

ARTICLE

Open Access

Loss of the candidate tumor suppressor *ZEB1* (*TCF8*, *ZFHX1A*) in Sézary syndrome

Elisabetta Caprini¹, Antonella Bresin¹, Cristina Cristofolletti¹, Mauro Helmer Citterich¹, Valeria Tocco², Enrico Scala¹, Alessandro Monopoli¹, Roberto Benucci¹, Maria Grazia Narducci¹ and Giandomenico Russo¹

Abstract

Cutaneous T-cell lymphoma is a group of incurable extranodal non-Hodgkin lymphomas that develop from the skin-homing CD4⁺ T cell. Mycosis fungoides and Sézary syndrome are the most common histological subtypes. Although next-generation sequencing data provided significant advances in the comprehension of the genetic basis of this lymphoma, there is not uniform consensus on the identity and prevalence of putative driver genes for this heterogeneous group of tumors. Additional studies may increase the knowledge about the complex genetic etiology characterizing this lymphoma. We used SNP6 arrays and GISTIC algorithm to prioritize a list of focal somatic copy-number alterations in a dataset of multiple sequential samples from 21 Sézary syndrome patients. Our results confirmed a prevalence of significant focal deletions over amplifications: single well-known tumor suppressors, such as *TP53*, *PTEN*, and *RB1*, are targeted by these aberrations. In our cohort, *ZEB1* (*TCF8*, *ZFHX1A*) spans a deletion having the highest level of significance. In a larger group of 43 patients, we found that *ZEB1* is affected by deletions and somatic inactivating mutations in 46.5% of cases; also, we found potentially relevant *ZEB1* germline variants. The survival analysis shows a worse clinical course for patients with *ZEB1* biallelic inactivation. Multiple abnormal expression signatures were found associated with *ZEB1* depletion in Sézary patients we verified that *ZEB1* exerts a role in oxidative response of Sézary cells. Our data confirm the importance of deletions in the pathogenesis of cutaneous T-cell lymphoma. The characterization of *ZEB1* abnormalities in Sézary syndrome fulfils the criteria of a canonical tumor suppressor gene. Although additional confirmations are needed, our findings suggest, for the first time, that *ZEB1* germline variants might contribute to the risk of developing this disease. Also, we provide evidence that *ZEB1* activity in Sézary cells, influencing the reactive oxygen species production, affects cell viability and apoptosis.

Introduction

Sézary syndrome (SS) is characterized by erythroderma, lymphadenopathy, and leukemic involvement of the peripheral blood. The circulating atypical lymphocytes usually are CD4⁺ T helper cells of memory phenotype¹. The prevalence of SS is around 0.3 cases per 100,000 people and it accounts for less than 5% of all cutaneous T-cell lymphoma

(CTCL). Despite numerous efforts to characterize the events of its pathogenesis, SS remains an incurable and fatal disease. Multiple genomic array-based studies have delineated a complex profile of chromosome aberrations characterizing SS genome with multiple sporadic genetic events of gains/losses in addition to recurrent abnormalities affecting mainly chromosome 8, 9, 10, and 17^{2–5}. The mutational landscape of CTCL, compiled with recent reports of next-generation sequencing (NGS) data, identified a broad range of genes variously affected by somatic copy-number alterations (SCNAs) and somatic single-nucleotide variants, involved, predominantly, in the T-cell activation and apoptosis, activation of NF-κB, JAK/STAT signaling, chromatin remodeling, and DNA damage

Correspondence: Maria Grazia Narducci (m.narducci@idi.it) or Giandomenico Russo (g.russo@idi.it)

¹Istituto Dermatologico dell'Immacolata, IDI-IRCCS, Rome, Italy

²Department of Paediatric Haematology and Oncology, Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) "Bambino Gesù" Children's Hospital, Rome, Italy

These authors contributed equally: Antonella Bresin, Cristina Cristofolletti
Edited by G. Melino

© The Author(s) 2018



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

response^{6–13}. Also, NGS data confirmed that a significant proportion of the oncogenic abnormalities identified in CTCL are SCNAs with an average of ~25 SCNAs observed per tumor;^{6,8,12} of these, focal deletions, resulting in frequent hemi-homozygous loss of target tumor suppressor gene loci, account for a significant proportion of the genetic aberrations, highlighting the significance of these genes as drivers of CTCL pathogenesis⁶. However, some discrepancies still exist concerning identities and frequencies of alterations found at these putative driver genes across the various studies. Particularly, some investigators reported *ZEB1* as variously affected by deletion and somatic mutation at 10p11.22 locus^{2,6,9,10}, whereas others did not^{7,8}. *ZEB1* is a zinc-finger-containing transcriptional repressor which regulates interleukin (IL)-stimulated cytokine signaling in normal T cells and it is essential for the correct development of T cells during hematopoiesis^{14–18}. Aberrantly expressed in a variety of human solid cancers, it is assumed to foster migration, invasion, and metastasis, promoting epithelial-to-mesenchymal transition (reviewed in ref. ¹⁹). On the contrary, *ZEB1* is a candidate tumor suppressor gene in adult T-cell leukemia/lymphoma (ATLL), where it contributes to TGF- β 1-mediated growth suppression resistance of malignant CD4⁺ T cells and *ZEB1* mutant mice frequently undergo spontaneous CD4⁺ T-cell lymphomas^{20,21}. Here, we report *ZEB1* as the target of a significant focal deletion with a correspondent transcript downregulation. DNA sequence analysis identified not only somatic inactivating mutations, but also germline variants of potential clinical impact. A tumor suppressor role for *ZEB1* in SS is further supported by our survival analysis showing a worse outcome for patients carrying *ZEB1* homozygous loss. Gene set enrichment analysis (GSEA) identified, among multiple expression signatures associated with *ZEB1* absence in SS patients, the cellular response to oxidative signal. Using *ZEB1* knockout (KO) cell lines, we verified that *ZEB1* has a role in controlling intracellular reactive oxygen species (ROS) production affecting viability and apoptosis of SS cells. Although our findings require additional functional investigations, they provide the basis to understand the onco-suppressive role of *ZEB1* in CTCLs.

Materials and methods

Patients and cell lines

Diagnosis of SS was based on described criteria²². Clinical characteristics of SS patients investigated with SNP6 arrays and GISTIC analysis are shown in Table S1. Hut78 (TIB161), H9 (HTB 176), and HH (CRL2105) cell lines established from peripheral blood of CTCL patients were obtained from American Type Culture Collection (ATCC) and grown in complete RPMI 1640 with 10% FBS (Sigma-Aldrich).

SNP6 DNA array

Tumor CD4⁺ T lymphocytes and matched granulocyte cells (normal counterpart) were isolated as previously described;³ genomic DNA was extracted using Archive-Pure DNA Cell/Tissue Kit (5Prime, Gaithersburg, MD) following the manufacturer instructions. DNA samples were profiled on Affymetrix Genome-Wide Human SNP 6.0 Array as per the manufacturer's protocol (Affymetrix Santa Clara, CA), and the CEL files were generated using the GeneChip Command Console software 4.0 (Affymetrix). Normalization and data modeling have been performed with dChip²³, and tumors were analyzed using paired normal samples as the reference. The calculated tumor-to-normal log₂ ratio data were then exported and segmented with the Circular Binary Segmentation algorithm (CBS; $\alpha = 0.01$, permutations = 10,000)²⁴ via GenePattern platform at <http://software.broadinstitute.org/cancer/software/genepattern/>. Thresholds for copy-number change were fixed at ± 0.3 for gains/losses. We used the same segmentation algorithm applied for SNP6 arrays also for re-analyzing our previous 10 K SNP dataset³ to estimate the regions of copy-number changes, similar to our previous work on PTEN analysis²⁵. With respect to this work, however, we have introduced additional follow-up samples/patients arrays (indicated with N superscript in Table 1) and, for those tumors with enough specimen material for new hybridizations, we have substituted the 10 K with SNP6 platform (S superscript in Table 1); finally, for more consistency of 10 K analysis results, we have eliminated SS patients with neoplastic clonal expansion equal or below 50% (i.e., SS44 and 46 of Table S1 in Cristofolletti et al.²⁵). For the identification of significant regions of SCNAs, we applied the standard GISTIC (Genomic Identification of Significant Targets in Cancer) algorithm²⁶ on the segmented SNP6 data using the GISTIC GenePattern module previous removal of genomic regions of structural copy-number variations identified in the Database of Genomic Variants (DGV, <http://dgv.tcag.ca/dgv/app/home>).

Real-time PCR

Total RNA extraction by TRIzol reagent (Invitrogen) was performed according to the manufacturer's instruction. Mature miR200c was analyzed using the TaqMan microRNA assay (Applied Biosystems) with RNU48 and miR16 as normalization references. *ZEB1* expression analysis was performed with GoScript Reverse Transcriptase and GoTaq qPCR master mix (Promega), using β -actin as normalization reference, on ABI PRISM 7000 SDS (Applied Biosystem). Relative expression was calculated using the comparative Ct method ($2^{-\Delta\Delta Ct}$)²⁷. The following primers were used: *ZEB1* forward: 5'-TGCC AACAGACCAGACAGTG-3' reverse: 5'-CTGTTCATCC

Table 1 Status of ZEB1 locus 10p11.22 (chr10:31,648,147–31,856,740)

Sample ID	Array Type	Genomic Alteration [§]	Mutation/Variation	Variation type	mRNA FC (Log2 Values)
P02T	10K/aCGH	Loss	WT		-0.24
P05T	10K/aCGH	Homo Del	Homo Del		-2.92
P08T1	10k	Loss	n.d.		n.d.
P08T2	10K/aCGH	Loss	WT		-1.68
P11T	10K/aCGH	No change	c.481+64A>G	G	n.d.
P15T	10K/aCGH	Loss	L689*	S	n.d.
P22T	10K/aCGH	Homo Del	Homo Del		-2.10
P23T1	SNP 6.0	No change	n.d.		n.d.
P23T2	SNP 6.0	Loss	n.d.		n.d.
P23T3	SNP 6.0	Loss	WT		-2.25
P25T	10K/aCGH	No change	WT		-2.14
P27T	10K/aCGH	No change	WT		-2.18
P28T	10K/aCGH	UPD	R301*	S	-0.85
P30T1	10K	No change	n.d.		1.01
P30T2	SNP 6.0 [§]	No change	n.d.		n.d.
P30T3	SNP 6.0 ^W	No change	WT		n.d.
P32T1	10K	Homo Del	Homo Del		-1.81
P32T2	SNP 6.0	Homo Del	Homo Del		-5.54
P33T	10K/aCGH	No change	V853I	G	1.41
P34T	10K/aCGH	Loss	WT		-0.80
P35T	10K/aCGH	No change	c.685-15G>A	G	1.08
P36T	10K/aCGH	No change	c.481+64A>G	G	0.38
P37T	10K/aCGH	No change	WT		-1.88
P38T1	10K/aCGH	Loss	n.d.		-1.79
P38T2	10K	Loss	n.d.		n.d.
P38T3	SNP 6.0	Loss	WT		n.d.
P39T	10K/aCGH	Loss	c.*25T>C	G	0.77
P39T2	SNP 6.0	Loss	n.d.		0.67
P40T3	SNP 6.0 [§]	No change	WT		-1.63
P41T2	10K/aCGH	No change	n.d.		n.d.
P41T3	SNP 6.0	No change	WT		1.72
P42T1	10K/aCGH	No change	WT		n.d.
P42T2	10K ^W	No change	n.d.		n.d.
P43T1	10k	No change	n.d.		n.d.
P43T2	10K/aCGH	No change	n.d.		-0.93
P43T3	SNP 6.0	No change	WT		2.99
P45T1	10K/aCGH	No change	n.d.		1.55
P45T2	SNP 6.0	No change	WT		n.d.
P48T1	10K/aCGH	Loss	WT		-3.67
P49T1	10K/aCGH	No change	c.58+20G>A	S	0.48
P50T1	SNP 6.0	No change	n.d.		n.d.
P50T2	SNP 6.0 [§]	No change	c.481+64A>G	G	-0.16
P51T1	10K/aCGH	No change	c.481+64A>G	G	1.86
P53T1	SNP 6.0	Loss	n.d.		-1.04
P53T2	SNP 6.0	Loss	WT		-1.39
P55T1	10k	Loss	T232T Q840P	G	0.3
P57T1	SNP 6.0	Loss	WT Y23Y	G	n.d.
P58T1	10k	Loss	c.481+64A>G	G	n.d.
P60T1	SNP 6.0	No change	n.d.		0.20
P60T2	SNP 6.0	No change	n.d.		0.64
P60T3	SNP 6.0 ^W	No change	c.481+64A>G	G	n.d.
P61T1	SNP 6.0 [§]	Loss	WT		-1.82
P61T2	SNP 6.0 [§]	Homo Del	Homo Del		-2.85
P61T3	SNP 6.0 ^W	Loss	n.d.		n.d.
P62T1	10k	No change	c.481+64A>G R1097K	G	-1.41
P63T1	SNP 6.0	No change	n.d.		1.45
P63T2	SNP 6.0	Loss	c.481+64A>G	G	-1.49
P64T1	SNP 6.0	Loss	c.481+64A>G	G	-4.84
P64T2	SNP 6.0	Loss	n.d.		-3.10
P65T1	SNP 6.0	No change	n.d.		n.d.
P66T1	10k	No change	WT		n.d.
P67T1	SNP 6.0	No change	c.685-15G>A	G	-1.61
P68T1	SNP 6.0 ^W	Loss	c.481+64A>G	G	-0.28
P68T2	SNP 6.0 ^W	Loss	n.d.		n.d.
P69T1	SNP 6.0	No change	WT		-1.12
P70T1	SNP 6.0	No change	n.d.		n.d.
HUT78	SNP 6.0	Loss	WT		-1.81
H9	SNP 6.0	Loss	c.481+64A>G	G	0.60
HH	SNP 6.0	Loss	WT		1.84

FC fold change, UPD uniparental disomy determined by dChip SNP software (Caprini, E et al. 2009), *Homo Del* homozygous deletion, WT wild type, n.d. not determined, G germline, S somatic

[§]Determined by Circular Segmentation Algorithm (Olshen et al.²⁴):

Loss = log2 value ≤ -0.3; Homo Del = log2 value ≤ -2.0; no change = -0.3 ≤ log2 value ≤ +0.3

Referred to the data set of Cristofaletti et al.²⁵; [§] indicates array sample substitution, ^N indicates new sample array

Note: light-gray shadow includes multiple samples from the same patient. Bold font indicates ZEB1 DNA mutation/variation

TCCCAGCAGTT-3'. *β-actin* forward: 5'-GATGAGAT TGGCATGGCTTT-3' reverse: 5'-GTCACCTTCACCGT TCCAGT-3'.

Mutation analysis

The nine coding exons of ZEB1 gene were amplified by genomic PCR using 12 couples of primers. PCR products were then subjected to direct Sanger nucleotide sequencing. Primers' sequence is reported in Table S2. ZEB1 sequence variants were assessed to determine potential effects on splicing using Human Splicing Finder Version 3.1 (available in the public domain at: <http://www.umd.be/HSF3/index.html>). The effects of amino acid substitutions on protein function were assessed using the computational prediction scores accessible through VarSome search engine available at: <https://varsome.com/>.

Survival and statistical analyses

A time-to-event analysis was performed using non-parametric Kaplan–Meier product-limit survival estimates, and differences between Kaplan–Meier survival curves were calculated using the Mantel–Haenszel log-rank test. Analyses were performed using survival curve and survival difference modules at: <http://software.broadinstitute.org/cancer/software/genepattern/>.

Gene set enrichment analysis

The GSE17601 dataset published in 2009³ was processed and analyzed by GenePattern (<http://software.broadinstitute.org/cancer/software/genepattern>) using GEOImporter module (v5) and GSEA module (v17) with the Molecular Signatures Database “Hallmarks” gene set collection; briefly, this method identifies predefined database gene sets associated with phenotypic differences utilizing *t*-test statistics²⁸. The default GSEA basic parameters were used with Gene Set as permutation-type option due to the small size of the dataset (<7); to find gene sets that correlate with ZEB1 expression profile (continuous phenotype label), Pearson metric was used for ranking genes.

CRISPR/Cas9 knockout

CRISPR–Cas9 technique was used to knock out ZEB1. All-in-one vectors (pCLIP-ALL-hCMV-Puro) expressing Cas9 nuclease and specific or control gRNA were purchased from Transomic Technologies (Huntsville, AL,

USA). For detailed information on the construction of *ZEB1* heterozygous and homozygous knockout cell lines, please refer to Supplementary Methods. Briefly, SS cell lines were first characterized for endogenous *ZEB1* level (see Figure S1), and H9 cell line was chosen for knockout. Then, Amaxa™ Nucleofector™ Technology (Lonza Cologne GmbH) was used for cell transfection, and selection was performed with 1 µg/mL of puromycin for 72 h. Next, clonal cells were generated by single-cell dilutions and screened by a custom-designed drop-off assay performed with droplet-digital PCR (Biorad) as illustrated in Figure S2. Finally, western blot analysis was used to assess the level of *ZEB1* protein as shown in Figure S3.

Western blot

Western blot (WB) analyses were conducted as previously described²⁵. Membranes were probed with primary antibodies for *ZEB1* (1:400; Santa Cruz) and β -actin (1:5,000; Santa Cruz) (Santa Cruz Biotechnology, Inc., Dallas, TX, USA). Immunodetection was performed with appropriate horseradish peroxidase-linked secondary antibodies and enhanced chemiluminescence detection reagents (GEHealthcare, Amersham Biosciences, Little Chalfont, UK). The film was scanned on a GS-710 Calibrated Imaging Densitometer and analyzed by means of Quantity One Software Version 4.1.1 (Bio-Rad Laboratories, Hercules, CA, USA) and by Image J open-source software (<https://imagej.net>).

ROS generation assay

Cells were stained with the oxidation-sensitive dye H2DCFDA (20 µmol/L) (Abcam) for 30 min. Then cells were treated with the indicated concentration of glucose oxidase (from *Aspergillus niger*, Sigma-Aldrich) for 4 hours, and ROS generation was determined by fluorescence-activated cell sorting (FacsCalibur, BD Biosciences, San Jose, CA, USA).

Cell viability and cell death assays

The cells were treated with increasing concentrations of glucose oxidase, as indicated, for 24 hours and measured by the methyl thiazolyl tetrazolium (MTT) assay using ELISA reader (Biorad). Each experiment was performed three times, and results are shown as means \pm SEM. Apoptosis was assessed by AnnexinV-FITC and propidium iodide (Sigma-Aldrich). Samples were acquired on a flow cytometer and analyzed using CellQuest software (BD Biosciences).

Statistical analysis

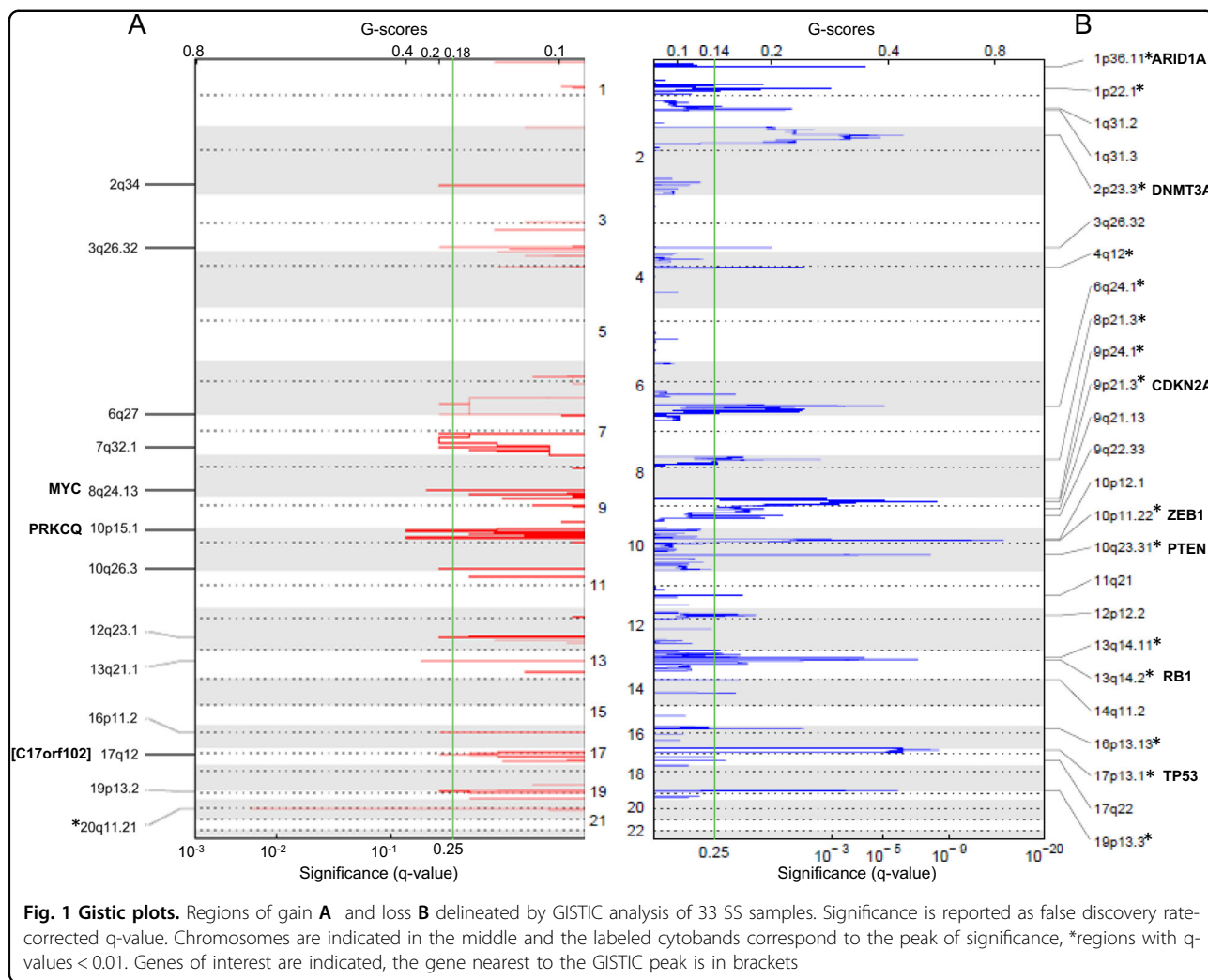
Where not otherwise specified, statistical analyses were performed with GraphPad PRISM 6 software (GraphPad Software Inc., La Jolla, CA). Differences were evaluated

with two-tailed Student's *t* test and Pearson correlation test. $P \leq .05$ was considered significant.

Results

GISTIC analysis of somatic DNA copy-number alterations in Sézary syndrome identified *ZEB1* gene as the target of a highly significant focal deletion of 10p11.22

We used SNP 6 arrays and GISTIC2.0 algorithm to identify genes targeted by focal SCNAs, likely representing candidate “drivers” for cancer growth^{26,29}, in a dataset composed of 33 samples, collected at different time points, from 21 SS patients and 3 CTCL cell lines (Hut78, H9, HH). Applying the default threshold for significance ($qv.thresh = 0.25$), we found 13 regions of gains and 25 genomic losses across our dataset (Fig. 1 and Table S3). Considering as significant all events with false discovery rate (FDR) q -values < 0.01 , only 20q11.21 could be taken into account among the focal amplifications; this locus, however, was found only in the Hut78 and H9 cell lines, but in none of the SS samples (data not shown), suggesting that it represents an alteration related to the in vitro replication. If the chromosomes most frequently affected by copy-number gains were considered (i.e., chromosome 10p, 8/8q, and 17q²⁻⁵), the following GISTIC segments may assume importance: 10p15.1 with 19% of SS tumors affected (4/21) spanning *PRKCQ* gene, 8q24.13, altering 62% of cases (13/21 and 3/3 cell lines) encompassing two genes, one of which is *MYC*, already cited as candidate gene in SS², and 17q12, involving 52% of tumors (11/21 and 2/3 cell lines), where the GISTIC peak falls in close proximity of an uncharacterized open-reading frame (*C17orf102*) (Fig. 1). On the other hand, we found 15 focal deletions having FDR q -values < 0.01 , four of which identified single consensus cancer drivers: 10p11.22 targeted *ZEB1* gene, which showed a deletion frequency of 47.6% (10/21 tumors and 3/3 cell lines); 17p13.1 interval, deleted in 81% of SS (17/21 and 1/3 cell line), identified *TP53*; 10q23.31 deletion fragment present in 57% of cases (12/21 and 3/3 cell lines) enclosed *PTEN*; and 13q14.2 deletion, affecting 19% of tumors (4/21 and 3/3 cell lines), involved *RBI*. The 9p21.3 GISTIC peak encompassed ten genes including *CDKN2A* as the main candidate cancer driver of this interval deleted in 33% of SS (7/21 and 3/3 cell lines) (Table S3). The high level of significance of *ZEB1* deletion in our cohort of SS cases, prompted us to re-examine the 10 K array dataset (see Materials and methods for details) and combine the 10 K and the SNP6 results to investigate the *ZEB1* locus alterations in a total of 66 specimens from 43 SS patients and 3 CTCL cell lines (Table 1). We found loss of genetic material (defined as $-2.0 < \text{Log}_2 \text{ values} \leq -0.3$) for about 35% (15/43) of SS cases and 3 cell lines, homozygous DNA deletions (defined as $\text{Log}_2 \text{ value} \leq -2.0$) in ~9% (4/43) of SS patients, one case (P28) of uniparental disomy



(UPD, i.e., a deletion followed by a re-duplication of the remaining allele³), for a total of about 46.5% (20/43) of patients, and 3 cell lines affected by *ZEB1* locus abnormalities (Table 1). Eight of ten cases with one or more follow-ups (P08, P32, P38, P39, P53, P61, P64, and P68) display *ZEB1* loss in all the samples examined, whereas in two patients (P23, P63), the deletion appears from the second follow-up (median of the first observation time 4.32 months, median of follow-up observation time 29.47 months). This suggests that *ZEB1* aberrations might be related to disease onset rather than to tumor progression. Altogether, our data identified *ZEB1* as the target of a highly significant SCNA within 10p11.22 locus affecting almost half of the SS patients and the CTCL cell lines analyzed.

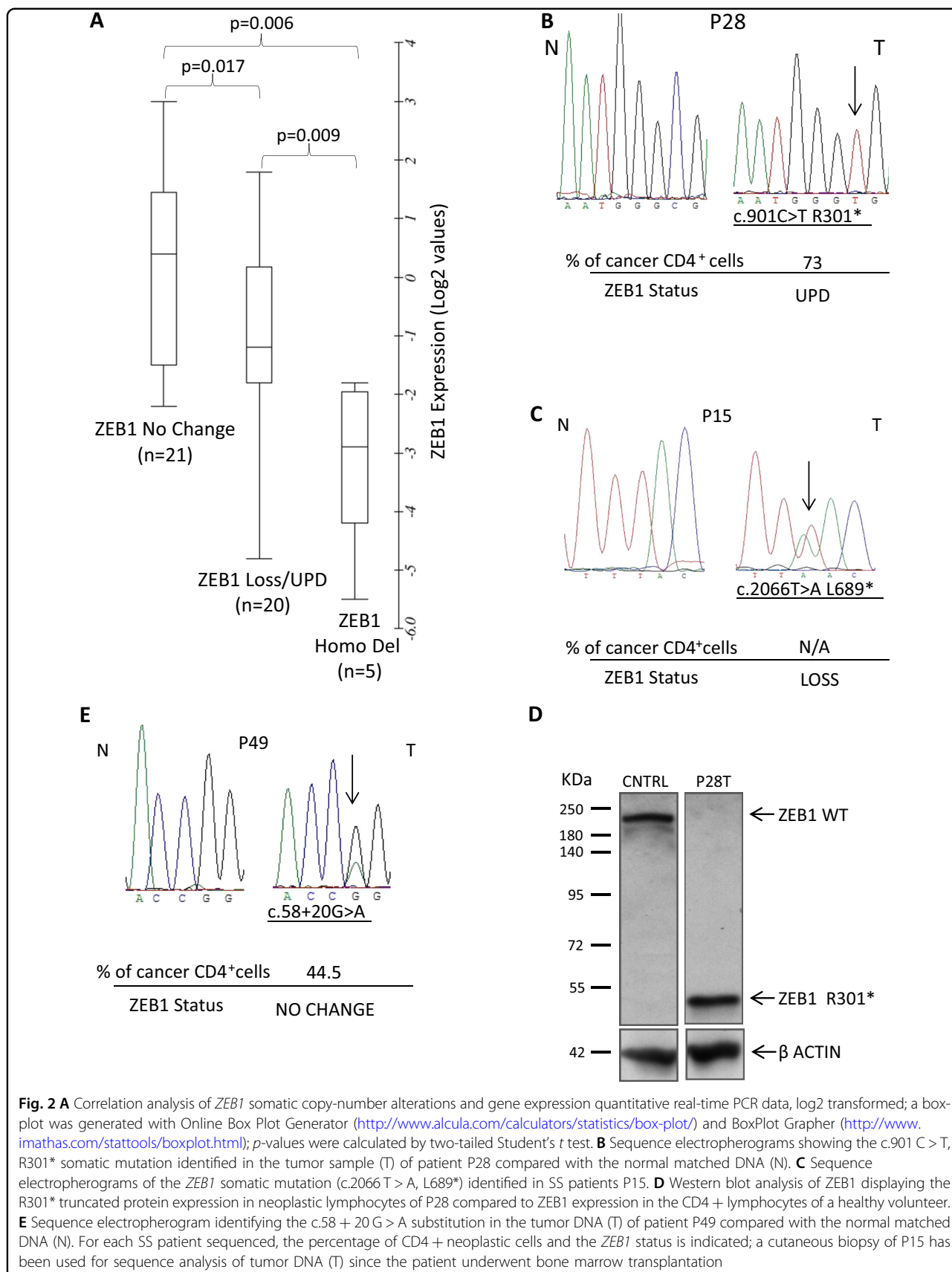
***ZEB1* mRNA expression in SS tumors is dependent on copy-number status**

To verify whether the high frequency of *ZEB1* SCNA led also to an altered transcript level, we analyzed the

mRNA expression in 43 samples, with available RNA, from 35 SS patients and 3 cell lines (Table 1); we used CD4⁺ T cells from four healthy controls to calibrate the fold-change (FC) values of *ZEB1* in quantitative PCR assays. Comparing samples showing loss (20/46) or homozygous deletion (5/46) with samples exhibiting no change of *ZEB1* copy number (21/46), we found *ZEB1* expression values proportional to copy-number data with significant statistical differences (Fig. 2A), indicating that *ZEB1* mRNA level correlates with its genomic locus status.

***ZEB1* mutation analysis in SS patients**

We then performed *ZEB1* DNA sequence analysis in 38 SS patients and 3 CTCL cell lines. We detected seven coding nucleotide substitutions in 6/38 cases and four noncoding variations affecting 14 patients and 1 cell line. The available normal matched DNA allowed us to establish that three were somatic and eight were germline substitutions (Table 1). Two somatic mutations were



nonsense predicting to generate *ZEB1*-truncated proteins: c.901 C > T (R301*) in P28 and c.2066 T > A (L689*) in P15 (Fig. 2B, C). These SS cases display also UPD and loss of *ZEB1* locus, respectively. Western blot analysis revealed that c.901 C > T mutated gene produces a single fragment of a lower molecular mass compared with the full-length protein (Fig. 2D); if, similarly, the c.2066 T > A mutated gene is translated in a shortened 689 amino acids, then both proteins will likely result in defective molecules (Fig. S4). The third somatic variation, identified in P49 associated with a normal copy number of *ZEB1*, is a transition (c.58 + 20 G > A) occurring within the first intron of the gene (Fig. 2E). Computational analysis predicted the nucleotide change as probably having no effect on splicing (see Materials and methods for mutation analyses). Since the eight other substitutions were germline, we sought to determine their pathogenicity using data from dbSNP and Genome Aggregation Database (gnomAD)³⁰ (Table 2). The intronic substitutions c.481 + 64 A > G and c.685–15 G > A, identified in 10/38 and 2/38 SS patients, respectively, are genetic variants reported as having minor allele frequencies (MAF) in dbSNP, or allele frequencies of the alternative (non-reference) allele (AF) in gnomAD, above 0.05, thereby likely representing a genetic variant of benign nature. Conversely, we found three rare genetic variations (MAF and AF < 0.05): one missense transversion (c.2519 A > C, Q840P) and two synonymous transitions (c.696 G > A, T232T and c.69 T > C, Y23Y). The missense substitution is reported as having clinical relevance owing to its association with late-onset Fuchs corneal dystrophy (FCD) and shown to cause loss of function (LoF) by an in vivo complementation assay³¹. Noticeably, the patient (P55) carrying this variant displays also concomitant loss of *ZEB1* locus (Fig. 3A, left panel) possibly resulting in the deficiency of a functional allele. The c.696 G > A, T232T was found in the same patient (P55) (Fig. 3A, right panel), and the c.69 T > C, Y23Y in P58 SS case (Fig. 3B, left panel) both combined with *ZEB1* deletions; these silent rare variations were tested for their effects on putative exonic splicing enhancers and silencers³², but the analysis did not detect any significant splicing motif alteration; it cannot be excluded, however, that they might have other consequences on mRNA stability and/or translation efficiency³³. Finally, we identified the following three germline variants: (1) the c.2557 G > A (V853I) missense substitution in P33 associated with no change of *ZEB1* locus (Fig. 3B, right panel), (2) the c.3290 G > A (R1097K), found in P62 carrying a normal *ZEB1* copy number, and (3) c.*25 T > C present in P39 with *ZEB1* loss (Fig. 3C), none of which was found in the dbSNP and gnomAD (Table 2). It is worth noting that these variations were all absent also from the COSMIC (Catalogue of Somatic Mutations in Cancer) database, suggesting that they do not occur in sporadic cancers.

Multiple lines of computational evidence suggest no impact of c.2557 G > A (V853I) and c.3290 G > A (R1097K) on gene or gene product; however, due to the absence from a large general population, the significance of these variant alleles remains uncertain.

Biallelic inactivation of *ZEB1* reduces survival in SS patients

We then evaluated if *ZEB1* deficiency might affect SS patients survival. In our series, median overall survival was 48 months for *ZEB1* no change or loss of copy number ($n = 34$), compared with 30 months for cases with biallelic loss ($n = 7$), including patients carrying inactivating (i.e., nonsense mutations) or Q840P LoF variation associated with loss. Overall survival analysis with the Kaplan–Meier estimator revealed reduced survival probability among patients with biallelic loss (log-rank test, $P = 0.0106$; Fig. 3D). Thus, beyond representing a potential oncogenetic event, lack of *ZEB1* function also affects SS prognosis.

Gene sets enriched in *ZEB1*-depleted SS samples

To gain further insights into *ZEB1* role in SS disease, we used our previously published gene expression (GE) dataset (GSE 17601³) to explore the transcriptional patterns of *ZEB1*-deficient and *ZEB1*-expressing patients. Considering the available 32 arrays in the GE dataset, we had corresponding *ZEB1* genetic and expression data for 19 samples (Table 1: P02T, P05T, P08T2, P22T, P23T3, P25T, P28T, P30T1, P32T1, P33T, P34T, P35T, P36T, P37T, P38T1, P39T1, P40T3, P43T2, P45T1). Clustering these SS cases for *ZEB1* expression, we had six samples showing a positive log₂ FC values (P30T1, P33T, P35T, P36T, P39T1, P45T1) that we labeled as *ZEB1* OVER class; in order to maximize the transcriptional differences, we compared this group of cases with a comparable size set of six *ZEB1*-depleted samples, i.e., exhibiting either homozygous deletion of *ZEB1* (P05T, P22T, P32T1), UPD (P28T) or strong mRNA down-regulation with log₂ FC expression level ≤ -2.0 (P23T3, P25T); for simplicity, we indicated this class as *ZEB1* HD (i.e., HomoDel). We used GSEA and the Molecular Signatures Database “Hallmarks” gene set collection³⁴ to investigate the molecular alterations enriched in the two different groups of patients; The analysis yielded 32 gene sets that were enriched, with FDR < 0.25, in the *ZEB1* HD class (Table 3A). To further verify our findings and select the Hallmarks signatures dependant on *ZEB1* expression, the same dataset of samples was used to identify lists of enriched gene sets negatively correlated with the expression profile of *ZEB1* (Affymetrix ID: 210875_s_at), i.e., gene sets made up of core enriched genes mostly up-regulated in *ZEB1*-deficient samples, but downregulated in *ZEB1*-expressing cases. A negative association was found for 11 gene sets (FDR < 0.25) all of which were included in

Table 2 ZEB1 germline substitutions

Nucleotide position ⁵	Amino acid change	Exon/ intron	No. of cases (%)	dbSNP rs* global MAF	Alternative allele frequencies (AF)	
					gnomAD ⁵ exome	gnomAD ⁵ whole genome
c.69T > C	Y23Y	Exon 2	1/38 (2.63)	rs775270565 C = 0.00002/2	C = 0.00001223	C = 0.0000323
c.481 + 64 A > G	–	Intron 4	10/38 (26.3)	rs2839664 G = 0.0787/394	–	G = 0.1308
c.685-15 G > A	–	Intron 5	2/38 (5.2)	rs220060 G = 0.0787/394	A = 0.9458	A = 0.9197
c.696 G > A	T232T	Exon 6	1/38 (2.63)	rs149166539 A = 0.0002/1	A = 0.001227	A = 0.0009054
c.2519 A > C	Q840P	Exon 7	1/38 (2.63)	rs118020901 C = 0.0030/15	C = 0.007583	C = 0.009168
c.2557 G > A	V853I	Exon 7	1/38 (2.63)	Absent	Allele not found despite > 20X coverage	Allele not found despite > 20X coverage
c.3290 G > A	R1097K	Exon 9	1/38 (2.63)	Absent	0.00	Allele not found despite > 20X coverage
c.*25 T > C	–	Exon 9	1/38 (2.63)	Absent	Allele not found despite > 20X coverage	Allele not found despite > 20X coverage

⁵Ref Seq NM_030751 (CDS), from 5' to 3' end

* <https://www.ncbi.nlm.nih.gov/snp> (Build 151). The current default global population is 1000Genome phase 3 genotype data from 2500 worldwide individuals, released in the May 2013 dataset. ⁵Genome aggregation database r2.0.2 accessed through <https://varsome.com/>. The data set spans 123,136 exome sequences and 15,496 whole-genome sequences from unrelated individuals

the *ZEB1* HD class of enriched signatures (Table 3B). Notably, the presence of the Hallmark of IL2_STAT5 signaling either in positive association with *ZEB1* HD class or in negative correlation with *ZEB1* expression support our analyses since *ZEB1* is a known repressor of IL2 transcription in T cells^{14,15}. The results identified the cell response to interferon (IFN) α and IFN γ as highly significant signatures associated with *ZEB1* depletion, suggesting it might have a role in the immune-regulatory activities mediated by these cytokines in SS T lymphocytes. Of note also the correlation with the tumor necrosis factor (TNF) α signaling via NF- κ B, one of the most important activated pathway conveying persistent survival signals in SS³⁵. Additionally, we found enriched gene sets, such as DNA repair and the P53 pathway, indicating a potential *ZEB1* control also in the regulation of genomic DNA integrity. To test this hypothesis, we compared the fraction of tumor genome affected by SCNAs for nine SS samples with *ZEB1* homo-deletions or strong mRNA under expression and eight SS tumors with *ZEB1*-positive expression calculated as the proportion of genome altered (PGA) described by Thu KL et al.³⁶ No significant results were found considering either all lesions (i.e., gains and losses) or losses and gains taken separately (Supplementary Figure S5), thus, *ZEB1*

depletion does not seem to increase genetic instability in SS. The more widespread changes in the transcriptional program observed comparing *ZEB1* HD with *ZEB1* OVER classes might be explained, at least in part, as the downstream effects of *ZEB1* lack of activities distinguishing the two groups. In support of this data interpretation, we found the upregulation of the oxidative phosphorylation enzymes in both types of analyses (Tables 3A and 3B), and the hallmarks of reactive oxygen species pathways and peroxisome associated with *ZEB1* HD class (Table 3A). This is of particular interest as *ZEB1* is indeed involved in ROS-induced cellular response as a crucial target of the miR-200 family members that are upregulated under oxidative stress conditions^{37–41}.

ZEB1 is not regulated by miR200c in SS

We tried to assess if any relationship existed between *ZEB1* and miR200 expression in SS. From microarray data of our previous study on miRNA transcription profiling in SS⁴², we identify miR200c as the family member having the highest variance in expression between patients and controls; thus we performed a real time quantitative assay to estimate miR200c FC values in 18 SS patients/cell lines (P25T, P30T1, P37T, P39T1, P39T2, P40T3, P43T3, P45T1, P48T1, P51T1, P62T1, P63T1, P67T1, P68T1,

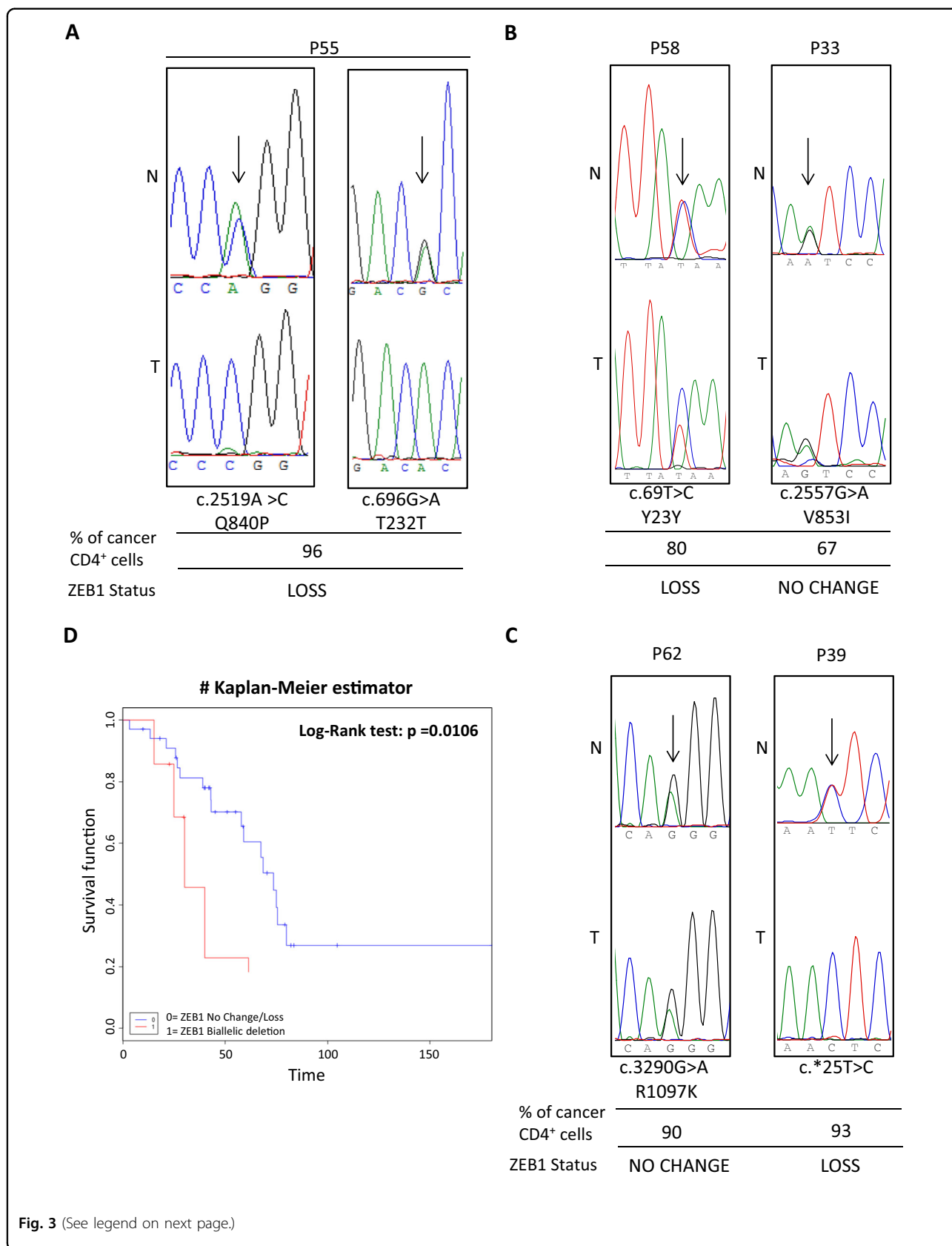


Fig. 3 (See legend on next page.)

(see figure on previous page)

Fig. 3 A Sequence electropherograms showing the germline heterozygous c.2519 A > C (Q840P) (left) and c.696 G > A (T232T) (right) rare variations identified in the normal DNA of P55 (N) which become hemizygous following the loss of the wild-type allele in the tumor sample (T). **B, C** Sequence electropherograms of DNA patients displaying rare c.69 T > C, Y23Y, and the newly identified c.2557 G > A, V853I; c.3290 G > A, R1097K; c.*25 T > C germline variants. The percentage of neoplastic CD4+ cells are indicated for each patient as well as the *ZEB1* status. **D** Comparison of survival time (months) between SS patients carrying no change or loss of copy number ($n = 34$, blue line) vs. biallelic deletion ($n = 7$, red line) of *ZEB1* gene. Patients with truncating (R301*, L689*) or loss-of-function mutations (Q840P) associated with loss were included in the biallelic deleted group. Forty-one out of 43 patients were used for this analysis because one was lost at follow-up and the other had an evolution of SS after a long stable MF disease

P69T, HUT78, H9, and HH cell lines) using the same set of healthy donors employed for *ZEB1* expression analysis. No significant correlation emerged between the log₂ FC values of *ZEB1* and miR200c in a Pearson's correlation analysis ($R = 0.0144$, p -value = 0.956). Also, dividing the same group of samples for *ZEB1* expression level, we calibrated the miR200c FC values of the under-expressing cases (P25T, P37T, P40T3, P48T1, P62T1, P67T1, P68T1, P69T, HUT78) on the average value of the over-expressing tumors (P30T1, P39T1, P39T2, P43T3, P45T1, P51T1, P63T1, H9, HH); in this case a weak negative correlation emerged that, however, did not reach the significance (Pearson's correlation $R = -0.426$, p -value = 0.25). Our results suggest that *ZEB1* role in oxidative phosphorylation response of SS cells is independent from miR200c expression.

Oxidative stress and *ZEB1* depletion in SS cells

The stimulation of T-cell receptor (TCR) generates an oxidative signal that is crucial to trigger the activation-induced cell death (AICD), a key process required to keep in balance the immune system^{43,44}. This prompted us to study the effects of *ZEB1* deficiency in SS context of oxidative condition. To this end, we used heterozygous and homozygous CRISPR/Cas9-generated *ZEB1* KO cell clones, derived from H9 SS cell line (see Supplementary Methods), treated with glucose oxidase (GOX) which reacts with glucose and oxygen for a continuous enzymatic generation of H₂O₂. This reagent, used at relatively low concentration (2–5 mU/mL), it is used to reproduce a physiological condition of constant low level H₂O₂-mediated oxidative signal⁴⁵. Flow cytometry analysis of cells cultured with increasing concentrations of GOX for 4 hours showed a sustained increase in intracellular ROS production, however, if compared with the WT CRISPR/Cas9 control cells (Ctr *ZEB1*^{wt}), *ZEB1*^{-/-} cell clone (C9 *ZEB1*^{-/-}) showed an appreciable lower amount of ROS at 3.5 and 4.5 mU/mL of GOX concentration, while the fluorescence intensity of *ZEB1*^{+/-} cells (B8 *ZEB1*^{+/-}) overlapped the WT control at 3.5 mU/mL and exhibited an intermediate profile between the two genotypes at 4.5 mU/mL of GOX (Fig. 4A). When ROS levels were calculated relatively to *ZEB1*^{wt} control cells, significant statistical decreases were measured for *ZEB1*^{-/-} cells at both

GOX concentrations, whereas no significant differences emerged in *ZEB1*^{+/-} clone (Fig. 4B). These results suggest that *ZEB1*-deficient SS cells are more competent to respond to H₂O₂-induced oxidative stress counterbalancing increasing intracellular ROS production. Cell viability assays of *ZEB1*^{wt} and KO clones showed that *ZEB1*^{-/-} cells exhibited higher cell viability with respect to *ZEB1*^{wt} and *ZEB1*^{+/-} following the exposure to increase GOX-induced oxidative stress with significant differences observed at 4.0 and 4.5 mU/mL (Fig. 4C). Flow cytometry, used to quantify cellular apoptosis in response to 3.5 and 4.5 mU/mL of GOX, revealed that *ZEB1*^{-/-} KO showed 29 and 64% of apoptotic cells, respectively; conversely, *ZEB1*^{wt} exhibited 64 and 94% of apoptosis at the corresponding GOX concentrations, comparable with the 67 and 82% values observed for *ZEB1*^{+/-} KO (Fig. 4D). Collectively, our data show that *ZEB1*-devoid SS cells are able to respond to oxidative stress counteracting ROS imbalance, this ability influences cell survival and apoptosis and it requires the absence of both *ZEB1* alleles.

Discussion

We have generated a high-resolution SNP6 data set of SS tumors; in the absence of hallmarks of disease progression, we used multiple follow-up samples to increase the detection of positively selected clonal DNA aberration in the disease course. The GISTIC algorithm delineated a profile of significant regions of gain and loss, with a large predominance of focal deletions over amplifications. This is in line with previously published studies and support the hypothesis that deletion events, targeting critical tumor suppressors, represent a mechanism of CTCL pathogenesis^{5,6}. We narrowed some of the most significant focal deletions to critical tumor suppressor targets, such as *TP53*, *PTEN*, and *RBI* (Fig. 1), thereby involving their loss as driver events in SS tumorigenesis. Recent studies, integrating copy-number and single-nucleotide variant analyses, identified these genes with variable frequencies in CTCL cohorts: *TP53* was affected in 21–92.5% of cases, *PTEN* ranged from 20 to 30%, and *RBI* and *CDKN2A* were targeted in 16–39% and 32.5–58% of tumors, respectively^{6,8,9,11,46}. Our data also confirmed targets such as *PRKCQ* found altered in 20–30% of CTCLs^{6,11}, and support other candidates like

Table 3 Gene sets enriched in ZEB1 HD class (A). Gene sets enriched in negative correlation with ZEB1 profile (B)

	NAME	SIZE	ES	NES	NOM p-val	FDR q-val
(A)						
1	HALLMARK_OXIDATIVE_PHOSPHORYLATION	188	0.67	2.67	0	0
2	HALLMARK_INTERFERON_ALPHA_RESPONSE	77	0.74	2.62	0	0
3	HALLMARK_INTERFERON_GAMMA_RESPONSE	171	0.65	2.58	0	0
4	HALLMARK_COMPLEMENT	181	0.58	2.31	0	0
5	HALLMARK_MYC_TARGETS_V1	171	0.57	2.29	0	0
6	HALLMARK_ADIPOGENESIS	169	0.53	2.08	0	0
7	HALLMARK_ALLOGRAFT_REJECTION	191	0.52	2.07	0	0
8	HALLMARK_REACTIVE_OXIGEN_SPECIES_PATHWAY	43	0.58	1.89	0	0.001
9	HALLMARK_APOPTOSIS	157	0.48	1.88	0	0.001
10	HALLMARK_IL6_JAK_STAT3_SIGNALING	85	0.51	1.85	0	0.001
11	HALLMARK_PROTEIN_SECRETION	88	0.5	1.83	0	0.001
12	HALLMARK_IL2_STAT5_SIGNALING	169	0.45	1.78	0	0.003
13	HALLMARK_MTORC1_SIGNALING	182	0.45	1.75	0	0.003
14	HALLMARK_FATTY_ACID_METABOLISM	140	0.45	1.73	0.001	0.004
15	HALLMARK_DNA_REPAIR	135	0.45	1.71	0.001	0.005
16	HALLMARK_KRAS_SIGNALING_UP	191	0.42	1.69	0	0.006
17	HALLMARK_XENOBIOTIC_METABOLISM	185	0.42	1.67	0.001	0.008
18	HALLMARK_INFLAMMATORY_RESPONSE	185	0.41	1.64	0	0.01
19	HALLMARK_HEME_METABOLISM	183	0.41	1.63	0.001	0.011
20	HALLMARK_TNFA_SIGNALING_VIA_NFKB	185	0.41	1.63	0.002	0.01
21	HALLMARK_G2M_CHECKPOINT	176	0.41	1.62	0.001	0.01
22	HALLMARK_UNFOLDED_PROTEIN_RESPONSE	104	0.42	1.55	0.011	0.018
23	HALLMARK_E2F_TARGETS	167	0.39	1.53	0.003	0.02
24	HALLMARK_COAGULATION	132	0.39	1.51	0.011	0.024
25	HALLMARK_CHOLESTEROL_HOMEOSTASIS	60	0.44	1.51	0.031	0.023
26	HALLMARK_PEROXISOME	94	0.4	1.45	0.041	0.04
27	HALLMARK_TGF_BETA_SIGNALING	49	0.43	1.45	0.043	0.041
28	HALLMARK_P53_PATHWAY	182	0.36	1.4	0.024	0.059
29	HALLMARK_ANDROGEN_RESPONSE	94	0.36	1.34	0.077	0.098
30	HALLMARK_UV_RESPONSE_UP	150	0.34	1.33	0.051	0.099
31	HALLMARK_ESTROGEN_RESPONSE_LATE	193	0.33	1.31	0.061	0.115
32	HALLMARK_PANCREAS_BETA_CELLS	36	0.42	1.3	0.129	0.114
(B)						
1	HALLMARK_INTERFERON_ALPHA_RESPONSE	77	-0.57	-2.86	0	0
2	HALLMARK_INTERFERON_GAMMA_RESPONSE	171	-0.47	-2.68	0	0
3	HALLMARK_TNFA_SIGNALING_VIA_NFKB	185	-0.28	-1.65	0	0.029
4	HALLMARK_IL2_STAT5_SIGNALING	169	-0.28	-1.6	0	0.034
5	HALLMARK_APOPTOSIS	157	-0.28	-1.58	0	0.031

Table 3 continued

	NAME	SIZE	ES	NES	NOM p-val	FDR q-val
6	HALLMARK_DNA_REPAIR	135	-0.28	-1.56	0.008	0.029
7	HALLMARK_P53_PATHWAY	182	-0.25	-1.45	0.005	0.059
8	HALLMARK_ADIPOGENESIS	169	-0.24	-1.39	0.014	0.081
9	HALLMARK_OXIDATIVE_PHOSPHORYLATION	188	-0.24	-1.37	0.02	0.089
10	HALLMARK_UV_RESPONSE_UP	150	-0.23	-1.28	0.059	0.153
11	HALLMARK_COMPLEMENT	181	-0.21	-1.21	0.073	0.239

SIZE size of gene set, ES enriched score, NES normalized enriched score, NOM p-val nominal p-value, FDR q-val, false discovery rate q-value
 Bold font denotes common gene sets enriched in both analyses

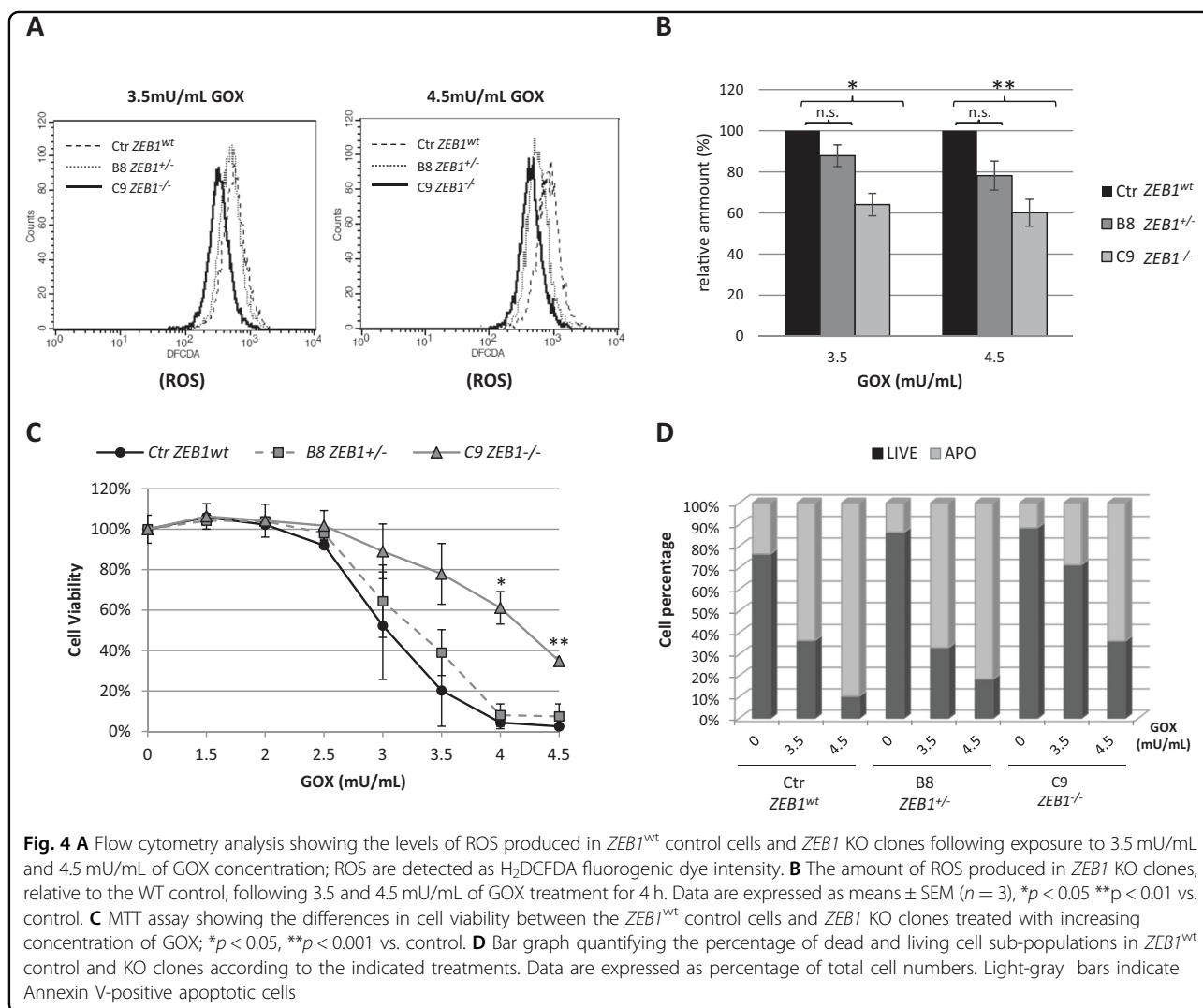


Fig. 4 **A** Flow cytometry analysis showing the levels of ROS produced in *ZEB1*^{wt} control cells and *ZEB1* KO clones following exposure to 3.5 μM/mL and 4.5 μM/mL of GOX concentration; ROS are detected as H₂DCFDA fluorogenic dye intensity. **B** The amount of ROS produced in *ZEB1* KO clones, relative to the WT control, following 3.5 and 4.5 μM/mL of GOX treatment for 4 h. Data are expressed as means ± SEM (n = 3), *p < 0.05 **p < 0.01 vs. control. **C** MTT assay showing the differences in cell viability between the *ZEB1*^{wt} control cells and *ZEB1* KO clones treated with increasing concentration of GOX; *p < 0.05, **p < 0.001 vs. control. **D** Bar graph quantifying the percentage of dead and living cell sub-populations in *ZEB1*^{wt} control and KO clones according to the indicated treatments. Data are expressed as percentage of total cell numbers. Light-gray bars indicate Annexin V-positive apoptotic cells

the chromatin-modifying genes *ARID1A*, implicated in 33–62.5% of CTCLs, and *DNMT3A* involved in 9–42.5% of cases^{6,8,9,11,12}. This great variability in the alteration frequencies might depend on sample sizes/compositions

and/or diversities of technical/analytical approaches. In our cohort, the GISTIC analysis identified 10p11.22 as the region most significantly affected by loss of genetic material and *ZEB1* as the target of this region; also, we

found the presence of nonsense somatic mutations, predicting to generate defective proteins coupled with *ZEB1* deletion. A combined burden of deletions and somatic mutations, affecting *ZEB1*, has been reported with frequencies ranging from 45 to 65% of CTCLs^{2,6,9,10}. Thus, our and previous findings support *ZEB1* as a strong candidate tumor suppressor gene in CTCL as already described in ATLL,²⁰ differently from canonical cancer genes like *TP53*, *PTEN*, *RBI*, and *CDKN2A* recurrently inactivated also in other cancer types, *ZEB1* seems to be a lineage-specific tumor suppressor. Here, we show, for the first time, that SS patients harbor not only *ZEB1* somatic mutations, but also germline variations; we found that 8/11 allelic variants were of germline origin (Table 1 and Table 2). We found substitutions reported as either rare variants (MAF/AF < 5%) or totally absent from dbSNPs or gnomAD, representing, thereby, novel sequence variations (Table 2). Importantly, the Q840P (c.2519 A > C) is a *ZEB1* LoF variation associated with a rare inherited disorder of the corneal endothelium³¹. Interestingly, Hidaka et al.²⁰, reported the identification of *ZEB1* c.232 A > C (N78T) variation in the Hut102 CTCL cell line, which, also, is recurrent in FCD patients and causes a protein impairment³¹. Although the germline origin of the N78T variation cannot be proved for Hut102, it is worth noting that it is absent from the COSMIC database, whereas it is cataloged as a variation with a global MAF < 5% in the dbSNP (rs80194531) and gnomAD, further supporting our findings. Rare genetic variations play an important role in human disorders, as they may cause many Mendelian and rare forms of common diseases⁴⁷. Evolutionary theory predicts that disease alleles are likely to be rare as a result of purifying selection^{47,48} and indeed, LoF variants are especially rare^{49,50}. Also, empirical evidences have shown that low frequency and rare variants are associated with complex diseases, including cancer^{51–53}. Moreover, emerging evidence suggests that synonymous SNPs could affect, besides mRNA splicing, also mRNA stability, protein expression, and enzymatic activity contributing to human disease risk and other complex traits (reviewed in ref. ⁵⁴). Although further observations and experimental validations are necessary to establish the functional consequences of *ZEB1* germline variations in SS, this is the first report to suggest a genetic risk for this rare neoplasm. Overall survival analysis, in our SS series, highlighted an association of poorer outcome with biallelic loss of *ZEB1*, which supports its recessive nature of tumor suppressor and underlines its potential role as a risk-stratification and prognostic marker. In CTCL, *ZEB1* acts as a transcriptional repressor contributing to an oncogenic regulatory mechanism involving IL15, HDAC1, HDAC6, and miR-21¹⁶. Here, we found that *ZEB1* absence is associated with multiple dysregulated pathways (Table 3), among which, the upregulation of IL2 signaling and the oxidative stress

cell response were already shown to depend on *ZEB1* expression. In oxidative stress condition, *ZEB1* down-regulation by miR-200c has been described; we could not confirm this relationship in SS, rather *ZEB1* expression is mainly dependant on the integrity of its genomic locus (Fig. 2A). We verified that *ZEB1*^{-/-} SS cells respond to oxidative stress with significant lower intracellular ROS levels, higher cell viability, and resistance to apoptosis. ROS mediate apoptosis in a number of cell types; in T cells, re-stimulation of previously activated T cells via the TCR increases ROS concentrations leading to AICD, an important regulatory mechanism of T-cell homeostasis⁵⁵. Indeed, CTCLs have been shown to be AICD-resistant, while inhibition of the NF- κ B pathway induces CTCL cell death via free intracellular iron and massive ROS production^{56,57}. NF- κ B exerts a negative control on ROS formation, which is central for the suppression of TNF α -induced apoptosis^{58–61}. Interestingly, we found that upregulation of TNF α signaling via NF- κ B is significantly correlated with *ZEB1* absence, suggesting a control of this pathway in SS disease. Recently, Morel et al. reported that *ZEB1* reduces cellular ROS in mammary stem cells following an oncogenic stimulus, while the expression of several detoxifying enzymes was increased; among them, they identified *MSRB3* as a transcriptional target of *ZEB1*, implying its direct role in the control of ROS production⁶². We did not find *MSRB3*, but *HIF1A*, another antioxidant gene found by Morel et al., positively correlated with *ZEB1* depletion. It would be interesting to further investigate if *ZEB1* is capable, not only to induce ROS-scavenging factors, but also to repress them in a cell-type-specific manner. Although we found molecular pathways of genomic instability associated with *ZEB1* absence, no significant differences emerged in our PGA analysis between *ZEB1*-depleted and *ZEB1*-expressing tumors; this might be due to additional independent factors contributing to DNA damage such as defects in genes deputed to genome maintenance¹¹. Also, it has been shown very recently that *ZEB1* mediates the epigenetic silencing of protective type III IFN response in airway mucosal epithelium⁶³. Here, we found *ZEB1* depletion highly correlated with IFN α and IFN γ response (Table 3). IFNs have been proposed to play an important role in regulating the abundance and persistence of memory T cell;^{64,65} particularly, previous studies showed that IFN α / β inhibited activated T-cell apoptosis directly, without induction of secondary mediators, promoting, among the diverse survival pathways, also the upregulation of intracellular antioxidant molecule glutathione^{66–68}. Although our observations await further confirmations and mechanistic experimental validations, they provide evidence for uncharacterized *ZEB1* roles in regulating T-cell homeostasis, paving the way for future developments and treatment opportunities.

Ethics approval and consent to participate

This study was conducted in accordance with Good Clinical Practice Guidelines and the Declaration of Helsinki and approved by the Ethical Committee of the Istituto Dermatologico dell'Immacolata (ID n. 4/CE/2015). A written consent was obtained from all patients.

Acknowledgements

The authors wish to thank the patients of I Clinical Dermatology Unit of the IDI-IRCCS, G. Alvetreti and E. Ferranti for nursing assistance and collaboration, G. Costanzi for computer technology assistance, and Dr. A. Magenta and Dr. S. Pastore for scientific support and helpful discussion. The authors wish to remember Dr. F. Carlomosti who generously gave initial contribution and prematurely passed away. This study was supported by Associazione Italiana Ricerca sul Cancro IG-15828 (GR), Associazione Italiana Ricerca sul Cancro IG-17748 (MGN), Fondi Ricerca Corrente from Italian Ministry of Health RC2016-17 (EC).

Authors contribution

E.C. conceived and supervised the study, performed sequence, bioinformatics and survival analyses, interpreted and analyzed data, and wrote the paper, G.R. co-designed the study, interpreted data, and revised the paper. A.B. and C.C. collected samples, acquired SNP array data, performed expression and biochemical experiments, cell biology and flow cytometry assays, KO cell lines generation, data acquisition, and interpretation; M.H.C. and V.T. collected samples, processed DNA for SNP array and sequencing, and performed real-time and droplet digital PCR assays; E.S. performed cytofluorimetric analyses of patients samples; A.M. and R.B. recruited eligible patients and collected clinical information; M.G.N. interpreted data and critically revised the paper. All authors read and approved the final paper.

Data availability

The datasets 10 K SNP Arrays (GSE17595) and Gene Expression Arrays (GSE17601) supporting the conclusions of this article are available in GEO public data repository under the super series GSE17602 at the following link: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE17602>, SNP6 Arrays dataset is under ethical evaluation for submission in GEO public data repository.

Conflict of interest

The authors declare that they have no conflict of interest.

Publisher's note

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

Supplementary Information accompanies this paper at (<https://doi.org/10.1038/s41419-018-1212-7>).

Received: 16 August 2018 Revised: 8 October 2018 Accepted: 7 November 2018

Published online: 05 December 2018

References

- Fierro, M. T. et al. Heterogeneity of circulating CD4+memory T-cell subsets in erythrodermic patients: CD27 analysis can help to distinguish cutaneous T-cell lymphomas from inflammatory erythroderma. *Dermatology* **216**, 213–221 (2008).
- Vermeer, M. H. et al. Novel and highly recurrent chromosomal alterations in Sézary syndrome. *Cancer Res.* **68**, 2689–2698 (2008).
- Caprini, E. et al. Identification of key regions and genes important in the pathogenesis of Sézary syndrome by combining genomic and expression microarrays. *Cancer Res.* **69**, 8438–8446 (2009).
- Laharanne, E. et al. Genome-wide analysis of cutaneous T-cell lymphomas identifies three clinically relevant classes. *J. Invest. Dermatol.* (2010). <https://doi.org/10.1038/jid.2010.8>
- Lin, W. M. et al. Characterization of the DNA copy-number genome in the blood of cutaneous T-cell lymphoma patients. *J. Invest. Dermatol.* **132**, 188–197 (2012).
- Choi, J. et al. Genomic landscape of cutaneous T cell lymphoma. *Nat. Genet.* **47**, 1–11 (2015).
- Ungewickell, A. et al. Genomic analysis of mycosis fungoides and Sézary syndrome identifies recurrent alterations in TNFR2. *Nat. Genet.* **47**, 1056–1060 (2015).
- da Silva Almeida, A. C. et al. The mutational landscape of cutaneous T cell lymphoma and Sézary syndrome. *Nat. Genet.* **47**, 1465–1470 (2015).
- Wang, L. et al. Genomic profiling of Sézary syndrome identifies alterations of key T cell signaling and differentiation genes. *Nat. Genet.* **47**, 1426–1434 (2015).
- McGirt, L. Y. et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. *Blood* **126**, 508–519 (2015).
- Woollard, W. J. et al. Candidate driver genes involved in genome maintenance and DNA repair in Sézary syndrome. *Blood* **127**, 3387–3397 (2016).
- Kiel, M. J. et al. Genomic analyses reveal recurrent mutations in epigenetic modifiers and the JAK-STAT pathway in Sézary syndrome. *Nat. Commun.* **6**, 8470 (2015).
- Park, J. et al. Genomic analysis of 220 CTCLs identifies a novel recurrent gain-of-function alteration in RLTPR (p.Q575E). *Blood* **130**, 1430–1440 (2017).
- Williams, T. M. et al. Identification of a zinc finger protein that inhibits IL-2 gene expression. *Sci. (80-)* **254**, 1791–1794 (1991).
- Wang, J. T. et al. The transcription repressor, ZEB1, cooperates with CtBP2 and HDAC1 to suppress IL-2 gene activation in T cells. *Int. Immunol.* **21**, 227–235 (2009).
- Mishra, A. et al. Mechanism, consequences, and therapeutic targeting of abnormal IL15 signaling in cutaneous T-cell lymphoma. *Cancer Discov.* **6**, 986–1005 (2016).
- Higashi, Y. et al. Impairment of T cell development in deltaEF1 mutant mice. *J. Exp. Med.* **185**, 1467–1479 (1997).
- Brabletz, T. et al. Negative regulation of CD4 expression in T cells by the transcriptional repressor ZEB. *Int. Immunol.* **11**, 1701–1708 (1999).
- Caramel, J., Ligier, M. & Puisieux, A. Pleiotropic roles for ZEB1 in cancer. *Cancer Res.* **78**, 30–35 (2018).
- Hidaka, T. et al. Down-regulation of TCF8 is involved in the leukemogenesis of adult T-cell leukemia/lymphoma. *Blood* **112**, 383–393 (2008).
- Nakahata, S., Yamazaki, S., Nakauchi, H. & Morishita, K. Downregulation of ZEB1 and overexpression of Smad7 contribute to resistance to TGF- β 1-mediated growth suppression in adult T-cell leukemia/lymphoma. *Oncogene* **29**, 4157–4169 (2010).
- Olsen, E. A. et al. Clinical end points and response criteria in mycosis fungoides and Sézary syndrome: a consensus statement of the international society for cutaneous lymphomas, the United States cutaneous lymphoma consortium, and the cutaneous lymphoma task force of the E. *J. Clin. Oncol.* **29**, 2598–2607 (2011).
- Lin, M. et al. dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. *Bioinformatics* **20**, 1233–1240 (2004).
- Olshen, A. B., Venkatraman, E. S., Lucito, R. & Wigler, M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* **5**, 557–572 (2004).
- Cristofolletti, C. et al. Comprehensive analysis of PTEN status in Sezary syndrome. *Blood* 3511–3520 (2013).
- Beoukheim, R. et al. Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. *Proc. Natl. Acad. Sci.* **104**, 20007–20012 (2007).
- Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402–408 (2001).
- Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).
- Mermel, C. H. et al. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **12**, R41 (2011).
- Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
- Riazuddin, S. A. et al. Missense mutations in TCF8 cause late-onset fuchs corneal dystrophy and interact with FCD4 on chromosome 9p. *Am. J. Hum. Genet.* **86**, 45–53 (2010).

32. Cartegni, L., Chew, S. L. & Krainer, A. R. Listening to silence and understanding nonsense: Exonic mutations that affect splicing. *Nat. Rev. Genet.* **3**, 285–298 (2002).
33. Tennesen, J. A. et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Sci. (80-)*. **336**, 64–69 (2012).
34. Liberzon, A. et al. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
35. Sors, A. et al. Down-regulating constitutive activation of the NF- κ B canonical pathway overcomes the resistance of cutaneous T-cell lymphoma to apoptosis. *Blood* (2006). <https://doi.org/10.1182/blood-2005-06-2536>
36. Thu, K. L. et al. Lung adenocarcinoma of never smokers and smokers harbor differential regions of genetic alteration and exhibit different levels of genomic instability. *PLoS. One.* **7**, e333003 (2012).
37. Magenta, A. et al. MiR-200c is upregulated by oxidative stress and induces endothelial cell apoptosis and senescence via ZEB1 inhibition. *Cell Death Differ.* **18**, 1628–1639 (2011).
38. Cortez, M. A. et al. Therapeutic delivery of miR-200c enhances radiosensitivity in lung cancer. *Mol. Ther.* **22**, 1494–1503 (2014).
39. Wu, Y., Gu, C. & Huang, X. Sevoflurane protects against hepatic ischemia/reperfusion injury by modulating microRNA-200c regulation in mice. *Biomed. Pharmacother.* **84**, 1126–1136 (2016).
40. D'Agostino, M. et al. Circulating miR-200c is up-regulated in paediatric patients with familial hypercholesterolaemia and correlates with miR-33a/b levels: implication of a ZEB1-dependent mechanism. *Clin. Sci. (Lond.)* **131**, 2397–2408 (2017).
41. Yi, Y. et al. Ultraviolet A irradiation induces senescence in human dermal fibroblasts by down-regulating DNMT1 via ZEB1. *Aging* **10**, 212–228 (2018).
42. Narducci, M. G. et al. MicroRNA profiling reveals that miR-21, miR486 and miR-214 are upregulated and involved in cell survival in Sézary syndrome. *Cell Death Dis.* **2**, e151 (2011).
43. Devadas, S., Zaritskaya, L., Rhee, S. G., Oberley, L. & Williams, M. S. Discrete generation of superoxide and hydrogen peroxide by T cell receptor stimulation. *J. Exp. Med.* **195**, 59–70 (2002).
44. Gulow, K. et al. HIV-1 trans-activator of transcription substitutes for oxidative signaling in activation-induced T cell death. *J. Immunol.* **174**, 5249–5260 (2005).
45. Kaczara, P., Sarna, T. & Burke, J. M. Dynamics of H₂O₂ availability to ARPE-19 cultures in models of oxidative stress. *Free Radic. Biol. Med.* **48**, 1064–1070 (2010).
46. Prasad, A. et al. Identification of gene mutations and fusion genes in patients with Sézary Syndrome. *J. Invest. Dermatol.* (2016). <https://doi.org/10.1016/j.jid.2016.03.024>
47. Gibson, G. Rare and common variants: twenty arguments. *Nat. Rev. Genet.* **13**, 135–145 (2012).
48. Kryukov, G. V., Shpunt, A., Stamatoyannopoulos, J. A. & Sunyaev, S. R. Power of deep, all-exon resequencing for discovery of human trait genes. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 3871–3876 (2009).
49. MacArthur, D. G. et al. A systematic survey of loss-of-function variants in human protein-coding genes. *Science* **335**, 823–828 (2012).
50. 1000 Genomes Project Consortium. et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
51. Rivas, M. A. et al. Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. *Nat. Genet.* **43**, 1066–1073 (2011).
52. Gudmundsson, J. et al. A study based on whole-genome sequencing yields a rare variant at 8q24 associated with prostate cancer. *Nat. Genet.* **44**, 1326–1329 (2012).
53. Jonsson, T. et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature* **488**, 96–99 (2012).
54. Sauna, Z. E. & Kimchi-Sarfaty, C. Understanding the contribution of synonymous mutations to human disease. *Nat. Rev. Genet.* **12**, 683–691 (2011).
55. Hildeman, D., Mitchell, T., Kappler, J. & Marrack, P. T cell apoptosis and reactive oxygen species. *J. Clin. Invest.* **111**, 575–581 (2003).
56. Klemke, C. D. et al. Lack of T-cell receptor-induced signaling is crucial for CD95 ligand up-regulation and protects cutaneous T-cell lymphoma cells from activation-induced cell death. *Cancer Res.* **69**, 4175–4183 (2009).
57. Kiessling, M. K. et al. Inhibition of constitutively activated nuclear factor- κ B induces reactive oxygen species- and iron-dependent cell death in cutaneous T-cell lymphoma. *Cancer Res.* **69**, 2365–2374 (2009).
58. Sakon, S. et al. NF- κ B inhibits TNF-induced accumulation of ROS that mediate prolonged MAPK activation and necrotic cell death. *EMBO J.* **22**, 3898–3909 (2003).
59. Pham, C. G. et al. Ferritin heavy chain upregulation by NF- κ B inhibits TNF α -induced apoptosis by suppressing reactive oxygen species. *Cell* **119**, 529–542 (2004).
60. Kamata, H. et al. Reactive oxygen species promote TNF α -induced death and sustained JNK activation by inhibiting MAP kinase phosphatases. *Cell* **120**, 649–661 (2005).
61. Papa, S. et al. The NF- κ B-mediated control of the JNK cascade in the antagonism of programmed cell death in health and disease. *Cell Death Differ.* **13**, 712–729 (2006).
62. Morel, A. P. et al. A stemness-related ZEB1-MSRB3 axis governs cellular plasticity and breast cancer genome stability. *Nat. Med.* **23**, 568–578 (2017).
63. Yang, J., Tian, B., Sun, H., Garofalo, R. P. & Brasier, A. R. Epigenetic silencing of IRF1 dysregulates type III interferon responses to respiratory virus infection in epithelial to mesenchymal transition. *Nat. Microbiol.* **2**, 17086 (2017).
64. Akbar, A. N., Lord, J. M. & Salmon, M. IFN- α and IFN- β : a link between immune memory and chronic inflammation. *Immunol. Today* **21**, 337–342 (2000).
65. Whitmire, J. K., Eam, B., Benning, N. & Whitton, J. L. Direct interferon-gamma signaling dramatically enhances CD4+ and CD8+ T cell memory. *J. Immunol.* **179**, 1190–1197 (2007).
66. Marrack, P., Kappler, J. & Mitchell, T. Type I interferons keep activated T cells alive. *J. Exp. Med.* **189**, 521–529 (1999).
67. Pilling, D. et al. Interferon-beta mediates stromal cell rescue of T cells from apoptosis. *Eur. J. Immunol.* **29**, 1041–1050 (1999).
68. Hyde, H., Borthwick, N. J., Janossy, G., Salmon, M. & Akbar, A. N. Upregulation of intracellular glutathione by fibroblast-derived factor(s): enhanced survival of activated T cells in the presence of low Bcl-2. *Blood* **89**, 2453–2460 (1997).