

# STAT3 Targets Suggest Mechanisms of Aggressive Tumorigenesis in Diffuse Large B-Cell Lymphoma

Jennifer Hardee,<sup>\*,1</sup> Zhengqing Ouyang,<sup>\*,1,2,3</sup> Yuping Zhang,<sup>\*,4</sup> Anshul Kundaje,<sup>\*,5</sup> Philippe Lacroute,<sup>\*</sup> and Michael Snyder<sup>\*,5</sup>

<sup>\*</sup>Department of Genetics, Stanford University School of Medicine, Stanford, California 94305, <sup>†</sup>Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, Connecticut 06520, and <sup>‡</sup>Department of Computer Science, Stanford University School of Engineering, Stanford, California 94305

**ABSTRACT** The signal transducer and activator of transcription 3 (STAT3) is a transcription factor that, when dysregulated, becomes a powerful oncogene found in many human cancers, including diffuse large B-cell lymphoma. Diffuse large B-cell lymphoma is the most common form of non-Hodgkin's lymphoma and has two major subtypes: germinal center B-cell–like and activated B-cell–like. Compared with the germinal center B-cell–like form, activated B-cell–like lymphomas respond much more poorly to current therapies and often exhibit overexpression or overactivation of STAT3. To investigate how STAT3 might contribute to this aggressive phenotype, we have integrated genome-wide studies of STAT3 DNA binding using chromatin immunoprecipitation-sequencing with whole-transcriptome profiling using RNA-sequencing. STAT3 binding sites are present near almost a third of all genes that differ in expression between the two subtypes, and examination of the affected genes identified previously undetected and clinically significant pathways downstream of STAT3 that drive oncogenesis. Novel treatments aimed at these pathways may increase the survivability of activated B-cell–like diffuse large B-cell lymphoma.

## KEYWORDS

genomics  
next-generation  
sequencing  
cancer signaling  
signal  
transduction  
oncogenic  
pathways

The transcription factor signal transducer and activator of transcription 3 (STAT3) is a key signal transducer that regulates gene expres-

sion in response to interleukins and cytokines, as well as a wide variety of other extracellular signals (Zhong *et al.* 1994; Baker *et al.* 2007; Minegishi 2009). The binding of one of these messengers to its receptor launches a tyrosine phosphorylation cascade that results in the cytosolic activation and dimerization of STAT3, which is then imported to the nucleus, where it binds its target sequences. STAT3 mediates the expression of a large number of genes and plays a key role in many cellular processes, especially those related to cell growth and apoptosis (Baker *et al.* 2007).

As the result of these proliferative and antiapoptotic effects, STAT3 is also a powerful oncogene (Alvarez and Frank 2004). Constitutively active STAT3, caused by upstream dysregulation, is found in a large number of human cancers and is generally associated with a poorer prognosis (Benekli *et al.* 2003; Turkson 2004; Hodge *et al.* 2005). In particular, overactive STAT3 is frequently found in diffuse large B-cell lymphoma (DLBCL) and is associated with poorer outcomes (Ding *et al.* 2008; Wu *et al.* 2011). DLBCL is the most common form of lymphoma and comprises at least two subtypes: germinal center B-cell-like (GCB) and activated B-cell-like (ABC) (Alizadeh *et al.* 2000; Rosenwald *et al.* 2002; Wright *et al.* 2003; American Cancer Society 2012). These two subtypes have significant differences in three-year survival, which is nearly 85% for GCB but only 65–70% for patients with ABC (Fu *et al.* 2008; Lenz *et al.* 2008). High levels of STAT3 generally are found only in the activated B-cell–like subtype.

Copyright © 2013 Hardee *et al.*

doi: 10.1534/g3.113.007674

Manuscript received July 15, 2013; accepted for publication October 5, 2013; published Early Online October 18, 2013.

This is an open-access article distributed under the terms of the Creative Commons Attribution Unported License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supporting information is available online at <http://www.g3journal.org/lookup/suppl/doi:10.1534/g3.113.007674/-/DC1>

<sup>1</sup>Present address: The Jackson Laboratory for Genomic Medicine, Farmington, CT 06030.

<sup>2</sup>Present address: Department of Biomedical Engineering, University of Connecticut, Storrs, CT 06269.

<sup>3</sup>Present address: Department of Genetics and Developmental Biology, University of Connecticut Health Center, Farmington, CT 06030.

<sup>4</sup>Present address: Department of Biostatistics, Yale School of Public Health, Yale University, New Haven, CT 06520.

<sup>5</sup>Corresponding author: Department of Genetics, MC: 5120, 300 Pasteur Drive, M-344, Stanford, CA 94305. E-mail: [mjsnyder@stanford.edu](mailto:mjsnyder@stanford.edu)

All ChIP-Seq and RNA-Seq data from this article have been deposited with the NCBI GEO public database under accession no. GSE50724. Within this accession, ChIP-Seq data are found in the subseries GSE50723, and RNA-Seq data are in the subseries GSE50721. Individual data files are assigned the consecutive accession nos. GSM1227193–GSM1227214.

In the present study, we sought to further understand the difference in STAT3 function between these two subtypes through mapping its binding regions (BRs) and analyzing gene expression in GCB and ABC patient tumor-derived cell lines. We performed ChIP-Seq (chromatin immunoprecipitation followed by DNA sequencing) experiments to map STAT3 binding sites and RNA-Seq to analyze the global gene expression patterns. We then synthesized these data to determine which genetic loci show both differential STAT3 binding and differential mRNA expression. We found that STAT3 likely up-regulates a number of oncogenic pathways to promote aggressive tumor growth and migration.

## MATERIALS AND METHODS

Cell lines were grown at 37° and 5% CO<sub>2</sub>. SU-DHL2, SU-DHL4, SU-DHL6, SU-DHL10, OCI-Ly7, and U-2932 were grown in RPMI 1640 media supplemented with 15% FBS and antibiotics. OCI-Ly3 and OCI-Ly10 were grown in IMDM media supplemented with 15% fetal bovine serum, antibiotics, and 55 μM beta-mercaptoethanol. Western blots were performed on whole-cell lysate, with equal protein loading in each lane, with the use of anti-STAT3 rabbit polyclonal antibody sc-482X (Santa Cruz Biotechnology, Inc.); anti-pSTAT3-Y705 mouse monoclonal antibody sc-8059X (Santa Cruz Biotechnology, Inc.); and anti-GAPDH mouse monoclonal antibody ab8245 (Abcam). ChIP-sequencing was performed with the anti-STAT3 antibody sc-482X on formaldehyde-crosslinked pellets of 1 × 10<sup>6</sup> cells. DNA was mechanically sheared using a Branson sonicator, then immunoprecipitated for 16 hr. Bound DNA was recovered on protein A-agarose beads and purified via ethanol precipitation. mRNA for RNA-sequencing was isolated directly from whole cell lysate using magnetic poly-dT beads (Dynabeads mRNA DIRECT Kit; Invitrogen), then chemically fragmented (RNA Fragmentation Reagents; Ambion). cDNA was synthesized using random hexamer primers. For library preparation, standard Illumina GA-IIx primers were ligated and gel purification was used to size-select DNA in the 150- to 300-bp range. Single-ended 36-bp reads were generated for both ChIP- and RNA-sequencing runs.

## Statistical analysis

Sequencing results were mapped to the human genome (hg19) using Bowtie (Langmead *et al.* 2009). STAT3 ChIP-sequencing peaks were compared with a non-IP'd genomic DNA control, and identified using the SPP peak caller (Kharchenko *et al.* 2008). Replicates were analyzed using irreproducible discovery rate analysis to identify strong, repeatable peaks for each cell line (Li *et al.* 2011). These lists were combined and any overlapping or abutting peaks were merged into broader BRs (Kasowski *et al.* 2010). BRs that occurred in only one cell line were eliminated from further analysis. The ChIP-sequencing data for each line were rescored to determine how many fragments mapped to each BR. RNA-sequencing mapped reads were intersected with the RefSeq database to generate read counts for each annotated gene. Both data sets were then normalized to library size with DESeq, and to contrast the two subtypes, the DESeq parameterized negative binomial model was used to obtain the *P*-values for each BR or gene (Anders and Huber 2010). To correct for multiple comparisons, *P*-values were adjusted with the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995). ChIP-Seq BRs were associated with genes that they might regulate via the Genomic Regions Enrichment of Annotations Tool using its default association settings (GREAT; McLean *et al.* 2010). This list of genes associated with significant changes in STAT3 binding was compared with the list of genes with significant expression changes using the online tool Galaxy (Giardine *et al.* 2005;

Blankenberg *et al.* 2010; Goecks *et al.* 2010). Extended materials and methods are available as Supporting Information, File S1, complete data tables are available as File S2. All ChIP-Seq and RNA-Seq data are available in the NCBI GEO public database as series accession no. GSE50724. Within this accession, ChIP-Seq data are found in the subseries GSE50723, and RNA-Seq data are in the subseries GSE50721. Individual data files are assigned the consecutive accession numbers GSM1227193-GSM1227214.

## RESULTS AND DISCUSSION

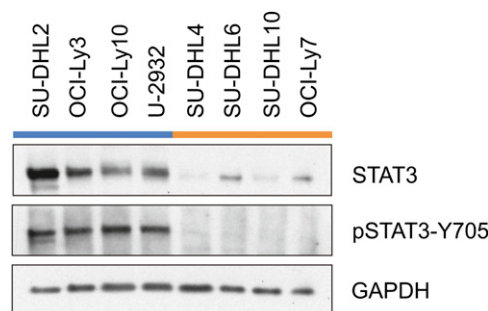
### STAT3 is overexpressed and overactive in ABC DLBCL

In normal B-cells, STAT3 is expressed at low endogenous levels in the cytosol. High levels of STAT3 expression and activation frequently are found in ABC DLBCL, suggesting that it plays a major role in the oncogenesis of this subtype (Ding *et al.* 2008; Wu *et al.* 2011). To investigate the levels of activation of STAT3 in the treatment-responsive GCB lymphoma relative to ABC lymphoma, we selected eight patient tumor-derived cell lines for study and examined levels of total and activated STAT3 by using immunoblot analysis. Four of these cell lines have been classified by Drexler (2010) as GCB (SU-DHL4, SU-DHL6, SU-DHL10, OCI-Ly7) and four as ABC (SU-DHL2, OCI-Ly3, OCI-Ly10, U-2932).

Whole-cell lysates from each line were prepared and probed with antibodies that recognize total STAT3 as well as those that recognize only the phosphorylated (Tyr705) activated form (Shuai *et al.* 1993; Wen *et al.* 1995; Zhang *et al.* 1995). STAT3 protein is present in all of the GCB and ABC cell lines; however, it is at much greater levels (7.2-fold) in all four ABC cell lines (Figure 1). Furthermore, the phosphorylated form of the protein exhibits an even larger difference (56-fold) between the GCB and ABC lines (Figure 1).

### STAT3 shows differential binding in ABC relative to GCB cells

To better understand how STAT3 induces aggressive behavior in ABC cells, we next sought to determine the chromosomal targets of STAT3 in all eight cell lines. We performed ChIP-Seq by using an antibody prepared against the protein's C-terminal region, the area most divergent among STAT family proteins (Baker *et al.* 2007). Immunoblot analyses indicate that this antibody recognizes a single band of the expected molecular mass, which was confirmed as STAT3 by immunoprecipitation followed by mass spectrometry (see Figure S1). We



**Figure 1** Western blot analysis of STAT3 in ABC and GCB cell lines. Western immunoblots were performed on whole-cell lysate from DLBCL cell lines probed with antibodies against STAT3, tyrosine-phosphorylated STAT3, and glyceraldehyde 3-phosphate dehydrogenase (GAPDH). Fold change was measured by densitometric quantification using NIH ImageJ software, normalized to a GAPDH loading control.

performed ChIP-Seq for all eight cell lines along with genomic DNA input controls. Two to nine biological replicates were carried out for each line; replicates were sequenced to an average depth of 18 million mapped reads (87% mapping; see Table S1). All replicates were used in the analysis.

The ChIP-Seq binding peaks were scored using the SPP peak caller and irreproducible discovery rate analysis (Kharchenko *et al.* 2008; Li *et al.* 2011), which resulted in the identification of 1000–6500 strongly reproducible peaks per cell line (Table 1). Overlapping or abutting peaks were merged into BRs to facilitate comparison between cell lines, yielding a total of 10,337 distinct STAT3 BRs that are occupied in at least two of the eight cell lines. These BRs occur throughout the genome, with no obvious hot spots.

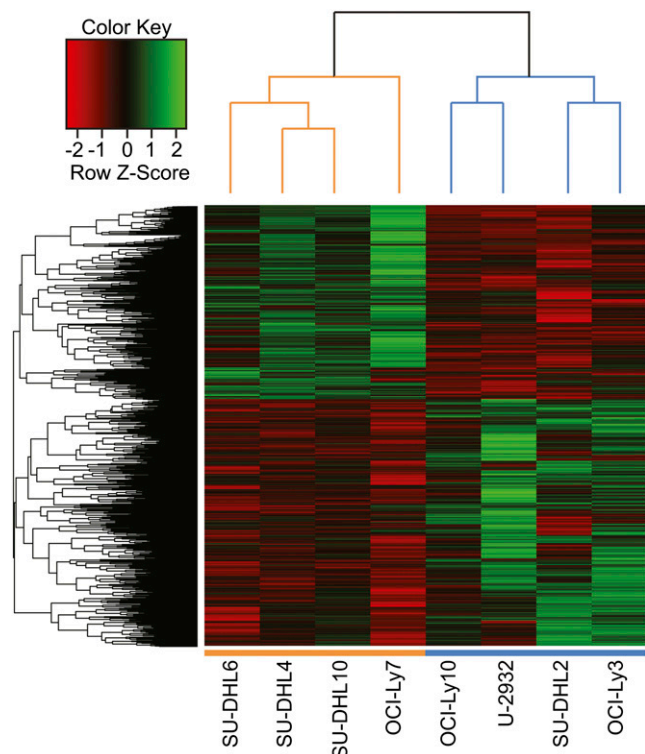
To identify target regions differentially bound by STAT3 in the ABC and GCB subtypes, we used the parameterized negative binomial model in DESeq (Anders and Huber 2010) to identify those BRs that displayed significantly different normalized fragment counts between the ABC and GCB cell lines, then corrected for multiple comparisons using the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995). Of the 10,337 high-confidence BRs, one third ( $n = 3524$ ) are differentially bound by STAT3 between the two subtypes at a false-discovery rate (FDR) < 0.05. (For a complete list of peaks by enriched subtype, see Table S4.) When these differentially bound peaks are clustered, the DLBCL cell lines cluster according to their GCB-ABC subtype (Figure 2). Consistently, more BRs are strongly bound in ABC than in GCB, although BRs with increased STAT3 binding occur surprisingly frequently in the GCB cell lines given their low level of STAT3 expression: 44% of differentially bound BRs ( $n = 1550$ ) show more STAT3 binding in GCB, whereas 56% ( $n = 1974$ ) are more strongly bound in ABC (Figure 3A). STAT3 protein levels are very low in GCB cell lines, so one possible explanation for its equivalent number of binding sites is simply that they represent lower occupancy regions. Indeed, we noticed the signal strength for the GCB lines was lower than that of the ABC cell lines and required additional ChIP-Seq replicate experiments to identify peaks.

STAT3 has a binding peak 14 bp upstream of the transcription start site (TSS) of *JAK3*, and this binding is much stronger (fold change = 12.47, FDR =  $5.21 \times 10^{-27}$ ) in the ABC subtype (Figure 3B). *JAK3* associates with the interleukin (IL)-2 family  $\gamma$ -chain receptors, transducing the signal from cytokines such as IL-2, IL-7, IL-9, IL-15, and IL-21 (Suzuki *et al.* 2000). Although *JAK2* is often regarded as the primary partner of STAT3, constitutively active *JAK3:STAT3* signaling has been reported as a poor prognosis indicator in anaplastic large-cell lymphoma and in colon cancer cell lines (Lin *et al.* 2005; Han *et al.* 2006). Our results suggest that STAT3 activates key regulators of cytokine signaling.

■ Table 1 STAT3 peaks identified by IDR analysis by cell line

Name	Peaks by IDR
ABC cell lines	
SU-DHL2	5375
OCI-Ly3	2885
OCI-Ly10	3258
U-2932	5180
GCB cell lines	
SU-DHL4	6535
SU-DHL6	4570
SU-DHL10	1067
OCI-Ly7	6262

STAT3, signal transducer and activator of transcription 3; IDR, irreproducible discovery rate; ABC, activated B-cell-like; GCB, germinal center B-cell-like.



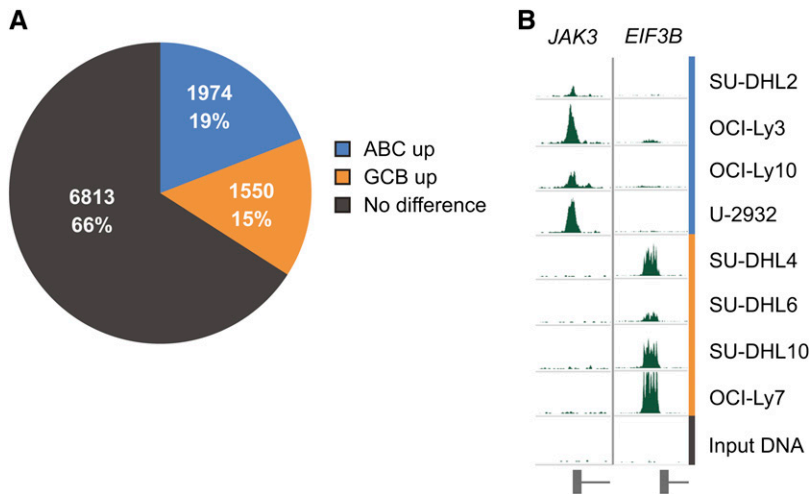
**Figure 2** Hierarchical clustering analysis of STAT3 BRs. Clustering was performed on the 3524 regions that are differentially bound by STAT3 at FDR < 0.05 in ABC cell lines vs. GCB cell lines. Green and red blocks, respectively, represent high and low number of fragments sequenced in that BR relative to the average, whereas black blocks indicate no difference in expression. Fragment counts were log transformed to approximate a normal distribution, standardized by cell line, and each BR was normalized with sum of squares set to unity. GCB cell lines are represented by orange bars and ABC cell lines by blue bars. Generated using the heatmap.2 function in the R package gplots (R Development Core Team 2008; Warnes *et al.* 2011).

Conversely, *EIF3B* shows increased STAT3 binding in the GCB subtype cell lines (fold change = 0.01, FDR =  $2.21 \times 10^{-3}$ ; Figure 3B). As with *JAK3*, the STAT3 binding peak is just upstream of the TSS, a common location for gene regulation by transcription factors. *EIF3B* is a major scaffolding subunit of the eukaryotic translation initiation factor 3 (eIF-3) complex, a complex required for the initiation of protein synthesis. Intriguingly, short hairpin RNA knockdown of eIF-3B induces apoptosis in glioblastoma (Liang *et al.* 2012), suggesting that eIF-3B may also play a role in GCB DLBCL.

**Systematic analyses reveal that STAT3 binding associates with oncogenic pathways in the ABC subtype:**

After the BRs with significant binding differences were identified, we associated them with nearby genes and then analyzed enrichment across a variety of ontologies using GREAT (McLean *et al.* 2010). All but 19 BRs fall within the putative regulatory domain of at least one gene, and the majority are associated with one ( $n = 4203$ ) or two ( $n = 6095$ ) nearby genes.

The Gene Ontology Biological Processes (GO BP) ontology shows that STAT3 BRs occupied in ABC DLBCL are located near genes involved in B-cell activation and mature B-cell differentiation (Figure 4A). This result is consistent with STAT3's role in B-cell maturation and with the theory that more mature B-cells give rise



**Figure 3** Differential STAT3 BRs. (A) STAT3 BRs divided by occupancy pattern. (B) Examples of differentially occupied STAT3 BRs. *JAK3* shows strong binding in the ABC cell lines: SU-DHL2, OCI-Ly3, OCI-Ly10, and U-2932. *EIF3B* shows strong binding in the GCB cell lines: SU-DHL4, SU-DHL6, SU-DHL10, and OCI-Ly7.

to the ABC subtype (Angelin-Duclos *et al.* 2000; Falini *et al.* 2000; Fornek *et al.* 2006; Avery *et al.* 2010). The Disease Ontology analysis reveals that the 1974 ABC-enriched STAT3 BRs fall near genes associated with many forms of cancer, especially lymphomas and leukemias (Figure 4C).

Although there are nearly as many STAT3 BRs with strong binding in the GCB cell lines as in the ABC lines, these BRs and their associated genes do not point to a clear oncogenic program like the ABC BRs do. Far fewer categories are enriched, and at lower significance levels; furthermore, the 1550 GCB-specific BRs are not enriched for any disease categories. The enriched PANTHER Pathway category “apoptosis signaling pathway” primarily contains genes that induce apoptosis for the GCB-bound BRs, but the ABC subtype list includes genes that both up- and down-regulate apoptosis (Figure 4B). Taken together, these data show that in the ABC subtype, overactive STAT3 substantially changes its binding behavior and promotes a genetic program that blocks apoptotic signaling.

**STAT3 binding sites in DLBCL are seen in other hematopoietic and cancer cell lines:** We compared the overall list of STAT3 BRs found in DLBCL with BRs identified in three other cell lines: GM12878, an EBV-transformed B lymphocyte line; HeLa S3, a suspension-tolerant clonal derivative of the HeLa cervical adenocarcinoma line; and LoVo, a colorectal adenocarcinoma cell line. GM12878 and HeLa S3 experiments were performed as part of the ENCODE project (The ENCODE Project Consortium *et al.* 2011) as GEO nos. GSM935557 and GSM935276, whereas the LoVo experiments were performed by Yan *et al.* (2013) as GEO no. GSM1208799.

On the basis of a 1-bp overlap, 3932 (38%) of the 10,337 STAT3 BRs observed in our DLBCL cell lines are also found in GM12878 cells, whereas 1143 (11%) are found in LoVo and 1063 (10%) in HeLa S3 (see Figure S2). It is not surprising that GM12878 has the greatest concordance with DLBCL, as it is the most similar cell type. DLBCL and GM12878 share 2958 BRs not found in either HeLa S3 or LoVo. As determined using GREAT (McLean *et al.* 2010), these BRs are found near genes enriched for type I interferon and cytokine signaling, many B-cell properties including activation and differentiation, DNA damage response, and the JAK/STAT pathway itself (including *JAK1* and *JAK3*; *STAT1*, 3, 4, 5A, and 5B; and the down-regulators *SOC31*, *PIAS2*, and *PTPRC*). More than one-third (36%,  $n = 1406$ ) of these regions are subtype-specific in DLBCL, and three quarters of those

(74%,  $n = 1049$ ) are more strongly bound in the ABC subtype. This bias toward ABC regions may be due to the presence of the Epstein-Barr virus used to transform GM12878; EBV is a well-known oncovirus that promotes the development of hematopoietic cancers (Pallesen *et al.* 1991; Yu *et al.* 1996; Siu *et al.* 2002; Brady *et al.* 2007; Castillo *et al.* 2008) and whose presence contributes to poorer survival in DLBCL specifically (Park *et al.* 2007).

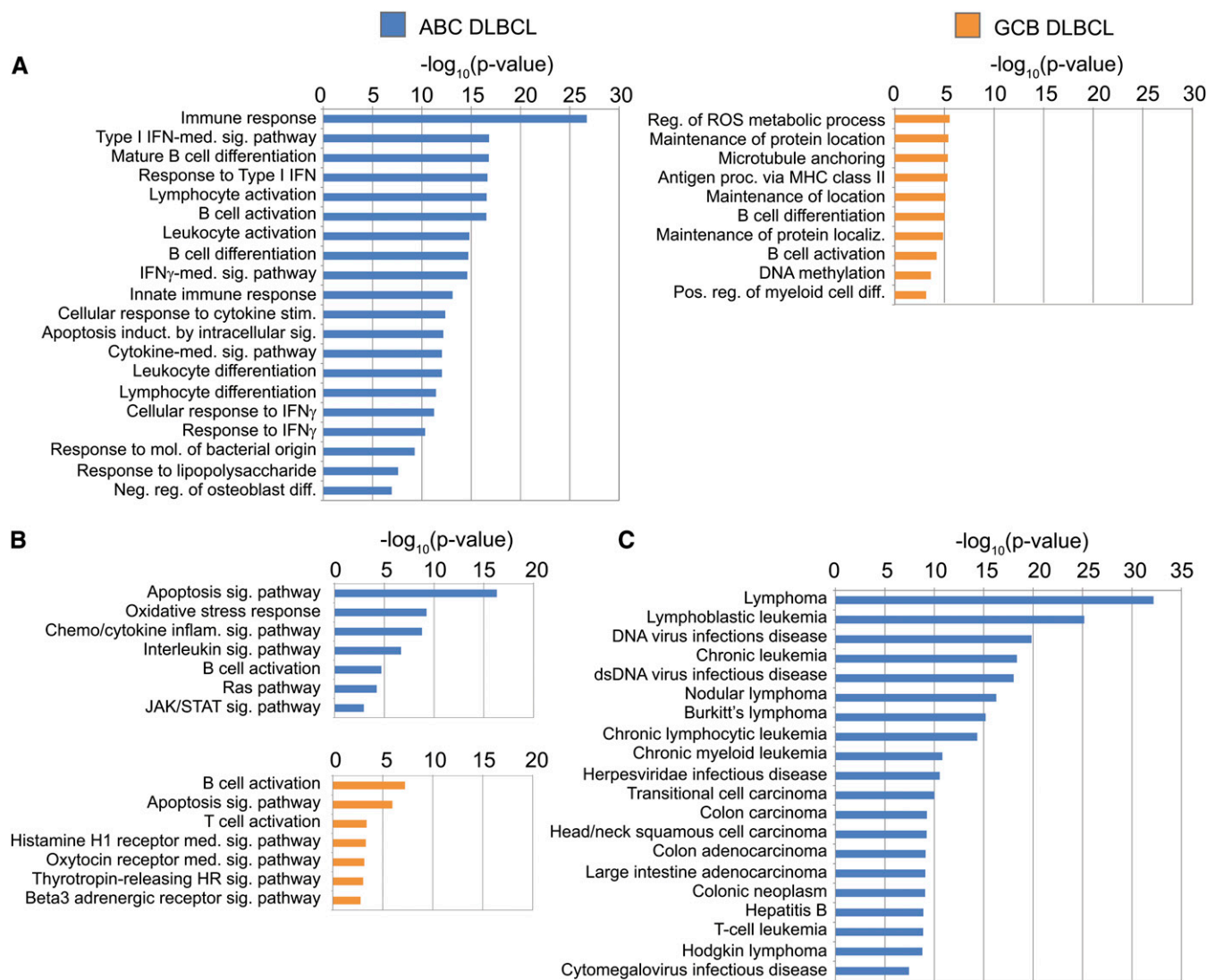
The 1063 STAT3 binding sites shared between DLBCL and HeLa S3 are enriched for genes involved in a number of cancer-related functions: chromatin structure, single-stranded DNA binding, DNA damage checkpoints, abnormal interferon levels, and the apoptotic signaling pathway. One third ( $n = 353$ ) of these regions are subtype-specific in DLBCL, and the vast majority (90%,  $n = 316$ ) are more strongly bound in the ABC subtype. The ABC-specific BRs in HeLa S3 and GM12878 fall near genes that enrich many of the same ontological categories, including immune response, the JAK/STAT pathway, interferon signaling and response, and cytokine signaling and response. They also share Disease Ontology categories that indicate immune system activation, such as leukemias of various types, allergies, hepatitis, and response to infection. In contrast, the 1143 shared LoVo BRs are primarily enriched for genes involved in RNA processing, splicing, and translation, rather than cancer-related categories. Bias toward ABC-specific BRs is not seen in the LoVo colon carcinoma line, where BRs specific to each DLBCL subtype are equally represented. Only 143 STAT3 BRs are found in all four cell types. These are generally not near genes enriched for any pathways or ontological categories. However, there is overlap with a curated gene sets for genes that are targets of *c-Myc*.

### Subtypes show differential gene expression

To better understand the biology of the GCB and ABC subtypes and to relate their gene expression profiles to identified STAT3 targets, we examined global mRNA expression in both subtypes using RNA-Seq analysis (Morin *et al.* 2008). RNA-Seq was performed on mRNA isolated from all eight cell lines. Two biological replicates per cell line were analyzed with an average of 22 million mapped reads (86% mapping; see Table S2).

To identify genes that are differentially expressed in each subtype, we used the parameterized negative binomial model in DESeq (Anders and Huber 2010) to identify annotated RefSeq genes that displayed significantly different normalized read counts between the





**Figure 4** ChIP-Seq BR enrichment for ontological categories. Scale is the  $-\log_{10}(P\text{-value})$  of the binomial  $P$ -value, equivalent to the exponent. (A) All enriched GO Biological Processes,  $P < 0.05$ . (B) All enriched PANTHER Pathways,  $P < 0.05$ . (C) All enriched Disease Ontology listings,  $P < 0.05$ . Data shown for ABC only, as GCB had no significantly enriched categories.

ABC and GCB cell lines, then corrected for multiple comparisons by using the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995). We found many genes that exhibit differential gene expression in the GCB and ABC cell types, including several genes that play a role in apoptosis and cancer. This analysis also successfully re-identified many genes that have previously been shown to be biomarkers for each DLBCL subtype (Rosenwald *et al.* 2002; Tirado *et al.* 2012), including *BCL2*, *IRF4*, *MME (CD10)*, and *BCL6* (Table 2). Surprisingly, the established ABC subtype biomarker *MUM1* (Hans *et al.* 2004; Sweetenham 2005; Anderson *et al.* 2009) showed no significant expression change (FDR = 0.92) but did have a STAT3 BR 344 bp upstream of its TSS with significantly more binding in GCB (fold change = 0.46, FDR =  $4.89 \times 10^{-3}$ ).

A total of 1545 genes are differentially expressed between subtypes at an FDR < 0.05 (Figure 5), which is approximately 7% of the transcriptome. The vast majority of these genes (81%,  $n = 1251$ ) are more highly expressed in the ABC cell lines. In contrast, only 294 genes (19%) are more highly expressed in the GCB cell lines. Interestingly, half ( $n = 765$ ) of the genes with greater ABC subtype expres-

sion demonstrate very low read counts (<5) in the GCB cell types. Conversely, only 21 genes that are more highly expressed in GCB are unique to that subtype (Figure 6A). The prevalence of such “on/off” genes indicates that the major differences between ABC and GCB DLBCL are due almost exclusively to additional gene expression in ABC, rather than the two subtypes having divergent but equally active genetic programs.

**Genes up-regulated in ABC subtype:** Genes with GO BP categories related to cell adhesion, motility, and chemotaxis are enriched in ABC DLBCL cells relative to GCB cells (Figure 7). The expression of these pathways would prime tumors to spread rapidly from their lymph nodes of origin, and may partially explain why ABC lymphomas are so much more aggressive than their GCB counterparts.

The underlying gene list includes a large number of chemokine ligands and receptors, including several with known roles in cancer (Table 3). *CCR7* shows expression limited to the ABC subtype (Figure 6B). Increased *CCR7* expression is clearly associated with metastasis to

■ **Table 2 Validated DLBCL subtype signature genes**

Gene Name	Gene Description	GCB Read Counts	ABC Read Counts	Fold Change	FDR
<b>ABC signature</b>					
<i>BCL2</i>	B-cell CLL/lymphoma 2	4,599.52	22,162.49	4.82	4.90E-05
<i>CCND2</i>	Cyclin D2	31.20	3692.67	118.60	1.56E-18
<i>CFLAR</i>	CASP8 and FADD-like apoptosis regulator	460.63	6463.37	14.03	5.52E-09
<i>FOXP1</i>	forkhead box P1	1202.37	6588.56	5.46	1.13E-04
<i>IRF4</i>	Interferon regulatory factor 4	976.01	21,475.28	22.01	2.43E-14
<i>NFKBIZ</i>	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta	962.34	8119.55	8.46	5.63E-06
<b>GCB signature</b>					
<i>BCL6</i>	B-cell CLL/lymphoma 6	15,105.68	5124.25	0.34	1.94E-02
<i>MAPK10</i>	Mitogen-activated protein kinase 10	1473.10	57.42	0.04	4.45E-05
<i>MME (CD10)</i>	Membrane metallo-endopeptidase	6163.64	663.16	0.11	6.01E-05
<i>MYBL1</i>	v-myb myeloblastosis viral oncogene homolog (avian)-like 1	2932.21	661.39	0.23	4.99E-02

DLBCL, diffuse large B-cell lymphoma; GCB, germinal center B-cell-like; ABC, activated B-cell-like; FDR, false-discovery rate; CLL, chronic lymphocytic leukemia.

the lymph nodes in a broad variety of cancers (Schimanski *et al.* 2006; Kodama *et al.* 2007; Nakata *et al.* 2008; Pan *et al.* 2008; Arigami *et al.* 2009; Ishida *et al.* 2009). Furthermore, *CCR7* and *CXCR5* contribute to apoptosis resistance in B-cell–derived acute and chronic lymphocytic leukemia (Hu *et al.* 2004; Burkle *et al.* 2007).

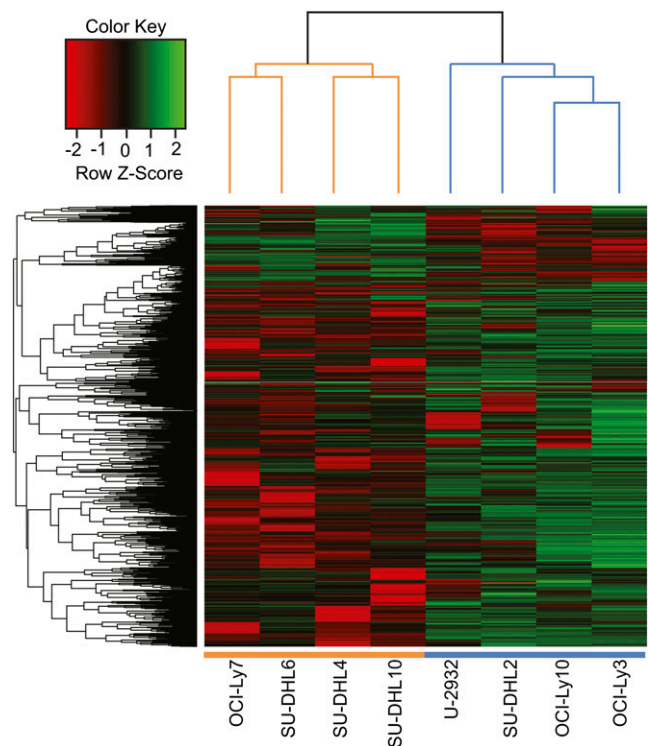
Also up-regulated in the ABC subtype is transforming growth factor alpha (*TGFA*), which is a ligand for the epidermal growth factor receptor. Epidermal growth factor receptor in turn activates a signaling pathway for cell proliferation, differentiation, and development that significantly affects the progression of many human cancers, including brain, lung, breast, ovarian, prostate, and pancreatic tumors (Lu *et al.* 2001; Yang *et al.* 2011).

Genes in GO BP categories related to angiogenesis and hypoxia are also substantially enriched in the ABC cell lines (Figure 7). Angiogenesis is known to be important and prognostic in malignant lymphomas (Koster and Raemaekers 2005). The factors up-regulated in ABC DLBCL include *FGF2*, a major lymphangiogenic growth factor that promotes tumor growth and metastasis (Cao *et al.* 2008; Da *et al.* 2008). A number of growth stimulation factors and receptors are also up-regulated, such as platelet-derived growth factor alpha (*PDGFA*) and the beta subunit of the PDGF receptor (*PDGFRB*), both of which play essential roles in blood vessel development. Hypoxia induced genes include *ARNT2*, which has been reported to promote *VEGF* expression in teratomas and to protect neuronal cells from hypoxia (Maltepe *et al.* 2000). *ARNT2* expression has not been reported previously as a mechanism in non-Hodgkin lymphoma. Thus, these analyses suggest new molecules that may be involved in ABC tumorigenesis.

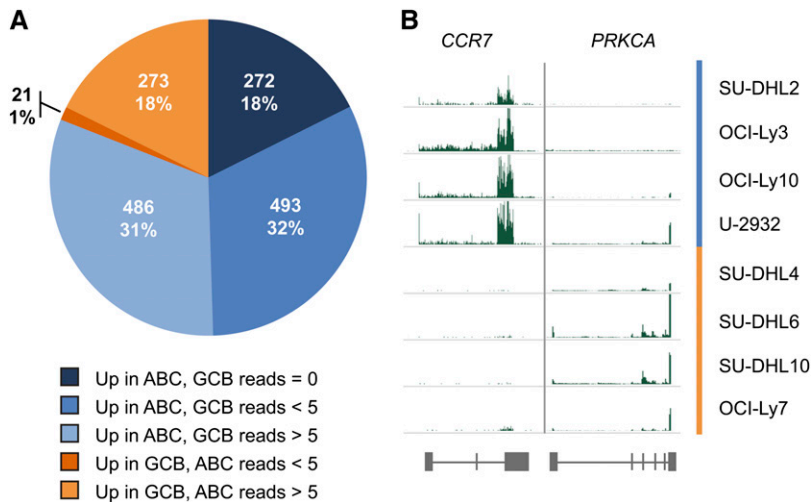
**Genes up-regulated in GCB subtype:** GO analysis of the differentially expressed genes reveals that the GCB-enriched transcripts are involved in activation of lymphocytes (Figure 7). This list includes *BCL6*, a major biomarker for the GCB DLBCL subtype. Another notable GCB up-regulated gene is protein kinase C alpha ( $PKC\alpha$ ), which is encoded by *PRKCA* (Figure 6B).  $PKC\alpha$  is a protein kinase that can act as an oncogene (Besson and Yong 2000; Huttemann *et al.* 2012). However, in response to inflammatory signaling,  $PKC\alpha$  also can negatively regulate NF- $\kappa$ B–induced genes by promoting the synthesis of the inhibitor  $I\kappa$ B $\alpha$  (Han *et al.* 1999). Constitutive NF- $\kappa$ B signaling is a prominent feature of ABC DLBCL and includes crosstalk with STAT3 via up-regulation of IL-6 and IL-10 (Motoyama *et al.* 2005). The high  $PKC\alpha$  and low NF- $\kappa$ B levels in our GCB cell lines suggests that its predominant role in DLBCL may not be as an oncogene but rather as an inhibitor of the NF- $\kappa$ B pathway.

### STAT3 induces an aggressive genetic program in DLBCL

To better understand the downstream consequences of STAT3 binding, we next integrated the binding and gene expression results. Genes associated with STAT3 BRs were correlated with their



**Figure 5** Hierarchical clustering analysis of gene expression. Clustering was performed on the 1545 genes that are differentially expressed at FDR < 0.05 in ABC cell lines vs. GCB cell lines. Green and red blocks, respectively, represent high and low signals in the BR relative to the average, whereas black blocks indicate no difference in expression. Read counts were log transformed to approximate a normal distribution, standardized by cell line, and each gene was normalized with sum of squares set to unity. GCB cell lines are represented by orange bars and ABC cell lines by blue bars. Generated using the heatmap.2 function in the R package gplots (R Development Core Team 2008; Warnes *et al.* 2011).



**Figure 6** Differential gene expression. (A) RefSeq genes with differential expression (FDR < 0.05) divided by expression pattern. (B) Examples of differentially expressed genes. *CCR7* shows greater expression in the ABC cell lines: SU-DHL2, OCI-Ly3, OCI-Ly10, and U-2932. *PRKCA* shows greater expression in the GCB cell lines: SU-DHL4, SU-DHL6, SU-DHL10, and OCI-Ly7.

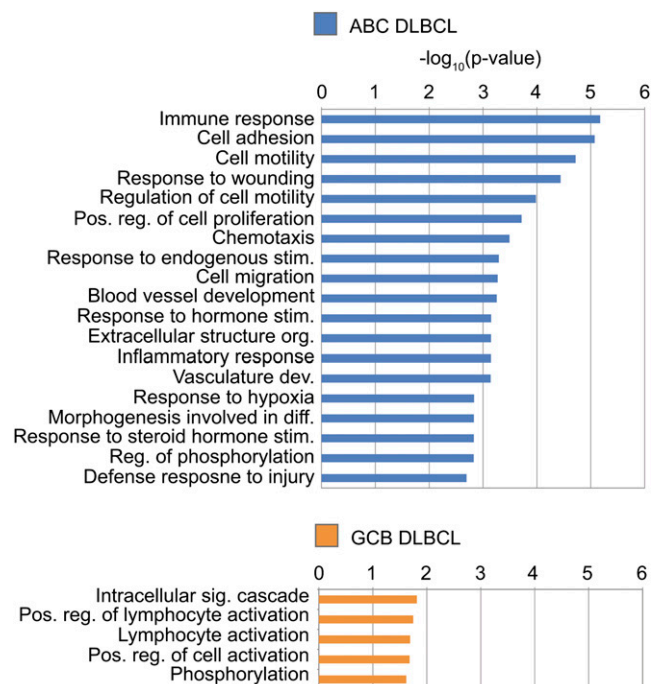
expression levels from the RNA-Seq data. Both the STAT3 binding and mRNA expression were evaluated for statistical significance, resulting in a list of genes that shows both nearby STAT3 binding changes and an accompanying relative mRNA expression level for the two cell types.

ChIP-Seq identified a total of 10,337 STAT3 BRs, whereas the RNA-Seq analysis compared the expression level of 20,992 genes. Using GREAT analysis, we found that more than half of these genes (60%,  $n = 12,631$ ) have a STAT3 BR associated with them. Many of these genes have multiple nearby BRs; likewise, many BRs are associated with two genes because they are located in a region in which putative regulatory domains overlap. Intersecting the ChIP-Seq and RNA-Seq datasets resulted in a total of 16,341 gene/BR matches (see Table S4). Forty percent ( $n = 6580$ ) showed significant differences (FDR < 0.05) in at least one category: either gene expression or STAT3 binding differed markedly between the two DLBCL subtypes. A total of 823 of these gene/BR matches, corresponding to 445 individual genes, were significantly different in both categories (Figure 8).

If STAT3 binding did not influence DLBCL gene expression, we would expect the nature of STAT3 binding near a gene to have no correlation with its expression level (Zhu *et al.* 2012). Instead, we observe that an increase in STAT3 binding is correlated with an increase in expression in most of these pairs (83%,  $n = 686$ ), an association that is highly significant at  $P = 2.3 \times 10^{-28}$  (Table 4). The strong association between binding and increased expression indicates that STAT3 binding is a significant contributor to the gene expression pattern in GCB and ABC DLBCL. The analysis indicates that STAT3 primarily functions as a transcriptional activator whereas a subset of its binding might contribute to transcriptional repression, which is consistent with its roles as previously described (Ouyang *et al.* 2009).

**High STAT3 binding and gene expression in ABC:** The most enriched category of BR/gene pairs consists of genes with greater STAT3 binding in the ABC subtype as well as increased gene expression in ABC: 531 pairs correspond to 287 different genes. *STAT3* itself is among these, as are several of its known targets, including *PRDM1* and *SOCS3* (Table 5). We also confirm the findings from array CGH studies that *NFKB1Z* and the oncogene *PIM2* are elevated in ABC lymphoma (Tirado *et al.* 2012), but with the added information that *STAT3* shows elevated levels of binding in the regulatory domains of these genes.

STAT3 also binds near many genes that are up-regulated in ABC cells relative to GCB cells; these genes have a wide range of functions that are pertinent to oncogenesis, especially apoptosis and cellular signaling. The *CXCR5* chemokine receptor, which contributes to apoptosis resistance in B-cell—derived leukemias, is 6.83-fold more expressed in the ABC subtype (FDR =  $8.83 \times 10^{-7}$ ) and also has a distinct STAT3 peak 43 bp upstream of its TSS that is occupied only in ABC cell lines (Figure 9). STAT3 is also characterized as regulating the *BCL2* protein family, which influences apoptosis. We specifically identify *BCL2A1* (formerly *BFL1*), which is associated with three strongly bound STAT3 BRs and demonstrates a 4.43-fold increase in expression. High levels of *BCL2A1* have been proposed as a mechanism of chemoresistance in chronic lymphocytic leukemia (Olsson *et al.* 2007).



**Figure 7** RNA-Seq gene enrichment for ontological categories. Scale is the  $-\log_{10}(P\text{-value})$  of the binomial  $P\text{-value}$ , equivalent to the exponent. All enriched GO Biological Processes categories are shown,  $P < 0.05$ .



■ **Table 3 Selected differentially expressed genes**

Gene Name	Gene Description	GCB Read Counts	ABC Read Counts	Fold Change	FDR
ABC up					
ARNT2	Aryl-hydrocarbon receptor nuclear translocator 2	0.00	386.14	N/A	2.25E-08
CCR7	Chemokine (C-C motif) receptor 7	46.30	8,252.02	178.25	4.72E-23
CXCR5	Chemokine (C-X-C motif) receptor 5	1,845.20	12,595.47	6.83	8.83E-07
FGF2	Fibroblast growth factor 2 (basic)	5.23	149.82	28.63	1.34E-02
PDGFA	Platelet-derived growth factor alpha polypeptide	3.25	106.66	32.83	1.59E-02
PDGFRB	Platelet-derived growth factor receptor, beta polypeptide	0.43	26.39	60.79	1.62E-02
TGFA	Transforming growth factor, alpha	4.50	132.35	29.42	2.46E-02
GCB up					
BCL6	B-cell CLL/lymphoma 6	15,105.68	5,124.25	0.34	1.94E-02
PRKCA	Protein kinase C, alpha	5,600.83	1,531.13	0.27	5.00E-04

GCB, germinal center B-cell-like; ABC, activated B-cell-like; FDR, false-discovery rate; N/A, not available; CLL, chronic lymphocytic leukemia.

STAT3 also has a binding peak just 7 bp upstream of the TSS of *BATF*, which is highly expressed in ABC cells. *BATF* belongs to the AP-1/ATF superfamily of transcription factors and dimerizes with the Jun family, acting as a negative regulator (Echlin *et al.* 2000). Although *BATF* typically promotes growth arrest and terminal differentiation (Liao *et al.* 2011a), chronically elevated *BATF* levels often are seen in lymphoid tumors and promote an impaired response to apoptotic signals, apparently mediated by the *BCL2* family (Logan *et al.* 2012). Because normal levels of *BATF* stop cell division, this pathway presents a compelling target for drug modification.

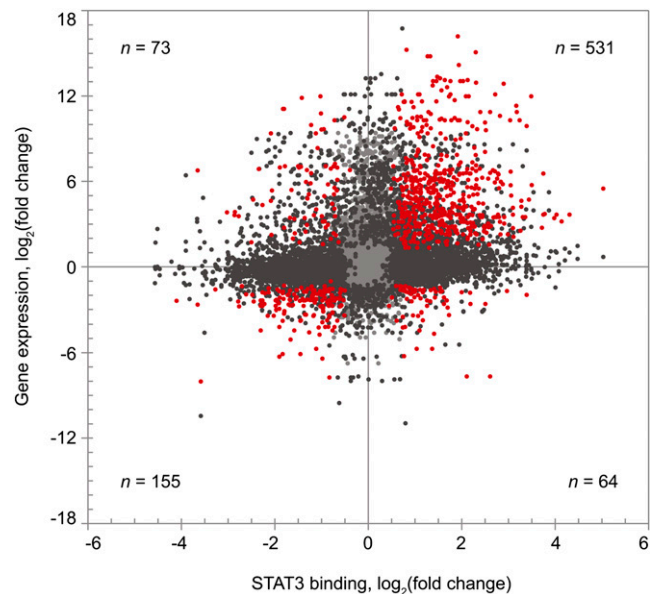
We also identify *APOL2* as a potential target of STAT3, suggesting a novel mechanism of apoptosis resistance in ABC DLBCL cells. A STAT3 BR is located 322 bp downstream of the *APOL2* TSS (fold change = 2.64, FDR =  $2.44 \times 10^{-6}$ ), and its mRNA expression level rises 36.76-fold in the ABC subtype (FDR =  $6.80 \times 10^{-14}$ ). *APOL2* has recently been found to protect endothelial cells from IFN $\gamma$  cytotoxic signaling (Liao *et al.* 2011b), and could play the same role in DLBCL in response to inflammatory signals.

Many interleukins and interleukin receptors show both elevated expression and STAT3 binding, including *IL-6*, *IL-10*, *IL-10RA*, and *IL-12A*. Both *IL-6* and *IL-10* are known to induce signaling via STAT3 upon binding their extracellular receptors; in turn, STAT3 up-regulates expression of the *IL-10* and *IL-6* genes to form an activation loop. *NFKBIZ*, which is also upregulated by STAT3, is a nuclear I $\kappa$ B protein that activates *IL-6* transcription and may also play a role in this loop (Motoyama *et al.* 2005). Polymorphisms in *IL-10* and its receptor subunit *IL-10RA*, also up-regulated in this study, have been implicated in the development of non-Hodgkin lymphoma generally and B-cell lymphoma specifically (Lan *et al.* 2006; Rothman *et al.* 2006), suggesting that *IL-10* pathway dysregulation promotes hematopoietic tumors.

*IL-2RB* produces the beta subunit of the IL-2 cytokine receptor. Its expression level increases 935.76-fold from GCB to ABC (FDR =  $6.45 \times 10^{-16}$ ) and is associated with increased STAT3 binding at a distal site about 27 kb downstream (FDR =  $1.33 \times 10^{-19}$ ). *IL-2RA*, which encodes the other half of the IL-2 receptor complex, is also up-regulated in the ABC subtype, although it lacks a proximal STAT3 peak (fold change = 8.00, FDR =  $1.95 \times 10^{-3}$ ). These findings point to a potential therapy target for refractory cases of ABC DLBCL. The existing anticancer treatment denileukin diftotox (trade name Ontak) is an engineered protein combining IL-2 and diphtheria toxin, which binds to IL-2 receptors and introduces the diphtheria toxin into cells, killing them (Lansigan *et al.* 2010). Ontak has been specifically approved by the Food and Drug Administration for the treatment of cutaneous T-cell lymphoma but should be effective in treating any

lymphoma that displays an elevated number of IL-2 receptors. Furthermore, increased expression of the *IL-2RB* gene is known to increase the efficacy of Ontak treatment (Shao *et al.* 2002).

**High STAT3 binding and gene expression in GCB:** A total of 95 genes have greater STAT3 binding and gene expression in GCB cells. Among these is *MYBL1*, a known biomarker for the GCB subtype (Blenk *et al.* 2007), which is overexpressed as expected in GCB cell lines and also has a statistically significant STAT3 peak 199 bp upstream of its TSS. Also included in this list are two tumor suppressor genes: *ABI3* and *TRIB2*. In thyroid tumors, *ABI3* expression blocks tumor progression on many fronts, including reducing anchorage-independent growth and *in vivo* tumor formation while increasing cellular senescence (Latini *et al.* 2011). In the GCB subtype, it demonstrates greater expression that corresponds to the presence of STAT3 424 bp downstream of its TSS (Figure 9). *TRIB2* inhibits



**Figure 8** STAT3 binding correlated with gene expression change. Red dots indicate STAT3 BRs that have significant differences in binding (FDR < 0.05) between subtypes and are also associated with genes that display significant differences in expression (FDR < 0.05); dark gray dots indicate STAT3 BRs that either show significant differences in binding or whose associated genes display a significant difference in expression; and light gray dots indicate other BRs.



■ **Table 4 STAT3 binding is correlated with gene expression by the  $\chi^2$  test**

STAT3 Binding	Gene Expression	Expected Frequency	Expected Counts	Observed Counts	$\frac{(\text{obs.} - \text{exp})^2}{\text{exp}}$
ABC+	ABC+	55% • 81% = 46.5%	382	531	57.19
ABC+	GCB+	55% • 19% = 9.5%	78	64	2.51
GCB+	ABC+	45% • 81% = 36.5%	300	155	70.08
GCB+	GCB+	45% • 19% = 7.5%	61	73	1.95
			823	823	$\chi^2 = 131.74$
					d.f. = 3
					$P = 2.3 \times 10^{-28}$

STAT3, signal transducer and activator of transcription 3; TSS, transcription start site; ChIP, chromatin immunoprecipitation; FDR, false-discovery rate; ABC, activated B-cell-like; FDR, false-discovery rate; obs, observed counts; exp, expected counts.

JNK, which ultimately reduces the antiapoptotic activity of BCL2 (Gilby *et al.* 2010). The loss of *TRIB2* expression in ABC DLBCL is particularly problematic, as STAT3 increases the expression of BCL2 and its family members.

**Anticorrelated STAT3 binding and gene expression:** Seventeen percent ( $n = 137$ ) of BR/gene pairs have strong STAT3 binding in one subtype but strong gene expression in the other. These 107 genes may be regulated by other factors, or STAT3 may serve as a repressor at these sites, potentially by blocking the binding of other factors. Fully 40% ( $n = 43$ ) of these potential anticorrelated genes occur due to a mixture of STAT3 binding behavior: the gene is associated with many BRs, some of which are more occupied by STAT3 in ABC while others show stronger binding GCB.

We observe greater *CD22* expression in GCB DLBCL, associated with a STAT3 peak 138 bp upstream of its TSS. However, *CD22* also has two STAT3 BRs near its 3' end that are more strongly bound in

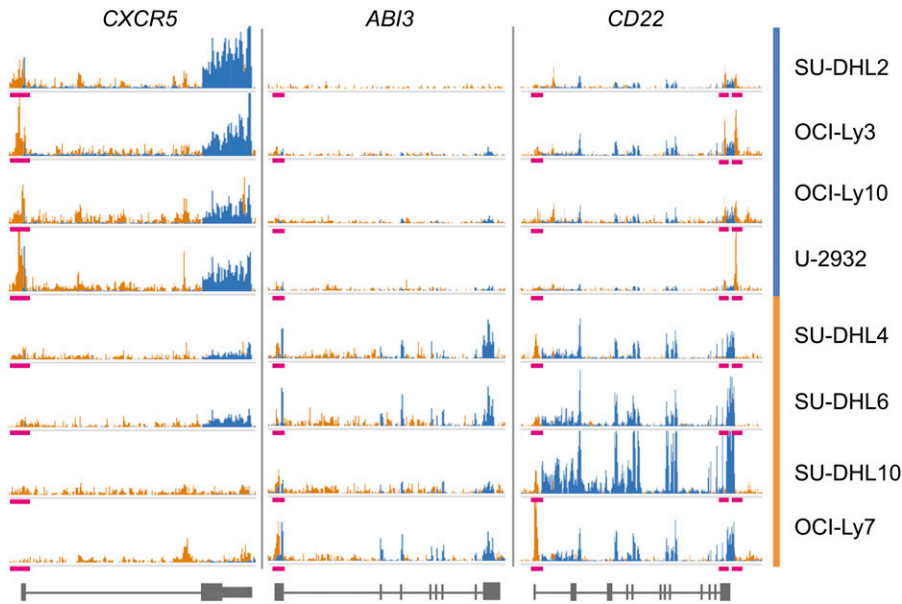
ABC (Figure 9). This finding suggests that STAT3 binding at the *CD22* promoter drives transcription, whereas binding at the two downstream sites suppresses it. *CD22* is a cell surface antigen expressed in a wide variety of non-Hodgkin lymphomas, and anti-*CD22* monoclonal antibody treatments currently are being developed (Lu *et al.* 2006; Li *et al.* 2013). However, they may be less successful in treating ABC DLBCL due to lowered expression of this antigen.

*BCL6* is the most extreme example of this mixed binding phenomenon. It is the primary biomarker for the GCB subtype of DLBCL and is known to be regulated by STAT3. We found that expression of *BCL6* is extremely high in GCB lines, three times the level found in our ABC cell lines. However, its regulation is complex: there are 24 STAT3 BRs within the putative regulatory domain of *BCL6*, which includes a large upstream gene desert that contains considerable STAT3 binding activity. Seven of these BRs show a significant binding differential between the subtypes. Two of them are more strongly bound in ABC, but the other five are more strongly bound by STAT3

■ **Table 5 Selected differentially expressed genes with associated STAT3 peaks**

Gene Name	Peak Proximity to TSS	ChIP Fold Change	ChIP FDR	Expression Fold Change	Expression FDR
Increased ABC binding and expression					
<i>APOL2</i>	322	2.64	2.44E-06	36.76	6.80E-14
<i>BATF</i>	-7868	4.14	2.50E-14	18.51	5.48E-10
<i>BATF</i>	-7	3.73	4.97E-13	18.51	5.48E-10
<i>BCL2A1</i>	-7594	5.50	3.07E-19	21.56	3.25E-09
<i>BCL2A1</i>	-6328	2.28	1.25E-04	21.56	3.25E-09
<i>BCL2A1</i>	59	2.69	5.14E-06	21.56	3.25E-09
<i>CXCR5</i>	-43	3.29	4.84E-07	6.82	8.83E-07
<i>IL-6</i>	-28625	2.79	1.96E-07	10369.08	7.03E-05
<i>IL-10</i>	5883	2.45	7.48E-04	222.86	2.37E-11
<i>IL-10RA</i>	917	2.00	1.91E-03	4.96	7.26E-04
<i>IL-12A</i>	-63	2.41	9.49E-05	28133.47	2.60E-07
<i>IL-2RB</i>	27623	4.69	1.33E-19	935.76	6.45E-16
<i>SOCS3</i>	-65	9.45	2.66E-33	1251.98	2.14E-15
<i>STAT3</i>	-308	17.63	1.05E-09	9.00	1.39E-07
<i>PIM2</i>	-25	1.78	2.74E-03	12.30	2.78E-10
<i>PRDM1</i>	11587	2.17	1.16E-03	127.12	3.21E-16
<i>NFKBIZ</i>	-95	4.59	7.14E-16	8.46	5.63E-06
Increased GCB binding and expression					
<i>ABI3</i>	424	0.61	4.9E-02	0.19	2.55E-02
<i>MYBL1</i>	-199	0.66	3.48E-02	0.23	4.99E-02
<i>TRIB2</i>	-144	0.37	6.71E-06	0.07	6.81E-08
Mixed binding and expression					
<i>CD22</i>	-138	0.30	1.21E-07	0.25	8.10E-04
<i>CD22</i>	17,622	2.91	8.27E-09	0.25	8.10E-04
<i>CD22</i>	18,556	4.65	7.74E-11	0.25	8.10E-04

STAT3, signal transducer and activator of transcription 3; TSS, transcription start site; ChIP, chromatin immunoprecipitation; FDR, false-discovery rate; ABC, activated B-cell-like; FDR, false-discovery rate.

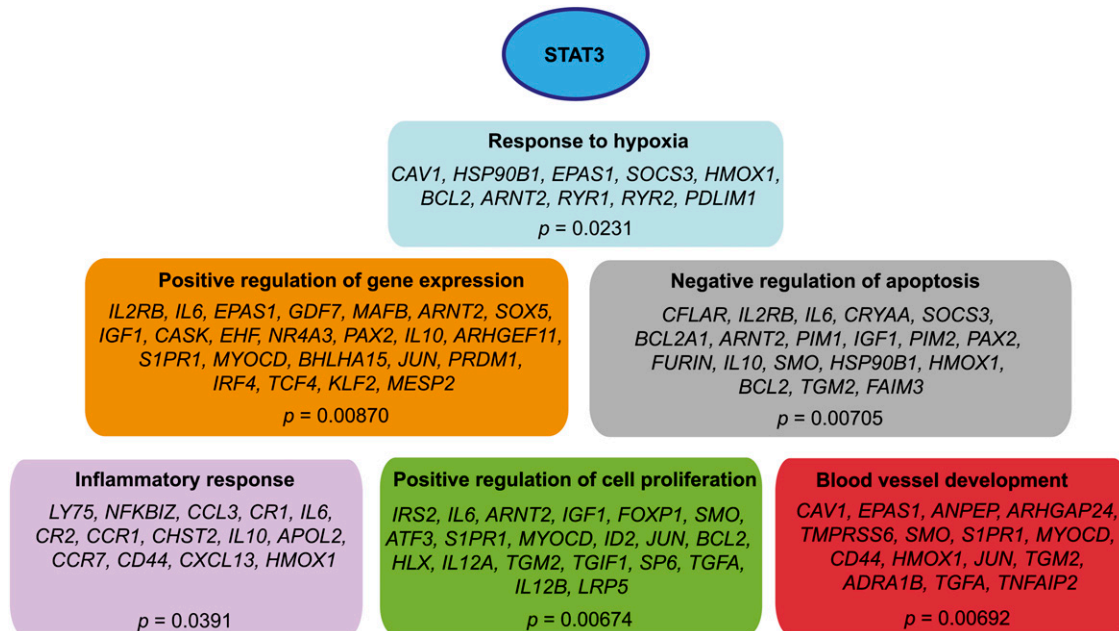


**Figure 9** STAT3 peaks associated with gene expression changes. Examples of differentially occupied STAT3 BRs are shown. Blue tracks represent gene expression; orange tracks represent STAT3 binding; pink bars show noteworthy STAT3 BRs. *CXCR5* shows increased gene expression correlated with STAT3 binding (chr11: 118,754,259–118,754,737) in the ABC cell lines: SU-DHL2, OCI-Ly3, OCI-Ly10, and U-2932. *ABI3* shows strong increased gene expression correlated with STAT3 binding (chr17: 47,287,858–47,288,168) in the GCB cell lines: SU-DHL4, SU-DHL6, SU-DHL10, and OCI-Ly7. *CD22* shows increased gene expression in the GCB cell lines, but a mixture of STAT3 binding behavior: STAT3 has more GCB binding near the TSS (chr19: 35,819,821–35,820,060) but more ABC binding at two sites near the 3' end (chr19: 35,837,564–35,837,836; chr19: 35,838,409–35,838,860).

in GCB (see Figure S3 and Table S3). This complexity suggests that the overall sum of STAT3 binding contributes to the final regulation of *BCL6*.

The combination of whole-genome profiling approaches shows differential STAT3 binding sites near almost a third of all genes that differ in expression between GCB and ABC DLBCL, leaving little doubt that STAT3 is a regulator of the aggressive clinical phenotype and drug resistance displayed by the ABC subtype. The STAT3 binding and expression results suggest possible models by which STAT3 regulates biological pathways to promote aggressive oncogenic behavior in ABC DLBCL. Ontological analysis of the genes up-regulated in ABC DLBCL with increased nearby STAT3 binding

indicates that STAT3 promotes cell migration, blood vessel development, and the inflammatory response (Figure 10). Furthermore, it ensures its own continued high expression by up-regulating its pathway partners and upstream signaling molecules, such as IL-6 and IL-10, and promotes cross-talk with the nuclear factor- $\kappa$ B pathway. Most critically, STAT3 strongly up-regulates many mechanisms of apoptosis resistance. These effects synergize and likely lead to increased cell proliferation and distributions of cells, which are hallmarks of aggressive lymphomas. These results provide novel insights into the genetics of oncogenesis in the ABC form of diffuse large B-cell lymphoma and identify a substantial number of plausible downstream mechanisms for the treatment-resistant phenotype of this subtype.



**Figure 10** Downstream pathways and genes induced by STAT3 in ABC DLBCL. Examples of the top GO Biological Processes categories enriched among genes with increased expression and increased STAT3 binding in ABC DLBCL. All *P*-values are corrected for multiple comparisons using the Benjamini-Hochberg procedure.

## ACKNOWLEDGMENTS

We thank the ENCODE Consortium for generating some of the data and techniques used in this work. We would also thank the laboratories of Dr. Louis Staudt at the National Cancer Institute and Dr. Ash Alizadeh at the Stanford School of Medicine for the generous donation of some of the DLBCL cell lines used in these experiments. This research was supported in part by the National Institutes of Health “Analysis of Human Genome Using Integrated Technologies” Centers of Excellence in Genomic Science Award Grant no. P50 HG002357 and by the ENCODE Consortium under National Institutes of Health grant NIH 5U54 HG 004558.

## LITERATURE CITED

- Alizadeh, A. A., M. B. Eisen, R. E. Davis, C. Ma, I. S. Lossos *et al.*, 2000 Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403: 503–511.
- Alvarez, J. V., and D. A. Frank, 2004 Genome-wide analysis of STAT target genes: elucidating the mechanism of STAT-mediated oncogenesis. *Cancer Biol. Ther.* 3: 1045–1050.
- American Cancer Society, 2012 *Cancer Facts & Figures—2012*. American Cancer Society, Atlanta.
- Anders, S., and W. Huber, 2010 Differential expression analysis for sequence count data. *Genome Biol.* 11: R106.
- Anderson, J. J., S. Fordham, L. Overman, H. Dignum, K. Wood *et al.*, 2009 Immunophenotyping of diffuse large B-cell lymphoma (DLBCL) defines multiple sub-groups of germinal centre-like tumours displaying different survival characteristics. *Int. J. Oncol.* 35: 961–971.
- Angelini-Duclos, C., G. Cattoretti, K. I. Lin, and K. Calame, 2000 Commitment of B lymphocytes to a plasma cell fate is associated with Blimp-1 expression in vivo. *J. Immunol.* 165: 5462–5471.
- Arigami, T., S. Natsugoe, Y. Uenosono, S. Yanagita, H. Arima *et al.*, 2009 CCR7 and CXCR4 expression predicts lymph node status including micrometastasis in gastric cancer. *Int. J. Oncol.* 35: 19–24.
- Avery, D. T., E. K. Deenick, C. S. Ma, S. Suryani, N. Simpson *et al.*, 2010 B cell-intrinsic signaling through IL-21 receptor and STAT3 is required for establishing long-lived antibody responses in humans. *J. Exp. Med.* 207: 155–171.
- Baker, S. J., S. G. Rane, and E. P. Reddy, 2007 Hematopoietic cytokine receptor signaling. *Oncogene* 26: 6724–6737.
- Benekli, M., M. R. Baer, H. Baumann, and M. Wetzler, 2003 Signal transducer and activator of transcription proteins in leukemias. *Blood* 101: 2940–2954.
- Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57: 289–300.
- Besson, A., and V. W. Yong, 2000 Involvement of p21(Waf1/Cip1) in protein kinase C alpha-induced cell cycle progression. *Mol. Cell. Biol.* 20: 4580–4590.
- Blankenberg, D., G. Von Kuster, N. Coraor, G. Ananda, R. Lazarus *et al.*, 2010 Galaxy: a web-based genome analysis tool for experimentalists. *Curr. Protoc. Mol. Biol.* Jan: Chapter 19: Unit 19.10.1–21.
- Blenk, S., J. Engelmann, M. Weniger, J. Schultz, M. Ditttrich *et al.*, 2007 Germinal center B cell-like (GCB) and activated B cell-like (ABC) type of diffuse large B cell lymphoma (DLBCL): analysis of molecular predictors, signatures, cell cycle state and patient survival. *Cancer Inform.* 3: 399–420.
- Brady, G., G. J. MacArthur, and P. J. Farrell, 2007 Epstein-Barr virus and Burkitt lymphoma. *J. Clin. Pathol.* 60: 1397–1402.
- Burkle, A., M. Niedermeier, A. Schmitt-Graff, W. G. Wierda, M. J. Keating *et al.*, 2007 Overexpression of the CXCR5 chemokine receptor, and its ligand, CXCL13 in B-cell chronic lymphocytic leukemia. *Blood* 110: 3316–3325.
- Cao, Y., R. Cao, and E. M. Hedlund, 2008 Regulation of tumor angiogenesis and metastasis by FGF and PDGF signaling pathways. *J. Mol. Med. (Berl.)* 86: 785–789.
- Castillo, J., L. Pantanowitz, and B. J. Dezube, 2008 HIV-associated plasmablastic lymphoma: lessons learned from 112 published cases. *Am. J. Hematol.* 83: 804–809.
- Da, M. X., Z. Wu, and H. W. Tian, 2008 Tumor lymphangiogenesis and lymphangiogenic growth factors. *Arch. Med. Res.* 39: 365–372.
- Ding, B. B., J. J. Yu, R. Y. Yu, L. M. Mendez, R. Shakhovich *et al.*, 2008 Constitutively activated STAT3 promotes cell proliferation and survival in the activated B-cell subtype of diffuse large B-cell lymphomas. *Blood* 111: 1515–1523.
- Drexler, H. G., 2010 *Guide to Leukemia-Lymphoma Cell Lines*. German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany.
- Echlin, D. R., H. J. Tae, N. Mitin, and E. J. Taparowsky, 2000 B-ATF functions as a negative regulator of AP-1 mediated transcription and blocks cellular transformation by Ras and Fos. *Oncogene* 19: 1752–1763.
- Falini, B., M. Fizzotti, A. Pucciarini, B. Bigerna, T. Marafioti *et al.*, 2000 A monoclonal antibody (MUM1p) detects expression of the MUM1/IRF4 protein in a subset of germinal center B cells, plasma cells, and activated T cells. *Blood* 95: 2084–2092.
- Fornek, J. L., L. T. Tygrett, T. J. Waldschmidt, V. Poli, R. C. Rickert *et al.*, 2006 Critical role for Stat3 in T-dependent terminal differentiation of IgG B cells. *Blood* 107: 1085–1091.
- Fu, K., D. D. Weisenburger, W. W. Choi, K. D. Perry, L. M. Smith *et al.*, 2008 Addition of rituximab to standard chemotherapy improves the survival of both the germinal center B-cell-like and non-germinal center B-cell-like subtypes of diffuse large B-cell lymphoma. *J. Clin. Oncol.* 26: 4587–4594.
- Giardine, B., C. Riemer, R. C. Hardison, R. Burhans, L. Elnitski *et al.*, 2005 Galaxy: a platform for interactive large-scale genome analysis. *Genome Res.* 15: 1451–1455.
- Gilby, D. C., H. Y. Sung, P. R. Winship, A. C. Goodeve, J. T. Reilly *et al.*, 2010 Tribbles-1 and -2 are tumour suppressors, down-regulated in human acute myeloid leukaemia. *Immunol. Lett.* 130: 115–124.
- Goecks, J., A. Nekrutenko, and J. Taylor, 2010 Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* 11: R86.
- Han, Y., T. Meng, N. R. Murray, A. P. Fields, and A. R. Brasier, 1999 Interleukin-1-induced nuclear factor-kappaB-IkappaBalpha autoregulatory feedback loop in hepatocytes. A role for protein kinase calpha in post-transcriptional regulation of ikappabalpha resynthesis. *J. Biol. Chem.* 274: 939–947.
- Han, Y., H. M. Amin, B. Franko, C. Frantz, X. Shi *et al.*, 2006 Loss of SHP1 enhances JAK3/STAT3 signaling and decreases proteasome degradation of JAK3 and NPM-ALK in ALK+ anaplastic large-cell lymphoma. *Blood* 108: 2796–2803.
- Hans, C. P., D. D. Weisenburger, T. C. Greiner, R. D. Gascoyne, J. Delabie *et al.*, 2004 Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood* 103: 275–282.
- Hodge, D. R., E. M. Hurt, and W. L. Farrar, 2005 The role of IL-6 and STAT3 in inflammation and cancer. *Eur. J. Cancer* 41: 2502–2512.
- Hu, C., J. Xiong, L. Zhang, B. Huang, Q. Zhang *et al.*, 2004 PEG10 activation by co-stimulation of CXCR5 and CCR7 essentially contributes to resistance to apoptosis in CD19+CD34+ B cells from patients with B cell lineage acute and chronic lymphocytic leukemia. *Cell. Mol. Immunol.* 1: 280–294.
- Huttemann, M., I. Lee, L. I. Grossman, J. W. Doan, and T. H. Sanderson, 2012 Phosphorylation of mammalian cytochrome c and cytochrome c oxidase in the regulation of cell destiny: respiration, apoptosis, and human disease. *Adv. Exp. Med. Biol.* 748: 237–264.
- Ishida, K., M. Iwahashi, M. Nakamori, M. Nakamura, S. Yokoyama *et al.*, 2009 High CCR7 mRNA expression of cancer cells is associated with lymph node involvement in patients with esophageal squamous cell carcinoma. *Int. J. Oncol.* 34: 915–922.
- Kasowski, M., F. Grubert, C. Heffelfinger, M. Hariharan, A. Asabere *et al.*, 2010 Variation in transcription factor binding among humans. *Science* 328: 232–235.
- Kharchenko, P. V., M. Y. Tolstorukov, and P. J. Park, 2008 Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nat. Biotechnol.* 26: 1351–1359.



- Kodama, J., T. Hasengaowa, N. Kusumoto, T. Seki, Matsuo *et al.*, 2007 Association of CXCR4 and CCR7 chemokine receptor expression and lymph node metastasis in human cervical cancer. *Ann. Oncol.* 18: 70–76.
- Koster, A., and J. M. Raemaekers, 2005 Angiogenesis in malignant lymphoma. *Curr. Opin. Oncol.* 17: 611–616.
- Lan, Q., T. Zheng, N. Rothman, Y. Zhang, S. S. Wang *et al.*, 2006 Cytokine polymorphisms in the Th1/Th2 pathway and susceptibility to non-Hodgkin lymphoma. *Blood* 107: 4101–4108.
- Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg, 2009 Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10: R25.
- Lansigan, F., D. M. Stearns, and F. Foss, 2010 Role of denileukin diftitox in the treatment of persistent or recurrent cutaneous T-cell lymphoma. *Cancer Manag Res* 2: 53–59.
- Latini, F. R., J. P. Hemerly, B. C. Freitas, G. Oler, G. J. Riggins *et al.*, 2011 AB13 ectopic expression reduces in vitro and in vivo cell growth properties while inducing senescence. *BMC Cancer* 11: 11.
- Lenz, G., G. Wright, S. S. Dave, W. Xiao, J. Powell *et al.*, 2008 Stromal gene signatures in large-B-cell lymphomas. *N. Engl. J. Med.* 359: 2313–2323.
- Li, D., K. Achilles-Poon, S. F. Yu, R. Dere, M. Go *et al.*, 2013 DCDT2980S, an anti-CD22-monomethyl auristatin e antibody-drug conjugate, is a potential treatment for non-Hodgkins lymphoma. *Mol Cancer Ther.* 12: 1255–1265.
- Li, Q., J. B. Brown, H. Huang, and P. J. Bickel, 2011 Measuring reproducibility of high-throughput experiments. *Ann. Appl. Stat.* 5: 1752–1779.
- Liang, H., X. Ding, C. Zhou, Y. Zhang, M. Xu *et al.*, 2012 Knockdown of eukaryotic translation initiation factors 3B (EIF3B) inhibits proliferation and promotes apoptosis in glioblastoma cells. *Neurol. Sci.* 33: 1057–1062.
- Liao, J., S. E. Humphrey, S. Poston, and E. J. Taparowsky, 2011a Batf promotes growth arrest and terminal differentiation of mouse myeloid leukemia cells. *Mol. Cancer Res.* 9: 350–363.
- Liao, W., F. Y. Goh, R. J. Betts, D. M. Kemeny, J. Tam *et al.*, 2011b A novel anti-apoptotic role for apolipoprotein L2 in IFN-gamma-induced cytotoxicity in human bronchial epithelial cells. *J. Cell. Physiol.* 226: 397–406.
- Lin, Q., R. Lai, L. R. Chiriac, C. Li, V. A. Thomazy *et al.*, 2005 Constitutive activation of JAK3/STAT3 in colon carcinoma tumors and cell lines: inhibition of JAK3/STAT3 signaling induces apoptosis and cell cycle arrest of colon carcinoma cells. *Am. J. Pathol.* 167: 969–980.
- Logan, M. R., K. L. Jordan-Williams, S. Poston, J. Liao, and E. J. Taparowsky, 2012 Overexpression of Batf induces an apoptotic defect and an associated lymphoproliferative disorder in mice. *Cell Death Dis.* 3: e310.
- Lu, P. P., Z. Y. Meng, M. X. Zhou, M. W. Wang, and G. F. Dou, 2006 Immunotherapy of non-Hodgkin's lymphomas (NHL) by anti-CD22 antibody. *Zhongguo Shi Yan Xue Ye Xue Za Zhi* 14: 1258–1261 (review in Chinese).
- Lu, Z., G. Jiang, P. Blume-Jensen, and T. Hunter, 2001 Epidermal growth factor-induced tumor cell invasion and metastasis initiated by dephosphorylation and downregulation of focal adhesion kinase. *Mol. Cell. Biol.* 21: 4016–4031.
- Maltepe, E., B. Keith, A. M. Arsham, J. R. Brorson, and M. C. Simon, 2000 The role of ARNT2 in tumor angiogenesis and the neural response to hypoxia. *Biochem. Biophys. Res. Commun.* 273: 231–238.
- McLean, C. Y., D. Bristol, M. Hiller, S. L. Clarke, B. T. Schaar *et al.*, 2010 GREAT improves functional interpretation of *cis*-regulatory regions. *Nat. Biotechnol.* 28: 495–501.
- Minegishi, Y., 2009 Hyper-IgE syndrome. *Curr. Opin. Immunol.* 21: 487–492.
- Morin, R., M. Bainbridge, A. Fejes, M. Hirst, M. Krzywinski *et al.*, 2008 Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. *Biotechniques* 45: 81–94.
- Motoyama, M., S. Yamazaki, A. Eto-Kimura, K. Takeshige, and T. Muta, 2005 Positive and negative regulation of nuclear factor-kappaB-mediated transcription by IkappaB-zeta, an inducible nuclear protein. *J. Biol. Chem.* 280: 7444–7451.
- Nakata, B., S. Fukunaga, E. Noda, R. Amano, N. Yamada *et al.*, 2008 Chemokine receptor CCR7 expression correlates with lymph node metastasis in pancreatic cancer. *Oncology* 74: 69–75.
- Olsson, A., M. Norberg, A. Okvist, K. Derkow, A. Choudhury *et al.*, 2007 Upregulation of bfl-1 is a potential mechanism of chemoresistance in B-cell chronic lymphocytic leukaemia. *Br. J. Cancer* 97: 769–777.
- Ouyang, Z., Q. Zhou, and W. H. Wong, 2009 ChIP-Seq of transcription factors predicts absolute and differential gene expression in embryonic stem cells. *Proc. Natl. Acad. Sci. USA* 106: 21521–21526.
- Pallesen, G., S. J. Hamilton-Dutoit, M. Rowe, and L. S. Young, 1991 Expression of Epstein-Barr virus latent gene products in tumour cells of Hodgkin's disease. *Lancet* 337: 320–322.
- Pan, M. R., M. F. Hou, H. C. Chang, and W. C. Hung, 2008 Cyclooxygenase-2 up-regulates CCR7 via EP2/EP4 receptor signaling pathways to enhance lymphatic invasion of breast cancer cells. *J. Biol. Chem.* 283: 11155–11163.
- Park, S., J. Lee, Y. H. Ko, A. Han, H. J. Jun *et al.*, 2007 The impact of Epstein-Barr virus status on clinical outcome in diffuse large B-cell lymphoma. *Blood* 110: 972–978.
- R Development Core Team, 2008 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rosenwald, A., G. Wright, W. C. Chan, J. M. Connors, E. Campo *et al.*, 2002 The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N. Engl. J. Med.* 346: 1937–1947.
- Rothman, N., C. F. Skibola, S. S. Wang, G. Morgan, Q. Lan *et al.*, 2006 Genetic variation in TNF and IL10 and risk of non-Hodgkin lymphoma: a report from the InterLymph Consortium. *Lancet Oncol.* 7: 27–38.
- Schimanski, C. C., R. Bahre, I. Gockel, T. Junginger, N. Simiantonaki *et al.*, 2006 Chemokine receptor CCR7 enhances intrahepatic and lymphatic dissemination of human hepatocellular cancer. *Oncol. Rep.* 16: 109–113.
- Shao, R. H., X. Tian, G. Gorgun, A. G. Urbano, and F. M. Foss, 2002 Arginine butyrate increases the cytotoxicity of DAB(389)IL-2 in leukemia and lymphoma cells by upregulation of IL-2Rbeta gene. *Leuk. Res.* 26: 1077–1083.
- Shuai, K., G. R. Stark, I. M. Kerr, and J. E. Darnell, Jr., 1993 A single phosphotyrosine residue of Stat91 required for gene activation by interferon-gamma. *Science* 261: 1744–1746.
- Siu, L. L., J. K. Chan, and Y. L. Kwong, 2002 Natural killer cell malignancies: clinicopathologic and molecular features. *Histol. Histopathol.* 17: 539–554.
- Suzuki, K., H. Nakajima, Y. Saito, T. Saito, W. J. Leonard *et al.*, 2000 Janus kinase 3 (Jak3) is essential for common cytokine receptor gamma chain (gamma(c))-dependent signaling: comparative analysis of gamma(c), Jak3, and gamma(c) and Jak3 double-deficient mice. *Int. Immunol.* 12: 123–132.
- Sweetenham, J. W., 2005 Diffuse large B-cell lymphoma: risk stratification and management of relapsed disease. *Hematology (Am Soc Hematol Educ Program)* 252–259.
- The ENCODE Project Consortium, R. M. Myers, J. Stamatoyannopoulos, M. Snyder, I. Dunham *et al.*, 2011 A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol.* 9: e1001046.
- Tirado, C. A., W. Chen, R. Garcia, K. A. Kohlman, and N. Rao, 2012 Genomic profiling using array comparative genomic hybridization define distinct subtypes of diffuse large b-cell lymphoma: a review of the literature. *J Hematol Oncol* 5: 54.
- Turkson, J., 2004 STAT proteins as novel targets for cancer drug discovery. *Expert Opin. Ther. Targets* 8: 409–422.
- Warnes, G. R., B. Bolker, L. Bonebakker, R. Gentleman, W. Huber *et al.*, 2011 gplots: Various R programming tools for plotting data. Available at: <http://cran.r-project.org/web/packages/gplots/index.html>. Accessed October 28, 2013.
- Wen, Z., Z. Zhong, and J. E. Darnell, Jr., 1995 Maximal activation of transcription by Stat1 and Stat3 requires both tyrosine and serine phosphorylation. *Cell* 82: 241–250.
- Wright, G., B. Tan, A. Rosenwald, E. H. Hurt, A. Wiestner *et al.*, 2003 A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc. Natl. Acad. Sci. USA* 100: 9991–9996.

- Wu, Z. L., Y. Q. Song, Y. F. Shi, and J. Zhu, 2011 High nuclear expression of STAT3 is associated with unfavorable prognosis in diffuse large B-cell lymphoma. *J Hematol Oncol* 4: 31.
- Yan, J., M. Enge, T. Whittington, K. Dave, J. Liu *et al.*, 2013 Transcription factor binding in human cells occurs in dense clusters formed around cohesin anchor sites. *Cell* 154: 801–813.
- Yang, S. Y., A. Miah, A. Pabari, and M. Winslet, 2011 Growth factors and their receptors in cancer metastases. *Front. Biosci.* 16: 531–538.
- Yu, G. H., K. T. Montone, D. Frias-Hidvegi, R. S. Cajulis, B. A. Brody *et al.*, 1996 Cytomorphology of primary CNS lymphoma: review of 23 cases and evidence for the role of EBV. *Diagn. Cytopathol.* 14: 114–120.
- Zhang, X., J. Blenis, H. C. Li, C. Schindler, and S. Chen-Kiang, 1995 Requirement of serine phosphorylation for formation of STAT-promoter complexes. *Science* 267: 1990–1994.
- Zhong, Z., Z. Wen, and J. E. Darnell, Jr., 1994 Stat3: a STAT family member activated by tyrosine phosphorylation in response to epidermal growth factor and interleukin-6. *Science* 264: 95–98.
- Zhu, B. M., K. Kang, J. H. Yu, W. Chen, H. E. Smith *et al.*, 2012 Genome-wide analyses reveal the extent of opportunistic STAT5 binding that does not yield transcriptional activation of neighboring genes. *Nucleic Acids Res.* 40: 4461–4472.

*Communicating editor: T. R. Hughes*