Reconstructing B cell receptor sequences from short-read single cell RNA-sequencing with BRAPeS

Shaked Afik[1] and Nir Yosef[1,2, 3, 4,*]

1. Center for Computational Biology, University of California, Berkeley, Berkeley, CA, 94720, USA

2. Department of Electrical Engineering and Computer Science, University of California, Berkeley, Berkeley, CA, 94720, USA

3. Ragon Institute of MGH, MIT and Harvard, Cambridge, MA, USA

4. Chan Zuckerberg Biohub, San Francisco, CA 94158, USA

* Corresponding author, niryosef@berkeley.edu.

ABSTRACT

RNA-sequencing of single B cells provides simultaneous measurements of the cell state and its binding specificity. However, in order to uncover the latter further reconstruction of the B cell receptor (BCR) sequence is needed. We present BRAPeS, an algorithm for reconstructing BCRs from short-read paired-end single cell RNA-sequencing. BRAPeS is accurate and achieves a high success rate even at very short (25bp) read length, which can decrease the cost and increase the number of cells that can be analyzed compared to long reads. BRAPeS is publicly available in the following link: https://github.com/YosefLab/BRAPeS.

BACKGROUND

B cells play a significant role in the adaptive immune system, providing protection against a wide range of pathogens. This diversity is due to the B cell receptor (BCR), which enables different cells to bind different pathogens [1]. Single cell RNA-sequencing (scRNA-seq) has emerged as one of the leading technologies to characterize and study heterogeneity in the immune system across cell types, development and dynamic processes [2,3]. Combining transcriptome analysis with BCR reconstruction in single cells can provide valuable insights to the relation between BCR and cell state, as was demonstrated by similar studies in T cells [4–6].

The BCR is comprised of two chains, a heavy chain (IgH) and a light chain (IgL, either a kappa or a lambda chain). Each chain is encoded in the germline by multiple segments of three types - variable (V), joining (J) and constant (C) segments (the heavy chain also includes a diversity (D) segment, see methods). The specificity of the BCRs comes from the V(D)J recombination process, in which for each chain one variable (V) and one joining (J) segments are recombined in a process which introduces mutations, insertions and deletion into the junction region between the segments, called the complementarity determining region 3 (CDR3) [7]. The resulting sequence is the main determinant of the cell's ability to recognize a specific antigen. Following B cell activation, somatic hypermutations are introduced in the complementarity determining regions of the BCR, and the constant region is replaced in a process termed isotype switching or class switching [8]. The random mutations make BCR reconstruction a challenging task. While methods to reconstruct BCR sequences from full length scRNA-seq are available [9–11] (as well as single cell V(D)J enriched libraries from 10x Genomics [12]), they were only tested on long reads (150bp and 50bp). The ability to reconstruct BCR sequences from short length (25-30bp) reads is important, as it can decrease cost which can, in turn, increase the number of cells which could be feasibly analyzed.

RESULTS AND DISCUSSION

We introduce BRAPeS ("BCR Reconstruction Algorithm for Paired-end Single cells"), an algorithm and software for BCR reconstruction. Conversely to other methods, BRAPeS was designed to work with short (25-30bp) reads, and indeed we demonstrate that under these settings it performs better than other methods. Furthermore, we show that the performance of BRAPeS when provided with short reads is similar to what can be achieved with much longer (50-150bp) reads from the same cells, suggesting that BCR reconstruction does not necessitate costly sequencing with many cycles.

BRAPes is an extension of the TCR reconstruction software TRAPeS [4] and takes advantage of the small scale of the reconstruction problem and performs a local assembly of the CDR3 (see methods and Figure S1 for full description of BRAPeS). Briefly, BRAPeS takes as input the alignment of the reads to the genome. BRAPeS first recognizes the possible V and J segments by finding reads with one mate mapping to a V segment and the other mate mapping to a J segment. Next, BRAPeS collects all unmapped reads whose mates were mapped to the V/J/C segments, assuming that most CDR3-originating reads will be unmapped when aligning to the reference genome. Then, BRAPeS reconstructs the CDR3 region with an iterative dynamic programming algorithm. At each step, BRAPeS aligns the unmapped reads to the edges of the V and J segments, using the sequence of the aligned reads to extend the V and J sequences until convergence. Finally, BRAPeS determines the BCR isotype by appending all possible constant segments to the reconstructed sequence and taking the most likely complete transcript based on transcriptomic alignment with RSEM [13]. BRAPeS determines the CDR3 sequences of all BCRs and its productivity based on the criteria established by the international ImMunoGeneTic information system (IMGT) [14,15]. BRAPeS reports all CDR3 sequences per cell, their productivity status, V/J/C segments and the number of reads mapped to the various segments of the BCRs (Additional File 1).

We evaluated BRAPeS' performance on 374 cells from two previously published datasets - 174 human B cells and 200 mouse B cells (methods) [9,16]. To evaluate BRAPes, we first trimmed the original reads (50bp for the human data and 150bp for the mouse data) and kept only the outer 25 or 30 bases. We compared BRAPeS' performance on the trimmed data to two other previously published methods - BASIC [9] and VDJPuzzle [10] applied either on the trimmed data or the original long reads.

When applied to 30bp reads, BRAPeS' success rates are similar to other methods for the light chain, but are higher for heavy chain reconstruction (Figure 1a, Table S1). BRAPeS reconstructs productive heavy chains in a total of 349 cells (93.3% of the cells across both datasets), and reconstructs productive light chains in 364 cells (97.3% of the cells). These results are in line with the success rates of BASIC and VDJPuzzle on the original long reads: BASIC reconstructs productive heavy and light chains in 352 (94.1%) and 364 (97.3%) cells, respectively, and VDJPuzzle reconstructs heavy chains in 346 (92.5%) cells and light chains in 368 (98.4%) cells. On 30bp reads, BASIC and VDJPuzzle achieve reconstruction rates similar and even slightly higher compared to long reads for the light chain (362 (96.8%) cells and 370 (98.9%) cells with a productive light chain in BASIC and VDJPuzzle, respectively). However, BASIC and VDJPuzzle see a decline in success rates for the heavy chain, reconstructing a productive heavy chain in only 273 (73%) cells for BASIC and 242 (64.7%) cells for VDJPuzzle (Figure 1a, Table S1).

BRAPeS is also able to maintain a high success rate on 25bp reads, reconstructing heavy chains in 331 (88.5%) cells and light chains in 357 (95.5%) cells (Figure 1b and Table S2). Yet, we observe a substantial decrease in the results of other methods. VDJPuzzle is unable to reconstruct any chains with 25bp reads. This is likely due to its use of the *De-novo* assembler Trinity [17] which requires a seed k-mer length of 25bp that is unsuitable for very short reads. Similarly to 30bp, BASIC is able to maintain a high reconstruction rate for light chains, with productive reconstructions in 363 (97.1%) cells, but is only able to reconstruct productive heavy chains in 204 (54.6%) cells (Figure 1b and Table S2). Moreover, BASIC only outputs fasta sequences, thus requiring further processing to annotate the BCR.

We next turn to evaluate the accuracy of the short-read based CDR3 reconstructions, by comparing the resulting sequences to those obtained with long reads (Figure 2, methods). We use the long-read based reconstruction of BASIC as a reference (we achieve similar results with VDJPuzzle on the long-read data; see Figure S2) and evaluate the accuracy in terms of sensitivity (how many of the CDR3 sequences in the long-read data have an identical reconstruction with the short-read data) and specificity (how many of the CDR3 sequences in the short-read data have an identical long-read reconstruction). In general, all methods show a high level of specificity, having almost all CDR3 sequences identical to the sequences reconstructed on long reads, whenever both read lengths produce a productive reconstruction (Figure 2a-b). In accordance with the higher success rate, BRAPeS shows a high sensitivity, with an average rate of 0.96 for 30bp data and 0.93 for 25bp data (Figure 2c-d). This is in line with the agreement of different methods on the original data, as VDJPuzzle on long reads has an average sensitivity rate of 0.96. On the trimmed data, BASIC and VDJPuzzle show a lower sensitivity rate - BASIC achieves sensitivity rates of 0.91 and 0.85 for 30bp and 25bp respectively, and VDJPuzzle has a sensitivity rate of 0.89 with the 30bp data.

Coupling BCR reconstruction with transcriptome analysis in single cells can provide valuable information about the effect of binding specificity and isotype to cellular heterogeneity. BRAPeS

is a software for BCR reconstruction which utilizes short-read scRNA-seq, allowing for decreased cost. BRAPeS is accurate and has a higher success rate on short reads compared to existing methods, especially for 25bp reads and heavy chains. BRAPeS is publicly available at https://github.com/YosefLab/BRAPeS

METHODS

The BRAPeS algorithm

The input given to BRAPeS is a directory where each subdirectory includes genomic alignments of a single cell.

The BRAPeS algorithm have several steps, performed separately for each chain in each cell:

1. **Identifying possible pairs of V and J segments:** BRAPeS searches for reads where one mate of the pair is mapped to a V segment and the other mate is mapped to a J segment. BRAPeS collects all possible V-J pairs and will try to reconstruct complete BCRs from all possible pairs. Since the D segment is very short, reads do not align to it, thus as part of the reconstruction step (step 3) BRAPeS will also reconstruct the sequence of the D segment. If no V-J pairs are found, BRAPeS will look for V-C and J-C pairs and will take all possible V/J pairing of the found V and J segments.
   In case of many possible V-J pairs (which can occur due to the similarity among the segments), the user can limit the number of V-J pairs to attempt reconstruction on. BRAPeS will rank the V-J pairs based on the number of reads mapped to them and take only the top few pairs (the exact number is a parameter controlled by the user).
2. **Collecting the set of putative CDR3-originating reads:** Next, BRAPeS collects the set of reads that are likely to originate from the CDR3 region. Those are the reads that are unmapped to the reference genome, but their mates are mapped to the V/J/C segments. In addition, since the first step of CDR3 reconstruction includes alignment to the ends of the genomic V and J sequences, BRAPeS also collects the reads mapping to the V and J segments.
3. **Reconstructing the CDR3 region:** For each V-J pair, BRAPeS extends the edges of the V and J segments with an iterative dynamic programming algorithm. BRAPeS starts

the reconstruction from the end bases of the V and J segment (3' end of the V segment and 5' end of the J segment). The number of bases is a parameter which can be controlled by the user, set by default to the length of the J segment. In each iteration, BRAPeS tries to align all the unmapped reads to the V and J segments separately with the Needleman-Wunsch algorithm with the following scoring scheme: +1 for match, -1 for mismatch, -20 for gap opening and -4 for gap extension. In addition, BRAPeS does not penalize having a read "flank" the genomic segment. All reads that passed a user defined threshold are considered successful alignments. BRAPeS will then build the extended V and J segments by taking for each position the base which appears in most reads. BRAPeS will continue to run this process for a given number of iterations or until the V and J segments overlap. BRAPeS can also run a "one-side" mode, where if an overlap was not found (e.g. due to assigning the wrong V segment), BRAPeS will attempt to determine the productivity of only the extended V and only of the extended J segment.

4. **Isotype determination:** to find the BCR class, for each V-J pair with a reconstructed CDR3, BRAPeS concatenates all possible constant segments. Then, BRAPeS runs RSEM [13] on all sequences using all the paired-end reads with at least one mate that was mapped the genomic V/J/C segments as input. For each V-J pair BRAPeS takes the constant region with the highest expected count as the chosen constant segment.
5. **Separating similar BCRs and determining chain productivity:** After selecting the top isotype for each V-J pair, BRAPeS determines if the reconstructed sequence is productive (i.e. the V and J are in the same frame with no stop codon in the CDR3) and annotates the CDR3 junction. If more than one V-J pair produces a CDR3 sequence (either due to having more than one recombined chain in the cell or due to similar V-J segments resulting in the same CDR3 sequence reconstruction), BRAPeS will rank the various reconstructions based on their expression values from RSEM.

The output for BRAPeS is the full ranked list of reconstructed chains, including the CDR3 sequences, V/J/C annotations and the number of reads mapped to each segment, as well as a summary file of the success rates across all cells.

BRAPeS is implemented in python. To increase performance, the dynamic programming algorithm is implemented in C++ using the SeqAn package [18]. Moreover, to decrease running time for deeply sequenced cells, BRAPeS has the option to randomly downsample the V-J aligning reads and the putative CDR3-originating reads to only 10,000 reads. BRAPeS is publicly available and can be downloaded in the following link:
https://github.com/YosefLab/BRAPeS

Data availability and preprocessing

Raw fastq files of mouse B cells were downloaded from Wu et al. (ArrayExpress E-MTAB-4825) [16]. All analysis was performed on the 200 cells that were available through ArrayExpress. Raw fastq files for the human data from Canzar et al. [9] were provided by the author. We excluded single-end cells and cells filtered out in the original study, leaving a total of 174 cells. Next, reads were trimmed to be 25 or 30bp paired-end with trimmomatic [19], keeping only the outer bases.

For BRAPeS, low quality reads were trimmed using trimmomatic with the following parameters: LEADING:15, TRAILING:15, SLIDINGWINDOW:4:15, MINLEN:16. The remaining reads were aligned to the genome (hg38 or mm10) using Tophat2 [20].

Running BRAPeS

For this study, BRAPeS was run using the following parameters for the human data: “-score 15 -top 6 -byExp -iterations 6 -downsample -oneSide”, and with the following parameters for the mouse data: “-score 15 -oneSide -byExp -top 10”. In addition, as some cells required a higher alignment score threshold, we ran BRAPeS with a scoring threshold of 21 for cells without a productive chain.

Running VDJPuzzle and BASIC

We ran VDJPuzzle using default parameters, providing VDJPuzzle with the hg38 genome and GRCh38.p2 annotation for human, and mm10 genome with the GRCm38.p4 annotation for mouse. We then considered only results which appeared in the “summary_corrected” folder as valid productive reconstructions.

BASIC was ran with default parameters. After running BASIC we collected all the output fasta files and ran them through IMGT/HighV-Quest [21,22]. Only sequences that resulted in productive CDR3 according to IMGT were considered successful reconstructions.

Comparison of sensitivity and specificity

To determine the accuracy of the methods, we compared the reconstructed CDR3 amino acid sequences to the reconstruction produced by running BASIC or VDJPuzzle on the long reads. Only CDR3s with amino acid sequences identical to the sequences reconstructed on the long-read data were considered accurate. In case of more than one reconstructed CDR3 sequence, if both methods had at least one identical CDR3 sequence it was considered an accurate reconstruction.

DECLARATIONS

Ethics approval and consent to participate

Not applicable

Availability of data and material

Mouse raw fastq files from Wu et al. were downloaded from ArrayExpress (E-MTAB-4825). Human raw fastq files from Canzar et al. were provided by the author. BRAPeS can be downloaded in the following link: https://github.com/YosefLab/BRAPeS

Competing interests

The authors declare they have no competing interests

Funding

This work was supported by US National Institute of Health, grant number 5U19AI090023-07

Authors' contributions

S.A. wrote BRAPeS, performed the analysis and wrote the manuscript. N.Y. designed and oversaw the study and wrote the manuscript.

ADDITIONAL FILES:

Additional file 1: BRAPeS output for datasets analyzed in this study. In addition to the standard output, the last column mentions the alignment threshold used for each reconstruction

FIGURES

Figure 1

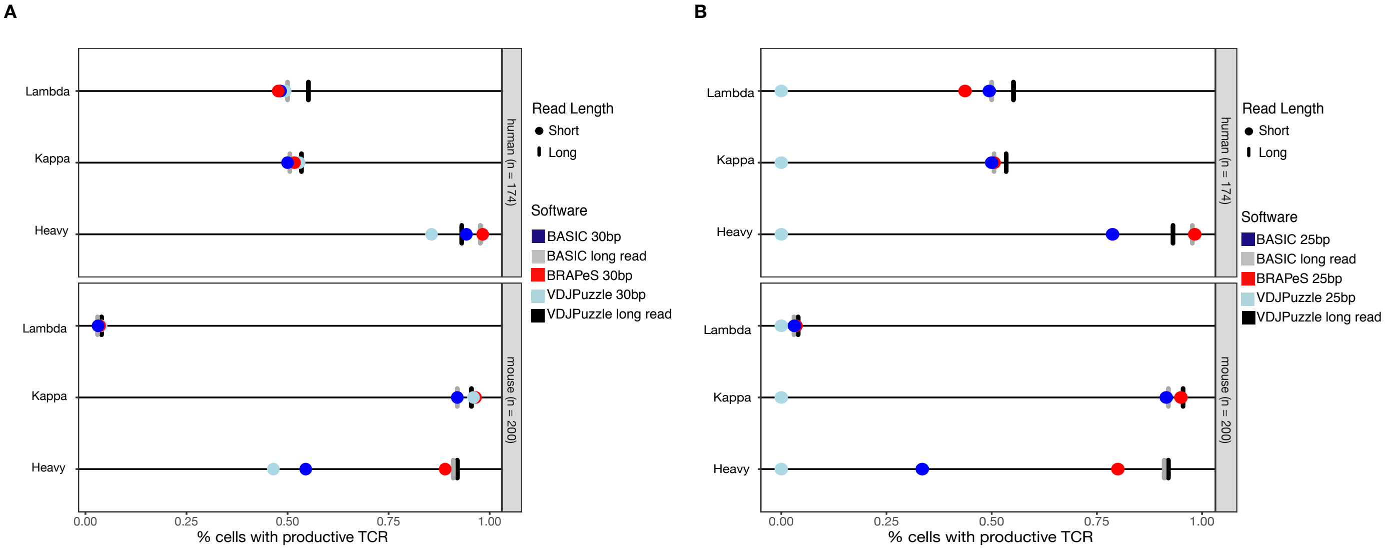

Figure 1: BRAPeS success rates. A) Fraction of cells with a successful reconstruction of a productive CDR3 in human and mouse B cells using the following methods: VDJPuzzle applied to the original, long-read data (black line) and the trimmed version of the data, trimmed to 30bp (light blue circle). BASIC applied to the long-read data (grey line) and the trimmed data (dark blue circle), and BRAPeS applied to the trimmed data (red circle). **B)** Same as A, but the trimmed version of the data was trimmed down to include only the outer 25bp, instead of 30bp.

Figure 2

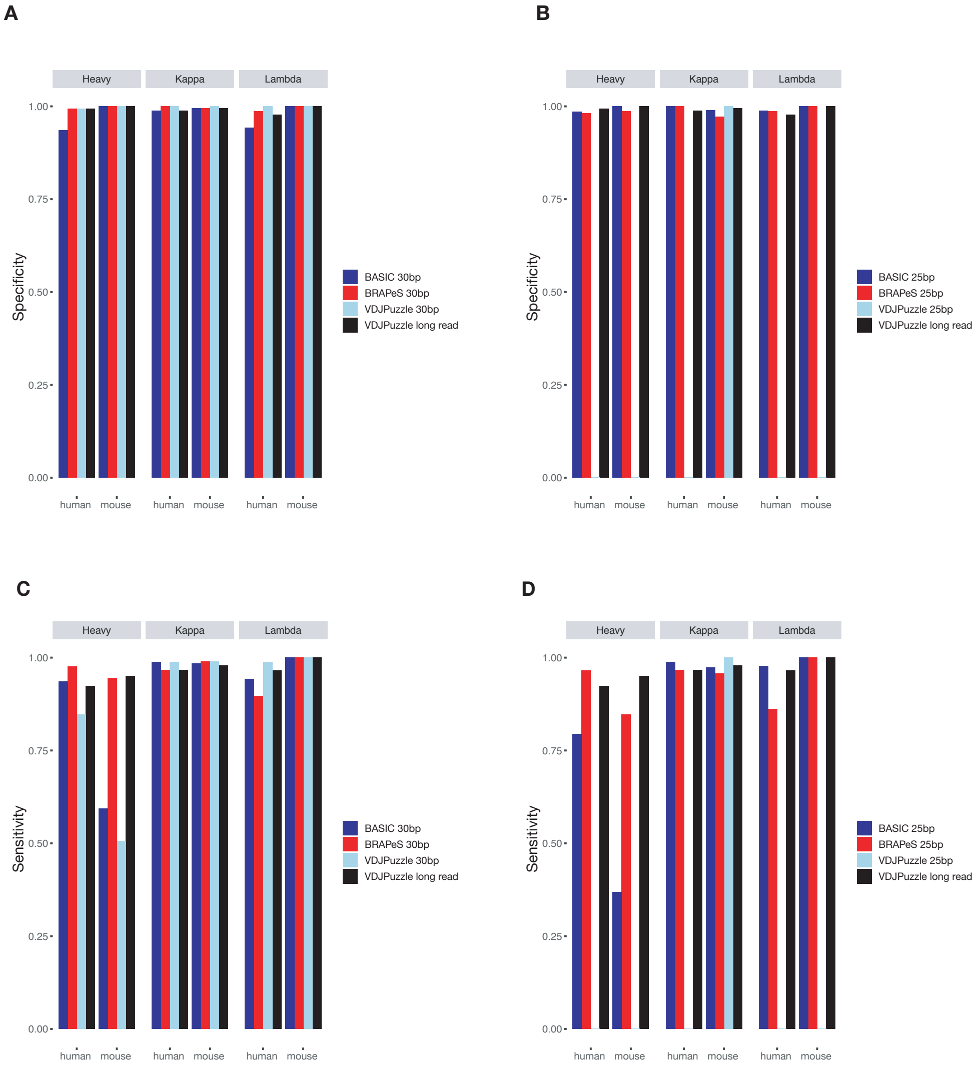

Figure 2: Sensitivity and specificity of BRAPeS. A) Specificity of BRAPeS for 30bp. The fraction of cells with a CDR3 sequence identical to the CDR3 reconstructed by BASIC on the long-read data, using the following methods: VDJPuzzle when applied to the long-read data (black), BRAPeS (red), BASIC (dark blue) and VDJPuzzle (light blue) applied to a version of the data trimmed to 30bp. The fraction is calculated only for cells that had a productive chain in both the long-read BASIC results and the other method. **B)** Specificity of BRAPeS for 25bp. Same as A, except the short-read version of the data was trimmed to include only the outer 25bp, instead of 30bp. **C)** Sensitivity of BRAPeS for 30bp. Same as A, except the fraction is calculated out of all the cells that had a productive chain when running BASIC on the long-read data. **D)** Sensitivity of BRAPeS for 25bp. Same as B, except the fraction is calculated out of all the cells that had a productive chain when running BASIC on the long-read data.

REFERENCES

1. Imkeller K, Wardemann H. Assessing human B cell repertoire diversity and convergence. Immunol Rev. Wiley Online Library; 2018;284:51–66.

2. Papalexi E, Satija R. Single-cell RNA sequencing to explore immune cell heterogeneity. Nat Rev Immunol. nature.com; 2018;18:35–45.

3. Villani A-C, Sarkizova S, Hacohen N. Systems Immunology: Learning the Rules of the Immune System. Annu Rev Immunol. 2018;36:813–42.

4. Afik S, Yates KB, Bi K, Darko S, Godec J, Gerdemann U, et al. Targeted reconstruction of T cell receptor sequence from single cell RNA-seq links CDR3 length to T cell differentiation state. Nucleic Acids Res. academic.oup.com; 2017;45:e148.

5. Stubbington MJT, Lönnberg T, Proserpio V, Clare S, Speak AO, Dougan G, et al. T cell fate and clonality inference from single-cell transcriptomes. Nat Methods. nature.com; 2016;13:329–32.

6. Eltahla AA, Rizzetto S, Pirozyan MR, Betz-Stablein BD, Venturi V, Kedzierska K, et al. Linking the T cell receptor to the single cell transcriptome in antigen-specific human T cells. Immunol Cell Biol. 2016;94:604–11.

7. Tonegawa S. Somatic generation of antibody diversity. Nature. Nature Publishing Group; 1983;302:575.

8. Di Noia JM, Neuberger MS. Molecular mechanisms of antibody somatic hypermutation. Annu Rev Biochem. 2007;76:1–22.

9. Canzar S, Neu KE, Tang Q, Wilson PC, Khan AA, Hancock J. BASIC: BCR assembly from single cells. Bioinformatics. Oxford University Press; 2017;33:425–7.

10. Rizzetto S, Koppstein DNP, Samir J, Singh M, Reed JH, Cai CH, et al. B-cell receptor reconstruction from single-cell RNA-seq with VDJPuzzle. Bioinformatics. 2018; Available from: http://dx.doi.org/10.1093/bioinformatics/bty203

11. Lindeman I, Emerton G, Mamanova L, Snir O, Polanski K, Qiao S-W, et al. BraCeR: B-cell-receptor reconstruction and clonality inference from single-cell RNA-seq. Nat Methods. 2018;15:563–5.

12. Single Cell Immune Profiling - 10x Genomics. 10x Genomics. 2018. https://www.10xgenomics.com/solutions/vdj/. Accessed 10 Aug 2018.

13. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12:323.

14. Lefranc M-P, Giudicelli V, Duroux P, Jabado-Michaloud J, Folch G, Aouinti S, et al. IMGT®, the international ImMunoGeneTics information system® 25 years on. Nucleic Acids Res. 2015;43:D413–22.

15. Lefranc M-P. Immunoglobulin and T Cell Receptor Genes: IMGT(®) and the Birth and Rise of Immunoinformatics. Front Immunol. 2014;5:22.

16. Wu YL, Stubbington MJT, Daly M, Teichmann SA, Rada C. Intrinsic transcriptional heterogeneity in B cells controls early class switching to IgE. J Exp Med. Rockefeller University Press; 2016;jem.20161056.

17. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29:644–52.

18. Döring A, Weese D, Rausch T, Reinert K. SeqAn An efficient, generic C++ library for sequence analysis. BMC Bioinformatics. 2008;9:11.

19. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.

20. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013;14:R36.

21. Alamyar E, Giudicelli V, Li S, Duroux P, Lefranc M-P. IMGT/HighV-QUEST: the IMGT® web portal for immunoglobulin (IG) or antibody and T cell receptor (TR) analysis from NGS high throughput and deep sequencing. Immunome Res. 2012;8:26.

22. Li S, Lefranc M-P, Miles JJ, Alamyar E, Giudicelli V, Duroux P, et al. IMGT/HighV QUEST paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire immunoprofiling. Nat Commun. 2013;4:2333.

Figure S1

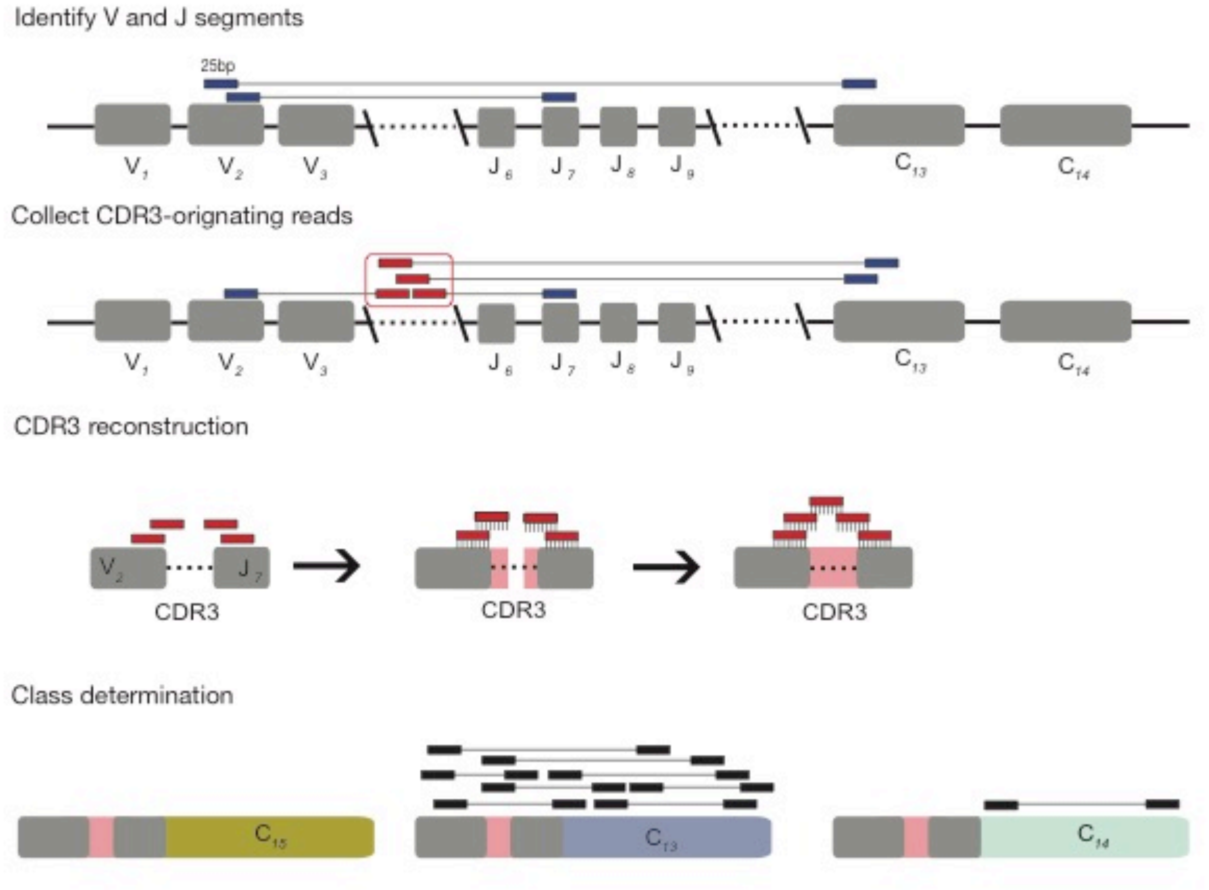

Figure S1: The BRAPeS algorithm. First, the V and J segments are selected by searching for paired reads with one read mapping to a V segment and its mate mapping to a J segment. Next, putative CDR3-originating reads are identified as the unmapped reads whose mates map to the V/J/C segments. Then, BRAPeS runs an iterative dynamic programming algorithm to align the CDR3-originating reads to the V and J segments and extend them until they overlap. Finally, BRAPeS runs RSEM on all possible full BCR transcripts (the reconstructed V-J segments combined with all possible constant segments) to determine the BCR isotype.

Figure S2

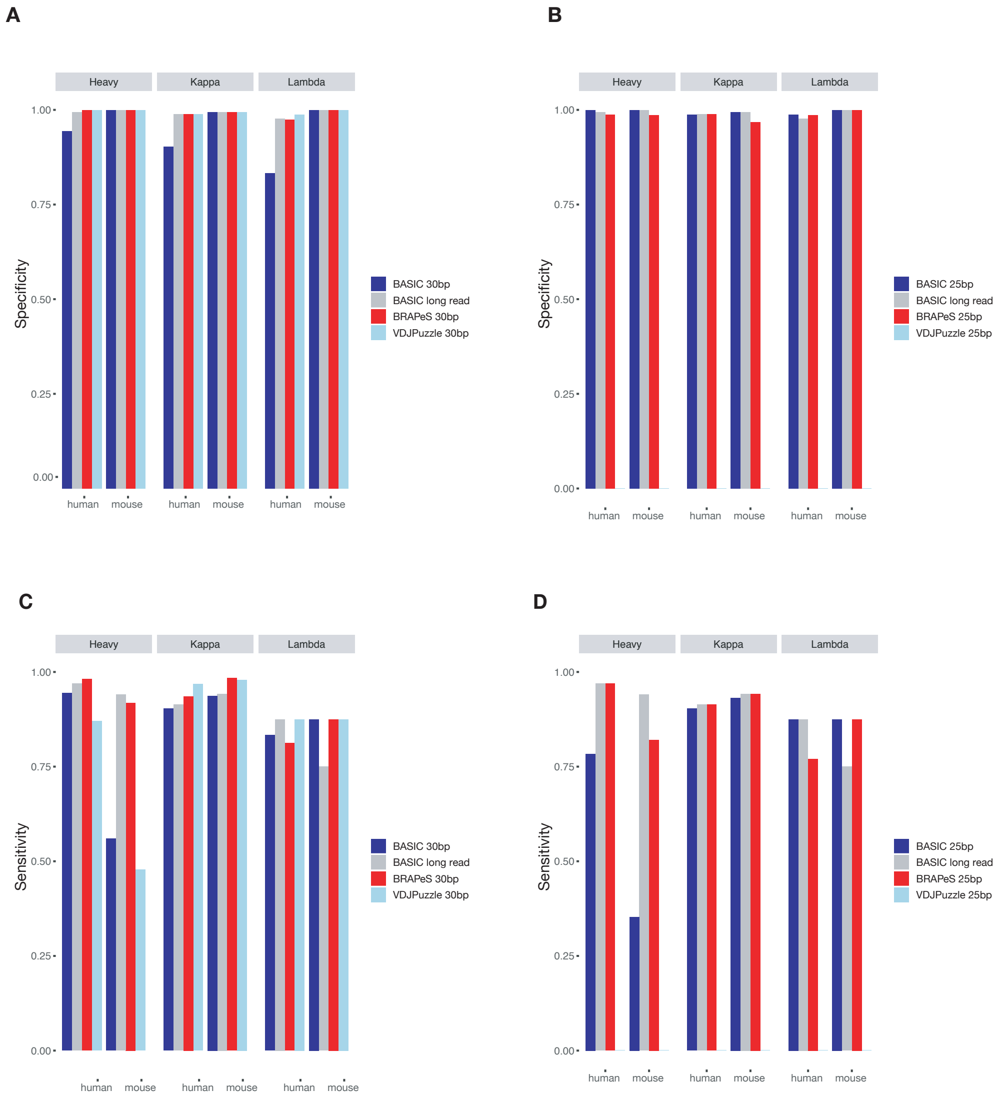

Figure S2: Sensitivity and specificity of BRAPeS compared to VDJPuzzle reconstructions on long-read data. **A**) Specificity of BRAPeS for 30bp. The fraction of cells with a CDR3 sequence identical to the CDR3 reconstructed by VDJPuzzle on the long-read data, using the following methods: BASIC when applied to the long-read data (grey), BRAPeS (red), BASIC (dark blue) and VDJPuzzle (light blue) applied to a version of the data trimmed to 30bp. The fraction is calculated only for cells that had a productive chain in both the long-read VDJPuzzle results and the other method. **B**) Specificity of BRAPeS for 25bp. Same as A, except the short-read version of the data was trimmed to include only the outer 25bp, instead of 30bp. **C**) Sensitivity of BRAPeS for 30bp. Same as A, except the fraction is calculated out of all the cells that had a productive chain when running VDJPuzzle on the long-read data. **D**) Sensitivity of BRAPeS for 25bp. Same as B, except the fraction is calculated out of all the cells that had a productive chain when running VDJPuzzle on the long-read data.

Table S1: Detailed description of the number of productive reconstructions for the original long reads and 30bp sequencing

	Human (n = 174)			**Mouse (n = 200)**		
	Heavy	**Light (kappa or Lambda)**	**Both kappa and Lambda**	**Heavy**	**Light (kappa or Lambda)**	**Both kappa and Lambda**
BASIC - long read	170 (97.7%)	174 (100%)	1 (0.6%)	182 (91%)	190 (95%)	0 (0%)
VDJPuzzle - long read	162 (93.1%)	172 (98.9%)	17 (9.77%)	184 (92%)	196 (98%)	3 (1.5%)
BRAPeS - 30bp	171 (98.3%)	165 (94.8%)	8 (4.6%)	178 (89%)	199 (99.5%)	1 (0.5%)
BASIC - 30bp	164 (94.3%)	171 (98.3%)	0 (0%)	109 (54.5%)	191 (95.5%)	0 (0%)
VDJPuzzle - 30bp	149 (85.6%)	172 (98.9%)	6 (3.45%)	93 (46.5%)	198 (99%)	1 (0.5%)

Table S2: Detailed description of the number of productive reconstructions for the original long reads and 25bp sequencing

	Human (n = 174)			**Mouse (n = 200)**		
	Heavy	**Light (kappa or Lambda)**	**Both kappa and Lambda**	**Heavy**	**Light (kappa or Lambda)**	**Both kappa and Lambda**
BASIC - long read	170 (97.7%)	174 (100%)	1 (0.6%)	182 (91%)	190 (95%)	0 (0%)
VDJPuzzle - long read	162 (93.1%)	172 (98.9%)	17 (9.77%)	184 (92%)	196 (98%)	3 (1.5%)
BRAPeS - 25bp	171 (98.3%)	161 (92.5%)	3 (1.72%)	160 (80%)	196 (98%)	1 (0.5%)
BASIC - 25bp	137 (78.7%)	173 (99.4%)	0 (0%)	67 (33.5%)	190 (95%)	0 (0%)
VDJPuzzle -25bp	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)