










ARTICLE

<https://doi.org/10.1038/s41467-019-13515-5>

OPEN

# ASCL1 is a MYCN- and LMO1-dependent member of the adrenergic neuroblastoma core regulatory circuitry

Lu Wang<sup>1,2</sup>, Tze King Tan <sup>1</sup>, Adam D. Durbin <sup>3,4,5</sup>, Mark W. Zimmerman <sup>3</sup>, Brian J. Abraham <sup>6,7</sup>, Shi Hao Tan<sup>1</sup>, Phuong Cao Thi Ngoc <sup>1</sup>, Nina Weichert-Leahey<sup>3,4</sup>, Koshi Akahane <sup>3,8</sup>, Lee N. Lawton<sup>6</sup>, Jo Lynne Rokita <sup>9,10</sup>, John M. Maris<sup>9,11</sup>, Richard A. Young <sup>6,12</sup>, A. Thomas Look<sup>3,4\*</sup> & Takaomi Sanda <sup>1,2\*</sup>

A heritable polymorphism within regulatory sequences of the *LMO1* gene is associated with its elevated expression and increased susceptibility to develop neuroblastoma, but the oncogenic pathways downstream of the LMO1 transcriptional co-regulatory protein are unknown. Our ChIP-seq and RNA-seq analyses reveal that a key gene directly regulated by LMO1 and MYCN is *ASCL1*, which encodes a basic helix-loop-helix transcription factor. Regulatory elements controlling *ASCL1* expression are bound by LMO1, MYCN and the transcription factors GATA3, HAND2, PHOX2B, TBX2 and ISL1—all members of the adrenergic (ADRN) neuroblastoma core regulatory circuitry (CRC). *ASCL1* is required for neuroblastoma cell growth and arrest of differentiation. *ASCL1* and *LMO1* directly regulate the expression of CRC genes, indicating that *ASCL1* is a member and *LMO1* is a coregulator of the ADRN neuroblastoma CRC.

<sup>1</sup>Cancer Science Institute of Singapore, National University of Singapore, Singapore 117599, Singapore. <sup>2</sup>Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117599, Singapore. <sup>3</sup>Department of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02216, USA. <sup>4</sup>Division of Pediatric Hematology/Oncology, Boston Children's Hospital, Boston, MA 02215, USA. <sup>5</sup>The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. <sup>6</sup>Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA. <sup>7</sup>Department of Computational Biology, St. Jude Children's Research Hospital, Memphis, TN 38102, USA. <sup>8</sup>Department of Pediatrics, School of Medicine, University of Yamanashi, Chuo 4093898, Japan. <sup>9</sup>Oncology Division, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA. <sup>10</sup>Department of Bioinformatics and Health Informatics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA. <sup>11</sup>Department of Pediatrics, University of Pennsylvania, Philadelphia, PA 19104, USA. <sup>12</sup>Biology Department, MIT, Cambridge, MA 02142, USA. \*email: [thomas\\_look@dfci.harvard.edu](mailto:thomas_look@dfci.harvard.edu); [takaomi\\_sanda@nus.edu.sg](mailto:takaomi_sanda@nus.edu.sg)

Neuroblastoma is an embryonal tumor of the peripheral sympathetic nervous system, accounting for 7.5% of all cancer diagnoses in children<sup>1–3</sup>. The majority of children with neuroblastoma have highly invasive tumors, which are often metastatic at diagnosis. The outcome for high-risk neuroblastoma cases is very poor, with long-term survival less than 50%<sup>3,4</sup>. Neuroblastoma can be classified into two subtypes based on gene expression and enhancer profiles, one with a committed adrenergic (ADRN) signature and the other with mesenchymal, migratory neural crest properties (MES)<sup>5</sup>. Genome-wide studies revealed several recurrent molecular abnormalities in primary neuroblastoma cases, including genetic amplification of *MYCN* and mutations of *ALK*, *ATRX*, and *PTPN11*<sup>6</sup>. *MYCN* amplification has been used as a risk factor that is associated with a poor prognosis<sup>3,4,7</sup>. In addition, our recent studies have implicated *LMO1* as a major predisposition gene that functions as an oncogene in neuroblastoma<sup>8–10</sup>.

LMO proteins (LMO1–4) are LIM-domain-containing transcriptional co-regulatory factors that lack DNA-binding domains<sup>11–13</sup>. LMO proteins function as adapters to form complexes between DNA-binding proteins such as the class I basic helix-loop-helix (bHLH) proteins, class II bHLH proteins, LDB1 and GATA proteins<sup>12,14</sup>. *LMO2* is an oncogene that is overexpressed in T-cell acute lymphoblastic leukemia (T-ALL) due to chromosomal translocation into the vicinity of the T-cell receptor locus<sup>12,14</sup>. Point mutations in the noncoding elements that generate an enhancer driving overexpression of *LMO2* have also been reported<sup>15</sup>. *LMO1*, another member of the LMO family, is also frequently overexpressed in T-ALL due to chromosomal translocation or mutations in the enhancer<sup>12,14</sup>. *LMO3* is overexpressed in some T-ALL cases due to enhancer hijacking mediated by chromosomal translocation<sup>16</sup>. *LMO1*, *LMO2*, and *LMO3* are considered to be functionally redundant oncogenes in T-ALL<sup>17</sup>.

In childhood neuroblastoma, our previous genome-wide association study (GWAS) has shown that polymorphisms at the *LMO1* gene locus are strongly associated with susceptibility to tumor formation<sup>8</sup>. Germline single nucleotide polymorphism (SNP) risk alleles are associated with increased *LMO1* expression in neuroblastoma cell lines and primary tumors. Genetic knockdown of *LMO1* inhibits the growth of neuroblastoma cells, whereas overexpression of *LMO1* enhances proliferation in cells with low *LMO1* expression<sup>8</sup>. The risk allele of SNP rs2168101 G>T, which is the most highly associated variant, creates a GATA motif, and GATA3 binds at this locus<sup>9</sup>. This GATA3 binding is essential for the creation of a super-enhancer that drives high levels of *LMO1* expression and increases the proliferative fraction of sympathetic neuroblasts<sup>9</sup>. Subsequent studies showed that *LMO1* overexpression significantly accelerates the latency, penetrance, and metastatic potential of *MYCN*-induced neuroblastomas in a zebrafish transgenic model<sup>10</sup>. These findings implicate *LMO1* as an oncogene that collaborates with *MYCN* in neuroblastoma pathogenesis, causing arrest of neuroblast differentiation into chromaffin cells or sympathetic ganglia within the adrenal medulla, and also driving rapid neuroblast proliferation<sup>10</sup>. However, molecular mechanisms by which *LMO1* alters transcription to drive cellular proliferation and differentiation block remain to be identified.

Recent work has suggested that a small set of transcription factors cooperate to dominate regulation of the expression program of a given cell identity through binding the majority of expressed genes/enhancers<sup>18</sup>. These factors form the core regulatory circuitry (CRC), which consists an interconnected auto-regulatory loop whereby their expression is driven by themselves and other members of the CRC<sup>5,19–21</sup>. CRC members can be identified as those that are associated with top ranked regulatory

elements by active histone marks such as Histone H3 lysine-27 acetylation (H3K27ac) signals<sup>18</sup>. As one of first examples, we have demonstrated that *TAL1*, *GATA3*, *RUNX1*, and *MYB* form the CRC in T-ALL cells<sup>22,23</sup>. In the ADRN subtype neuroblastoma, *PHOX2B*, *HAND2*, *TBX2*, *ISL1*, and *GATA3* have been implicated as CRC members<sup>5,19–21</sup>. Meanwhile, *MYCN* serves as an additional amplifier of the CRC<sup>24</sup>. However, the involvement of *LMO1* in the neuroblastoma CRC has not been elucidated yet.

Here, we identify the transcriptional signature driven by *LMO1* in combination with *MYCN* in neuroblastoma cells. *LMO1* regulates genes in a tissue-specific manner and co-binds to the enhancers of genes regulated by the CRC transcription factors that underlie cell state in ADRN neuroblastoma. Intriguingly, *LMO1* and the CRC members bind to enhancer elements and directly upregulate the *ASCL1* gene, resulting in promotion of cell growth and repression of neuronal differentiation.

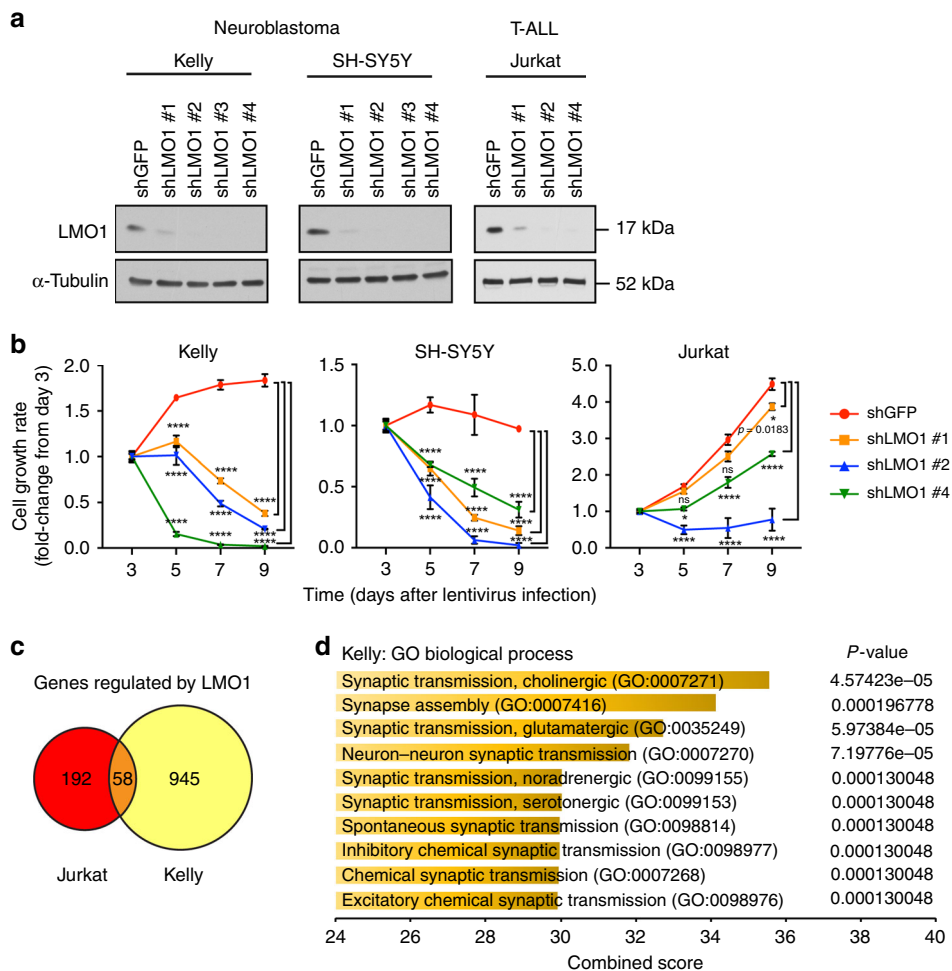
## Results

### *LMO1* regulates gene expression in a tumor-specific manner.

To begin to understand the role of *LMO1*-mediated transcriptional regulation, we first designed short-hairpin RNAs (shRNAs) to specifically knock down the expression of *LMO1* (Fig. 1a and Supplementary Fig. 1a). We chose two representative neuroblastoma cell lines (Kelly and SH-SY5Y), both of which express high levels of *LMO1* protein and harbor permissive SNP alleles<sup>8,9</sup>. To compare gene expression profiles of *LMO1*-associated malignancies, we included the T-ALL cell line Jurkat, in which *LMO1* is overexpressed due to a somatically acquired point mutation in the enhancer and participates with *TAL1* to promote an oncogenic transcriptional regulatory program<sup>22</sup>. In this setting, knockdown of *LMO1* significantly reduced cell growth in all three cell lines examined (Fig. 1b), supporting earlier findings<sup>8</sup>. *LMO1* expression level was positively associated with the cellular growth rate, in which *LMO1* shRNA #2 showed the greatest reduction of *LMO1* expression levels (Fig. 1a) and the strongest inhibition of cell growth (Fig. 1b). The growth of neuroblastoma cell lines was more strongly inhibited by *LMO1* knockdown than that of T-ALL cells, indicating that neuroblastoma cells are highly dependent on *LMO1* expression for cell growth and survival.

To analyze *LMO1*-regulated genes and pathways, we then performed RNA-sequencing (RNA-seq) analysis after knockdown of *LMO1* in Kelly and Jurkat cells using the sh*LMO1* #2. We first selected genes that were significantly downregulated or upregulated after *LMO1* knockdown compared with control samples (Supplementary Data 1). This analysis demonstrated that differentially expressed genes had limited overlap between Kelly and Jurkat cells (Fig. 1c). This is not surprising because *LMO1* does not bind directly to promoter-enhancer DNAs but rather serves as a transcriptional adapter<sup>13,25–28</sup>. Thus, target loci of *LMO1* depend on its binding to transcription factor partners that are expressed in a tissue-specific manner. Indeed, gene ontology (GO) analysis revealed that genes downregulated by *LMO1* knockdown in Kelly cells were enriched in the GO biological processes associated with synaptic transmission (Fig. 1d), while none of these biological processes were enriched in the analysis of Jurkat cells (Supplementary Fig. 1b). In contrast, different GO biological processes such as positive regulation of cell proliferation were enriched in genes differentially expressed in Jurkat cells (Supplementary Fig. 1b). These results indicate that *LMO1* regulates different sets of genes in neuroblastoma and T-ALL cells.

***LMO1* co-occupies targets with the members of the ADRN CRC.** Since *LMO1* is known to bind to transcription factors as an adapter to elicit oncogenic effects, we next examined the genome-

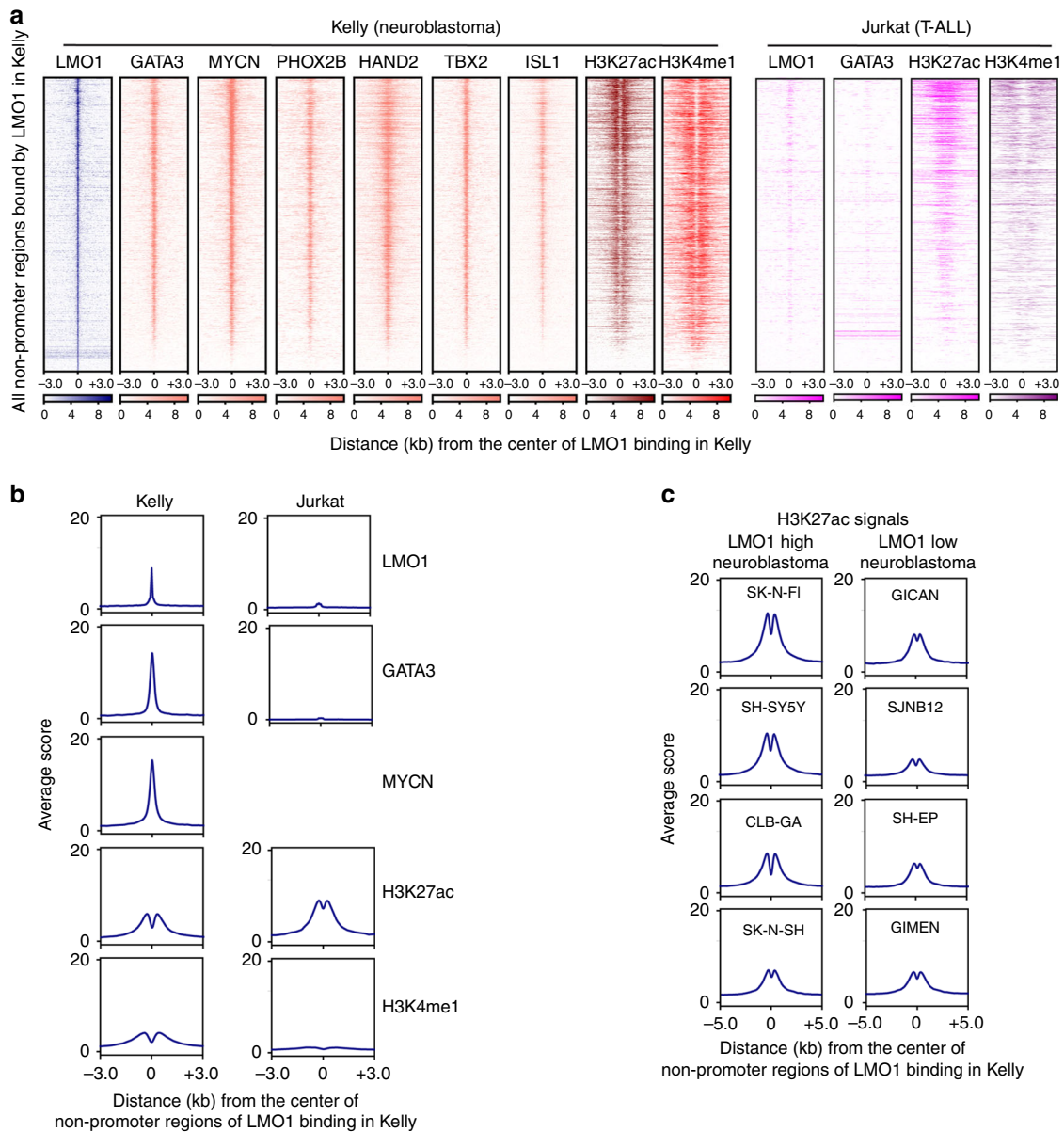


**Fig. 1** LMO1 regulates gene expressions in a tumor-specific manner. **a** Four independent shRNAs targeting *LMO1* (shLMO1 #1, 2, 3, and 4) as well as a control shRNA targeting *GFP* (shGFP) were transduced by lentivirus infection in two neuroblastoma cell lines (Kelly and SH-SY5Y) and one T-ALL cell line (Jurkat). Whole cell protein extract was harvested after 3 days of virus infection and were subjected to western blot analysis using antibodies specific for LMO1 or  $\alpha$ -tubulin (internal control). **b** Cell viability was measured after 3, 5, 7, and 9 days of lentiviral transduction of shRNA. The growth rate (fold-change) over 9 days compared with day 3 was indicated ( $n = 3$  per group). Data are represented as means  $\pm$  standard deviation (SD) for technical triplicates. The  $p$  values by two-way ANOVA (repeated measurements) followed by Tukey's multiple comparisons post hoc test are indicated. \*\*\*\* $p$  value  $< 0.0001$ . ns not significant. **c** The shRNA targeting *LMO1* or control (shGFP) was transduced into Kelly and Jurkat cells by lentiviral infection. Experiments were done in biological duplicates. Total RNAs were harvested after 3 days of infection and were subjected to RNA-seq analysis to compare gene expression profiles between two controls and two *LMO1* knockdown (KD) samples. Genes differentially expressed were selected based on the following criteria: adjusted  $p$  value  $< 0.05$ ,  $\log_2$  fold-change  $< -0.5$  or  $> 0.5$ , and TPM  $> 3$ . A Venn diagram represents the number of genes which were significantly downregulated by *LMO1* knockdown in each cell line. **d** Gene ontology analysis was performed using the gene list selected above. Top ten terms are shown with combined score. The  $p$  values by the Fisher exact test are indicated.

wide occupancy of LMO1 by chromatin immunoprecipitation sequencing (ChIP-seq) analysis in Kelly and Jurkat cells. Since a relatively high level of background signal was observed with the LMO1 antibody, we combined the results from two independent ChIP-seq experiments (Supplementary Fig. 2a). Analysis of the top DNA-binding motifs enriched at the apex of the LMO1-binding peaks indicated that these regions typically contained GATA binding motifs (Supplementary Fig. 2b). Using a co-immunoprecipitation approach, we demonstrated that FLAG-tagged LMO1 could interact with both GATA3 and LDB1 proteins when overexpressed in 293T cells (Supplementary Fig. 2c). Consistently, FLAG-tagged LMO1 could interact with endogenous GATA3 and LDB1 proteins when overexpressed in Kelly neuroblastoma cells (Supplementary Fig. 2d). This is analogous to the mechanism in T-ALL, in which LMO1 or LMO2 binds to GATA3 and LDB1<sup>25</sup>. Because GATA3 is a known member of the

CRC that underlies cell state in both T-ALL and neuroblastoma, and is required for cell growth in both tumor types<sup>19,22</sup>, we next analyzed the co-occupancy of LMO1 with GATA3 by ChIP-seq.

We first selected all nonpromoter regions bound by LMO1 in Kelly cells and then analyzed the occupancy of GATA3 at the same regions. We also compared the result with DNA occupancy of MYCN, which is a collaborating oncogenic transcription factor of LMO1<sup>10,19</sup>. Strikingly, we observed a remarkable concordance of DNA occupancy by LMO1, GATA3, and MYCN proteins in Kelly cells (Fig. 2a, left; Fig. 2b, left). In addition, LMO1 occupancy was associated with enrichment of transcription factors that have previously been shown as members of the ADRN neuroblastoma CRC, including PHOX2B, HAND2, TBX2, and ISL1<sup>5,19-21</sup> (Fig. 2a, left). This result suggested that LMO1 collaborates with these factors to co-occupy the regulatory region of the same target genes. However, there was very little overlap in



**Fig. 2** LMO1 co-occupies targets with the members of the ADRN CRC. **a** The LMO1-bound nonpromoter loci were first identified in Kelly cells. Density plots show the distribution of LMO1, GATA3, MYCN, PHOX2B, HAND2, TBX2 and ISL1, H3K27ac and H3K4me1 signals at the LMO1-bound regions ( $\pm 3$  kb from binding sites) in Kelly and LMO1, GATA3, H3K27ac and H3K4me1 signals at the same regions in Jurkat cells. The color scale shows the intensity of the distribution signal. **b** Metagene plots show distribution of LMO1, GATA3, MYCN, H3K27ac, and H3K4me1 signals at the LMO1-bound regions ( $\pm 3$  kb from binding sites) in Kelly and Jurkat cells. **c** Metagene plots show distribution of H3K27ac signals at the LMO1-bound regions ( $\pm 5$  kb from binding sites) in various neuroblastoma cell lines. Cell lines were classified based on the expression level of *LMO1* mRNA.

these LMO1-bound regions between Kelly neuroblastoma and Jurkat T-ALL cells (Fig. 2a, right; Fig. 2b, right; and Supplementary Fig. 2e), indicating that LMO1 and GATA3 bind different lineage-specific enhancers.

We next analyzed the chromatin status at LMO1 target loci in neuroblastoma cells and found that LMO1-enriched regions were frequently associated with active histone marks (H3K27ac and H3K4me1) in Kelly cells (Fig. 2a, b). We also compared H3K27ac signals at regions bound by LMO1 in Kelly cells across a panel of neuroblastoma cell lines without *MYCN* amplification that expressed different levels of *LMO1* (Fig. 2c and Supplementary Fig. 2f). This analysis revealed that cell lines that expressed high levels of *LMO1* exhibited higher levels of H3K27ac signal at LMO1-bound regions, suggesting that these cell lines have a more similar cell identity to Kelly cells, as compared with the cell lines

which express low levels of *LMO1*. These results indicated that LMO1 associates with distinct genomic regions in neuroblastoma and T-ALL cells, which primarily coincide with chromatin marks of activated gene expression.

***ASCL1* is a downstream target gene regulated by LMO1 and MYCN.** We next sought to identify the individual target genes directly regulated by LMO1 in neuroblastoma cells. By integrating the results of ChIP-seq and RNA-seq analyses in Kelly cells, we predicted genes that were likely direct targets of LMO1 in which they show significant downregulation by RNA-seq after knock-down of *LMO1* and were also associated with binding sites for LMO1 and GATA3 proteins with an active histone mark (H3K27ac) by ChIP-seq (Supplementary Data 2). To confirm

whether the expression of LMO1-bound genes was regulated by LMO1 in other cell line, we also performed gene expression profiling in SH-SY5Y cells by microarray after knockdown of LMO1 using shLMO1#2. The gene set enrichment analysis (GSEA) demonstrated that many predicted LMO1 targets that were downregulated or upregulated by LMO1 knockdown in Kelly cells were also downregulated or upregulated by LMO1 knockdown in SH-SY5Y cells (Supplementary Fig. 3a, b). This result showed a high level of correlation in gene expression changes between two cell lines.

Next, to identify genes most strongly dependent on MYCN for their expression among all direct targets of LMO1, we performed RNA-seq after MYCN knockdown (Supplementary Fig. 3c, d) and predicted genes that were directly regulated by MYCN in Kelly cells using the same criteria as for LMO1 predicted targets (Supplementary Data 2). By combining these filters, we derived a short list of genes ( $n = 42$ ), which are displayed rank-ordered by fold-change after LMO1 knockdown as an expression heatmap in Fig. 3a. Interestingly, the transcription factor gene with the highest fold-change in response to both MYCN and LMO1 depletion was *ASCL1* (Fig. 3a). *ASCL1* encodes a bHLH transcription factor that has previously been implicated in neuronal cell development and neuroblastoma pathogenesis<sup>28,29</sup>. *ASCL1* expression is more highly expressed in neuroblastoma and small cell lung cancer cell lines than other cancer types in the CCLE database (Supplementary Fig. 3e). In our recent CRISPR/Cas9 screen, the *ASCL1* gene also showed selective dependency in neuroblastoma cell lines<sup>19</sup>. This gene was also significantly downregulated by LMO1 knockdown in SH-SY5Y cells (Supplementary Fig. 3f). In support of these findings, neuroblastoma cell lines displayed high level expression of ASCL1 together with LMO1 by western blot analysis; however, ASCL1 was not expressed by any T-ALL cell lines examined (Fig. 3b). It is noteworthy that ASCL1 protein expression was higher in the MYCN-amplified neuroblastoma cell lines, Kelly and CHP-134, than in the nonamplified cell lines, SK-N-SH and its subclone SH-SY5Y (Fig. 3b), which overexpress MYC (*c-MYC*)<sup>30</sup>. Consistently, analysis of primary neuroblastoma samples demonstrated that MYCN-amplified cases showed a higher level of ASCL1 expression than nonamplified cases (Fig. 3c). These results indicated that ASCL1 is a high-confidence target gene downstream of LMO1 and MYCN in neuroblastoma cells.

**High ASCL1 expression is associated with inferior survival.** Notably, analysis using the R2 database indicated that high ASCL1 expression was significantly associated with inferior survival of neuroblastoma patients in a cohort reported by Kocak et al.<sup>31</sup> (Fig. 3d). Two other cohorts reported by Versteeg et al.<sup>32</sup> and the Neuroblastoma Research Consortium (NRC) also showed the same trend (Supplementary Fig. 3g, h), although these datasets did not show high statistical significance due to the small size of samples in cases with high ASCL1 expression. Importantly, high ASCL1 expression was significantly associated with inferior survival even in the absence of MYCN amplification (Fig. 3e). These data implicate ASCL1 as a prognostic factor in neuroblastoma.

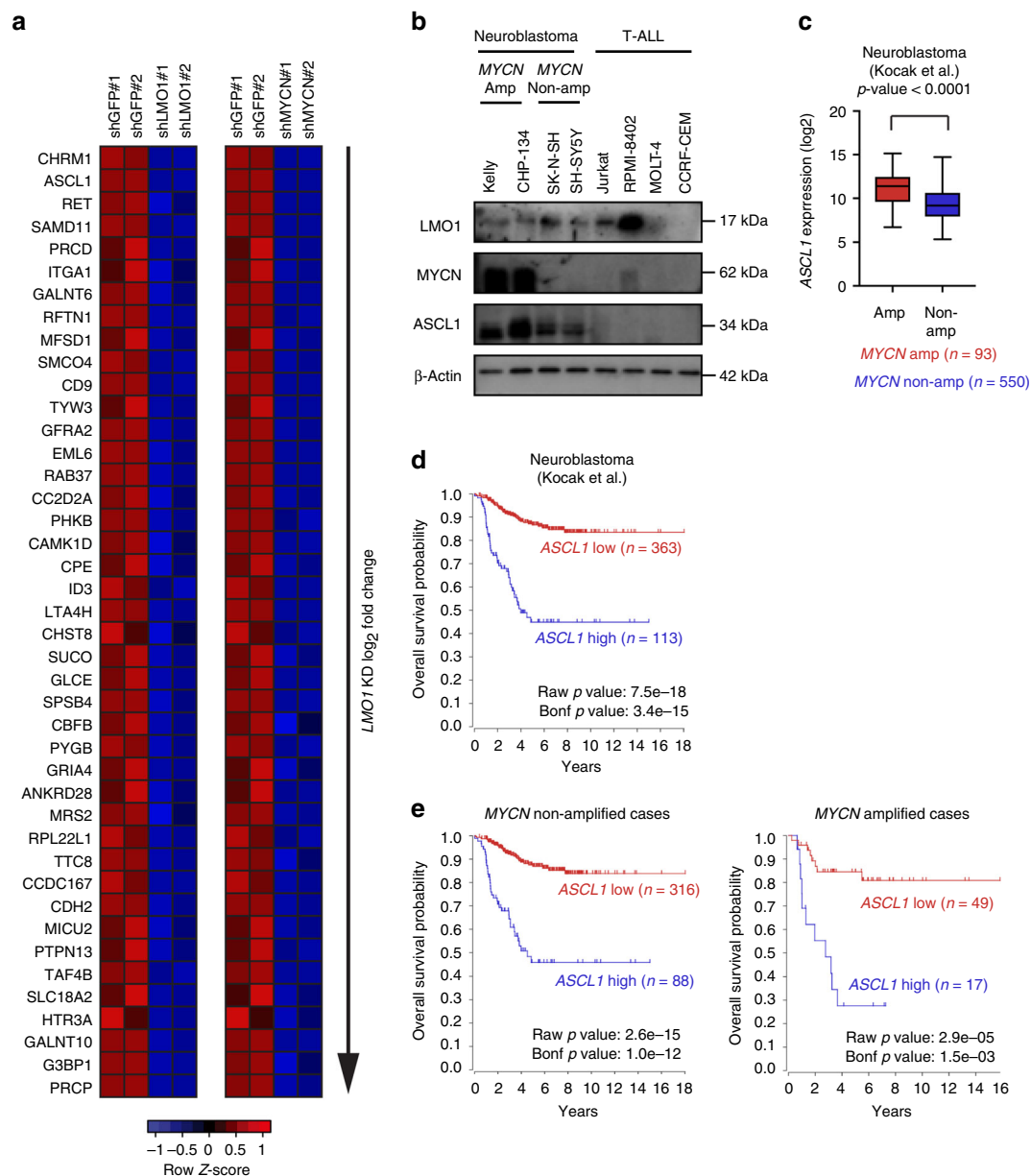
**LMO1, GATA3, and MYCN occupy the regulatory element of ASCL1.** To verify that ASCL1 is directly regulated by LMO1, we analyzed the genomic loci bound by LMO1 in neuroblastoma cells. By ChIP-seq analysis, we observed four peaks of LMO1 (+55, +93, +141, and +143 kb) located 3' to the ASCL1 gene and associated with H3K27ac and H3K4me1 signals in Kelly cells (Fig. 4a, bottom). Similar H3K27ac profiles were observed at the same loci in multiple ADRN subtype neuroblastoma cell lines with high level of LMO1 expression but not in the MES subtype

neuroblastoma cell lines with low level of LMO1 expression or in the T-ALL Jurkat cell line (Supplementary Fig. 4a). Notably, only the +141 and +143 kb elements (peak 3a and 3b) were exceptionally enriched in binding of bound by LMO1, GATA3, and MYCN proteins (Fig. 4a, middle). The same loci were also co-occupied by several other transcription factor proteins (HAND2, PHOX2B, TBX2, and ISL1) that represent CRC members of the ADRN neuroblastoma<sup>5,19–21</sup>. We have confirmed the enrichment of GATA3 protein at the same element by ChIP-PCR in Kelly and SH-SY5Y cells (Supplementary Fig. 4b).

To further clarify whether these regions serve as distal enhancers regulating ASCL1 expression, we engineered Kelly cells stably expressing the catalytically-dead Cas9 (dCas9) protein fused to a transcriptional repressor peptide (KRAB) (Fig. 4b). We then transduced this clone with doxycycline (DOX)-inducible single guide RNAs (sgRNAs) targeting each of these peaks. This system permits guidance to blockage of specific genomic regions by dCas9 protein coupled with repression of transcription in surrounding regions by the KRAB repressor activity. We first confirmed that dCas9 protein bound to specific sites, directed by sgRNAs, by ChIP-qPCR (Supplementary Fig. 4c). In this setting, the blockage of each enhancer region resulted in the reduction of ASCL1 expression at the protein (Fig. 4c) and mRNA levels (Fig. 4d), indicating that each of these four enhancer regions contributes to ASCL1 expression. Knockdown of LMO1, GATA3, and MYCN each separately downregulated the expression of ASCL1 mRNA (Fig. 4e and Supplementary Fig. 4d). Downregulation of ASCL1 protein expression was observed in multiple LMO1-positive neuroblastoma cell lines after LMO1 knockdown (Fig. 4f). Conversely, overexpression of LMO1 upregulated ASCL1 expression (Supplementary Fig. 4e, f). Taken together, these results indicated that LMO1 positively regulates the expression of ASCL1 in concert with the CRC members GATA3 and MYCN.

**LMO1 and MYCN directly regulate the RET gene.** Besides ASCL1, we found that the receptor tyrosine kinase gene *RET* was positively regulated by LMO1 and MYCN in neuroblastoma cells (Fig. 3a). *RET* has been implicated in neuroblastoma and shown to promote cell proliferation<sup>33,34</sup>. In our ChIP-seq analysis, multiple binding events of LMO1, GATA3, and MYCN proteins were observed upstream of the *RET* gene locus (Fig. 5a). The expression of *RET* was downregulated after knockdown of LMO1 at mRNA (Fig. 5b) and protein (Fig. 5c) levels in Kelly and SH-SY5Y cells. Phosphorylation of ERK1/2, which is a downstream effector of the MAP kinase pathway activated by RET kinase phosphorylation<sup>35</sup>, was decreased after knockdown of *RET* in the Kelly cells (Fig. 5d, e). Furthermore, knockdown of *RET* inhibited the growth and increased the fraction of Sub-G1 population in Kelly cells (Supplementary Fig. 5a, b), consistent with observations reported by others<sup>34</sup>. These results indicated that LMO1 and MYCN can promote cell growth and proliferation through upregulation of *RET*. However, it is noted that several LMO1-high cell lines did not express RET, whereas some LMO1-low cell lines also expressed this protein (Supplementary Fig. 5c). This suggested that LMO1 or MYCN may not be the primary determinant to induce *RET* expression, but they can positively regulate this gene when cooperating with other factors. Therefore, we highlighted ASCL1 for the downstream analysis.

**ASCL1 is a member of the CRC of ADRN neuroblastoma.** Since LMO1 directly regulated ASCL1 expression, we sought to understand the role of ASCL1 in mediating LMO1-driven oncogenesis. ASCL1 has been known to be regulated in a stage-specific manner during neuronal differentiation<sup>28,29</sup> (Fig. 6a). Analysis of available gene expression datasets indicated that

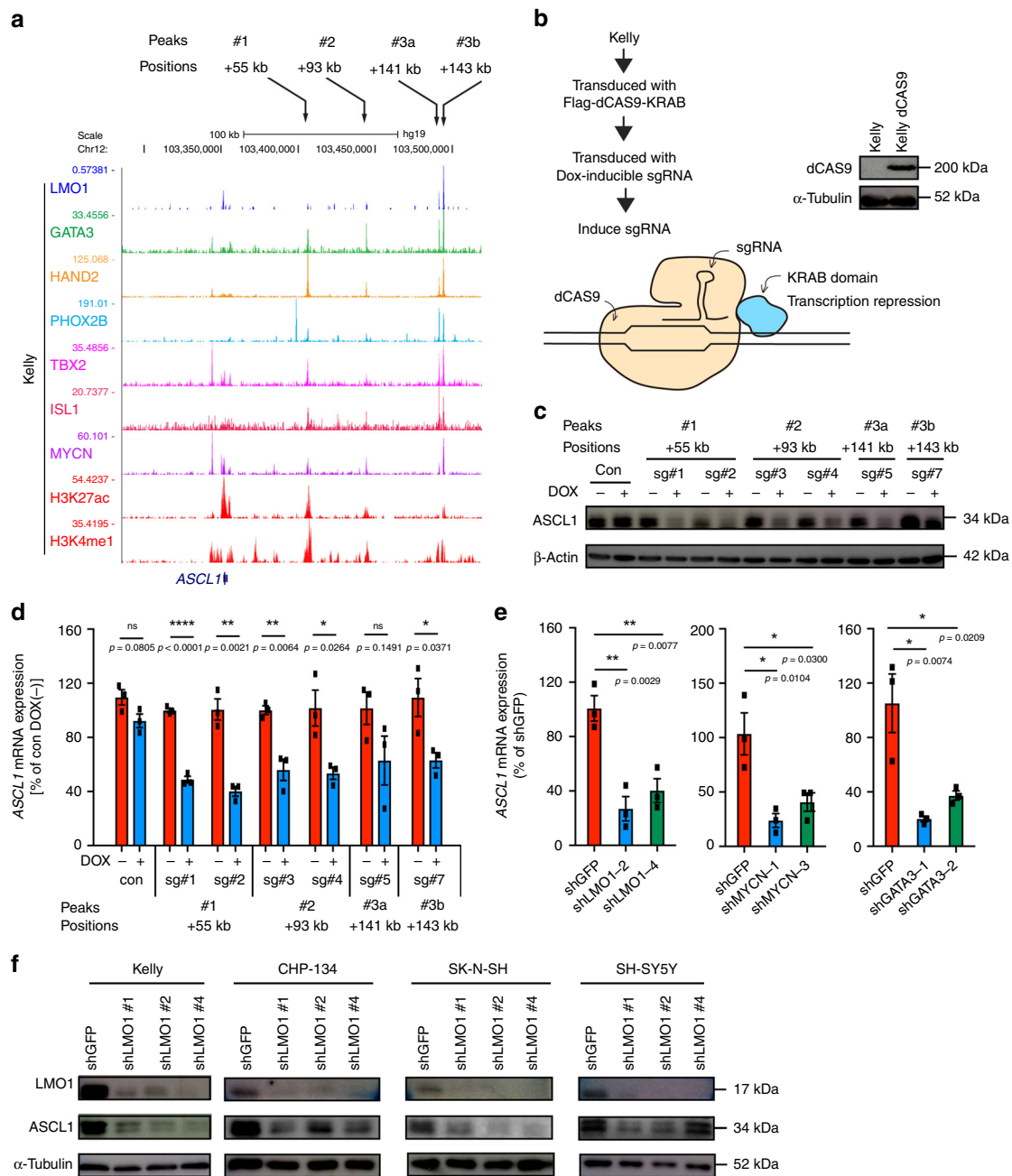


**Fig. 3** *ASCL1* a high-confidence *LMO1* target in neuroblastoma cells. **a** Heatmap image represents gene expression changes of selected genes after knockdown of *LMO1* or *MYCN* in Kelly cells. The genes were ordered based on log<sub>2</sub> fold-change values between control and *LMO1* knockdown. The color scale represents row z-scores. **b** Protein expression of *LMO1*, *MYCN*, and *ASCL1* in a panel of neuroblastoma and T-ALL cell lines were analyzed by western blot.  $\beta$ -actin was used as an internal control. *MYCN*-amplified (amp) and nonamplified (non-amp) cell lines are indicated. **c** A total of 643 primary neuroblastoma cases in the Kocak cohort<sup>31</sup> were classified into two groups: with *MYCN* amplification (n = 93), and without *MYCN* amplification (n = 550). *ASCL1* mRNA expression were shown in each group. Data are represented as box plots where the middle line indicates the median, the lower and upper hinges correspond to the first and third quartiles, the lowest datum indicates the minimum (within the 1.5 IQR of the lower quartile), and the highest datum indicates the maximum (within the 1.5 IQR of the upper quartile). The p values by the Mann-Whitney test are indicated. **d** The survival curve analysis for primary neuroblastoma samples in the Kocak cohort<sup>31</sup> was done using the R2 database. The samples for which the prognostic data are available were classified into two groups (*ASCL1*-high and low) by the Kaplan Scan method, which calculates the optimum cutoff based on statistical testing. The raw p value and Bonferroni-corrected p value are shown. The Kaplan-Meier curve analysis for each group was done using the R2 database. **e** Primary neuroblastoma cases for which the prognostic data are available from the Kocak cohort<sup>31</sup> were classified into two groups: without *MYCN* amplification (n = 404; left), and with *MYCN* amplification (n = 66; right).

human *ASCL1* expression was elevated in neural progenitor cells, compared with less differentiated neural crest cells or more differentiated chromaffin cells in the mature adrenal gland (Fig. 6b). Furlan et al. also reported that mouse *Ascl1* expression begins before the lineage specification that gives rise to progenitor cells that differentiate into sympathoblasts and chromaffin cells<sup>36</sup>. At the stage of E11.5, sympathoblasts start downregulating the

expression of *Ascl1*, whereas chromaffin cells start downregulating the expression of *Ascl1* after E13.5. These observations indicated that *ASCL1* is expressed at higher levels in neural progenitor cells.

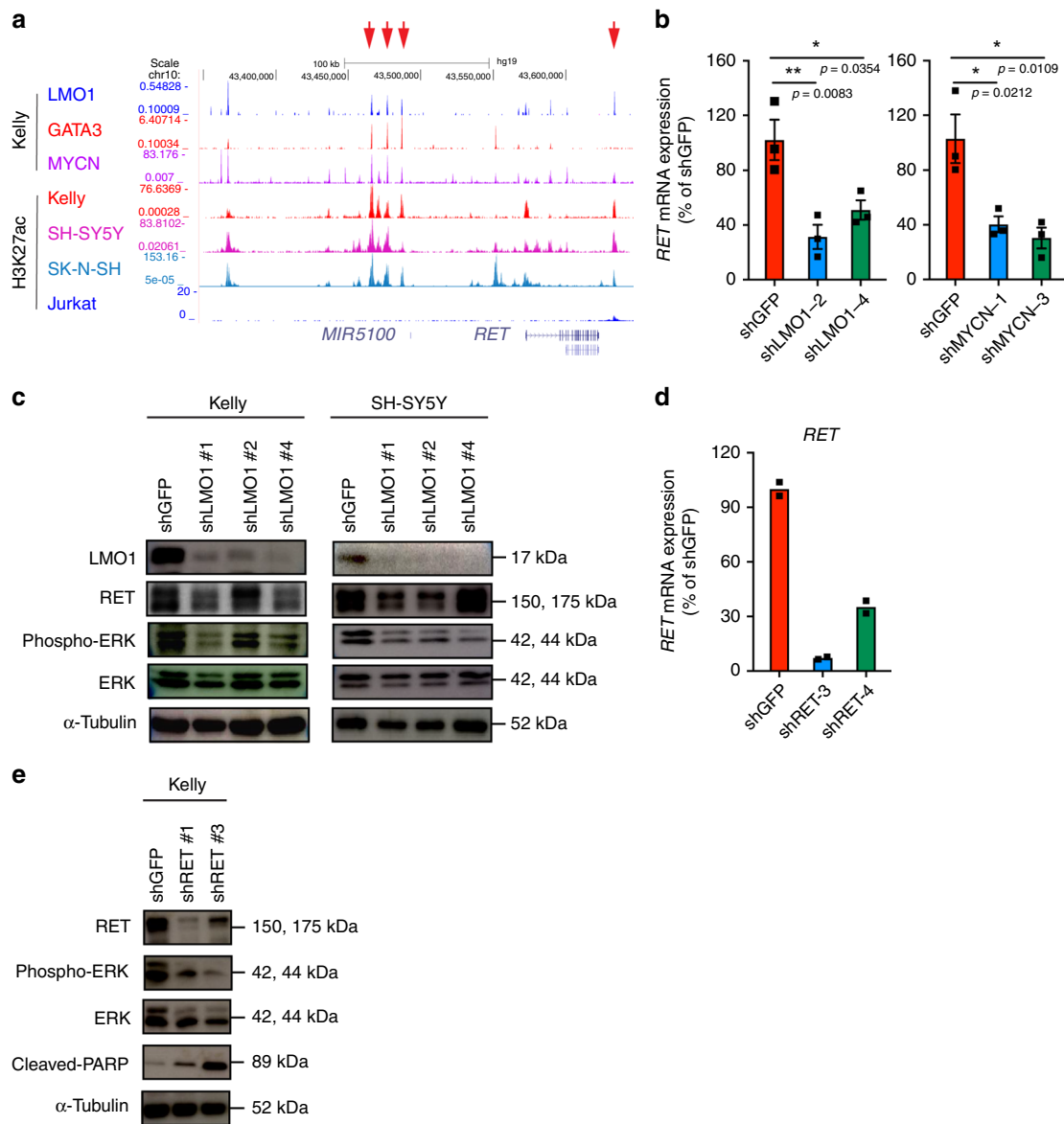
Importantly, *ASCL1* expression was high in primary neuroblastoma cells (Fig. 6b), consistent with the origin of most neuroblastomas from sympathetic nervous system progenitors.



**Fig. 4** ASCL1 is directly regulated by LMO1 and MYCN in neuroblastoma cells. **a** ChIP-seq gene tracks showing the binding locations of various transcription factors at the *ASCL1* gene locus in Kelly cells. Black arrows (top) indicate regions associated with H3K27ac signals in Kelly cells (peaks 1, 2, 3a, and 3b). **b** Kelly cells were first transduced with the dCas9-KRAB protein. sgRNAs (sg#1-7) targeting each of peaks bound by LMO1 (peaks 1, 2, 3a and 3b) were then induced by the treatment with doxycycline (DOX) (right). Protein expression of dCas9 was confirmed by western blot (left).  $\alpha$ -tubulin was used as an internal control. **c** Protein expression of ASCL1 after the induction of each sgRNA was analyzed by western blot. **d** mRNA expression of *ASCL1* after the induction of each sgRNA was analyzed by qRT-PCR. Expression was normalized to *spike-in* control RNA and shown as a percent relative to untreated control samples. Data are represented as means  $\pm$  standard error of the mean (SEM) for three biological replicates. The *p* values by two-tailed unpaired *t* test are indicated. **e** mRNA expression of *ASCL1* after knockdown of *LMO1*, *MYCN*, and *GATA3* in Kelly cells were analyzed by qRT-PCR. Expression was normalized to *spike-in* control RNA and shown as a percent relative to control shGFP. Data are represented as means  $\pm$  SEM for three biological replicates. The *p* value by one-way ANOVA followed by Tukey's multiple comparisons post hoc test are indicated. **f** Protein expression of LMO1 and ASCL1 were analyzed after transduction of different LMO1 shRNAs in five different neuroblastoma cell lines.

This was further supported by the fact that *ASCL1*, *LMO1*, and *MYCN* were each more highly expressed in the ADRN subtype of neuroblastoma cell lines corresponding to sympathetic progenitors, but less in cell lines of the MES subtype, which appear to retain features of earlier migratory neural crest cells (Supplementary Fig. 6a). ADRN cases expressed higher levels of CRC

member genes (*HAND2*, *PHOX2B*, *TBX2*, and *ISL1*) than the MES subtype, whereas MES cases commonly expressed *MYC* (*c-MYC*). Consistent with these findings, analysis of our previously established neuroblastoma zebrafish models<sup>30</sup> demonstrated that mRNA expression levels of the zebrafish orthologues, namely *ascl1a*, *lmo1*, *phox2bb*, *hand2*, *tbx2a*, and *isl1l*, were higher in



**Fig. 5** *RET* is directly regulated by *LMO1* in neuroblastoma cells. **a** ChIP-seq gene tracks showing the binding locations of various transcription factors at the *RET* gene locus in Kelly cells. Red arrows (top) indicate regions associated with H3K27ac signals in Kelly cells. **b** mRNA expressions of *RET* after knockdown of *LMO1* and *MYCN* in Kelly cells were analyzed by qRT-PCR. Data are represented as means  $\pm$  SEM for three biological replicates. The *p* value by one-way ANOVA followed by Tukey's multiple comparisons post hoc test are indicated. **c** Protein expressions of *RET*, *LMO1*, p-ERK, and total ERK were analyzed after transduction of different *LMO1* shRNAs in three different neuroblastoma cell lines.  $\alpha$ -tubulin was used as an internal control. **d** The mRNA expression of *RET* after *RET* knockdown were measured by qRT-PCR. Data are represented as mean of technical duplicates. The *p* values by one-way ANOVA followed by Tukey's multiple comparisons post hoc test are indicated. **e** Protein expressions of *RET*, p-ERK, total ERK, and cleaved PARP were analyzed after transduction of different *RET* shRNAs in Kelly neuroblastoma cell line.  $\alpha$ -tubulin was used as an internal control.

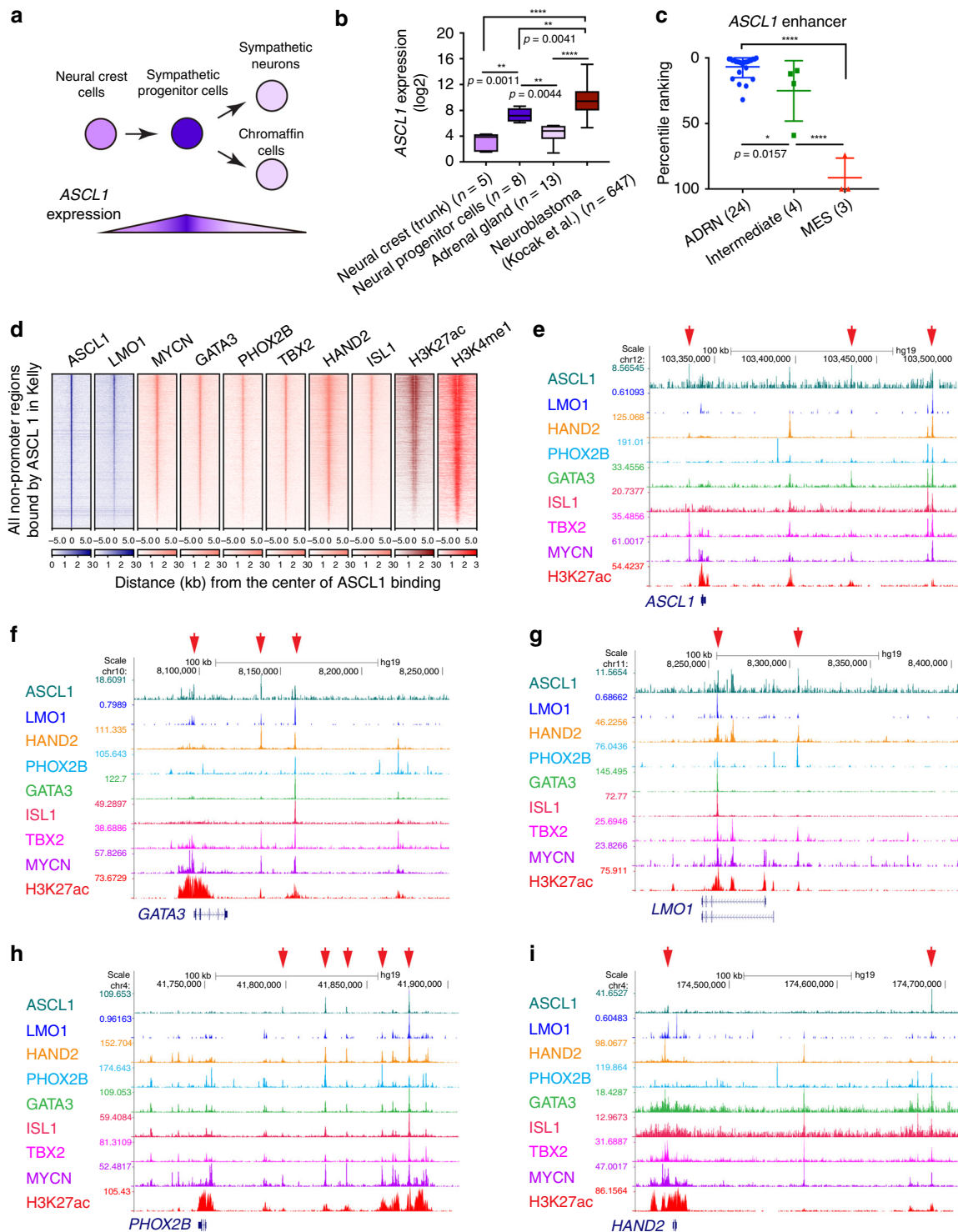
*MYCN*-induced tumors compared with the *c-MYC*-induced tumors (Supplementary Fig. 6b). Furthermore, the *ASCL1* gene locus itself was associated with a regulatory region highly enriched for H3K27ac, indicative of an enhancer ranked in top 10% of enhancers in ADRN subtype cell lines but not in MES subtype cell lines (Fig. 6c). This result suggested that *ASCL1* is possibly involved in the ADRN CRC in neuroblastoma cells.

In fact, ChIP-seq analysis demonstrated that *ASCL1* frequently co-occupies regions with *LMO1*, *MYCN* and members of the ADRN CRC in Kelly cells (Fig. 6d). *ASCL1* is bound to regions within regulatory elements of CRC genes at sites that were also bound by *LMO1*, *HAND2*, *PHOX2B*, *GATA3*, *ISL1*, *TBX2*, and *MYCN* in Kelly cells (Fig. 6e-i). The E-box motif that is recognized by type II bHLH proteins, such as *ASCL1*, as heterodimers with

type I bHLH proteins, was significantly enriched in the *ASCL1*-bound regions (Supplementary Fig. 6c). These results indicated that *ASCL1* is a member of CRC in ADRN neuroblastoma.

Of note, analysis of a dataset from single cells during neural crest development by Furlan et al.<sup>36</sup>, showed that mouse *Ascl1* is expressed by a large fraction of the bridging cell population between Schwann cell precursors and chromaffin cells (Supplementary Fig. 6d). There was some overlap of *Phox2b*, *Hand2*, and *Isl1* expression with cells expressing *Ascl1*, although there was not a significant correlation between cells expressing the highest expression of *Ascl1* and expression levels of ADRN CRC genes. *Lmo1*, *Mycn*, *Tbx2*, and *Gata3* were only expressed by a small fraction of these cells. Thus, this bridging cell population is not governed by the ADRN neuroblastoma CRC. Other regulatory





**Fig. 6** *ASCL1* is a member of CRC in the ADRN subtype of neuroblastoma. **a** Schematic image for neural differentiation and *ASCL1* expression. **b** The mRNA expression of *ASCL1* in normal neural crest, neural progenitor cells, adrenal gland, and neuroblastoma tumors (Kocak cohort<sup>31</sup>) were analyzed by the R2 database. The sample numbers are indicated. Data are presented as box plots where the middle line indicates the median, the lower, and upper hinges correspond to the first and third quartiles, the lowest datum indicates the minimum (within the 1.5 IQR of the lower quartile), and the highest datum indicate the maximum (within the 1.5 IQR of the upper quartile). The *p* value by one-way ANOVA followed by Tukey's multiple comparisons post hoc test are indicated in **b** and **c**. \*\*\*\**p* value < 0.0001. **c** The dot plot for enhancer ranking in neuroblastoma cell lines (24 ADRN subtype neuroblastoma cell lines, 4 intermediate subtype neuroblastoma cell lines, 3 MES subtype neuroblastoma cell lines). Enhancers were ranked based on H3K27ac signal from high to low. Data are represented as means ± SEM. **d** The *ASCL1*-bound gene loci were first selected in Kelly cells. Density plots show the distribution of *LMO1*, *MYCN*, *GATA3*, *H3K27ac*, and *H3K4me1* signals at the *ASCL1*-bound regions (±5 kb from binding sites) in Kelly cells. The color scale shows the intensity of the distribution signal. ChIP-seq gene tracks showing the binding locations of *ASCL1* together with various transcription factors at the *ASCL1* (**e**), *GATA3* (**f**), *LMO1* (**g**), *PHOX2B* (**h**), and *HAND2* (**i**) gene loci in Kelly cells. Red arrows (top) indicate *ASCL1*-bound peak regions, which were determined by peak calling.

mechanisms may drive high levels of *Ascl1* expression in this bridging cell population.

**ASCL1 is involved in neuroblastoma differentiation status.** Lastly, we analyzed genes and pathways controlled by ASCL1 in neuroblastoma cells. We performed an RNA-seq analysis after *ASCL1* knockdown in Kelly cells (Fig. 7a) and selected genes that were significantly downregulated after *ASCL1* knockdown compared with control samples. GO term analysis revealed that many pathways involved in the sympathetic nervous system development and noradrenergic lineage commitment were enriched for the genes downregulated by *ASCL1* knockdown (Fig. 7b).

Importantly, knockdown of *ASCL1* resulted in the increased expression of genes that are normally upregulated during the transition from neural progenitor to differentiated neuronal cells, such as *NTRK1/TrkA* and *NPY*. The result was independently validated by qRT-PCR (Fig. 7c). As a positive control, proneuronal differentiation stimuli, such as all-*trans* retinoic acid (ATRA)<sup>37</sup>, also resulted in upregulation of these genes (Supplementary Fig. 7a). In addition, *IGF2*, which has been reported as a target of ASCL1<sup>38</sup>, was upregulated after *ASCL1* knockdown (Supplementary Fig. 7b). In contrast, *ASCL1* knockdown resulted in a downregulation of several genes that are highly expressed in normal neuronal progenitor cells and are members of the neuroblastoma CRC, such as *PHOX2B* and *HAND2*. The result was independently validated by qRT-PCR (Fig. 7d). These findings indicate that ASCL1 affects the cellular differentiation status in neuroblastoma cells.

Previous studies have shown that phosphorylation also modulates the function of ASCL1<sup>36,44</sup>, and that the phosphorylated form of ASCL1 is associated with undifferentiated neuronal cells. Indeed, ASCL1 protein bands migrated faster after the treatment with lambda phosphatase in all cell lines examined (Supplementary Fig. 7c), suggesting that the predominant fraction is likely phosphorylated supporting the previous study by others<sup>36,44</sup>. Conversely, the treatment of Kelly cells with Calyculin A, a protein phosphatase inhibitor, increased this fraction (Supplementary Fig. 7d). These results suggested that the ASCL1 protein is constitutively phosphorylated in neuroblastoma cells, affecting cell differentiation status and resulting in differentiation arrest of neuroblastoma cells.

In fact, knockdown of *ASCL1* promoted axonal extension in the presence of NGF (Fig. 7e), which is a primary ligand for NTRK1. This phenotype was intensified with the presence of low dose of ATRA. *ASCL1* knockdown moderately inhibited the growth of Kelly cells (Fig. 7f), which was accompanied with a relative increase in percentage of cells in sub-G1 (Supplementary Fig. 7e). Similarly, the growth of SH-SY5Y cells, which does not possess *MYCN* amplification, was also inhibited by *ASCL1* knockdown (Supplementary Fig. 7f, g). However, *ASCL1* overexpression did not rescue the cell growth inhibition caused by *LMO1* knockdown (Supplementary Fig. 7h–j), indicating that the effects of the CRC in maintaining neuroblastoma cell growth and survival depend on high levels of expression of each member of the CRC and the *LMO1* CRC coregulator acting in concert.

Taken together, our results indicate that ASCL1 contributes to the CRC with the other members to ensure viability and establish the ADRN neuroblastoma cell state (Fig. 8). In this context, *LMO1* is essential for the formation of CRC and ASCL1 expression. Because *LMO1* does not contain a DNA-binding domain, it does not qualify as a *bona fide* CRC transcription factor<sup>18</sup>. Rather, it appears to operate as an auxiliary transcriptional cofactor involved in protein–protein interactions essential for CRC activities in the regulation of cell identity and the malignant cell state in neuroblastoma cells.

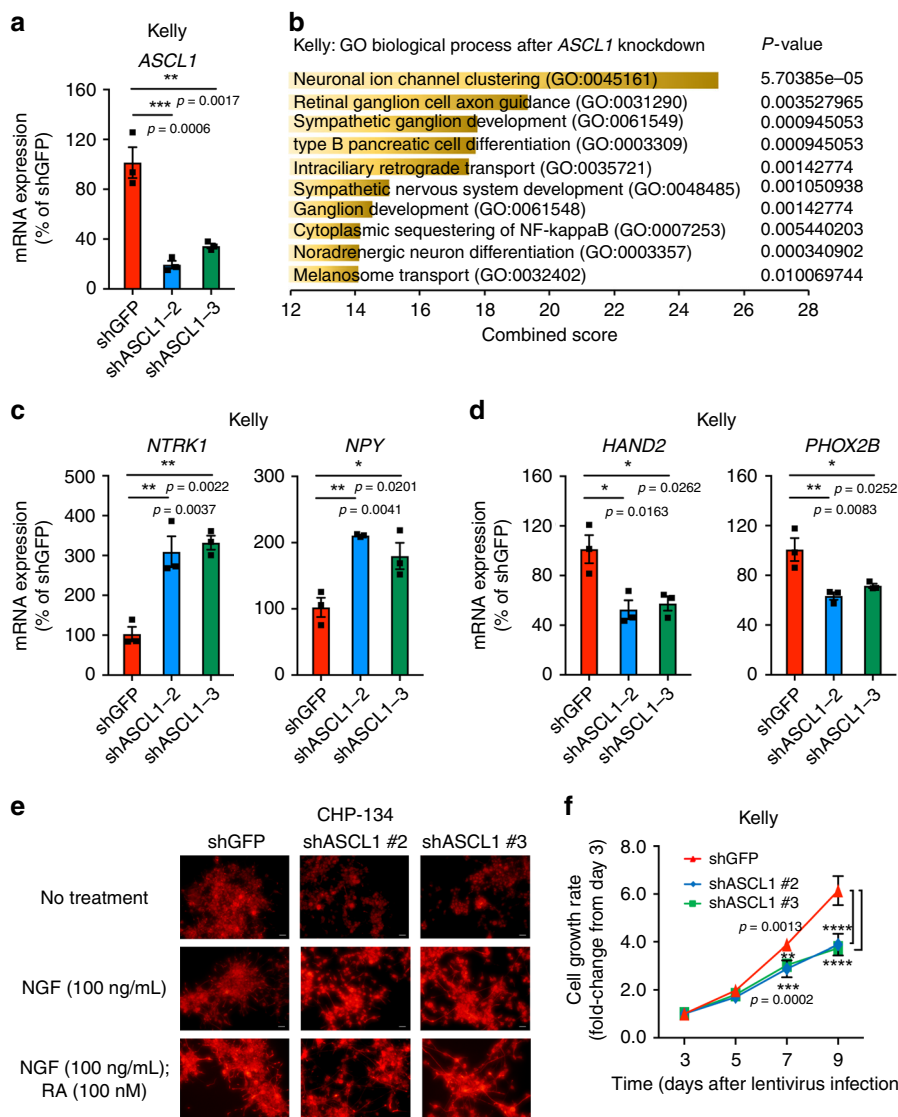
## Discussion

Accumulating evidence implicates *LMO1* as a critical factor in the ADRN subtype of neuroblastoma. A large-scale neuroblastoma GWAS study demonstrated that a polymorphism within the *LMO1* gene locus is strongly associated with the susceptibility of children to develop neuroblastoma<sup>8</sup>. The crucial SNP mediating this neuroblastoma phenotype consists of an ancestral G permissive allele that forms a GATA DNA-binding motif<sup>9</sup>. GATA3 then participates with other CRC transcription factors to generate a potent enhancer that activates *LMO1* gene expression. The protective T allele at this SNP, which appeared during evolution in human populations, does not support GATA3 binding. As a member of the ADRN CRC, GATA3 binding is essential for the super-enhancer to form that drives high levels of *LMO1* expression. Because *LMO1* lacks a DNA-binding domain, it cannot be a member of the ADRN CRC in neuroblastoma, but rather it serves as an essential transcriptional regulatory cofactor for formation of the autoregulatory loop that comprises the CRC, and the absence of *LMO1* expression leads to decreased levels of each of the ADRN CRC transcription factors.

Importantly, here we provide an insight into the mechanisms by which *LMO1* contributes to the ADRN neuroblastoma pathogenesis. Our study shows that *LMO1* is essential for proper expression of *ASCL1*, a newly identified ADRN CRC member. It is noteworthy that we and other groups independently identified *PHOX2B* and *HAND2* as key components of the CRC in ADRN neuroblastoma cases<sup>5,19,21</sup>. Our current results demonstrate that ASCL1 binds to the associated enhancers and regulates the expression of CRC members—*PHOX2B*, *HAND2*, *GATA3*, *TBX2*, and *ISL1*—in concert with *LMO1* and *MYCN*. Each of these transcription factors are highly expressed in the ADRN subtype of neuroblastoma as compared with the MES subtype. ASCL1 co-binds with *LMO1* at the loci of binding of the CRC transcriptional protein complex within regulatory elements of each CRC gene. This mechanism likely relates to the established proclivity of *LMO1* to participate in transcriptional complexes containing the class II bHLH proteins (*TAL1*, *TAL2*, or *LYL1*) observed in T-ALL<sup>13</sup>. Given that ASCL1 is also a selective neuroblastoma dependency gene whose locus is bound by the CRC delineated by our prior work<sup>19</sup>, and whose product binds coordinately with the ADRN CRC members to enhancers controlling each of the CRC members and genes of the extended regulatory network, we propose that ASCL1 is also an essential transcription factor and member of the feed-forward autoregulatory loop that comprises the CRC in the ADRN subtype of neuroblastoma.

In this context, *MYCN* further promotes gene expression program induced by the ADRN CRC in neuroblastoma cells. A recent study suggests that “enhancer invasion” shapes *MYCN*-dependent transcriptional upregulation in neuroblastomas with *MYCN* gene amplification<sup>24</sup>. Loss of *MYCN* leads to a global reduction in transcription, which is more evident at the *MYCN* target genes with the highest enhancer occupancy. In particular, tissue-specific enhancers define highly tumor-specific *MYCN* target genes. In fact, we observed that *MYCN* frequently co-occupies the enhancers of target genes with *LMO1*, including those of CRC members, which characterize the ADRN subtype neuroblastoma. Hence, when expressed at high levels, *MYCN* serves as an “enhancer invader” that reinforces the gene expression program of the entire adrenergic CRC, including *LMO1* as a coregulator of this CRC.

During normal peripheral nervous system development, *ASCL1* is expressed in sympathetic progenitor cells and extinguished when cells become differentiated into neurons or chromaffin cells<sup>28,29</sup>. Narayanan et al. reported that ASCL1 not only promotes the acquisition of a proneural phenotype in glioblastoma cells but also represses mesenchymal features<sup>39</sup>. A

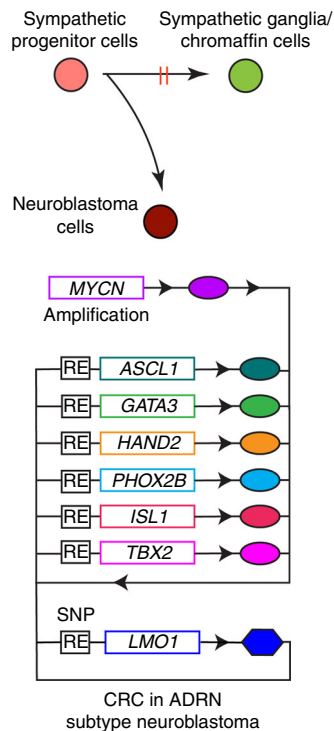


**Fig. 7** ASCL1 is required for cell growth and regulates neural differentiation status. **a** Two different shRNAs targeting *ASCL1* (#2, 3) were transduced into Kelly cells. The mRNA expression of *ASCL1* was measured by qRT-PCR, and shown as a percent relative control shGFP. Data are represented as means  $\pm$  SEM for three biological replicates. The *p* value by one-way ANOVA followed by Tukey’s multiple comparisons post hoc test are indicated in **a**, **c**, **d**, and **f**.  $***p$  value  $< 0.0001$ . **b** The shRNA targeting *ASCL1* or control (*GFP*) was transduced into Kelly by lentiviral infection. Experiments were done in biological duplicates. Total RNAs were harvested after 3 days of infection and were subjected to RNA-seq analysis to compare gene expression profiles between two controls and two *ASCL1* knockdown (KD) samples. Genes differentially expressed (adjusted *p* value  $< 0.05$ ,  $\log_2$  fold-change  $< -0.5$  and TPM  $> 3$ ) were selected for gene ontology analysis. Top ten terms were shown with combined score. The *p* values by the Fisher exact test are indicated. **c**, **d** The mRNA expression of genes involved in neural differentiation and mRNA expression of CRC members were analyzed by qRT-PCR after *ASCL1* knockdown in Kelly cells. Expression was shown as a percent relative to shGFP control. Data are represented as means  $\pm$  SEM for three biological replicates. **e** CHP-134 cells were treated with or without NGF (100 ng/mL) or/and ATRA (100 nM) after being transduced with shRNAs. The cells were then stained with an antineurofilament antibody, with images representative of three fields taken at  $\times 40$  original magnification. Scale bars were shown in white at bottom right, 20  $\mu$ m. **f** Cell viability was measured after 3, 5, 7, and 9 days of infection in Kelly cells ( $n = 3$  per group). The growth rate (fold-change) over 9 days compared with day 3 was indicated. Values represent means  $\pm$  SD for technical triplicates.

similar relationship may exist in neuroblastoma cells, in which the more neural crest-like mesenchymal cells do not express *ASCL1*, and *ASCL1* expression is specifically upregulated as cells differentiate along the sympathetic lineage and express members of the ADRN CRC. In fact, our study demonstrated that *ASCL1* can positively regulate CRC genes that are expressed in neuronal progenitor cells such as *PHOX2B* and *HAND2* and also negatively regulates genes expressed in more differentiated cells such as *NTRK1* and *NPY*. Thus, high levels of *ASCL1* expression, along with *LMO1* and *MYCN*, helps to mediate a differentiation arrest at the neural progenitor stage, which is an essential component of

childhood neuroblastoma. Of note, comparisons shown in Supplementary Fig. 6d argue that, although many of the cells bridging between Schwann cell precursors and chromaffin cells express high levels of *Ascl1*, this expression is not regulated by the CRC of the ADRN subset of human neuroblastoma. This finding is compatible with these bridging cells contributing to neuroblastoma through a different developmental CRC.

In this study, we also compared genomic loci bound by *LMO1* and gene expression profiles after *LMO1* knockdown between neuroblastoma and T-ALL cells. Although *LMO1* frequently co-occupies target loci with *GATA3* in both neuroblastoma and T-



**Fig. 8** Schematic image for CRC members in ADRN type of neuroblastoma. RE regulatory element.

ALL cells, there was little overlap among the regulated enhancers between these two cell types. Similarly, genes and pathways regulated by *LMO1* knockdown in neuroblastoma cells were distinct from those in T-ALL cells. These results indicate that *LMO1* regulates gene expression in a tissue/tumor-type specific manner, consistent with the fact that it acts in cooperation with type II bHLH proteins, which are known to play key roles in establishing cell lineage<sup>13,25–28</sup>. Because *LMO1* alone cannot directly bind to DNA, the target sites of *LMO1* largely depend on the binding of its protein–protein interaction partners, which are DNA-binding transcription factors expressed in specific tissues, whose binding may also be affected by epigenetic alterations and chromatin status. It is intriguing that *LMO1* participates as a cofactor in CRCs containing *GATA3* in both T-ALL and neuroblastoma; however, the remainder of the CRC of T-ALL and neuroblastoma remain completely distinct. The extensive studies of the roles of LMO proteins in T-ALL and the control of hematopoiesis<sup>13</sup> provide a model for future studies to define the transcriptional complexes that mediate gene expression through the ADRN CRC in neuroblastoma.

Taken together, our study shows that *ASCL1* is an integral member of the neuroblastoma CRC, essential for the formation of the feed-forward autoregulatory loop that drives the oncogenic transformation in concert with *MYCN*. In this context, *LMO1* is a critical cofactor of the CRC. Our data provide evidence that the neuroblastoma CRC, operating through an *LMO1*-*ASCL1* axis, not only promotes the growth and survival of malignant neuroblasts, but also blocks the differentiation of sympathetic progenitor cells.

## Methods

**Cell culture and reagents.** All neuroblastoma cell lines (Kelly, SH-SY5Y, CHP-134, and SK-N-SH) and T-ALL cell lines (Jurkat, RPMI-8402, MOLT-4, and CCRF-CEM) were cultured in RPMI-1640 medium (Biowest) supplemented with 10% FBS (Biowest) and 1% penicillin/streptomycin (Thermo). All cell lines were confirmed by DNA fingerprinting using the PowerPlex 1.2 system (Promega, Madison, WI, USA) and were regularly tested for mycoplasma contamination.

293T cells were cultured in DMEM medium (Biowest) supplemented with 10% FBS (Biowest) and 1% penicillin/streptomycin (Thermo). Cells were incubated in 37 °C with 5% CO<sub>2</sub>. ATRA, NGF, and DOX hydrochloride were purchased from Sigma and dissolved in DMSO.

**Gene knockdown.** The shRNA sequences were designed according to the RNA Consortium's recommendation (<http://www.broadinstitute.org/rnai/trc>) and cloned into the lentiviral vector pLKO.1-puro at the Age I and EcoR I digestion site. Each individual shRNA vector (500 ng) was co-transfected into 293T cells with the packaging plasmids pMDLg/pRRE (250 ng) and pRSV-Rev (250 ng) together with the envelope plasmid pMD2.G (250 ng) into 293T cells (200,000 cells) seeded into six-well plate 24 h before transfection, using the FuGENE 6 transfection reagent (Roche) and Opti-MEM, according to the manufacturer's instruction. Medium was changed 24 h after transfection. Supernatants containing lentivirus was then collected after 48 and 72 h post transfection, combined, and filtered through a 0.45 mm nitrocellulose filter (Thermo). The Jurkat cells were infected with lentivirus in the presence of polybrene (8 µg/mL; Millipore) by centrifugation at 1300 rcf for 1.5 h. Neuroblastoma cells were transduced with lentivirus in the presence of polybrene (8 µg/mL; Millipore) for 2 h incubation followed by topping up fresh medium. Approximately 500,000 cells were infected with 1 mL of virus medium. The cells were then selected by the addition of puromycin (0.7 µg/mL for Jurkat cell line; 0.5 µg/mL for neuroblastoma cell lines) for at least 36 h after infection. shRNA sequences are shown in Supplementary Data 3.

**Gene overexpression.** The *LMO1* cDNA (CDS region of *LMO1* transcript, NM\_001270428.1) was amplified and cloned into the pMSCV-IRES-GFP retrovirus vector. Retroviral vector (1000 ng) was co-transfected into 293T cells (200,000 cells) seeded into six-well plate 24 h before transfection with the packaging plasmid pMD-MLV (250 ng) and the envelope plasmid VSV-G (250 ng) using FuGENE 6 reagent (Roche) and Opti-MEM. Approximately 500,000 cells were transduced with 1 mL retrovirus media in the presence of polybrene (8 µg/mL; Millipore) for 2 h incubation followed by topping up of 1 mL fresh medium. The cells that express GFP were then sorted by flow cytometry using the BD FACSAria II (BD Biosciences).

**Cell viability assay.** Neuroblastoma cells were seeded into 96-well plates (2500 cells/well) at 24 h post virus infection. Cell viability was measured after 3, 5, 7, and 9 days, based on luminescence by the Cell Titer Glo assay (Promega) using the Tecan Infinite 200 PRO plate reader (Tecan).

**RNA extraction and quantitative reverse transcriptase PCR (qRT-PCR).** Total RNA was extracted using the NucleoSpin RNA kit (Macherey-Nagel). A total of 1000 ng of the purified RNA was reverse-transcribed into cDNA in 10 µL reaction and finally diluted to 50 µL, using the QuantiTect kit (QIAGEN) according to manufacturer guidelines. PCR mix contains 2 µL cDNA, 10 µL Power SYBR Green PCR Master Mix (Roche), and 1 µL of each forward primer and reverse primer (final concentration of 500 nM). The mRNA expression levels of the genes of interest were evaluated by  $\Delta\Delta$  Ct method by performing qPCR analysis the Quant Studio 3 Real-Time PCR System (Thermo Fisher Scientific) with following cycling condition: 2 min at 50 °C, 10 min at 95 °C, and 1 min at 60 °C, with ramping temperature 1.6 °C/s. Primers were tested for standard curve with serial diluted DNA to ensure efficiency between 90 and 110% and R<sup>2</sup> for standard curve greater than 0.95 and single peak for melt curve to ensure the primer specificity. Primer sequences for each individual gene are shown in Supplementary Data 4.

**Protein extraction and western blot analysis.** The cells were lysed in RIPA Buffer (20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM Na<sub>2</sub>EDTA, 1 mM EGTA, 1% NP-40, 1% sodium deoxycholate, 2.5 mM sodium pyrophosphate, 1 mM beta-glycerophosphate, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 1 µg/mL leupeptin) (#9806, Cell Signaling Technology) with protease inhibitors (Roche). Equal amounts of protein was diluted in Laemmli sample buffer (Bio-Rad) with 10%  $\beta$ -mercaptoethanol and boiled for 10 min at 95 °C. Protein samples were resolved by SDS-PAGE gel using Bio-Rad system, with running buffer (Bio-Rad), and subsequently blotted onto a PVDF membrane (Bio-Rad). Membranes were block with 5% nonfat milk for 1 h and washed with washing buffer (TBS-Tween 0.1%) following incubation with primary antibody diluted in 5% BSA (Sigma) dissolved in washing buffer, overnight at 4 °C. Incubation with secondary antibodies were performed after washing for 1 h at room temperature. HRP-labeled anti-rabbit (7074, Cell Signaling Technologies, 1:10,000 dilution) and anti-mouse (7076P2, Cell Signaling, 1:10,000 dilution) antibodies were used. Proteins were visualized by enhanced chemiluminescence (Thermo). Antibodies used for immunoblotting were as follows: *LMO1* (Bethyl Laboratories, #A300314A, 1:1000 dilution), *ASCL1* (Santa Cruz, 1:200 dilution, #sc-390794), *MYCN* (Cell Signaling Technologies, 1:1000 dilution, #84406), *RET* (Cell Signaling Technologies, 1:1000 dilution, #14299), *ERK* (Cell Signaling Technologies, 1:1000 dilution, #9102), phosphor-ERK (Cell Signaling Technologies, 1:1000 dilution, #9101),  $\alpha$ -tubulin (Cell Signaling Technologies, 1:1000 dilution, #2144) and  $\beta$ -actin (Cell Signaling Technologies, 1:1000 dilution, #8457). Original uncropped images are shown in Supplementary Fig. 8.

**Blocking of genomic DNA elements.** To express dCas9-KRAB protein, lentivirus was made using pLV hU6-sgRNA hUbc-dCas9-KRAB-T2a-Puro lentivirus vector (500 ng; Addgene plasmid #71236) together with packaging plasmids pMDLg/pRRE (250 ng), pRSV-Rev (250 ng) and the envelope plasmid pMD2.G (250 ng). Kelly cells were infected with virus medium using the same method mentioned in previous section. The cells were selected by puromycin (0.5 µg/mL) for 3 days. Subsequently, dCas9-KRAB expressing cells were transduced with an inducible sgRNA and a GFP gene. The sgRNA sequences were designed and cloned into FgH1UTG lentivirus vector (Addgene#70183) at the BmsBI digestion site. Lentivirus were made using the vector by the method mentioned above. After 4 days of transduction of the sgRNA vector, the cells were selected by GFP signal by flow cytometry using the BD FACSAria II (BD Biosciences). The cells were then treated with 1000 ng/mL DOX to induce a sgRNA expression. The sgRNA sequences were designed using CRISPR Design Tool (<http://crispr.mit.edu/>) and shown in Supplementary Data 5.

**Chromatin immunoprecipitation-quantitative PCR (ChIP-qPCR).** Approximately 10 million cells were crosslinked with 1% formaldehyde (final concentration) for 10 min by inverting flasks at room temperature and quenched with 250 mM Glycine. The cells pellets were washed in PBS and then stored at -80 °C. The pellets were lysed in lysis buffer I (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8.0, 10% Glycerol, 0.5% NP-40, 0.25% Triton X-100, and complemented with protease inhibitor cocktail) for 10 min. After centrifugation, supernatant was discarded and the pellet was lysed in lysis buffer II (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0, and complemented with protease inhibitor cocktail) for 10 min. After centrifugation, supernatant was discarded and the pellet was lysed in lysis buffer III (10 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA pH 8.0, 0.1% SDS buffer, 1% Triton-X, and complemented with protease inhibitor cocktail) and subject to sonication. Chromatin was sheared by the Bioruptor Plus sonicator (Diagenode) at high power for 50 cycles of 30 s, with a 30-s pause between cycles. Sheared chromatin was incubated with primary antibody bound to the Dynabeads Protein G (Thermo Fisher Scientific) overnight, followed by elution and reverse cross-linking at 65 °C for 15 h. TE-buffer was added in DNA elution buffer followed by RNase treatment (0.45 mg/mL) at 37 °C for 2 h and by proteinase K treatment (0.3 mg/mL) at 55 °C for 30 min. DNA was isolated using 400 µL of the separated lower phase of the Phenol-Chloroform-isoamyl alcohol in MaXtract High Density 2 mL gel tubes (Qiagen) by centrifugation at full speed for 7 min. The aqueous layer was then transferred to a new tube and subsequently added NaCl solution (100 mM final concentration), glycogen solution (20 µg) and 800 µL 100% ethanol, and incubated at -80 °C for 1 h. DNA was precipitated by centrifuge at full speed for 20 min and washed with 75% of ethanol and pelleted. Pelleted DNA was dissolved in RNase/DNase free water. Antibodies used for ChIP-PCR were as follows: LMO1 (Bethyl Laboratories, #A300314A), Flag (Sigma, #F1804), H3K27ac (Abcam), and IgG (Santa Cruz). After the purification of DNA, quantitative PCR was performed using the Quant Studio 3 Real-Time PCR System (Thermo Fisher Scientific) with purified DNA, Power SYBR Green PCR Master Mix (Roche) and PCR primers against the genomic regions of interests. PCR primer sequences were shown in Supplementary Data 6.

**Chromatin immunoprecipitation sequencing (ChIP-seq).** For each ChIP, 10 µg of antibody was conjugated to 2 mg of M270 epoxy beads and added to 3 mL of sonicated nuclear extract. Anti-LMO1 antibody (Bethyl Laboratories, #A300314A) and anti-ASCL1 antibody (Santa Cruz, #sc-390794) were used for ChIP. Purified DNA was prepared for sequencing according to a modified version of the Solexa Genomic DNA protocol. Fragmented DNA was end repaired and subjected to linker-mediated-PCR using oligos from Illumina. Amplified fragments were purified and applied to the flow-cell using the Solexa Cluster Station fluidics device. Samples were then subjected to sequencing according to Illumina's standard protocols. All ChIP-seq reads were mapped to the hg19 human reference genome using bowtie2 with default parameters. Duplicates were removed from aligned ChIP-seq reads using the samtools rmdup package. ChIP-seq peaks were called with MACS14 software version 1.4.2. The output bedgraph were normalized subtracting corresponding background using MACS2 bdgcmp. LMO1 peaks ( $P < 1e-4$ ) and ASCL1 peaks ( $P < 1e-9$ ) called by MACS14 were selected and filtered for nonpromoter binding peaks based on annotatePeaks from the HOMER package. Co-occurrence heatmap and metagene were plotted using deepTools version 3.1.2. Each of the signal matrix was calculated using bigwig output generated using UCSC bedGraphToBigWig of each ChIP-seq sample by deepTools computerMatrix. The signal matrix was then used to plot heatmap using deepTools plotHeatmap package<sup>40</sup>.

**Enhancer ranking analysis.** Enhancers were defined as regions with H3K27ac ChIP-seq enrichment. H3K7ac peaks were called using MACS2 version 1.4.2 ( $q$ -value  $< 5e-2$  for narrow region and  $q$ -value  $< 1e-1$  for broad region) with  $-keepdup = 1$  and  $-broad$ . Enhancers were identified using ROSE v0.1 by merging H3K27ac peaks within 12.5 kb between each other, excluded those were fully contained within  $\pm 2$  kb from TSS and ranked them along the  $x$ -axis based on H3K27ac enrichment plotted on the  $y$ -axis (<https://bitbucket.org/youngcomputation/rose>). All enhancers were assigned responsible to the nearest RefSeq genes. The percentile of ASCL1 enhancer was measured by ordering the

ranking index and calculating the top ranking of ASCL1 enhancer among all other enhancers in each cell line.

**RNA-sequencing (RNA-seq).** Total RNA was extracted from control and knockdown samples using the QIAzol lysis reagent (Qiagen) and cleaned using the RNeasy kit (Qiagen). Samples were treated with the TURBO DNase (TURBO DNA-free Kit; Ambion) and cleaned using the RNeasy MinElute Cleanup kit (Qiagen). Strand-specific library construction and sequencing of ~100 M paired-end 100-bp-long reads by Illumina HiSeq or BGISEQ were performed at the BGI Biotech Solutions Co. Ltd. (Hong Kong, China). RNA-seq reads were aligned to the hg19 human reference genome using STAR 2.5.2a with out-FilterMultimapNmax set to 1. Total mapped reads were quantified using htseq-count version 0.6.1, and count tables were generated based on Ensembl hg19 gene annotation gtf files. Differential expression analysis was conducted using the Bioconductor package DESeq2 version 1.12.4. Gene expression for each neuroblastoma cell lines was estimated in transcripts per million (TPM) using Kallisto software version 0.43.1. All RNA-Seq data were normalized using the Sleuth R package from Patchter lab<sup>41</sup>. Selected gene expression from sleuth normalized TPM were used to generate heatmap using web-based heatmap tools, heatmapper<sup>42</sup>. Genes differentially expressed after knockdown of each transcription factor were selected using the following criteria: adjusted  $p$  value  $< 0.05$ , log 2 fold-change  $< -0.5$  or  $> 0.5$ , and TPM  $> 3$ .

**GO analysis.** GO analysis was performed using the Enrichr tool (<http://amp.pharm.mssm.edu/Enrichr/>). Combined scores were calculated by multiplying the  $p$  value computed by Fisher's exact test by the  $z$ -score of the deviation from the expected rank.

**Immunostaining.** CHP-134 cells were treated with or without NGF (100 ng/mL) or ATRA (100 nM) after transduction of shRNAs. The cells were stained with an antineurofilament antibody (Cell Signaling, #2837).

**Statistics and reproducibility.** All the statistical analyses were done in GraphPad Prism software. A  $p$  value less than 0.05 was considered statistically significant. The details of methods used can be found in each figure legend. Experiments in Figs. 1a, b, 3b, 4b, d, e, 5b, 7a, c-f were repeated three or more times. Experiments in Figs. 4c, f, 5c-e were repeated two times. The numbers of experiments conducted in Supplementary Figures were stated in Supplementary Methods section and Supplementary Figure Legends.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

ChIP-seq dataset for LMO1 and ASCL1 in Kelly cells have been deposited in the GEO database (GSE120074). ChIP-seq datasets of PHOX2B, HAND2, GATA3, MYCN, ISL1, and TBX2, H3K27ac and H3K4me1 in Kelly cells were obtained from the GEO database (GSE94824 and GSE62726)<sup>19,43</sup>. ChIP-seq datasets of H3K27ac in various neuroblastoma cell lines were obtained from the GEO database (GSE86672 and GSE90683)<sup>21,44</sup>. ChIP-seq datasets of LMO1, GATA3, H3K27ac and H3K4me1 in Jurkat cells were obtained from the GEO database (GSE94391, GSE68976, and GSE50622, GSE119439)<sup>45-48</sup>. RNA-seq dataset for Kelly cells after LMO1 knockdown and ASCL1 knockdown were deposited in the GEO database (GSE132760). Microarray dataset for SH-SY5Y cells after LMO1 knockdown was deposited in the GEO database (GSE130747). RNA-seq dataset for Jurkat after LMO1 knockdown has been reported by us and deposited in the GEO database (GSE97514)<sup>48</sup>. RNA-seq datasets for zebrafish neuroblastoma samples reported by us<sup>30</sup> were obtained from the GEO database (GSE107518). RNA-seq dataset for various neuroblastoma cell lines was obtained from the GEO database (GSE90683)<sup>21</sup>. Single cell sequencing dataset for mouse neuronal cells was obtained from the GEO database (GSE99933)<sup>36</sup>. The microarray datasets for primary neuroblastoma cases reported by the Kocak et al.<sup>31</sup> (GSE45547), Versteeg et al.<sup>32</sup> (GSE16476) and the NRC (GSE85047) were analyzed by the R2 database (<http://hgserver1.amc.nl/cgi-bin/r2/main.cgi>). The cancer cell line dataset is derived from CCLE database (<https://portals.broadinstitute.org/ccle>). Detailed information is also shown in Supplementary Data 7. All other methods used for Supplementary Figures are described in the Supplementary Methods section.

Received: 8 January 2019; Accepted: 11 November 2019;

Published online: 09 December 2019

## References

1. Maris, J. M., Hogarty, M. D., Bagatell, R. & Cohn, S. L. Neuroblastoma. *Lancet* **369**, 2106–2120 (2007).

2. Brodeur, G. M. Neuroblastoma: biological insights into a clinical enigma. *Nat. Rev. Cancer* **3**, 203–216 (2003).
3. Maris, J. M. Recent advances in neuroblastoma. *N. Engl. J. Med* **362**, 2202–2211 (2010).
4. Pinto, N. R. et al. Advances in risk classification and treatment strategies for neuroblastoma. *J. Clin. Oncol.* **33**, 3008–3017 (2015).
5. van Groningen, T. et al. Neuroblastoma is composed of two super-enhancer-associated differentiation states. *Nat. Genet* **49**, 1261–1266 (2017).
6. Pugh, T. J. et al. The genetic landscape of high-risk neuroblastoma. *Nat. Genet* **45**, 279–284 (2013).
7. Brodeur, G. M., Seeger, R. C., Schwab, M., Varmus, H. E. & Bishop, J. M. Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage. *Science* **224**, 1121–1124 (1984).
8. Wang, K. et al. Integrative genomics identifies LMO1 as a neuroblastoma oncogene. *Nature* **469**, 216–220 (2011).
9. Oldridge, D. A. et al. Genetic predisposition to neuroblastoma mediated by a LMO1 super-enhancer polymorphism. *Nature* **528**, 418–421 (2015).
10. Zhu, S. et al. LMO1 synergizes with MYCN to promote neuroblastoma initiation and metastasis. *Cancer Cell* **32**, 310–323 e315 (2017).
11. Matthews, J. M., Lester, K., Joseph, S. & Curtis, D. J. LIM-domain-only proteins in cancer. *Nat. Rev. Cancer* **13**, 111–122 (2013).
12. Rabbitts, T. H. LMO T-cell translocation oncogenes typify genes activated by chromosomal translocations that alter transcription and developmental processes. *Genes Dev.* **12**, 2651–2657 (1998).
13. Chambers, J. & Rabbitts, T. H. LMO2 at 25 years: a paradigm of chromosomal translocation proteins. *Open Biol.* **5**, 150062 (2015).
14. Murre, C. Intertwining proteins in thymocyte development and cancer. *Nat. Immunol.* **1**, 97–98 (2000).
15. Abraham, B. J. et al. Small genomic insertions form enhancers that misregulate oncogenes. *Nat. Commun.* **8**, 14385 (2017).
16. Simonis, M. et al. High-resolution identification of balanced and complex chromosomal rearrangements by 4C technology. *Nat. Methods* **6**, 837–842 (2009).
17. Tremblay, M. et al. Modeling T-cell acute lymphoblastic leukemia induced by the SCL and LMO1 oncogenes. *Genes Dev.* **24**, 1093–1105 (2010).
18. Saint-Andre, V. et al. Models of human core transcriptional regulatory circuitries. *Genome Res.* **26**, 385–396 (2016).
19. Durbin, A. D. et al. Selective gene dependencies in MYCN-amplified neuroblastoma include the core transcriptional regulatory circuitry. *Nat. Genetics* **50**:1240–1246 (2018).
20. Decaestecker, B. et al. TBX2 is a neuroblastoma core regulatory circuitry component enhancing MYCN/FOXM1 reactivation of DREAM targets. *Nat. Commun.* **9**, 4866 (2018).
21. Boeva, V. et al. Heterogeneity of neuroblastoma cell identity defined by transcriptional circuitries. *Nat. Genet* **49**, 1408–1413 (2017).
22. Sanda, T. et al. Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia. *Cancer Cell* **22**, 209–221 (2012).
23. Mansour, M. R. et al. Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science* **346**, 1373–1377 (2014).
24. Zeid, R. et al. Enhancer invasion shapes MYCN-dependent transcriptional amplification in neuroblastoma. *Nat. Genet* **50**, 515–523 (2018).
25. El Omari, K. et al. Structural basis for LMO2-driven recruitment of the SCL: E47bHLH heterodimer to hematopoietic-specific transcriptional targets. *Cell Rep.* **4**, 135–147 (2013).
26. Murre, C. Helix-loop-helix proteins and lymphocyte development. *Nat. Immunol.* **6**, 1079–1086 (2005).
27. Meierhan, D. et al. DNA binding specificity of the basic-helix-loop-helix protein MASH-1. *Biochemistry* **34**, 11026–11036 (1995).
28. Guillemot, F. et al. Mammalian achaete-scute homolog 1 is required for the early development of olfactory and autonomic neurons. *Cell* **75**, 463–476 (1993).
29. Lo, L. C., Johnson, J. E., Wuenschell, C. W., Saito, T. & Anderson, D. J. Mammalian achaete-scute homolog 1 is transiently expressed by spatially restricted subsets of early neuroepithelial and neural crest cells. *Genes Dev.* **5**, 1524–1537 (1991).
30. Zimmerman, M. W. et al. MYC drives a subset of high-risk pediatric neuroblastomas and is activated through mechanisms including enhancer hijacking and focal enhancer amplification. *Cancer Discov.* **8**, 320–335 (2018).
31. Kocak, H. et al. Hox-C9 activates the intrinsic pathway of apoptosis and is associated with spontaneous regression in neuroblastoma. *Cell Death Dis.* **4**, e586 (2013).
32. Molenaar, J. J. et al. Sequencing of neuroblastoma identifies chromothripsis and defects in neurogenesis genes. *Nature* **483**, 589–593 (2012).
33. Cazes, A. et al. Activated Alk triggers prolonged neurogenesis and Ret upregulation providing a therapeutic target in ALK-mutated neuroblastoma. *Oncotarget* **5**, 2688–2702 (2014).
34. Lambert, I. et al. Upregulation of MAPK negative feedback regulators and RET in mutant ALK neuroblastoma: implications for targeted treatment. *Clin. Cancer Res.* **21**, 3327–3339 (2015).
35. Lopez-Delisle, L. et al. Activated ALK signals through the ERK-ETV5-RET pathway to drive neuroblastoma oncogenesis. *Oncogene* **37**, 1417–1429 (2018).
36. Furlan, A. et al. Multipotent peripheral glial cells generate neuroendocrine cells of the adrenal medulla. *Science* **357**, eaal3753 (2017).
37. Edsjo, A. et al. Neuroblastoma cells with overexpressed MYCN retain their capacity to undergo neuronal differentiation. *Lab Invest.* **84**, 406–417 (2004).
38. Li, J., Neumann, I., Volkmer, I. & Staeger, M. S. Down-regulation of achaete-scute complex homolog 1 (ASCL1) in neuroblastoma cells induces up-regulation of insulin-like growth factor 2 (IGF2). *Mol. Biol. Rep.* **38**, 1515–1521 (2011).
39. Narayanan A. et al. The proneural gene ASCL1 governs the transcriptional subgroup affiliation in glioblastoma stem cells by directly repressing the mesenchymal gene NDRG1. *Cell Death Differ.* **9**, 1813–1831 (2018).
40. Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
41. Pimentel, H., Bray, N. L., Puente, S., Melsted, P. & Pachter, L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat. Methods* **14**, 687 (2017).
42. Babicki, S. et al. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res.* **44**, W147–W153 (2016).
43. Chipumuro, E. et al. CDK7 inhibition suppresses super-enhancer-linked oncogenic transcription in MYCN-driven cancer. *Cell* **159**, 1126–1139 (2014).
44. Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
45. Li, Z. et al. APOBEC signature mutation generates an oncogenic enhancer that drives LMO1 expression in T-ALL. *Leukemia* **31**, 2057–2064 (2017).
46. Hnisz, D. et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454–1458 (2016).
47. Kwiatkowski, N. et al. Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. *Nature* **511**, 616–620 (2014).
48. Leong, W. Z. et al. ARID5B as a critical downstream target of the TAL1 complex that activates the oncogenic transcriptional program and promotes T-cell leukemogenesis. *Genes Dev.* **31**, 2343–2360 (2017).

## Acknowledgements

We thank members of the Young, Maris, Look and Sanda laboratories for discussions. We thank Shuo-Chien Ling and Wan Yun Ho for helping immunostaining. We are grateful to the Whitehead Genome Technology Core for assistance with ChIP-Seq. We acknowledge the RNAi Consortium for providing lentivirus shRNA constructs. This work is supported by grants from the National Cancer Institute, National Institute of Health, including K99 CA157951 (T.S.) and R01 CA180692 (J.M.M./A.T.L.), and the American Lebanese Syrian Associated Charities (ALSAC). T.S. is supported by the National Research Foundation (NRF), Prime Minister's Office, Singapore under its Competitive Research Programme (NRF-NRFF2013-02), and Ministry of Education (MOE), Singapore (FRC T1-2015 Sep-07). This work is also supported by the NRF and MOE under its Research Centres of Excellence initiative. A.D.D. and M.W.Z. are Damon Runyon–Sohn Pediatric Fellows supported by the Damon Runyon Cancer Research Foundation, DRG-24-18 (A.D.D.), DRSG-9-14 (M.W.Z.). A.D.D., M.W.Z., J.R. are recipients of Alex's Lemonade Stand Foundation Young Investigator Awards. A.D.D. is the recipient of an American Society for Clinical Oncology Young Investigator Award, and a Young Investigator Award from the Rally Foundation for Childhood Cancer Research. M.W.Z. is supported by the Claudia Adams Barr Program for Innovative Cancer Research.

## Author contributions

L.W., A.D.D., M.W.Z., A.T.L., and T.S. designed this study. L.W., A.D.D., M.W.Z., S.H.T., N.W., K.A., and L.N.L. performed experiments. T.K.T., B.J.A., P.C.T.N., and J.R. performed bioinformatics analysis. J.M.M., R.A.Y., and A.T.L. supervised the study.

## Competing interests

R.A.Y. is a founder and shareholder of Syros Pharmaceuticals, which is discovering and developing therapeutics directed at transcriptional pathways in cancer. B.J.A. is a shareholder of Syros Pharmaceuticals. No other potential conflicts of interest are declared.

**Additional information**

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-019-13515-5>.

**Correspondence** and requests for materials should be addressed to A.T.L. or T.S.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



**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/>.

© The Author(s) 2019