isotype specified in each row and lack thereof, respectively.

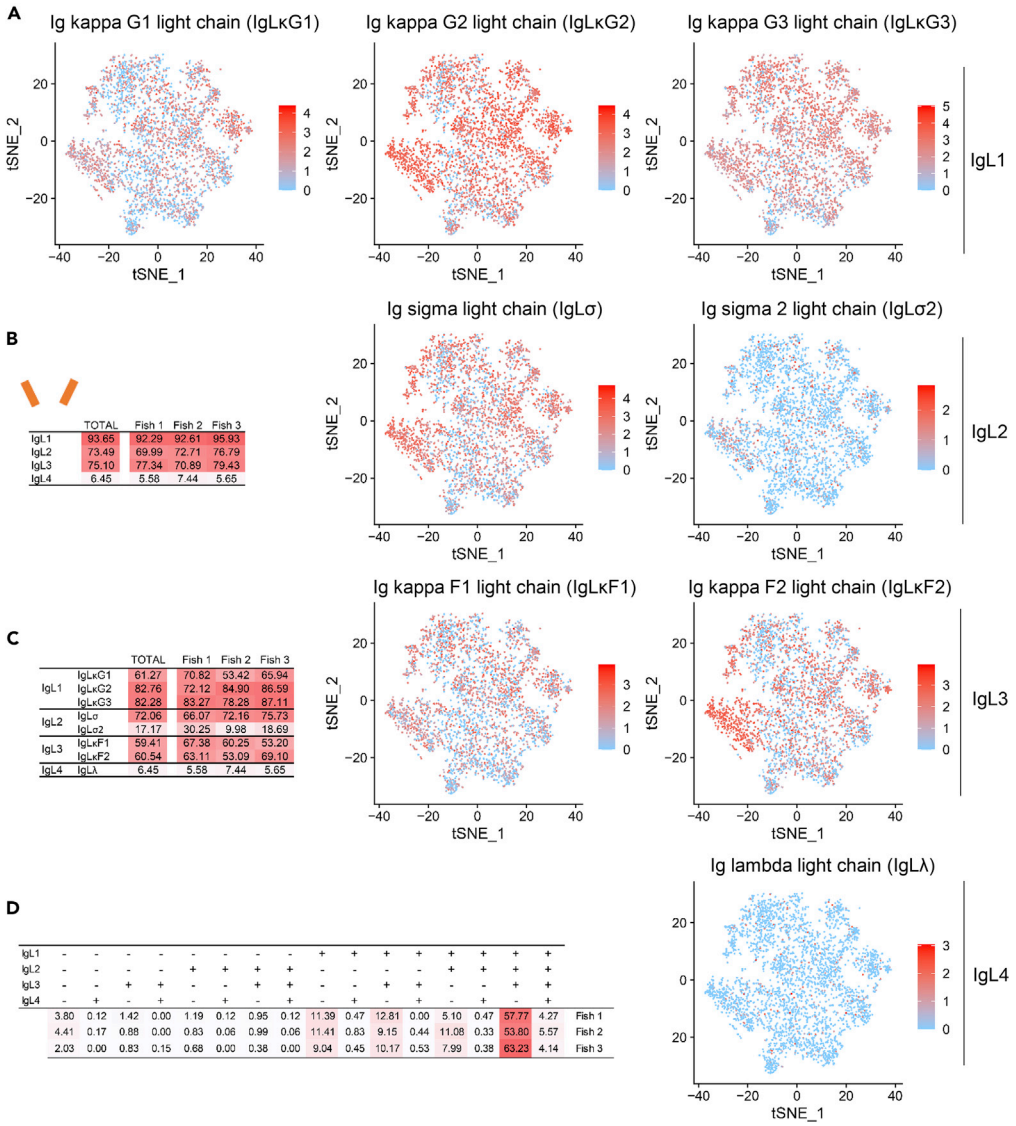
combinations was generally under 1% (Figure 2C). Finally, 16%, 18%, and 8.8% of B cells from fish 1, 2, and 3, respectively, showed no IgH expression (Figure 2C). These cells as well as those exhibiting unusual IgH combinations were excluded from further analysis.

IgL chain expression in single B cells from rainbow trout

B cells from rainbow trout express four major IgL chain isotypes termed IgL1 (IgLκG), IgL2 (IgLσ), IgL3 (IgLκF), and IgL4 (IgLλ). Within them, eight IgL sub-isotypes encoded by eight independent C_L genes termed IgLκG1, IgLκG2, IgLκG3, IgLσ, IgLσ2, IgLκF1, IgLκF2, and IgLλ are identified (Rego et al., 2020a; Zhang et al., 2016). Hence, we included each of these C_L regions in the mapping of the IgL chain transcriptional activity of single B cells.

All IgL sub-isotypes were transcriptionally active, with most being broadly expressed by single-sorted cells in t-SNE projections (Figure 3A). IgL1 was the most widely expressed IgL isotype, as it was identified in 92.3%, 92.6% and 95.9% of cells from fish 1, 2 and 3, respectively. In contrast, IgL2 and IgL3 were expressed by ~73.5% and ~75% of cells and had a comparable distribution in each fish. Finally, IgL4 was the least expressed IgL isotype, as it was detected in only ~6.5% of cells (Figure 3B).

When results were analyzed at a sub-isotype level, IgLκG2 and IgLκG3 were identified in ~82% of cells, whereas IgLκG1 was recognized in ~61.3% of cells, with percentages quite variable in each fish. While IgLσ was expressed in 66.1%, 72.2%, and 75.7% of cells in fish 1, 2, and 3, respectively, a quite variable frequency of cells, averaging about 17%, transcribed IgLσ2 in each fish, reaching a maximal frequency of 30.3% in fish 1. Finally, a comparable frequency of cells averaging to about 60% expressed IgLκF1 and IgLκF2 (Figure 3C).



B

	TOTAL	Fish 1	Fish 2	Fish 3
IgL1	93.65	92.29	92.61	95.93
IgL2	73.49	69.99	72.71	76.79
IgL3	75.10	77.34	70.89	79.43
IgL4	6.45	5.58	7.44	5.65

C

		TOTAL	Fish 1	Fish 2	Fish 3
IgL1	IgLκG1	61.27	70.82	53.42	65.94
	IgLκG2	82.76	72.12	84.90	86.59
	IgLκG3	82.28	83.27	78.28	87.11
IgL2	IgLσ	72.06	66.07	72.16	75.73
	IgLσ2	17.17	30.25	9.98	18.69
IgL3	IgLκF1	59.41	67.38	60.25	53.20
	IgLκF2	60.54	63.11	53.09	69.10
IgL4	IgLλ	6.45	5.58	7.44	5.65

D

IgL1	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	
IgL2	-	-	-	-	+	+	+	+	-	-	-	-	+	+	+	+	
IgL3	-	-	+	+	-	-	+	+	-	-	+	+	-	-	+	+	
IgL4	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	
	3.80	0.12	1.42	0.00	1.19	0.12	0.95	0.12	11.39	0.47	12.81	0.00	5.10	0.47	57.77	4.27	Fish 1
	4.41	0.17	0.88	0.00	0.83	0.06	0.99	0.06	11.41	0.83	9.15	0.44	11.08	0.33	53.80	5.57	Fish 2
	2.03	0.00	0.83	0.15	0.68	0.00	0.38	0.00	9.04	0.45	10.17	0.53	7.99	0.38	63.23	4.14	Fish 3

Figure 3. IgL chain expression in individual B cells from the peripheral blood of rainbow trout

(A) t-SNE plots showing cells expressing each IgL sub-isotype from rainbow trout. Color bar scale shows log2 of SCTransformed expression values.

(B) Heatmap showing the frequency of cells expressing each IgL isotype from the total number of cells analyzed and in each fish analyzed.

(C) Heatmap showing the frequency of cells expressing each IgL sub-isotype from the total number of cells analyzed and in each fish analyzed.

(D) Heatmap showing the frequency of cells expressing multiple IgL isotypes within each fish analyzed. + and – indicate expression of the isotype specified in each row and lack thereof, respectively.

Of note, several IgL combinations were identified at a single cell level. The most common expression pattern corresponded to IgL1, IgL2, and IgL3 co-expression, which was observed in 58%, 54%, and 63% of cells from fish 1, 2, and 3, respectively (Figure 3D). The co-expression of IgL1 and 2, as well as IgL1 and 3 was observed in ~10% of cells (Figure 3D). Among cells expressing only one IgL isotype, cells expressing only IgL1 were the most common and could be identified in 11%, 11%, and 9% of cells from fish 1, 2, and 3, respectively (Figure 3D). Remarkably, 4.3%, 5.6%, and 4.1% of cells from fish 1, 2 and 3, respectively, showed transcriptional activity of all four IgL isotypes (Figure 3D). The frequency of cells with no IgL expression ranged from 2% to 4.4% (Figure 3D), a number lower than that of cells with no IgH expression.

IgH and IgL chain co-expression in single B cells from rainbow trout

Having defined IgH and IgL chain expression in individual B cells, the combined expression of IgH and IgL chains was analyzed to further elucidate the whole BCR configuration at a single cell level. An initial analysis established that, in all major B cell subsets defined according to their surface IgH expression pattern, the presence of a single IgL chain was rare, because most cells expressed more than one IgL chain (Figure 4A). The most common pattern involved the co-expression of at least three IgL isotypes, which was observed in 342, 421, and 483 IgM^+IgD^+ cells from fish 1, 2, and 3, respectively (Figure 4A). This was also the most common expression pattern in IgM^+ cells and IgD^+ cells (Figure 4A). Interestingly, this dominant clonal co-expression of three IgL isotypes was not obvious in IgT^+ cell subsets, in which cells expressing two IgL or only one IgL isotype were also frequent (Figure 4A). The broad co-expression of three IgL isotypes by IgM^+IgD^+ cells became particularly evident in t-SNE projections (Figure 4B).

To further explore the preferential IgH and IgL chains combinations used by individual B cells, we established the IgL isotypes and sub-isotypes used by cells from each major B cell subset. Most IgM^+IgD^+ B cells co-expressed IgL1, 2 and 3 isotypes, which were identified in 339, 418, and 479 cells from fish 1, 2, and 3, respectively (Figure 4C). This IgL combination was also observed in 62, 137, and 116 IgM^+ cells from fish 1, 2, and 3, respectively, in 30, 83, and 58 IgD^+ cells from fish 1, 2, and 3, respectively, and in 51 and 32 $IgT2^+$ cells from fish 2 and 3, respectively. When two distinct IgL isotypes were co-expressed, these were mainly IgL1 and IgL2 or IgL1 and IgL3 (Figure 4C). Cells expressing these IgL combinations were present in comparable numbers within the major B cell subsets identified based on specific surface IgH patterns (Figure 4C). Remarkably, 28, 60, and 35 IgM^+IgD^+ cells from fish 1, 2, and 3, respectively, co-expressed transcripts encoding all the four IgL isotypes (Figure 4C). Other IgL isotype combinations were largely negligible. Finally, when cells expressed only one IgL isotype, this was mainly IgL1 (Figure 4C).

Next, the individual expression of the eight IgL sub-isotypes was dissected. This analysis revealed a common co-expression pattern that encompassed the transcription of all 7 sub-isotypes but IgLλ (Figure 4D). The number of IgM^+IgD^+ that expressed this pattern accounted for 148, 67, and 86 cells from fish 1, 2 and 3, respectively (Figure 4D). Cells co-expressing IgLκG1, IgLκG2, IgLκG3, IgLσ, IgLκF1, and IgLκF2 were also identified in 129, 209, and 172 IgM^+IgD^+ cells from fish 1, 2, and 3, respectively (Figure 4D). The co-expression of these six IgL sub-isotypes was the most common IgL expression pattern also in IgM^+ cells and IgD^+ cells (Figure 4D). In fish 3, the number of IgM^+IgD^+ clones that co-expressed five IgL sub-isotypes, including IgLκG1, IgLκG2, IgLκG3, IgLσ, and IgLκF2, was also high (80) (Figure 4D). Remarkably, the number of IgM^+IgD^+ clones showing co-transcription of eight IgL sub-isotypes was as low as 11, 15, and 6 cells from fish 1, 2, and 3, respectively, and negligible in other B cell subsets (Figure 4D). Finally, the expression pattern of different IgL sub-isotypes was more variable in IgT^+ B cell subsets, with many cells expressing only one IgL sub-isotype (Figure 4D).

Co-expression of multiple V_L and C_L transcripts in single IgM^+IgD^+ B cells from rainbow trout

Having unequivocally established that most IgM^+IgD^+ B cells transcribe the C_L region from multiple IgL chains, and considering that V_L and J_L gene segments undergo random recombination, expression of multiple IgL chains encoded by distinct C_L genes seemed to predict the transcription of multiple VL regions in individual B cells. Thus, our next step was to ascertain whether rainbow trout IgM^+IgD^+ B cells also transcribe the VL region from these IgL chains and whether these transcripts were productively rearranged to eventually produce a $V_LJ_LC_L$ protein. Our initial analysis using gene expression libraries demonstrated a significant over-representation of C_L genes in relation to V_L genes (Figure S1). Although this points to a high number of sterile IgL transcripts in B cells, we undertook a further analysis to investigate if more than one productively rearranged IgL chain could be identified in these cells. For this purpose, an amplification was performed using primers specific for the C_L 5′ end of the different IgL sub-isoforms up to the transcription star site, thus producing larger amplicons containing the full VL region, the cell barcode (CB) and the unique molecular identifier (UMI). These amplified products were sequenced and the specific repertoire sequencing data mapped against a database containing the full germline of V_L genes previously identified by Rego et al. (2020b) along the *O. mykiss* genome (v1.0). As predicted, the results obtained revealed that in a high proportion of cells (approximately 46%), no rearrangement of any IgL cluster was observed. However, the comparison of the CB:UMI combinations in the gene expression libraries mapped against V_L genes versus the CB:UMI combinations identified in the specific repertoire sequencing data highlighted the presence of several rearranged $V_LJ_LC_L$ transcripts in individual cells. Overall, among those individual cells that had at least one rearranged transcript, a higher proportion of cells contained more than one rearranged IgL cluster (approximately 60%), when compared to cells with only one rearranged IgL cluster

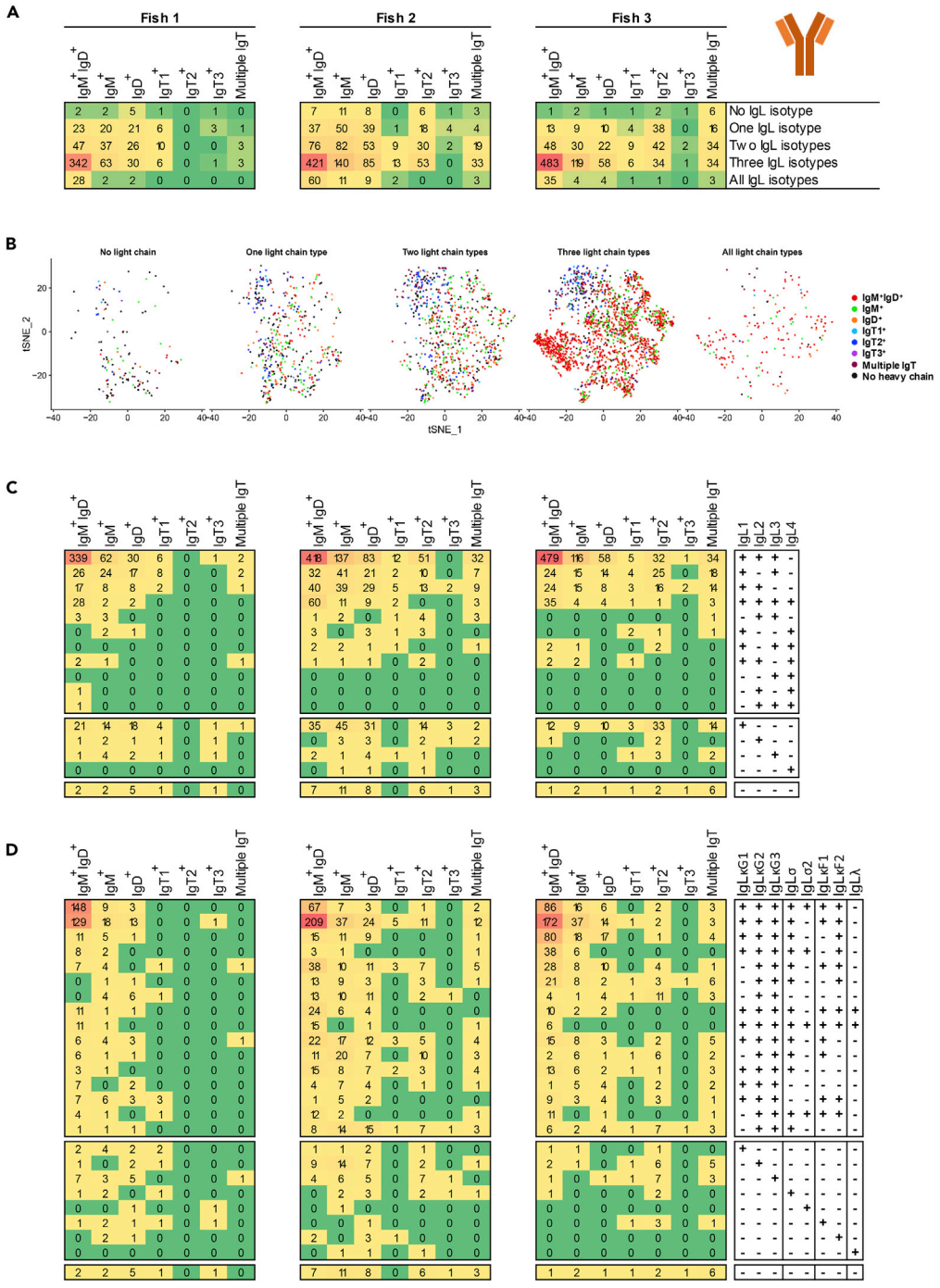


Figure 4. IgH and IgL chain co-expression in individual B cells from the peripheral blood of rainbow trout

(A) Heatmap showing the number of cells from each major B cell subset grouped based on the number of co-expressed IgL isotypes.

(B) t-SNE plots showing cells from each major B cell subset grouped based on the number of co-expressed IgL isotypes.

(C) Heatmap showing the number of cells from each major B cell subset co-expressing any possible IgL isotype combination.

(D) Heatmap depicting the number of cells from each major B cell subset co-expressing any possible IgL sub-isotype combination.

(approximately 40%) (Figure S2). Similar proportions were observed when cells were analyzed in individual fish or in randomly selected groups of cells, in which an average of 37% of cells had only one rearranged IgL cluster, whereas an average of 28%, 21%, and 14% of cells showed rearrangement in two, three, and at least four IgL clusters, respectively (Figure S2).

To further explore the presence of different rearranged transcripts in cells expressing several C_L genes, we chose individual IgM^+IgD^+ B cells expressing seven IgL sub-isotypes corresponding to four IgL isotypes (11, 24 and 10 cells from each fish) or eight IgL sub-isotypes (11, 15, and 6 cells from each fish). The 77 cells analyzed showed reads mapped to at least one V_L gene. Remarkably, in 16 of 22 cells (~72%) from fish 1, 29 of 39 cells (~74%) from fish 2, and 9 of 16 cells (56%) from fish 3, we obtained reads from V_L genes located at least in three different genome clusters commonly associated to distinct IgL sub-isotypes (Figure 5A). Thus, in 8 of 22 cells (~36%) from fish 1, 20 of 39 cells (~51%) from fish 2, and 8 of 16 cells (50%) from fish 3 germline rearrangements to produce at least two different rearranged $V_LJ_LC_L$ transcripts were identified (Figure 5A).

To further illustrate the transcription of multiple IgL chains in single cells resulting in diverse productive VL domains, we analyzed further in depth a set of B cells that showed multiple rearrangements, providing their CDR3 sequences (Figure 5B). For instance, in fish 1, a B cell identified with CB AAGGTTCCACATTCGA mainly transcribed an IgLkG1 rearranged transcript from the first cluster of chromosome 21. However, two other rearranged transcripts for productive IgLkG1 and IgLkF1 located in chromosome 15 were identified in this cell. Another cell from this fish (CB GACCTGGGTTACTGAC) transcribed up to five different productively rearranged transcripts. Two for IgLkG1 sub-isotypes from chromosome 15 and 21, respectively, both highly represented by different UMIs, and three other productive sequences encoding two IgLkF1 and IgLλ. The last cell analyzed from fish 1 (CB TGACGGCCAGTCGTGC) transcribed rearranged and productive transcripts for IgLκF1, IgLκF2, and IgLλ (Figure 5B). Similar patterns were identified in cells from fish 2 and fish 3 in which several rearranged IgL transcripts resulted in multiple combinations of productive IgL chains at single cell level (Figure 5B). These results evidence that rainbow trout B cells have the capacity to transcribe several rearranged IgL transcripts at the same time in a single cell.

DISCUSSION

In the current work, we have shown that B cell clones from rainbow trout showed unexpected IgH and IgL gene expression patterns, including co-expression of multiple V_L and C_L genes from distinct IgL sub-isotypes. Our findings highlight the laxity of isotype exclusion in teleosts, and suggest that clonal expression of multiple IgLs could facilitate the initial selection of specific B cells by antigen, thereby promoting the rapid initiation of protective antibody responses.

In 1955, Jerne formulated "The natural-selection theory of antibody formation". This theory postulates that Igs are continuously synthetized in a wide variety of configurations. Whenever an antigen enters the host, this antigen selects a unique antibody with complementary configuration among a vast array of pre-existing molecules, thus enhancing the production of this unique antibody (Jerne, 1955). In 1957, two independent studies lent further support to this theory, but introduced modifications emphasizing that antigen-mediated selection targets the cell expressing the Ig and not the Ig itself (Burnet, 1976; Talmage, 1957). This theory was later defined as "the clonal-selection theory of antibody diversity", which implies that individual B cells can only produce Igs of a unique specificity. This theory led to the subsequent formulation of what is today known as the "one cell-one antibody" paradigm (Viret and Gurr, 2009).

To verify the "one cell-one antibody" hypothesis, several studies were undertaken to identify cells with double reactivity after immunization with two different antigens. Most of these studies were unable to identify single B cells reactive against the two antigens (Green et al., 1967; Nossal and Lederberg, 1958; Nossal and Makela, 1962; White, 1958) or detected them in very small amounts (Nossal and Makela, 1962), but some exceptions were also reported. For example, immunization of rabbits with two serologically-unrelated bacteriophages, rendered approximately 15% of B cells with dual reactivity (Attardi et al., 1959). Despite these results, the consensus was that B cells only expressed one specificity, leading immunologists to assume the "one cell-one antibody" rule (Viret and Gurr, 2009).

Single-cell sequencing technologies have rapidly evolved over the past few years, revolutionizing several fields of research. Among the available technologies, 10 × Genomics has designed a methodology through which single-cell suspensions are introduced into nanoliter-scale droplets containing uniquely

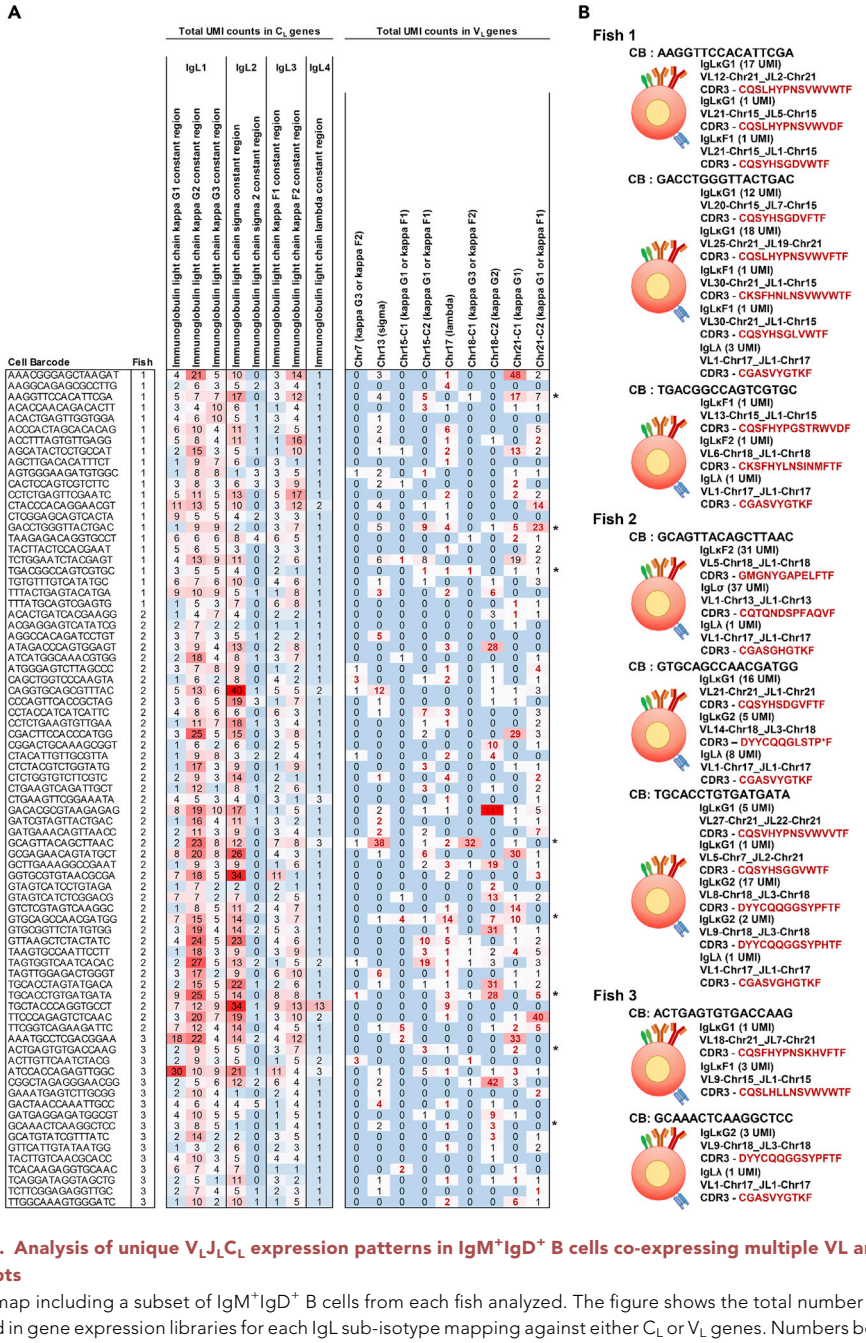


Figure 5. Analysis of unique $V_LJ_LC_L$ expression patterns in IgM^+IgD^+ B cells co-expressing multiple VL and CL transcripts

(A) Heatmap including a subset of IgM^+IgD^+ B cells from each fish analyzed. The figure shows the total number of UMIs identified in gene expression libraries for each IgL sub-isotype mapping against either C_L or V_L genes. Numbers bolded in red indicate the identification of shared CB:UMI codes between the gene expression libraries and the specific IgL repertoire sequencing libraries which reveal the presence of a rearranged $V_LJ_LC_L$ sequence.

(B) $V_LJ_LC_L$ recombination patterns and CDR3 sequences obtained for a subset of cells identified with an asterisk in (A). The specific sub-isotypes, the total number of UMIs identified, the V_LJ_L assignment and the CDR3 sequences are shown for each cell.

barcoded beads called GEMs (Gel Bead-In EMulsions), which can then be used to prepare next-generation sequencing libraries. State-of-the-art methodological approaches such as 10 × Genomics are allowing the revision of common accepted paradigms, including the BCR composition of individual B cells. Indeed, recent human studies have demonstrated the transcription of multiple BCRs in single B cells (Shi et al., 2019). In particular, two or more recombined V_HDJ_H and V_LJ_L genes have been identified in single human B cells in the order of hundreds or thousands of them.

In this work, we applied the 10 × Genomics technology to analyze B cells from fish, focusing our analysis on IgL gene transcription. First, it must be considered that, following the manufacturer's instructions, we adjusted the number of cells introduced into the chip to maximize the lack of doublets within GEMs. In our experimental setting, the ratio of doublets is expected to be below 0.8%, and most of these doublets are discarded during the filtering steps undertaken in response to aberrant gene number or read counts. Thus, the results obtained unequivocally established that a large fraction of cells from any major rainbow trout B cell subset transcribes multiple IgL sub-isotypes, most commonly 6–7 of them. Interestingly, this fraction is lower in the case of IgT^+ B cell subsets, which showed a large variability in the pattern of IgL chain expression. Given that IgT^+ B cells are thought to constitute a B cell linage independent of B cell subsets expressing IgM and/or IgD (Zhang et al., 2010), the lower frequency of IgT^+ clones with multiple IgL chain expression seems an interesting differential trait that should be further investigated both in systemic and in mucosal compartments. As IgT has been reported as a dedicated mucosal Ig (Zhang et al., 2010), it could be possible that the results obtained in a mucosal tissue are slightly different to those reported by us in blood.

Considering that V_L and J_L gene segments undergo random recombination, expression of multiple IgL chains encoded by distinct C_L genes seems to predict the expression of IgL chains with different CDRs. Nevertheless, our sequencing results revealed that transcription of C_L genes was very high in relation to that of V_L genes, suggesting that a high percentage of C_L transcripts are sterile non-rearranged transcripts. Despite this, the specific repertoire sequencing undertaken unequivocally demonstrates that rainbow trout individual B cells are capable of co-transcribing various productively rearranged IgL chains resulting in a variety of CDR3 sequences with different antigenic specificity, thus challenging the "one cell-one antibody" rule. Intriguingly, this observation is incompatible with commonly accepted isotype exclusion mechanisms, which entail the cessation of any V_L–J_L gene recombination upon the productive synthesis of a IgL chain, at least in mammals (Neuberger et al., 1989). The high percentage of cells identified with multiple rearranged IgL clusters seems to indicate that this is a common phenomenon in teleost, not just a consequence of errors in the isotype exclusion process. This flexibility has possibly been favored by the presence of multiple IgL locus throughout the genome, that have resulted after the successive whole-genome duplications suffered by these species.

Although our analysis was focused on IgL chains, we also found that a large fraction of IgT^+ B cells expressed multiple IgT sub-isotypes. Considering that each of these IgT genes have a dedicated set of VDJ gene segments that undergo random recombination, this finding also raises the possibility that at least some IgT^+ B cells may express multiple IgT specificities. However, we must also consider that salmonids such as the rainbow trout are tetraploid in nature and thus harbor two pairs of IgM and IgD loci in different alleles, which might change IgH allelic exclusion compared to how it occurs in mammals. To unequivocally conclude that single teleost B cells express multiple antibodies of different specificities, our single-cell transcriptomics-based findings should be confirmed at the protein level. However, the technology currently available in fish does not permit this study.

In summary, we propose that individual B cells from teleosts may express multiple productively recombined IgL chains with different specificity. Our results extend similar findings published previously in human B cells (Shi et al., 2019). Although our evidence needs solid confirmatory studies conducted at a protein level, the transcriptional evidence of more than one productively rearranged IgL chain in single teleost B cells, seems sufficient to justify additional state-of-the-art single-cell investigations aimed at validating the "one cell-one antibody" paradigm, at least in relevant species. Indeed, B cells co-expressing multiple Ig specificities may be more frequent in animal groups harboring multiple copies of a given antibody gene. In principle, this co-expression could increase the efficiency of the initial phase of protective humoral responses, as it would facilitate the targeting of B cell clones expressing multiple low-affinity BCRs over the targeting of rare B cell clones bearing a single matching BCR.

Limitation of the study

In this work, we show that rainbow trout individual B cells commonly transcribe more than one IgL chain. In most of these cells, more than one of these IgL isoforms is productively rearranged, demonstrating that antibodies of different specificities are transcribed by a high number of single B cells in this species. The co-expression of these multiple VL and CL genes has not been revealed at protein level in this study, as to date the techniques required for this are not available in rainbow trout.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.102615.

ACKOWLEDGMENTS

The authors want to thank Dr. Andrea Cerutti for critically reading the manuscript. Lucia González is also greatly acknowledged for technical assistance. This work was supported by the European Research Council (ERC Consolidator Grant 2016 725061 TEMUBLYM), by the Spanish Ministry of Science, Innovation and Universities (project AGL2017-85494-C2-1-R) and by the *Comunidad de Madrid* (grant 2016-T1/BIO-1672).

AUTHOR CONTRIBUTIONS

C.T. conceived and supervised the study. P.D.-R. obtained the leukocyte populations. E.M. performed the cell sorting. P.P. carried out all other experimental procedures, analyzed the sequencing data, prepared the figures and wrote an initial manuscript draft. C.T. wrote the final manuscript with inputs from all other authors.

DECLARATION OF INTEREST

The authors declare no competing interests.

Received: November 4, 2020
Revised: April 14, 2021
Accepted: May 19, 2021
Published: June 25, 2021

REFERENCES

Abos, B., Bird, S., Granja, A.G., Morel, E., More Bayona, J.A., Barreda, D.R., and Tafalla, C. (2018a). Identification of the first teleost CD5 molecule: additional evidence on phenotypical and functional similarities between fish IgM(+) B cells and mammalian B1 cells. J. Immunol. *201*, 465–480.

Abos, B., Estensoro, I., Perdiguero, P., Faber, M., Hu, Y., Diaz Rosales, P., Granja, A.G., Secombes, C.J., Holland, J.W., and Tafalla, C. (2018b). Dysregulation of B cell activity during proliferative kidney disease in rainbow trout. Front. Immunol. *9*, 1203.

Attardi, G., Cohn, M., Horibata, K., and Lennox, E.S. (1959). Symposium on the biology of cells modified by viruses or antigens. II. On the analysis of antibody synthesis at the cellular level. Bacteriol. Rev. *23*, 213–223.

Burnet, F.M. (1976). A modification of Jerne's theory of antibody production using the concept of clonal selection. CA Cancer J. Clin. *26*, 119–121.

Castro, R., Jouneau, L., Pham, H.P., Bouchez, O., Giudicelli, V., Lefranc, M.P., Quillet, E., Benmansour, A., Cazals, F., Six, A., et al. (2013). Teleost fish mount complex clonal IgM and IgT responses in spleen upon systemic viral infection. PLoS Pathog *9*, e1003098.

Daggfeldt, A., Bengten, E., and Pilstrom, L. (1993). A cluster type organization of the loci of the immunoglobulin light chain in Atlantic cod (*Gadus morhua* L.) and rainbow trout (*Oncorhynchus mykiss* Walbaum) indicated by nucleotide sequences of cDNAs and hybridization analysis. Immunogenetics *38*, 199–209.

Das, S., Nikolaidis, N., Klein, J., and Nei, M. (2008). Evolutionary redefinition of immunoglobulin light chain isotypes in tetrapods using molecular markers. Proc. Natl. Acad. Sci. U S A. *105*, 16647–16652.

Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. *30*, 207–210.

Edholm, E.S., Bengten, E., Stafford, J.L., Sahoo, M., Taylor, E.B., Miller, N.W., and Wilson, M. (2010). Identification of two IgD+ B cell

populations in channel catfish, *Ictalurus punctatus*. J. Immunol. *185*, 4082–4094.

Edholm, E.S., Wilson, M., Sahoo, M., Miller, N.W., Pilstrom, L., Wermenstam, N.E., and Bengten, E. (2009). Identification of Igsigma and Iglambda in channel catfish, *Ictalurus punctatus*, and Iglambda in Atlantic cod, *Gadus morhua*. Immunogenetics *61*, 353–370.

Fillatreau, S., Six, A., Magadan, S., Castro, R., Sunyer, J.O., and Boudinot, P. (2013). The astonishing diversity of Ig classes and B cell repertoires in teleost fish. Front. Immunol. *4*, 28.

Frippiat, J.P., Williams, S.C., Tomlinson, I.M., Cook, G.P., Cherif, D., Le Paslier, D., Collins, J.E., Dunham, I., Winter, G., and Lefranc, M.P. (1995). Organization of the human immunoglobulin lambda light-chain locus on chromosome 22q11.2. Hum. Mol. Genet. *4*, 983–991.

Geisberger, R., Lamers, M., and Achatz, G. (2006). The riddle of the dual expression of IgM and IgD. Immunology *118*, 429–437.

Ghaffari, S.H., and Lobb, C.J. (1993). Structure and genomic organization of immunoglobulin light chain in the channel catfish. An unusual genomic organizational pattern of segmental genes. J. Immunol. *151*, 6900–6912.

Glasauer, S.M., and Neuhauss, S.C. (2014). Whole-genome duplication in teleost fishes and its evolutionary consequences. Mol. Genet. Genomics *289*, 1045–1060.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., and Zeng, Q. (2011). Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nature Biotechnol. *29*, 644.

Granja, A.G., Leal, E., Pignatelli, J., Castro, R., Abos, B., Kato, G., Fischer, U., and Tafalla, C. (2015). Identification of teleost skin CD8alpha+ dendritic-like cells, representing a potential common ancestor for mammalian cross-presenting dendritic cells. J. Immunol. *195*, 1825–1837.

Granja, A.G., and Tafalla, C. (2019). Different IgM(+) B cell subpopulations residing within the peritoneal cavity of vaccinated rainbow trout are differently regulated by BAFF. Fish Shellfish Immunol. *85*, 9–17.

Green, I., Vassalli, P., Nussenzweig, V., and Benacerraf, B. (1967). Specificity of the antibodies produced by single cells following immunization with antigens bearing two types of antigenic determinants. J. Exp. Med. *125*, 511–526.

Guselnikov, S.V., Baranov, K.O., Najakshin, A.M., Mechetina, L.V., Chikaev, N.A., Makunin, A.I., Kulemzin, S.V., Andreyushkova, D.A., Stock, M., Wuertz, S., et al. (2018). Diversity of immunoglobulin light chain genes in non-teleost ray-finned fish uncovers IgL subdivision into five ancient isotypes. Front. Immunol. *9*, 1079.

Jerne, N.K. (1955). The natural-selection theory of antibody formation. Proc. Natl. Acad. Sci. U S A. *41*, 849–857.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357–359.

Li, H., Handsaker, B., WysokerFennell, A.T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079.

Neuberger, M.S., Caskey, H.M., Pettersson, S., Williams, G.T., and Surani, M.A. (1989). Isotype exclusion and transgene down-regulation in immunoglobulin-lambda transgenic mice. Nature *338*, 350–352.

Nossal, G.J., and Lederberg, J. (1958). Antibody production by single cells. Nature *181*, 1419–1420.

Nossal, G.J., and Makela, O. (1962). Kinetic studies on the incidence of cells appearing to form two antibodies. J. Immunol. *88*, 604–612.

Partula, S., Schwager, J., Timmusk, S., Pilstrom, L., and Charlemagne, J. (1996). A second immunoglobulin light chain isotype in the rainbow trout. Immunogenetics *45*, 44–51.

Penaranda, M.M.D., Jensen, I., Tollersrud, L.G., Bruun, J.A., and Jorgensen, J.B. (2019). Profiling the Atlantic salmon IgM(+) B cell surface proteome: novel information on teleost fish B cell protein repertoire and identification of potential B cell markers. Front. Immunol. *10*, 37.

Perdiguero, P., Martin-Martin, A., Benedicenti, O., Diaz-Rosales, P., Morel, E., Munoz-Atienza, E., Garcia-Flores, M., Simon, R., Soleto, I., Cerutti, A., and Tafalla, C. (2019). Teleost IgD(+)IgM(-) B cells mount clonally expanded and mildly mutated intestinal IgD responses in the absence of lymphoid follicles. Cell Rep *29*, 4223–4235, e4225.

Rego, K., Bengten, E., Wilson, M., Hansen, J.D., and Bromage, E.S. (2020a). Characterization of immunoglobulin light chain utilization and variable family diversity in rainbow trout. Dev. Comp. Immunol. *104*, 103566.

Rego, K., Hansen, J.D., and Bromage, E.S. (2020b). Genomic architecture and repertoire of the rainbow trout immunoglobulin light chain genes. Dev. Comp. Immunol. *113*, 103776.

Rodgers, K.K. (2017). Riches in RAGs: revealing the V(D)J recombinase through high-resolution structures. Trends Biochem. Sci. *42*, 72–84.

Rombel, I.T., Sykes, K.F., Rayner, S., and Johnston, S.A. (2002). ORF-FINDER: a vector for high-throughput gene identification. Gene *282*, 33–41.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. *33*, 495–502.

Schroeder, H.W., Jr., and Cavacini, L. (2010). Structure and function of immunoglobulins. J. Allergy Clin. Immunol. *125*, S41–S52.

Shi, Z., Zhang, Q., Yan, H., Yang, Y., Wang, P., Zhang, Y., Deng, Z., Yu, M., Zhou, W., Wang, Q., et al. (2019). More than one antibody of individual B cells revealed by single-cell immune profiling. Cell Discov *5*, 64.

Simon, R., Diaz-Rosales, P., Morel, E., Martin, D., Granja, A.G., and Tafalla, C. (2019). CpG oligodeoxynucleotides modulate innate and adaptive functions of IgM(+) B cells in rainbow trout. Front. Immunol. *10*, 584.

Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res. *27*, 491–499.

Talmage, D.W. (1957). Allergy and immunology. Annu. Rev. Med. *8*, 239–256.

Vettermann, C., and Schlissel, M.S. (2010). Allelic exclusion of immunoglobulin genes: models and mechanisms. Immunol. Rev. *237*, 22–42.

Viret, C., and Gurr, W. (2009). The origin of the "one cell-one antibody" rule. J. Immunol. *182*, 1229–1230.

White, R.G. (1958). Antibody production by single cells. Nature *182*, 1383–1384.

Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics *30*, 614–620.

Zhang, N., Zhang, X.J., Chen, D.D., Oriol Sunyer, J., and Zhang, Y.A. (2017). Molecular characterization and expression analysis of three subclasses of IgT in rainbow trout (*Oncorhynchus mykiss*). Dev. Comp. Immunol. *70*, 94–105.

Zhang, N., Zhang, X.J., Song, Y.L., Lu, X.B., Chen, D.D., Xia, X.Q., Sunyer, J.O., and Zhang, Y.A. (2016). Preferential combination between the light and heavy chain isotypes of fish immunoglobulins. Dev. Comp. Immunol. *61*, 169–179.

Zhang, Y.A., Salinas, I., Li, J., Parra, D., Bjork, S., Xu, Z., LaPatra, S.E., Bartholomew, J., and Sunyer, J.O. (2010). IgT, a primitive immunoglobulin class specialized in mucosal immunity. Nat. Immunol. *11*, 827–835.

Zheng, G.X., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat. Commun. *8*, 14049.

Zimmerman, A.M., Yeo, G., Howe, K., Maddox, B.J., and Steiner, L.A. (2008). Immunoglobulin light chain (IgL) genes in zebrafish: genomic configurations and inversional rearrangements between (V(L)-J(L)-C(L)) gene clusters. Dev. Comp. Immunol. *32*, 421–434.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rainbow trout anti-MHCII β-chain specific monoclonal Ab	Granja et al. (2015)	NA
AF647 goat anti-mouse Ig (H+L)	Invitrogen	Cat#A21236; RRID AB_2535805
Biological samples		
Healthy rainbow trout (*Oncorhynchus mykiss*), adults of 70–100 g	*Piscifactoria Cifuentes* (Cifuentes, Guadalajara, Spain)	NA
Chemicals, peptides, and recombinant proteins		
Benzocaine	Sigma-Aldrich	Cat#E1501
Heparin	Sigma-Aldrich	Cat#H3149
Leibovitz's medium (L-15)	Gibco	Cat#11415-049
Penicillin Streptomycin solution	Gibco	Cat#11548876
Fetal calf serum	Gibco	Cat#10270-106
Percoll	GE Healthcare	Cat#17-0891-01
APC Lightning-Link labeling kits	Innova Biosciences	Cat#705-0010
YO-PRO dye	Invitrogen	Cat#Y3603
Critical commercial assays		
Chromium™ Single Cell 5′ Gel Beads Kit	10× Genomics	Cat#PN-1000014
SPRIselect magnetic beads	Beckman Coulter	Cat#B23317
QuBit DNA quantification system	Invitrogen	Cat#Q33238
Agilent 2100 Bioanalyzer	Agilent Technologies	Cat#G2939BA
Q5 High-Fidelity DNA Polymerase	New England Biolabs	Cat#M0491S
NucleoSpin Gel and PCR Clean-up Kit	Macherey-Nagel	Cat#22740609.50
Deposited data		
Processed data 5′ gene expression libraries	This paper	GEO:GSE158102
Raw data gene expression libraries Illumina NextSeq	This paper	SRA BioProject:PRJNA664014SRA ACC:SRR12659396SRR12659397 SRR12659398
Raw data IgL repertoire libraries Illumina MiSeq	This paper	SRA BioProject:PRJNA664014SRA ACC:SRR14482288SRR14482289 SRR14482290
Oligonucleotides		
Primer Target Enrichment 1 FWF: AATGA TACGGCGACCACCGAGATCTACACTCT TTCCCTACACGACGCTC	This paper	NA
Primer target enrichment 1 IgLkG1 R1 R: CAAAGCAGAGGAACCAGCAG	This paper	NA
Primer target enrichment 1 IgLkG2 R1 R: GAGGTGCAGAGTAAAAACAGTGGA	This paper	NA
Primer target enrichment 1 IgLkG3 R1 R: TGCAGAATCCCAGATTGATGA	This paper	NA
Primer target enrichment 1 IgLσ R1 R: CATATATGCAGGTGAAACTTGTAATGA	This paper	NA

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Primer target enrichment 1 IgLσ2 R1 R: AGCTGCACTCATTCTGAAGCA	This paper	NA
Primer target enrichment 1 IgLkF1 R1 R: GCACCATGTGGGGAATCTCT	This paper	NA
Primer target enrichment 1 IgLkF2 R1 R: ACTCCTGCTTCCCCAGACAG	This paper	NA
Primer target enrichment 1 IgLλ R1 R: TCCCCTCCCTCAATCCTAATC	This paper	NA
Primer Target Enrichment 2 FW F: AATGATACGGCGACCACCGAGATCT	This paper	NA
Primer target enrichment 2 IgLkG1-2-3 R2R: GTGACTGGAGTTCAGACGTGTGCTCTTC CGATCTGGAAGCCCYTGTTGGCCAG	This paper	NA
Primer target enrichment IgLσ-2 R2R: GTGACTGGAGTTCAGACGTGTGCTCTTC CGATCTARSTCWTCGSTGGAYGGAGG	This paper	NA
Primer target enrichment IgLkF1-2 R2R: GTGACTGGAGTTCAGACGTGTGCTCTTC CGATCTCCAGAGAGCTGCTCAGAGGA	This paper	NA
Primer target enrichment IgLλ R2R: GTGACTGGAGTTCAGACGTGTGCTCTTC CGATCTCACACCAGGGTGGTCTTGTC	This paper	NA
Primer Sample Index FWF: AATGATA CGGCGACCACCGAGATCTACACT CTTTCCCTACACGACGCTC	This paper	NA
Primer Sample Index 4 Fish1R: CAAGC AGAAGACGGCATACGAGATTGGTC AGTGACTGGAGTTCAGACGTGT	This paper	NA
Primer Sample Index 6 Fish2R: CAAGC AGAAGACGGCATACGAGATATTGG CGTGACTGGAGTTCAGACGTGT	This paper	NA
Primer Sample Index 12 Fish3R: CAAGC AGAAGACGGCATACGAGATTACAA GGTGACTGGAGTTCAGACGTGT	This paper	NA
Software and algorithms		
BD FACSDiva™ software	BD Biosciences	www.bd.com
FlowJo® v.10	FlowJo LLC, Tree Star	https://www.flowjo.com/
Adobe Photoshop CS6 software	Adobe Systems	https://www.adobe.com/
NextSeq and MiSeq Analysis pipeline	Illumina	https://www.illumina.com/
Cell ranger software	10× Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest
Seurat software	Satija et al., 2015	https://satijalab.org/seurat/articles/install.html
Umitools software	Smith et al. (2017)	https://github.com/CGATOxford/UMI-tools
FASTX toolkit (Trimmer)		http://hannonlab.cshl.edu/fastx_toolkit/
Bowtie2	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Samtools	Li et al., 2009	http://www.htslib.org/download/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
PEAR software	Zhang et al. (2014)	https://cme.h-its.org/exelixis/web/software/pear/
Trinity software	Grabherr et al. (2011)	https://github.com/trinityrnaseq/trinityrnaseq/wiki
ORF finder software	Rombel et al. (2002)	https://www.ncbi.nlm.nih.gov/orffinder/
Office Excel 2010	Microsoft	https://www.microsoft.com

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Carolina Tafalla (tafalla@inia.es).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (Edgar et al., 2002) and are accessible through GEO Series accession number GSE158102. Raw data are accessible at the NCBI Sequence Read Archive (SRA) under BioProject PRJNA664014 with accession numbers SRR12659396, SRR12659397 and SRR12659398 for gene expression libraries and SRR14482288, SRR14482289 and SRR14482290 for IgL repertoire libraries.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Experimental fish

Female rainbow trout (*Oncorhynchus mykiss*) of approximately 70-100 g were obtained from *Piscifactoría Cifuentes* (Guadalajara, Spain). Fish were maintained at the Animal Health Research Center (CISA-INIA) laboratory at 14°C in a re-circulating water system with 12:12 h light:dark photoperiod. Fish were fed twice a day with a commercial diet (Skretting, Spain). Prior to sampling, fish were acclimatized to laboratory conditions for 2 weeks and during this period no clinical signs were ever observed. Procedures described comply with the Guidelines of the European Union Council (2010/63/EU) for the use of laboratory animals and were previously approved by the Ethics committee from the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA; Code CEEA PROEX002/17).

METHOD DETAILS

PBL isolation

Rainbow trout were killed by benzocaine (Sigma) overdose. Blood was extracted with a heparinized needle from the caudal vein and diluted 10 times with Leibovitz medium (L-15, Thermo Fisher Scientific) supplemented with 100 IU/ml penicillin and 100 μg/ml streptomycin (P/S, Thermo Fisher Scientific), 5% fetal calf serum (FCS, Thermo Fisher Scientific) and 10 IU/ml heparin (Sigma). PBLs were obtained by centrifugation (500 x *g* for 30 min at 4 °C) of diluted blood on 51% continuous Percoll (GE Healthcare) density gradients. The interface cells were collected, washed twice in L-15 containing antibiotics and 5% FCS and adjusted to 2 x 10^6 cells/ml.

Sorting of blood MHC II^+ cells

PBLs were washed in FACS staining buffer (phenol red-free L-15 medium supplemented with 2% FCS) and incubated with a monoclonal antibody specific for rainbow trout MHC II β-chain (mAb mouse IgG1 coupled to allophycocyanin, 2 μg/ml) previously characterized (Granja et al., 2015) for 30 min at 4 oC. After that time, cells were washed with FACS staining buffer and YO-PRO dye (0.05 μM) added to the cell suspension for dead cell exclusion. Lymphoid (small, low complexity) MHC $IIβ^+$ $YO\text{-}PRO^-$ (live) cells were then isolated in a FACSAria™ III flow cytometer (BD Biosciences) equipped with BD FACSDiva™ software (BD Biosciences).

The purity of the sorted population (above 98%) was confirmed in a FACS Celesta flow cytometer (BD Biosciences).

5′single cell library construction and sequencing

Isolated MHC IIβ$^+$ cells gently pipetted and diluted to a concentration of 700 cells/μl were used for cell isolation on a 10x Genomics Chromium Controller instrument (Zheng et al., 2017). All steps including PBL isolation, sorting and Chromium™ Single Cell isolation were carried out the same morning to avoid cell death. A total of 2,500 cells per donor were loaded into the chips of the Chromium™ Single Cell 5′ Gel Beads Kit (10x Genomics) and subjected to the Chromium Controller instrument to generate single cell Gel Bead-In Emulsions (GEMs) following manufacturer's instructions. Next, GEMs were subjected to library construction using the Chromium™ Single Cell 5′ Library Kit v1 (10x Genomics). As a first step, reverse transcription was performed, resulting in cDNA tagged with a cell-specific barcode and unique molecular index (UMI) per transcript. Fragments were then size selected using SPRIselect magnetic beads (Beckman Coulter). Next, Illumina sequencing adapters were ligated to the size-selected fragments and cleaned up using SPRIselect magnetic beads (Beckman Coulter). Finally, sample indices were selected and amplified, followed by a double sided size selection using SPRIselect magnetic beads (Beckman Coulter). Final library quality was assessed using an Agilent 2100 Bioanalyzer (Agilent technologies). Samples were then sequenced using a NextSeq instrument (Illumina) with 150PE chemistry.

Alignment and initial processing of sequencing data

The Cell Ranger software (10x Genomics, v3.1) was used to process the sequenced libraries. A rainbow trout reference transcriptome was constructed from RefSeq *Oncorhynchus mykiss* genome v1.0 using the "Cell Ranger mkref" tool. The sequences encoding constant regions from rainbow trout Igs were added to the fasta file prior to the construction of the reference transcriptome. For this, IgM and IgD constant regions were extracted from the IMGT database, whereas IgT1, IgT2 and IgT3 constant regions were defined according to Zhang *et al*. (Zhang et al., 2017). Additionally, sequences encoding the constant regions from the eight IgL sub-isotypes (IgLκG1, IgLκG2, IgLκG3, IgLκF1, IgLκF2, IgLσ1, IgLσ2, IgLλ) belonging to four IgL isotypes (IgLκG, IgLκF, IgLσ, and IgLλ), previously described in rainbow trout were included (Zhang et al., 2016). The complementary DNA reads from each donor were mapped against this reference using the "Cell Ranger count" tool. Through this system, filtered UMI expression matrices from each donor were generated. As a result, raw expression data was obtained containing transcriptomes for single MHC II$^+$ cells from the three different donor fish.

Quality filtration of cells

In accordance with published pipelines and quality control standards, abnormal cells in all datasets were uniformly filtered out on the basis of their gene expression distribution using the Seurat package (v3.1). Cells with at least 200 detected genes, and only those genes that appeared in at least three cells were included in an initial matrix for each fish. A cell was considered to be abnormal if any of the following criteria were met: (i) detected gene number >2,500; (ii) detected count number >15,000 or (iii) >25% of reads in a cell mapped to mitochondrial genes.

Sample integration and data reduction analysis

The SCTransform method from the Seurat software was applied in order to normalize the three filtered single-cell datasets from different fish. The percentages of mitochondrial and ribosomal proteins previously calculated for each cell were included as variables to regress data. Filtered and normalized datasets were then integrated using the PrepSCTIntegration tool to avoid a batch effect, allowing the systematic comparison between the three fish. The merged data was subjected to dimensionality reduction using the principal component analysis (PCA) followed by the t-distributed stochastic neighbor embedding (t-SNE) tool.

Specific IgL repertoire sequencing and $V_LJ_LC_L$ rearrangement analysis

An IgL repertoire analysis was performed following an adaptation of the Chromium Single Cell V(D)J Enrichment protocol, using specific primers in two successive PCR amplification steps. Briefly, 2.5 μl of the single cell amplified cDNAs from each fish were used as templates in a first target enrichment PCR amplification that included a cocktail of 8 specific reverse primers (final concentration 0.2 μM each) covering the 8 IgL rainbow trout sub-isotypes (20-50 bp after the stop codon) and the target enrichment FW1 primer from 10x Genomics (final concentration 0.2 μM). PCR was performed in a final volume of

50 μl with the Q5 High-Fidelity DNA Polymerase (0.02 U/μl) (New England Biolabs) and dNTPs (final concentration 0.2 mM each) with the following PCR program: a denaturalization cycle of 90s at 95°C; 8 cycles of 95°C 10s, annealing at 60°C during 20s followed by extension at 72°C during 40s; and a final extension cycle at 72°C during 4 min. PCR products were purified with the Macherey-Nagel™ NucleoSpin™ Gel and PCR Clean-up Kit following the manufacturer's instructions. A second PCR was performed using 2 μl of these purified PCR products as templates, including a cocktail of 4 specific reverse primers (final concentration 0.2 μM each) able to amplify all rainbow trout sub-isotypes from the 5' end of the C_L gene and the target enrichment FW2 primer from 10x Genomics (final concentration 0.2 μM). PCR conditions were the same as those in the first PCR, in this case with 10 cycles. PCR products were purified as mentioned above. Finally, the purified PCRs were subjected to a new amplification in order to include a Sample Index at the 3'end together with an Illumina adapter, this time using the target enrichment FW1 primer from 10x Genomics. These final libraries obtained were purified as mentioned above. Fragments bigger than 500 bp were further selected from an agarose gel to avoid the sequencing of undesired small PCR fragments. The purified libraries were sequenced using Illumina MiSeq 300PE.

In order to identify IgL rearranged transcripts, R1 sequences from gene expression libraries as well as from the specific repertoire libraries, were mapped independently against a specific reference containing the full germline of V_L genes identified along the rainbow trout genome described by Rego et al. (2020b). For this purpose, raw reads from both types of libraries were initially processed with the Umitools software (Smith et al., 2017) in order to extract the CB:UMI combination associated with each read. At this step, cells were filtered including as a whitelist the high confidence cell barcodes previously identified during the Seurat analysis. Filtered reads were mapped against the V_L reference using the Bowtie2 software (Langmead and Salzberg, 2012) allowing only one alignment for each read. Through this methodology, two raw UMI expression matrices with V_L assignation for both types of libraries were obtained by using the Umitools count tool. CB:UMI combinations of mapped reads from gene expression libraries were then compared with the CB:UMI combinations of mapped reads from the specific repertoire libraries. Shared CB:UMI combinations were selected as potential rearranged transcripts. All available reads for each CB:UMI combination from a set of cells were further extracted from the raw data sequence files. R1 and R2 reads from repertoire sequencing libraries were merged using the Pear software (Zhang et al., 2014) in order to obtain the full length sequence for each read. Then, merged reads from the repertoire sequencing libraries as well as R1 and R2 reads from the expression libraries for each CB:UMI code were independently assembled using the Trinity software (Grabherr et al., 2011) by which a high confidence consensus sequence for each CB:UMI combination was obtained. Reads showing high variability were discarded as potential PCR errors or cross amplifications. CB:UMI consensus sequences were compared with the $V_LJ_LC_L$ germline for a final assignment using the Blastn software. Consensus sequences were checked with ORF finder (Rombel et al., 2002) to determine if they are in frame. Finally, a V_LJ_L junction analysis was performed to identify potential CDR3 amino acid sequences.

QUANTIFICATION AND STATISTICAL ANALYSIS

All quantifications were performed as described in the Method Details section of the STAR Methods. No statistical analysis was applied to the data.