



CORRESPONDENCE

Comment on: “UGT2B17 modifies drug response in chronic lymphocytic leukaemia”

British Journal of Cancer (2020) 123:1345–1346;
https://doi.org/10.1038/s41416-020-1005-5

We read with great interest the article by Allain et al.¹ and were intrigued by the UGT2B17-mediated direct drug inactivation or indirect anti-leukaemic response modification of various therapeutic agents in chronic lymphocytic leukaemia (CLL). In human, UGT2B17 gene expression is detected in hepatic tissues as well as in steroid target tissues.² Nonetheless, shown elegantly in the study by Allain et al.¹ is that in lymphoid cell models and in CLL patients occurs a lineage-inappropriate UGT2B17 overexpression, predominantly mediated from the alternative UGT2B17_n2 transcript that comprises the additional exon 1c, extending the 5'-untranslated region. Lower UGT2B17 expression from the alternative UGT2B17_n3 and n4 transcripts, comprising the alternative 1b exon, was also reported. Of note, all novel transcripts identified encode a functional UGT2B17 protein.¹ Since UGT2B17 represents a main conjugating enzyme for fludarabine glucuronidation and consequently UGT2B17 misexpression could affect primary

response to first-line treatment with fludarabine in CLL patients, we fully agree with Allain et al.¹ that the molecular machinery underlying high UGT2B17 expression in lymphoid cells warrants further clarification. Herein, we present evidence that both UGT2B17 exons 1b and 1c were evolutionarily derived from endogenous-retrovirus (ERV) sequences and that UGT2B17 overexpression in B cells of CLL patients is driven by the in cis aberrant activation of long terminal repeats (LTRs) of the ERV1 family.

ERVs represent heritable provirus insertions into the host genome DNA, remnants of exogenous infectious retroviruses.³ The typical genomic structure of a retrovirus consists of a gag, pro, pol, and env genes flanked by two LTRs that naturally comprise core transcription regulatory elements and transcription factor (TF) binding sites.³ The majority of ERVs in human genome has undergone recombination events between the 5'- and 3'-LTRs, resulting in solitary LTRs.⁴ Interestingly, many solitary LTRs have preserved their ancestral promoter function and, when positioned in the sense orientation of an adjacent host gene, could drive ectopic transcription initiation.⁴ Accordingly, because epigenetic silencing represents the predominant path to ERV transcription

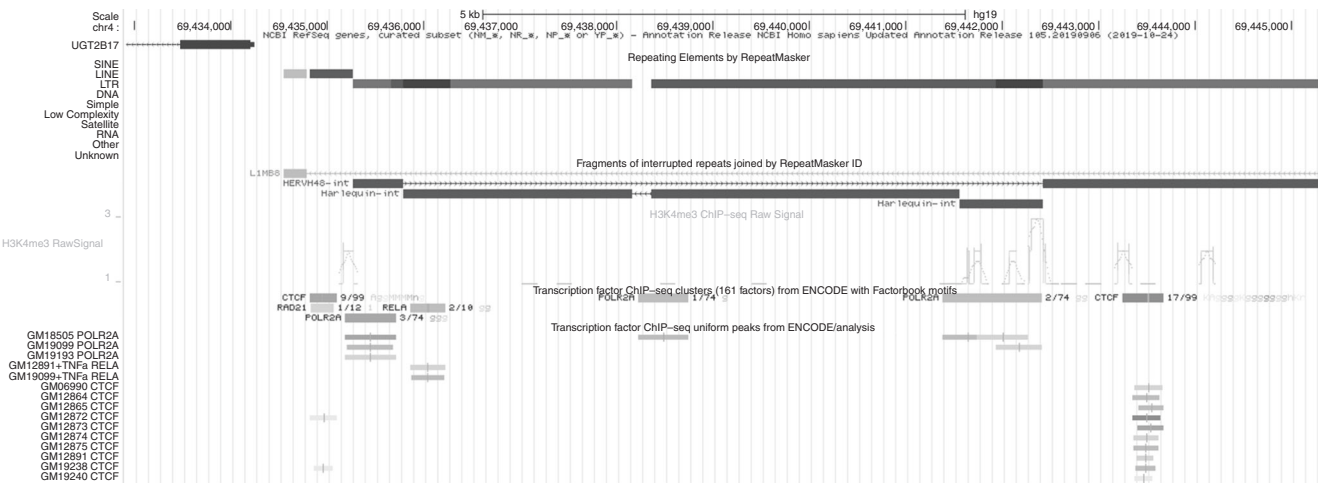


Fig. 1 ERV onco-exaptation drives UGT2B17 ectopic transcription. A screenshot from the UCSC Genome Browser Database (<https://genome.ucsc.edu/>); shown in the top left corner is the canonical UGT2B17 exon 1. Upper part: RepeatMasker software analysis⁶ reveals that the genomic segment upstream of UGT2B17 locus comprises a large cluster of transposable elements. Histone H3 lysine 4 trimethylation (H3K4me3) ChIP-seq data, that is, data for a histone modification highly enriched at active promoters,⁷ were added for the purposes of the study indicating the location of UGT2B17_n2 (ChIP-seq peak cluster) and UGT2B17_n3/UGT2B17_n4 transcriptional start sites, at -7466 and -1290, respectively (UGT2B17 translational start codon = +1). Proximal promoters of the transcripts as well as the corresponding exons 1c and 1b are comprised within Harlequin-int LTR and HERV48-int LTR retroviral sequences, respectively. Lower part: Depicted are TF ChIP-seq data in 91 cell types, generated by the ENCODE project.⁷ Grey boxes delimit the genomic sites occupied by TFs in the ChIP-seq experiments; darkness of the boxes is proportional to the level of enrichment at the site. An orange vertical line within a box reveals the point source of the peak, a green vertical bar denotes the presence of a DNA binding motif from ENCODE Factorbook repository. The symbol(s) and quota appearing next to each box define the cell line(s) where the given TF was found to interact with the genomic sequence. The POLR2A bindings sites, referred to in the text, co-localise with the H3K4me3 ChIP-seq peak clusters and are functional specifically in B-lineage cells; “g” represents a common ENCODE abbreviation for several lymphoid cell lines.⁷ ENCODE computational pipeline identifies also high-scoring binding motifs for RELA and CTCF in the retroelement-shaped genomic segment.

Received: 3 June 2020 Revised: 1 July 2020 Accepted: 9 July 2020
Published online: 24 July 2020

inactivation, alterations of the epigenetic landscape occurring during cellular transformation events could likely lead to ERV derepression with a significant impact on host gene transcriptional networks, a process designated as onco-exaptation.⁵

Interestingly, molecular evolutionary analysis⁶ of the genomic segment comprising *UGT2B17* exon 1c and its non-canonical proximal promoter sequence reveals that these genetic elements are comprised within a Harlequin-int LTR/ERV1 sequence, of sense orientation to *UGT2B17* gene (Fig. 1). Accordingly, *UGT2B17* exon 1b and the corresponding proximal promoter were also evolutionary derived from an HERVH48-int LTR/ERV1 sequence. However, because the HERVH48-int LTR is of anti-sense orientation (Fig. 1), a less potent promoter activity compared to Harlequin-int LTR would be expected,⁴ as is the case in the study by Allain et al.¹

TF chromatin immunoprecipitation and sequencing (ChIP-seq) data in 91 cell types from the ENCODE Project⁷ reveal that the Harlequin-int LTR sequence comprises a binding site for RNA polymerase II subunit A (POLR2A), which, among 74 cell types examined, is functional specifically in the lymphoid cell lines (Fig. 1). Likewise, the HERVH48-int LTR comprises a POLR2A site that is also specific for B cells. Interestingly, short upstream of the HERVH48-int POLR2A site located is a B-lineage functional binding site for RELA proto-oncogene (Fig. 1). In this sense, the identification in this retroelement-derived genomic segment of DNA binding sites for the CCCTC-binding factor (CTCF) is also of importance (Fig. 1). CTCF has the potential to act as transcriptional insulator disabling the interaction between enhancers and non-canonical promoters;⁸ therefore, in a tantalising scenario, a B-lineage-specific aberration of CTCF function could allow the RELA binding site to serve as an unrestricted enhancer of *UGT2B17* inordinate transcription in B-cancerous cells.⁹ These findings are likely in accordance with the results from the study by Allain et al.,¹ where a co-expression signature of *UGT2B17* with several nuclear factor- κ B (NF- κ B)-regulated genes was documented, subsequently pointing to NF- κ B as a key regulatory “hub point” that targets the non-canonical promoter and drives *UGT2B17* misexpression in CLL cells.

It has been shown that the positive selection of novel hypomethylation motifs, which could subsequently allow the epigenetic derepression of ERV transcription, entails co-evolution of genetic sub-clonal complexity in CLL.¹⁰ Subsequently, it would be tempting to deduce that the efficacy of *UGT2B17* expression as a prognostic marker of high-risk CLL² is mechanistically linked to the retroviral origin of its ectopic promoters. Most importantly, since the LTR-driven misexpression of *UGT2B17* may impact the effectiveness of fludarabine-based interventions, quantitative PCR-based assays of LTR-driven *UGT2B17* ectopic transcription in cancerous B cells would help tailor fludarabine dosage to better stratified CLL patients.

ACKNOWLEDGEMENTS

S.I.P. gratefully acknowledges Professor George Kassiotis at the Francis Crick Institute for the enlightening discussion on endogenous retroviruses during the EMBO|EMBL Symposium “The Mobile Genome” and the fruitful mentoring provided ever since.

AUTHOR CONTRIBUTIONS

Conception and design of the work: S.I.P. Data collection: S.I.P. Data analysis and interpretation: S.I.P. Drafting of the paper: S.I.P., C.J. Supervision of the work: C.J. Critical revision of the article: S.I.P., C.J.

ADDITIONAL INFORMATION

Ethics approval and consent to participate Not applicable.

Data availability The ChIP-seq data reported in this study represent publicly available material, generated from the Encyclopaedia of DNA Elements (ENCODE) Consortium, and are archived at the UCSC Genome Browser Database (<https://genome.ucsc.edu/>).

Competing interests The authors declare no competing interests.

Funding information The authors have not received any funding.

Note This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution 4.0 International (CC BY 4.0).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Spyros I. Papamichos¹ and Christof Jungbauer¹
¹Blood Service for Vienna, Lower Austria and Burgenland, Austrian
 Red Cross, Vienna, Austria
 Correspondence: Spyros I. Papamichos
 (spyros.papamichos@roteskreuz.at)

REFERENCES

- Allain, E. P., Rouleau, M., Vanura, K., Tremblay, S., Vaillancourt, J., Bat V. et al. *UGT2B17* modifies drug response in chronic lymphocytic leukaemia. *Br. J. Cancer* <https://doi.org/10.1038/s41416-020-0887-6> (2020).
- Gruber, M., Bellemare, J., Hoermann, G., Gleiss, A., Porpacz, E., Bilban, M. et al. Overexpression of uridine diphospho glucuronosyltransferase 2B17 in high-risk chronic lymphocytic leukemia. *Blood* **121**, 1175–1183 (2013).
- Kassiotis, G. Endogenous retroviruses and the development of cancer. *J. Immunol.* **192**, 1343–1349 (2014).
- van de Lagemaat, L. N., Landry, J. R., Mager, D. L. & Medstrand, P. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet.* **19**, 530–536 (2003).
- Babaian, A. & Mager, D. L. Endogenous retroviral promoter exaptation in human cancer. *Mob. DNA* **7**, 24 (2016).
- Smit, A., Hubley, R. & Green P. RepeatMasker Open-3.0. <http://www.repeatmasker.org> (1996–2010).
- ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
- Phillips, J. E. & Corces, V. G. CTCF: master weaver of the genome. *Cell* **137**, 1194–1211 (2009).
- Furman, R. R., Asgry, Z., Mascarenhas, J. O., Liou, H. C. & Schattner, E. J. Modulation of NF-kappa B activity and apoptosis in chronic lymphocytic leukemia B cells. *J. Immunol.* **164**, 2200–2206 (2000).
- Oakes, C. C., Claus, R., Gu, L., Assenov, Y., Hullein, J., Zucknick, M. et al. Evolution of DNA methylation is linked to genetic aberrations in chronic lymphocytic leukemia. *Cancer Discov.* **4**, 348–361 (2014).