Title: Response to Abdo, C. et al 2020.

Clara Bueno[1,2], Paola Ballerini[3], Ignacio Varela[4], Pablo Menendez[1,2,5,] and Rachael Bashford-Rogers[6]

[1]Josep Carreras Leukemia Research Institute and Department of Biomedicine, School of Medicine. University of Barcelona. Barcelona. Spain.
[2]Centro de Investigación Biomédica en Red de Cáncer (CIBER-ONC), ISCIII, Barcelona, Spain.
[3]Service d'Hématologie Biologique, Hospital Armand Trousseau, Paris, France.
[4]Instituto de Biomedicina y Biotecnología de Cantabria, Universidad de Cantabria-CSIC, Santander, Spain.
[5]Instituciò Catalana de Recerca i Estudis Avançats (ICREA). Barcelona. Spain.
[6]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK.

We thank Abdo and colleagues for their interest in Bueno *et al*.[1] and for their insights into the cell-of-origin of TCF3-ZNF384 and PTPN11 mutations in monozygotic twins with concordant BCP-ALL. We are delighted that the conclusions reached by their independent analyses are suggestive of the same underlying conclusions as Bueno et al 2019 that the inferred cell-of-origin (COO) is around the onset of bone marrow IGH DJ rearrangement, likely in a pre-DJ progenitor. Abdo, C. et al rise some important points about the study to which we can comment on and extend (Figure 1a).

Firstly, the complex rearrangement leading to the *TCF3-ZNF384* fusion resulted from a 3-way t(12 ;14) and t(14 ;19) translocation, and we confirmed that the chromosome 14 breakpoint involved the end of the IGH locus, spanning from position 107,283,683 onwards (whole locus encompasses from position 106,03,2614 to 107,288,051). Chromosomal breakpoints are identical in both twins as depicted in the graphical representation of the breakpoint shown in Figure 1b. The involvement of IGH locus in rearrangements is well recognized in BCP-ALL[2].

Secondly, quantification of IGH VDJ rearrangements in Abdo, C. et al demonstrated 2-3 significantly clonal IGH VDJ rearrangements (comprising >10% of BCRs) within each twin, some of which share a common DJ incomplete rearrangement. This observation was not observed in the BCR data from Bueno et al for two reasons: firstly, Bueno et al performed BCR sequencing analysis on RNA, whereas Abdo and colleagues performed their analyses on DNA, where the analyses can provide different results dependent on RNA expression of leukemic cells and the presence of non-expressed allele defective-rearrangements present in the genomic DNA[2]. A second, and more significant, difference between the two approaches was that out-of-frame non-functional BCRs were filtered out from the Bueno et al 2019 dataset to reduce the potential of retaining reads with sequencing or PCR errors, which is an accepted practice for analysing functional BCRs[3]. As pointed out by Abdo and colleagues, the common rearrangements from these twins were out-of-frame.

Therefore, a re-analysis of the data from Bueno et al 2019 with the modified filtering criteria including all BCRs irrespective of the presence of "stop" codons was performed. In accordance with Abdo and colleagues, 57.24% and 54% of the BCRs from this new filtering strategy contained "stop" codons for twin A and B respectively (Table 1). From this, we could identify expanded 3 clones from twin A matching the descriptions from Abdo, C. *et al.*, and 2 clones from twin B (Table 2). These IGH VDJ rearrangements are all out-of-frame, containing the "stop" codon within the conserved imputed IGHD region of the rearrangements.

In agreement with Abdo, C. *et al*., we identify a major DJ rearrangement common to the most frequent clones in both twins comprised of IGHD6-13*01 and IGHJ4*02, but different IGHV genes (Table 2 and Figure S1). This rearrangement contains no N nucleotide additions. Indeed, the presence of a common

DJ and multiple IGHV genes is suggestive of secondary rearrangements, which are very common in B-ALL[4,5]. Secondary rearrangements are thought to be a result of aberrant activity of recombination-activating genes promoting genomic rearrangements critical to B-ALL pathogenesis[6,7]. IgHD-J combinations (including junctional regions), known as 'stem sequences', are stable in instances of secondary rearrangements[8]. The false-positive detection rates for detecting these secondary rearrangements is estimated to be 9.245×10^{-6}([2]).

The largest clone in twin A contains the IGHD7-27*01 and IGHJ4*02 rearrangement (14.95% of BCRs and also including a "stop" codon) and contains N nucleotide additions, and also observed in Abdo, C. et al 2020 at a high frequency. Here, we also detected multiple clones related by secondary rearrangement sharing the stem DJ sequence in twin A (Figure S2). It appears that these two expanded clones are, in fact, from distinct rearrangements rather than being derived from a common pre-DJ precursor. Indeed, IGHJ4 is the most commonly rearranged IGHJ gene (comprising ~50% BCRs from healthy individuals[9]) and so the common IGHJ rearrangement may have occurred independently by chance.

Abdo et al suggests that the polyclonal BCR repertoires described by Bueno et al. are likely to correspond to non-leukemic IGH VDJ rearrangements. However, BCR repertoire was performed on highly FACS-purified (purity>99%) CD34+CD19+CD10- pro-B ALL cells from PB. Such pro-B normal counterparts do not exist in PB, ruling out the possibility that polyclonal BCR repertories are non-leukemic. The aforementioned presence of frequent and ongoing secondary rearrangements in cells containing both these DJ rearrangements provides further evidence that these are both derived from *bona fide* ALL cells, and that there are multiple DJ rearrangements associated with the leukemic cell population. Given that there is only a single IGH DJ stem region shared between both twins, this suggests that the COO must be prior to the full IGH VDJ rearrangement and after the IGH DJ rearrangement. In support of this, the expanded ALL clones observed only in one twin containing distinct IGH DJ stem regions may arise from a primary IG DJ rearrangement that can be replaced by secondary recombination of an upstream D to downstream J[10]. This means that a polyclonal repertoire of distinct IGH VDJ rearrangements may be generated from a ALL population containing a partial IG DJ rearrangement (Figure 1a). Both our data and that of Abdo, C. et al jointly supports the conclusion that the predicted COO shared between the two twins is at the IGH DJ recombination stage, rather than the containing an unrecombined (germline) or fully recombined IGH locus, as the cell of origin of TCF3-ZNF384 and PTPN11 mutations.

AUTHOR CONTRIBUTION: R.B-R analyzed data. C.B, P.B, I.V., P.M. and R.B-R contributed to artwork and wrote the manuscript.

CONFLICT OF INTEREST DISCLOSURE: *R.J.M.B.-R. is a co-founder and consultant for Alchemab Therapeutics Ltd and has been a consultant for Imperial College London and VHSquared.*

Correspondence, please contact:
Rachael Bashford-Rogers: rbr1@well.ox.ac.uk
Pablo Menendez: pmenendez@carrerasresearch.org

References:

1. C. Bueno *et al.*, Natural history and cell-of-origin of TCF3-ZNF384 and PTPN11 mutations in monozygotic twins with concordant BCP-ALL. *Blood* 2019; 134:900-905.

2. R. J. Bashford-Rogers *et al.*, Eye on the B-ALL: B-cell receptor repertoires reveal persistence of numerous B-lymphoblastic leukemia subclones from diagnosis to relapse. *Leukemia* 2016; 30:2312-2321.

3. G. Yaari, S. H. Kleinstein, Practical guidelines for B-cell receptor repertoire sequencing analysis. *Genome Med* 2016; 7:121.

4. A. Beishuizen *et al.*, Multiple rearranged immunoglobulin genes in childhood acute lymphoblastic leukemia of precursor B-cell origin. *Leukemia* 1991*;* 5:657-667.

5. A. Dupuis *et al.*, Biclonal and biallelic deletions occur in 20% of B-ALL cases with IKZF1 mutations. *Leukemia* 2013; 27:503-507.

6. E. Papaemmanuil *et al.*, RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nat Genet* 2014; 46:116-125.

7. L. Holmfeldt *et al.*, The genomic landscape of hypodiploid acute lymphoblastic leukemia. *Nat Genet* 2013, 45:242-252.

8. V. H. van der Velden *et al.*, Analysis of minimal residual disease by Ig/TCR gene rearrangements: guidelines for interpretation of real-time quantitative PCR data. *Leukemia* 2007; 21:604-611.

9. V. N. Petrova *et al.*, Combined Influence of B-Cell Receptor Rearrangement and Somatic Hypermutation on B-Cell Class-Switch Fate in Health and in Chronic Lymphocytic Leukemia. *Frontiers in Immunology* 2018; 9:1784.

10. E. Edry, D. Melamed, Receptor editing in positive and negative selection of B lymphopoiesis. *J Immunol 2004;* 173:4265-4271.

Figure legends:

Figure 1. a) Schematic of the possible IGH locus rearrangements present in the cell-of-origin (COO): Scenario 1 has the COO at the stage prior to IGH VDJ rearrangement. Some of these cancer cells may undergo full IGH VDJ rearrangements and subsequent secondary rearrangements, however these would not be shared between the twins. Scenario 2 has the COO at the stage with a partial IGH DJ rearrangement that is shared between the twins. Some of these cancer cells can undergo further IG V-DJ and/or secondary rearrangements (where a primary IG DJ rearrangement can be replaced by an upstream D to downstream J) that may be present only in one twin. Scenario 3 has the COO with a full IGH VDJ rearrangement. Some of these cancer cells may undergo further secondary IG V rearrangements. Only scenario 2 is supported by the data. b) Schematic representation depicting the genomic breakpoints determined by WG-seq at the base pair level for the t(14;19) and t(12;14).

Table 1. Total numbers of BCRs after different filtering strategies.

Sample	Excluding BCRs with STOP codons	Including BCRs with STOP codons	% BCRs with STOP codons
Twin A	9562	22362	57.24
Twin B	22790	49541	54.00

Table 2. Characterisation of clones greater than 2% in the BCR repertoires including stop codons.

#Id	% of Repertoire	V gene	D gene	J gene	SHM in V gene	STOP codon present?	CDR3 sequence (nucleotides)	Predicted CDR3 sequence based on in-frame IGHV gene (amino acids)
Twin A	14.95	IGHV3-33*01/06	IGHD7-27*01	IGHJ4*02	0	Yes	GCGATCCTAACTGGGGACCTAGTACTACTTTGACTAC	AILTGDLVLL*L
Twin A	14.41	IGHV3-9*01	IGHD6-13*01	IGHJ4*02	0	Yes	GCAAAAGATATAAGGGGTAGCAGCAGCTTTTGACTAC	AKDIRGSSSF*L
Twin A	2.45	IGHV3-21*01/02	IGHD6-13*01	IGHJ4*02	0	Yes	GCGAGAGATCCGCCGTAGCAGCAGCTTTTGACTAC	ARDPP*QQLLTT
Twin B	28.36	IGHV3-74*01/02	IGHD6-13*01	IGHJ4*02	0	Yes	GCAACCCGTGGGAGCAGCAGCTTTTGACTAC	ATRGSSSF*L
Twin B	3.12	IGHV3-30*04/03	IGHD6-13*01	IGHJ4*02	0	Yes	GCGAGAGATCCCCCCTCCGTATAGCAGCAGCTTTTGACTAC	ARDPPSV*QQLLTT