LETTER

LYMPHOMA

Insights into the mechanisms underlying aberrant *SOX11* oncogene expression in mantle cell lymphoma

Roser Vilarrasa-Blasi[1,2]✉, Núria Verdaguer-Dot[1], Laura Belver[3,4], Paula Soler-Vila[5], Renée Beekman[1], Vicente Chapaprieta[1], Marta Kulis[1], Ana C. Queirós[1], Maribel Parra[4], María José Calasanz[6,7], Xabier Agirre[6,7], Felipe Prosper[6,7,8], Sílvia Beà[1,2,7], Dolors Colomer[1,2,7], Marc A. Marti-Renom[5,9], Adolfo Ferrando[3], Elías Campo[1,2,7] and José Ignacio Martin-Subero[1,2,7,9]✉

Leukemia; https://doi.org/10.1038/s41375-021-01389-w

TO THE EDITOR:

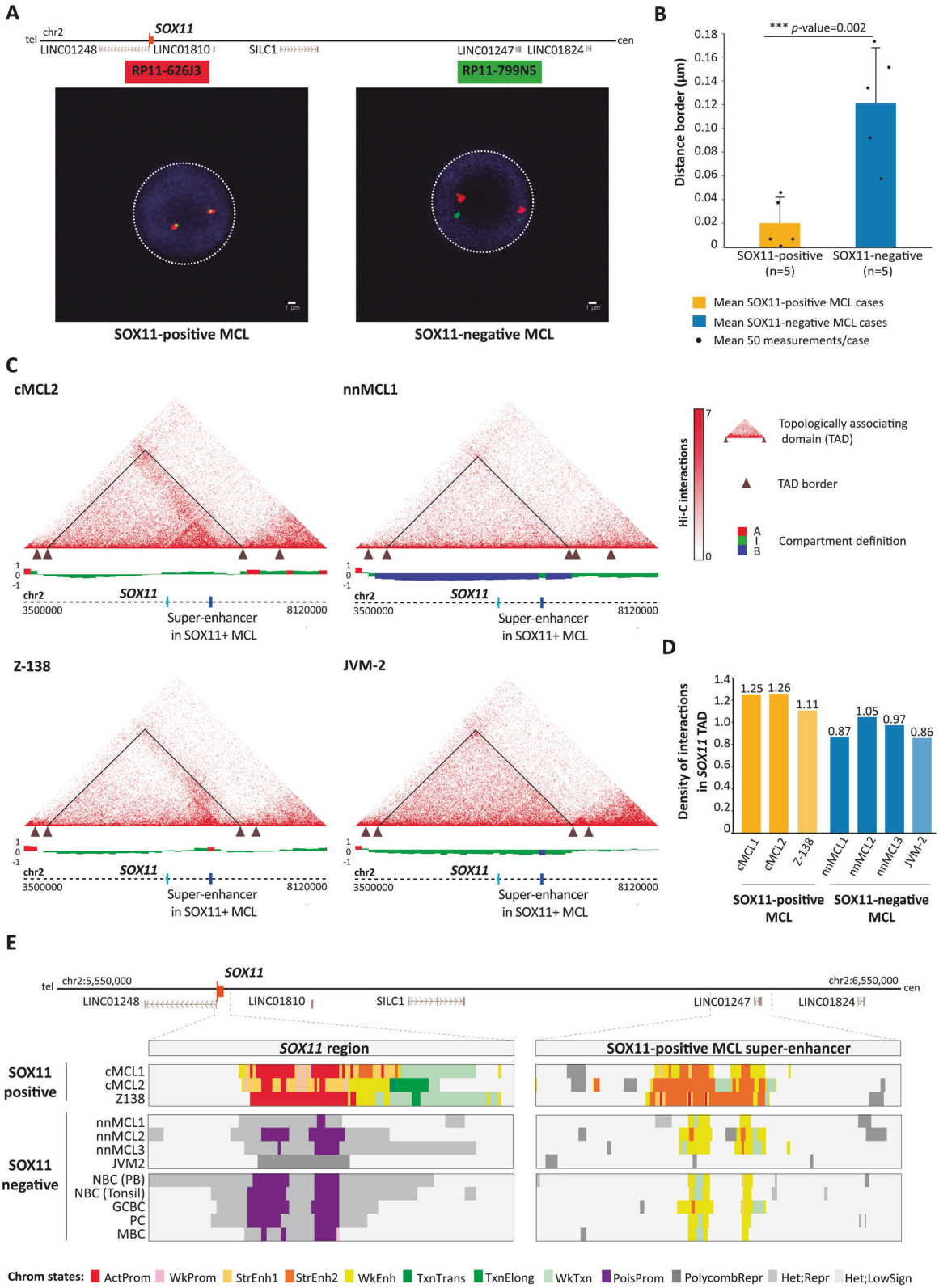
Mantle cell lymphoma is characterized by the primary CCND1 deregulation caused by the t(11;14)(q13;q32) translocation. In addition, SOX11 had been described as one of the most relevant oncogenes contributing to the pathogenesis of mantle cell lymphoma (MCL) [1]. In particular, high SOX11 expression is restricted to conventional MCL (cMCL) cases, and represents a diagnostic biomarker to differentiate this MCL subtype from the less aggressive leukemic non-nodal MCL (nnMCL) and other lymphoproliferative disorders [2]. Although the consequences of SOX11 overexpression are well described [1], its causes remain widely unknown, as *SOX11* is not targeted by any genetic alteration [3]. From the epigenetic perspective, the *SOX11* promoter is characterized by the expected presence of activating histone modifications [4]. In 2016, we identified a distant super-enhancer 624–653 kilobases (Kb) downstream *SOX11* which becomes activated and creates a three-dimensional (3D) loop structure with the *SOX11* locus [5]. Here, we have performed a detailed epigenetic dissection of this distant super-enhancer to shed additional light into the mechanisms underlying SOX11 expression in cMCL.

Although the proximity of *SOX11* and its super-enhancer in the 3D space has been previously detected by chromosome conformation capture experiments [5], we further explored it at the single-allele level in individual cells by 3D fluorescence in situ hybridization (FISH). We applied two differentially labeled DNA probes spanning the *SOX11* oncogene (in red) and the super-enhancer (in green) (Fig. 1A). In SOX11-positive MCL cases signals were fused in both alleles, whereas in SOX11-negative MCLs, the distance between the labeled probes was significantly larger (Fig. 1B, Supplementary Fig. 1). This finding indicates that the physical proximity between *SOX11* and its super-enhancer takes place in a biallelic fashion in SOX11 expressing MCL cells. Next, we wondered if the looping structure in SOX11-positive MCL cases was associated with changes in topologically associating domains (TADs). These structures are considered as regulatory units for transcriptional activity, and alterations in TAD boundaries have been described as a mechanism of aberrant oncogene activation in cancer [6]. Hence, we defined the TAD structures in SOX11-positive and negative MCLs. We generated in situ Hi–C data from two MCL cell lines as disease models, Z-138 (SOX11-positive) and JVM-2 (SOX11-negative) (Supplementary Table 1), and we mined recently published in situ Hi–C data from 2 SOX11-positive and 3 SOX11-negative MCLs [7]. We observed that the TAD structure at 20 Kb resolution in the *SOX11* locus and the distant super-enhancer region was similar in *SOX11* expressing and non-expressing MCLs (Fig. 1C and Supplementary Fig. 2). Despite TAD border conservation, we detected a significant overall increase of 3D interactions within this TAD (relative to the rest of chromosome 2) in SOX11-positive MCL (mean density of interactions = 1.2: SD = 0.084) as compared to SOX11-negative MCLs (mean density of interactions = 0.94: SD = 0.091) (Fig. 1D). This increase seemed to be mostly caused by a sharp increase in interactions in the intraTAD region that contains *SOX11* and its super-enhancer (Supplementary Fig. 3). Overall, these results rule out the presence of a large-scale structural disruption and favors a chromatin loop reconfiguration between *SOX11* and its super-enhancer as the potential cause of *SOX11* deregulation.

We next performed a multi-layer epigenomic profiling of the *SOX11* super-enhancer to elucidate its internal structure and identify specific regulatory elements. We mined chromatin states using ChIP-seq of six different histone marks, chromatin accessibility maps by ATAC-seq and DNA methylation by whole-genome bisulfite sequencing in MCL primary cases,

[1]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain. [2]Hospital Clínic de Barcelona and Departament de Fonaments Clínics, Facultat de Medicina, Universitat de Barcelona, Barcelona, Spain. [3]Institute for Cancer Genetics, Columbia University Medical Center, New York, NY, USA. [4]Josep Carreras Leukaemia Research Institute, IJC Building, Campus ICO-Germans Trias i Pujol, Barcelona, Spain. [5]CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain. [6]Área de Oncología, Centro de Investigación Médica Aplicada (CIMA), Instituto de Investigación Sanitaria de Navarra (IdiSNA), Universidad de Navarra, Pamplona, Spain. [7]Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain. [8]Servicio de Hematología, Clínica Universidad de Navarra, Pamplona, Spain. [9]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain. ✉email: rvilarrasa@clinic.cat; imartins@clinic.cat

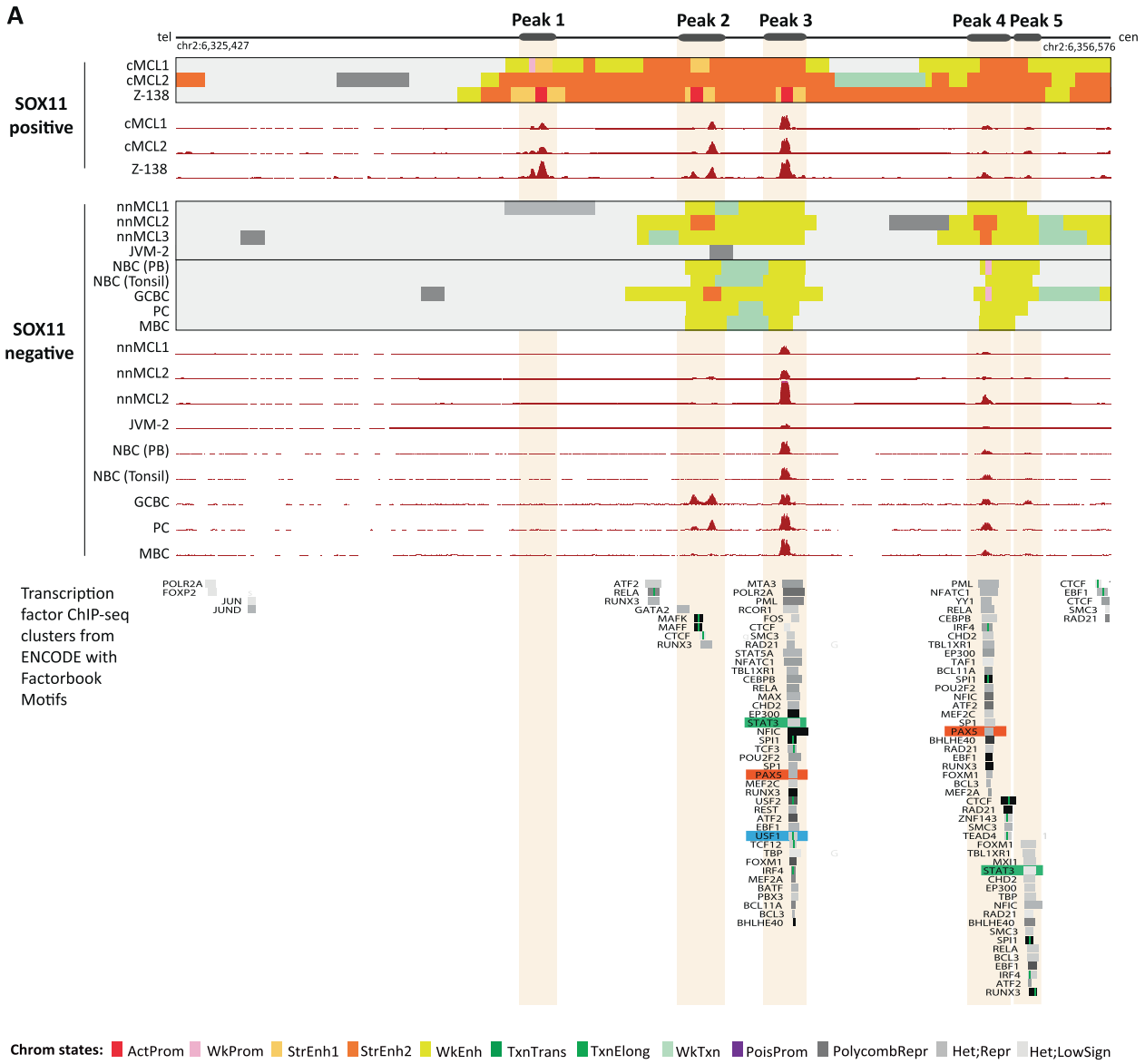
Received: 30 April 2021 Revised: 11 August 2021 Accepted: 13 August 2021
Published online: 28 August 2021

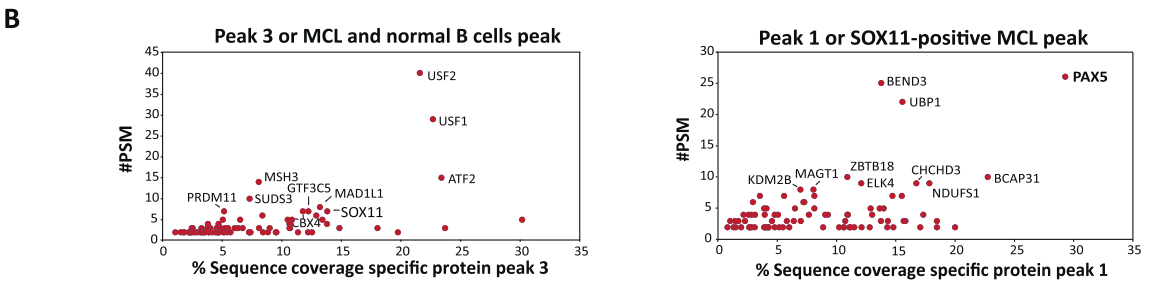


MCL cell lines and normal B cells, as well as histone mark profiling other B-cell neoplasms and cell lines from the ENCODE Consortium [8, 9]. We initially analyzed chromatin states and corroborated that in SOX11-positive MCLs, the super-enhancer was active and correlated with active chromatin at the distant *SOX11* locus. In contrast, in SOX11-negative

Fig. 1 Chromosome conformation and chromatin activity at the SOX11 locus in MCL cases. A Scheme of the 3D FISH experiment (*top*). Two BAC clones were used, one for the *SOX11* locus (red) and another for the SOX11 super-enhancer region (green). Two representative 3D FISH images (*bottom*) from SOX11-positive (*left*) and SOX11-negative (*right*) MCL cases are shown. Nuclei, which are highlighted with a dashed white line, were stained with DAPI. **B** Mean distances between *SOX11* locus and super-enhancer region were analyzed for 50 different signal pairs per each case (mean ± SD, *t* test for independent samples) (5 SOX11-positive and 5 SOX11-negative MCLs) **C** *Top*: Hi–C contact matrices for one conventional MCL case (SOX11-positive, cMCL2, *left*) and one leukemic non-nodal MCL cases (SOX11-negative, nnMCL1, *right*). *Bottom:* Hi–C contact matrices for the SOX11-positive MCL cell line Z-138 (*left*), and the SOX11-negative MCL cell line JVM-2 (*right*). Color code indicated the Hi–C interactions. The topologically associating domain (TAD) contacting *SOX11* locus and the SOX11-positive MCL super-enhancer are marked with a line. Below, TAD borders and the location of the *SOX11* locus (light blue) and the super-enhancer in SOX11-positive MCL cases (dark blue) were indicated per sample. The coordinates of the represented region are chr2:3,500,000-8,120,000 (GRCh38). **D** Density of interactions calculated at the TAD containing *SOX11* locus and SOX11-positive MCL super-enhancer among SOX11-positive MCL (two primary cases and one cell line) and SOX11-negative MCL (three primary cases and one cell line) samples (*t* test for independent samples). **E** Chromatin states related to the *SOX11* locus and the SOX11 positive MCL super-enhancer in SOX11-positive (two cMCL cases and one cell line, Z-138) and SOX11-negative (three nnMCL cases, one cell line, JVM-2, and five normal B cells; a representative B-cell subpopulation was chosen from the three replicates) samples. cMCL conventional MCL, nnMCL leukemic non-nodal MCL, NBC naive B cell, GCBC germinal center B cell, PC plasma cell, MBC memory B cell, PB peripheral blood. ActProm Active Promoter, WkProm Weak Promoter, StrEnh1 Strong Enhancer 1, StrEnh2 Strong Enhancer 2, WkEnh Weak Enhancer, TxnTrans Transcription Transition, TxnElong Transcription Elongation, WkTxn Weak Transcription, PoisProm Poised promoter, PolycombRepr Polycomb repressed, Het;Repr Heterochromatin-Repressed, Het;LowSign Heterochromatin-Low Signal. The coordinates of the whole represented region are chr2:5,550,000-6,550,000, being the *SOX11* region located at chr2:5,683,948-5,710,104, and the super-enhancer at chr2:6,314,751-6,381,194. All coordinates are given in GRCh38.

MCLs and normal B cells the super-enhancer region was mostly heterochromatic with two isolated weak enhancer regions, and the *SOX11* gene promoter was kept in a poised state (Fig. 1E). We then analyzed the chromatin accessibility profile, which marks nucleosome-free regions where transcription factors (TF) bind. We defined five different accessible regions embedded within the super-enhancer in SOX11-positive MCL cases, named peaks 1–5 (Fig. 2A). Peak 1 was remarkably interesting as it was active only in the SOX11-positive MCLs but not in MCLs lacking SOX11, other B-cell neoplasias or normal B cells (Supplementary Figs. 4 and 5). Moreover, in non-lymphoid cell lines expressing *SOX11* such as H1, HSMM, and NHLF, the enhancer was inactive further supporting its cMCL specificity (Supplementary Fig. 4). Peaks 2 and 5 were present in SOX11-positive MCL cells and germinal center B cells (GCBC), being peak 2 also present in plasma cells (PC). Besides, peaks 3 and 4 were detected in all samples regardless of their SOX11 expression status (Fig. 2A). Furthermore, analyzing the DNA methylome we observed that the SOX11-positive MCLs with accessible chromatin were in general demethylated as compared to SOX11-negative MCLs (Supplementary Fig. 6). We subsequently mined the TF ChIP-seq data obtained in multiple cell lines (not including MCLs) by the ENCODE Consortium [9] to identify candidate TF that might be involved in activating the *SOX11* super-enhancer (Fig. 2A). Peaks 2–5 were enriched in a variety of TFs, in part related to B-cell differentiation, which is in line with the fact that these peaks are also accessible in normal B-cell subpopulations. Peak 1 lacked TF-binding sites in cell lines from the ENCODE, supporting the finding that this accessible site is highly specific for SOX11-positive MCLs. Thus, this peak may contain clues to understand the mechanisms triggering the SOX11 expression in aggressive MCLs. To obtain insights into which proteins may be binding to peak 1, we performed targeted mass spectrometry (MS) experiments. As bait, we used biotinylated sequences both from the de novo accessible peak 1 and the pan-B accessible peak 3, and we incubated them with protein extracts of the SOX11-positive MCL cell line Z-138. Among the proteins associated with peak 3, USF was detected, which is a regulator of transcription for many genes during cell differentiation [10]. Furthermore, the SOX11 protein was located in this peak 3 suggesting that SOX11 might regulate itself in MCL by binding to this specific chromatin accessible peak (Fig. 2B left and Supplementary Table 2), suggesting the presence of a SOX11 positive feedback loop [1]. In the case of peak 1, which is accessible in SOX11-positive MCLs but not in normal B cells, we found the B-cell TF PAX5 to be highly enriched (Fig. 2B right and Supplementary Table 3). To further analyze this finding, we screened for potential PAX5 binding motifs across all the accessible peaks of the super-enhancer and detected potential binding motifs for PAX5 both at peak 1 and peak 3, and at peak 4 as well (Supplementary Table 4). Remarkably, the *SOX11* locus itself also contains PAX5 binding motifs (Supplementary Table 4). Overall, these findings suggest that PAX5 may play a role in the activation of SOX11 in MCL. Interestingly, PAX5 may not only influence SOX11 gene expression, but SOX11 itself has been shown to contribute to PAX5 expression regulation [11]. The identification of PAX5-binding sites at the SOX11 gene and its super-enhancer, and the PAX5 binding to the MCL-specific peak 1, is indeed intriguing. PAX5 is a key TF for B-cell commitment and identity [12], and SOX11 is not expressed at any B-cell maturation state. How PAX5 contributes to activate the SOX11 regulatory elements in MCL remains puzzling. We speculate that PAX5 might aberrantly bind to peak 1 of the super-enhancer leading to DNA demethylation and chromatin activation [13], and eventually to *SOX11* upregulation. This process may occur in a pre-B stage where the initial oncogenic event takes place, i.e., the t(11;14)(q13;32) leading to cyclin D1 overexpression. However, it is unlikely that cyclin D1 overexpression itself induces SOX11 activation, as the *CCND1*-IGH juxtaposition also takes place in SOX11-negative MCLs, and cyclin D1 ChIP-seq data did not reveal binding to the SOX11-positive MCL super-enhancer region [14]. In spite of these evidences, a recent report proposed a link between CCND1 and SOX11 in MCL suggesting that cyclin D1 might sequester histone deacetylase 1 (HDAC1) or HDAC2 from the *SOX11* locus, resulting in SOX11 upregulation [15]. The same study identified that pSTAT3 could repress SOX11 transcription [15]. However, our detailed dissection of the SOX11 super-enhancer localized STAT3 at the normal B cell-related peaks 3 and 5 but our MS analyses of peak 1 did not reveal binding of any member of the STAT family (Fig. 2). This result suggests that the described function of STAT3 might be exerted at the chromatin accessible regions present at the SOX11 positive and negative MCL cases





as well as normal B cells, and therefore, may not be directly involved in the *de novo* overexpression of SOX11.

Overall, our study dissects the SOX11 super-enhancer and identifies a small accessible region exclusively active in SOX11-positive MCL. These findings represent an additional step in the path toward understanding the causes underlying SOX11 upregulation in MCL, although further studies are required to completely elucidate the factors and chain of events associated with this oncogenic activation.

Fig. 2 Multi-omic dissection of the SOX11 super-enhancer region. A For both SOX11-positive and negative samples (including normal and neoplastic samples) the following tracks are shown, from up to bottom: chromatin states and chromatin accessibility (based on ATAC-seq signals, *y*-axis signal from 0 to 100). The chromatin accessibility signal shown from each B-cell subpopulation is the median from three different replicates. Below, transcription factor ChIP-seq clusters from ENCODE with the factorbook motifs track. ChIP-seq peaks derived from the ENCODE TFBS data, which contains data from 161 transcription factors in 91 different cell types. The transcription factors highlighted are USF2 in blue, PAX5 in orange, and STAT3 in green. ActProm Active Promoter, WkProm Weak Promoter, StrEnh1 Strong Enhancer 1, StrEnh2 Strong Enhancer 2, WkEnh Weak Enhancer, TxnTrans Transcription Transition, TxnElong Transcription Elongation, WkTxn Weak Transcription, PoisProm Poised promoter, PolycombRepr Polycomb repressed, Het;Repr Heterochromatin-Repressed, Het;LowSign Heterochromatin-Low Signal. The coordinates of the represented region are chr2:6,325,427-6,356,576 (GRCh38). **B** Results of the targeted MS experiments showing the proteins potentially binding to specific chromatin accessible peaks. Peptide spectrum matches (PSM), which is the total number of identified peptide sequences for each protein is shown at the *x*-axis, and the percentage of the protein sequence identified in the analysis or coverage is shown at the *y*-axis. Specific proteins in peak 3 are shown on the left, whereas specific proteins in peak 1 are displayed on the right. The names of the top ten protein genes are ordered from higher to lower PSM and coverage is listed. A protein can be estimated to be more abundant if the sequence coverage is high and the peptide is abundant at the sample (high PSM).

REFERENCES

1. Beekman R, Amador V, Campo E. SOX11, a key oncogenic factor in mantle cell lymphoma. Curr Opin Hematol. 2018;25:299–306.
2. Navarro A, Clot G, Martínez-Trillos A, Pinyol M, Jares P, González-Farré B, et al. Improved classification of leukemic B-cell lymphoproliferative disorders using a transcriptional and genetic classifier. Haematologica. 2017;102:e360–e363.
3. Nadeu F, Martin-Garcia D, Clot G, Díaz-Navarro A, Duran-Ferrer M, Navarro A, et al. Genomic and epigenomic insights into the origin, pathogenesis, and clinical behavior of mantle cell lymphoma subtypes. Blood. 2020;136:1419–32.
4. Vegliante MC, Royo C, Palomero J, Salaverria I, Balint B, Martín-Guerrero I, et al. Epigenetic activation of SOX11 in lymphoid neoplasms by histone modifications. PLoS ONE. 2011;6:e21382.
5. Queirós AC, Beekman R, Vilarrasa-Blasi R, Duran-Ferrer M, Clot G, Merkel A, et al. Decoding the DNA methylome of mantle cell lymphoma in the light of the entire B cell lineage. Cancer Cell. 2016;30:806–21.
6. Flavahan WA, Drier Y, Liau BB, Gillespie SM, Venteicher AS, Stemmer-Rachamimov AO, et al. Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature. 2016;529:110–4.
7. Vilarrasa-Blasi R, Soler-Vila P, Verdaguer-Dot N, Russiñol N, Di Stefano M, Chapaprieta V, et al. Dynamics of genome architecture and chromatin function during human B cell differentiation and neoplastic transformation. Nat Commun. 2021;12:651.
8. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. Nat Methods. 2012;9:215–6.
9. Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489:57–74.
10. Anantharaman A, Lin I-J, Barrow J, Liang SY, Masannat J, Strouboulis J, et al. Role of Helix-Loop-Helix proteins during differentiation of erythroid cells. Mol Cell Biol. 2011;31:1332–43.
11. Vegliante MC, Palomero J, Perez-Galan P, Roue G, Castellano G, Navarro A, et al. SOX11 regulates PAX5 expression and blocks terminal B-cell differentiation in aggressive mantle cell lymphoma. Blood. 2013;121:2175–85.
12. Cobaleda C, Schebesta A, Delogu A, Busslinger M. Pax5: the guardian of B cell identity and function. Nat Immunol. 2007;8:463–70.
13. Walter K, Bonifer C, Tagoh H. Stem cell-specific epigenetic priming and B cell–specific transcriptional activation at the mouse Cd19 locus. Blood. 2008;112:1673–82.
14. Albero R, Enjuanes A, Demajo S, Castellano G, Pinyol M, García N, et al. Cyclin D1 overexpression induces global transcriptional downregulation in lymphoid neoplasms. J Clin Investig. 2018;128:4132–47.
15. Mohanty A, Sandoval N, Phan A, Nguyen TV, Chen RW, Budde E, et al. Regulation of SOX11 expression through CCND1 and STAT3 in mantle cell lymphoma. Blood. 2019;133:306–18.

ACKNOWLEDGEMENTS

This work was supported by research funding from the Spanish Ministerio de Ciencia, Innovación y Universidades through SAF2017-86126-R (to JIM-S), RTI2018-094274-B-I00 and PMP15/00007 which is part of Plan Nacional de I + D + I and co-financed by the ISCIII-Sub-Directorate General for Evaluation and the European Regional Development Fund (FEDER-"Una manera de Hacer Europa") (to EC), and the National Institutes of Health, National Cancer Institute "Molecular Diagnosis, Prognosis, and Therapeutic Targets in Mantle Cell Lymphoma" (P01CA229100 to EC). Furthermore, the authors would like to thank the support of the Generalitat de Catalunya Suport Grups de Recerca AGAUR 2017-SGR-736 (to JIM-S), 2017-SGR-1142 (to EC) and 2017-SGR-468 (to MAMR), the Accelerator award CRUK/AIRC/AECC joint funder-partnership, the CERCA Programme/Generalitat de Catalunya and CIBERONC (CB16/12/00225, CB16/12/00334 and CB16/12/00489). RV-B (BES-2013-064328) was supported by a predoctoral FPI Fellowship from the Spanish Government. EC is an Academia researcher of the Catalan Institution for Research and Advanced Studies (ICREA). The authors thank the Barcelona Supercomputing Center for access to computational resources. This work was developed at the Centro Esther Koplowitz (CEK, Barcelona, Spain).

AUTHOR CONTRIBUTIONS

RV-B, coordinated and performed FISH, in situ Hi–C, targeted MS, histone mark and ATAC-seq data generation and computational data analysis. NV-D, performed and supported FISH, in situ Hi–C, histone mark, and ATAC-seq data generation. LB, and AF, supported targeted MS data generation and analysis. PS-V, and MAM-R, supported in situ Hi–C computational data analysis. VC contributed in histone marks and ATAC-seq data analysis. MK, ACQ, and RB participated in DNA methylome data generation and analyses as well as in the study data interpretation. MP, MJC, XA, FP, SB, DC, EC, contributed with key reagents as well as sample collection and their biological and clinical annotation. RV-B and JIM-S directed the research and wrote the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41375-021-01389-w.

Correspondence and requests for materials should be addressed to R.V-B. or J.I.M-S.