Reverse Engineering the Neuroblastoma Regulatory Network Uncovers MAX as One of the Master Regulators of Tumor Progression

Ricardo D'Oliveira Albanus*, Rodrigo Juliani Siqueira Dalmolin, Mauro Antônio Alves Castro, Matheus Augusto de Bittencourt Pasquali, Vitor de Miranda Ramos, Daniel Pens Gelain, José Cláudio Fonseca Moreira

Departamento de Bioquímica, Núcleo de Bioinformática, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brasil

Abstract

Neuroblastoma is the most common extracranial tumor and a major cause of infant cancer mortality worldwide. Despite its importance, little is known about its molecular mechanisms. A striking feature of this tumor is its clinical heterogeneity. Possible outcomes range from aggressive invasion to other tissues, causing patient death, to spontaneous disease regression or differentiation into benign ganglioneuromas. Several efforts have been made in order to find tumor progression markers. In this work, we have reconstructed the neuroblastoma regulatory network using an information-theoretic approach in order to find genes involved in tumor progression and that could be used as outcome predictors or as therapeutic targets. We have queried the reconstructed neuroblastoma regulatory network using an aggressive neuroblastoma metastasis gene signature in order to find its master regulators (MRs). MRs expression profiles were then investigated in other neuroblastoma datasets so as to detect possible clinical significance. Our analysis pointed MAX as one of the MRs of neuroblastoma progression. We have found that higher MAX expression correlated with favorable patient outcomes. We have also found that MAX expression and protein levels were increased during neuroblastoma SH-SY5Y cells differentiation. We propose that MAX is involved in neuroblastoma progression, possibly increasing cell differentiation by means of regulating the availability of MYC:MAX heterodimers. This mechanism is consistent with the results found in our SH-SY5Y differentiation protocol, suggesting that MAX has a more central role in these cells differentiation than previously reported. Overexpression of MAX has been identified as anti-tumorigenic in other works, but, to our knowledge, this is the first time that the link between the expression of this gene and malignancy was verified under physiological conditions.

Introduction

Neuroblastoma is the most common extra-cranial solid tumor and one of the leading causes of cancer mortality in children worldwide [1–4]. These tumors are originated from embryonic elements of the neural crest and sympathetic nervous system, usually developing in the adrenal glands, but also arising in nervous tissues of the neck, thorax, abdomen, and pelvis. The metastatic capacity of this type of cancer is notable, being able of compromising almost any tissue in the human body [2,5]. Another distinguishing feature of this disease is its clinical heterogeneity. The possible endpoints span from complete remission to patient death, even with advanced multimodal

therapy [6]. Numerous efforts have been made in order to sort neuroblastoma patients in separate risk groups, such as the International Neuroblastoma Risk Group (INRG) staging system [7], which is based on clinical factors and imaging studies, and the Children's Oncology Group (COG) risk stratification schema, which complements the former with molecular aspects. The molecular approaches for assessing patient clinical outcome are made by detecting DNA copy number alterations or by searching for more specific segmental aberrations, such as chromosomes 11q loss and 17q gain [8,9]. The most reliable molecular classifier, however, is the amplification of the MYCN oncogene, which is linked to grim prognosis in the majority of cases [10–12].

Lower grade neuroblastoma patients can be treated with surgical resection alone, and may even be subject to spontaneous remission without any intervention. Patients with metastatic MYCN amplified tumors have the highest mortality rate and are usually unresponsive to advanced multimodal treatment [6,7,13,14]. Patients with MYCN non-amplified metastatic tumors, however, present highly variable outcomes, but few endpoint predictors have been described for this group. Currently, the best known prognostic indicator is the age at which the tumor is diagnosed [7]. Patients younger than 18 months usually have better prognosis, with more than 90% 6-year event-free survival for patients younger than one year. Patients older than 18 months, on the other hand, suffer higher mortality rates and may present less than 25% 6-years event-free survival [15,16]. Several high-throughput studies have been made over the last years to understand the biology underlying this clinical variability, and at least one of them was aimed exclusively at these metastatic MYCN non-amplified patients [17]. Despite intensive study, few predictors have been brought forward, demanding increased efforts to understand this remarkable disease.

In recent years, the availability of gene expression studies has allowed novel strategies for understanding cancer biology. One of such is the use of mutual information models for inferring the regulatory networks of transcription factors (TFs) and their transcriptional targets (or regulons) in the gene expression profile of a given set of samples [18–20]. This methodology allows the detection of potential causal relationships between TFs and specific cancer signatures. As these statistics depend on rather large sample sizes (n>80), only recently are they being successfully applied in biology [21–23]. In this paper, we have reconstructed the neuroblastoma regulatory network in order to find TFs involved in the transition from primary tumors to highly aggressive bone marrow metastasis. We have found evidences that MAX is one of the master regulators of tumor progression, possibly by being an additional regulatory step for the availability of MYC:MAX heterodimers to regulate transcription and increase proliferation. We have also found evidences that MAX plays a more prominent role in SH-SY5Y cells differentiation than previously described.

Results

Neuroblastoma regulatory network and master regulator analysis

Through our regulatory network reconstruction workflow (Figure 1), we have identified 15,713 targets for 1,363 TFs in the first reconstructed neuroblastoma regulatory network (GSE16476), and 4,039 targets for 705 TFs in the second (GSE3960) (Figure 2). We have found eight master regulators (MRs) common to both networks (Table S1) using as query an aggressive neuroblastoma metastatic gene signature. We have chosen only regulons common to both networks so as to improve the specificity of our MRs analysis. However, by this criterion alone, there was still an elevated chance of finding nonspecific TFs. To overcome this problem, we have used the metastatic neuroblastoma gene signature to query a healthy

mammary tissue network (GSE10780) [24] and use it as a positive control to detect MRs that are not tissue-specific to neuroblastoma. We then re-ran the MR analysis within the two neuroblastoma networks using as query the meta-PCNA proliferative gene signature [25]. This last step was performed in order to exclude MRs that were directly involved with proliferation. Of the eight MRs common to both networks, only MAX, TFEC, and ZNF101 remained after these specificity tests.

Regulon overlap and motif analysis

In order to validate the power of our pipeline for detecting biologically significant regulons, we have compared the composition of each regulon across the two neuroblastoma networks and verified whether their members were enriched with the binding site motif of its MR. We have found less than 1% overlap in the genes potentially regulated by MAX and ZNF101 in both networks, and 25% for TFEC. We proceeded to search for regulatory motifs in the targets sequences flanking regions. Because there are no available motifs for TFEC and ZNF101 in public databases, we could only query for the known MAX motif V$MAX, obtained at the JASPAR database [26,27]. We have found significant enrichment for V$MAX motif in both networks (p=0.038), indicating that our pipeline predicted correctly both MAX regulons, albeit their different compositions. Although there is no public TFEC motif available, this transcription factor is a member of the MITF bHLH family, which is closely related to MAX and MYC families [28]. Members of these families have similar CACGTG binding motifs, meaning that we can query the TFEC regulon using V$MAX motif to determine whether it is regulated by a bHLH transcription factor. TFEC regulon was significantly enriched with this motif (p=0.003), suggesting that it is indeed regulated by a bHLH factor.

Given that the MYC/MAX/MAD network may regulate up to 15% of the human genome [29–31], one would not expect finding substantial overlap in both networks, particularly when taking into account the significant genetic heterogeneity across tumors of the same type, which could greatly affect the specificity of such an important regulatory system. Taking this into consideration and the fact that both MAX regulons were significantly enriched with V$MAX binding motif, we chose to carry our analysis with this gene. ZNF101 was not analyzed further because it did not have corresponding probes the datasets below.

Clinical relevance of MAX and TFEC expression in neuroblastoma

To access whether MAX and TFEC expression could be related to patient outcome, we have analyzed the GSE3446 dataset [17]. This dataset consists of expression profiles of primary tumor biopsies obtained at diagnosis from neuroblastoma patients who (i) either had relapse after five years (n=46), (ii) did not present disease progression in the same period (n=56), and (iii) from tumors obtained at progression (n=12 + 3 obtained both at diagnosis and relapse). All patients in this study had untreated metastatic MYCN non-amplified tumors. 74 of the 102 patients studied at diagnosis
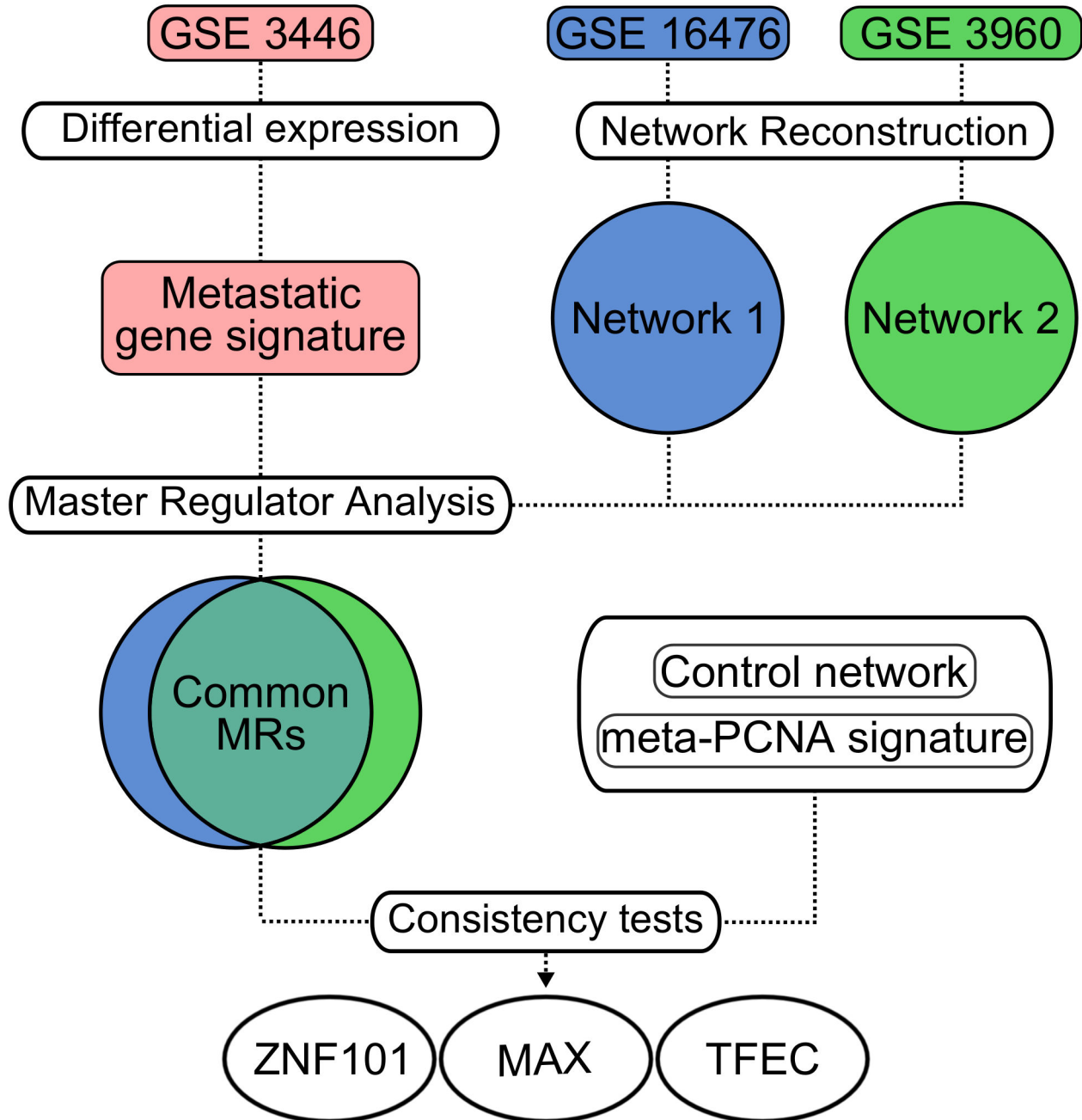
Figure 1: Schematic representation of the workflow used for reconstructing the neuroblastoma network and searching for master regulators of a metastatic gene signature.

Figure 1. Schematic representation of the workflow used for reconstructing the neuroblastoma network and searching for master regulators of a metastatic gene signature.

were considered high risk by COG risk stratification, and the rest was ranked as intermediate risk. We have found that MAX expression was significantly higher in patients who did not present disease setback in five years as opposed to those who had or were already in relapse (Figure 3), indicating that this
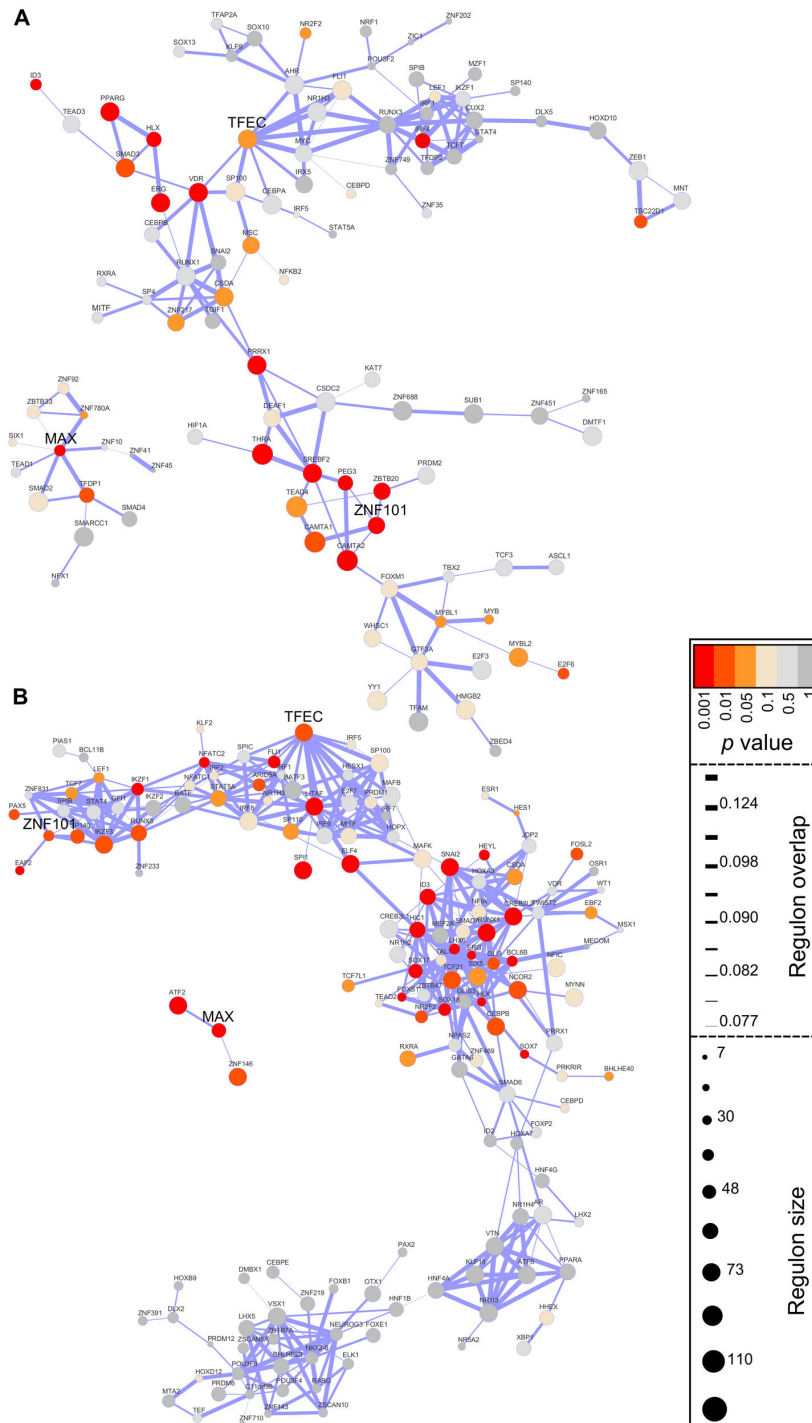
Figure 2: The neuroblastoma reconstructed regulatory network. Principal components of the regulatory networks inferred using GSE3996 (A) and GSE16476 (B) datasets. Each node represents a regulon, which is named by its regulator transcription factor. Node sizes are proportional to the number of regulon members, node colors are representative of enrichment significance, and edges widths are proportional to regulon overlap (using Jaccard similarity).

Figure 2. The neuroblastoma reconstructed regulatory network. Principal components of the regulatory networks inferred using GSE3996 (A) and GSE16476 (B) datasets. Each node represents a regulon, which is named by its regulator transcription factor. Node sizes are proportional to the number of regulon members, node colors are representative of enrichment significance, and edges widths are proportional to regulon overlap (using Jaccard similarity).

gene plays an important role in disease progression. We did not find any significant alterations in TFEC expression.

We have also analyzed two neuroblastoma survival cohorts with gene expression data. In the first cohort (E-TABM-38, n=130) [32], we have found that lower MAX expression correlated with decreased patient survival (Figure 4a), corroborating our previous result. However, this was not the case with the second cohort (E-MTAB-179, n=478) [33], in which higher MAX expression significantly correlated with poor survival (Figure 4b). We could not analyze TFEC because these studies were made with custom array platforms that did not contain probes for this gene. This was also the case with the regulons members themselves, which were poorly represented in these platforms and could not be analyzed further. We chose to analyze these genes in other types of cancers so as to detect if there was a general trend for their expression which could confirm our previous results with MAX and shed some light at TFEC. Using the Kaplan-Meier Plotter web tool [34,35], we have found that higher MAX expression was associated with improved prognosis in breast cancer (Figure S1), in lung cancer (Figure S2), and not related to prognosis in ovarian cancer cohorts (Figure S3). As for TFEC, we have found association with higher expression of this gene and improved lung cancer survival (Figure S4). We have not found associations with TFEC expression in breast and ovarian cancer outcome (Figures S5 and S6, respectively). Detailed results from this analysis are presented in Table 1.

Role of MAX and TFEC in SH-SY5Y differentiation

To understand whether the alterations detected previously in MAX expression could be involved in neuroblastoma cells differentiation and, as such, provide an explanation to why there seems to be a correlation with patient outcome, we chose to study a dataset of neuroblastoma SH-SY5Y cells undergoing differentiation (GSE9169) [36]. In this study, the authors differentiated cells by treatment with retinoic acid for 8 days, with further addition of brain-derived neurotrophic factor (BDNF) after the 5th day. We have found a significant increase in MAX expression starting at the 3rd and 5th days, which lasted until the end of the differentiation protocol (Figure 5). We have not found any significant alterations in TFEC expression during the course of this experiment (data not shown).

We have also performed a similar SH-SY5Y differentiation protocol in our laboratory and quantified MAX protein content during differentiation. We have found an increased amount of MAX after the 4th day of experiment (Figure 6), demonstrating that the changes in mRNA expression are indeed reflected at protein level in neuroblastoma cells.

Discussion

Neuroblastoma is clinically very heterogenic, and one of the possible outcomes for this disease is spontaneous tumor regression [1,2]. There are reported cases of differentiation of these tumors into benign ganglioneuromas, but the mechanisms by which this processes can occur are still unknown [8,37,38]. We have queried our reconstructed neuroblastoma network with a genetic signature related to

bone marrow metastases, which are considered grim clinical prognosis for neuroblastoma patients [3,39]. We have observed that MAX regulon was significantly enriched with differentially expressed genes, suggesting its role as a master regulator of the transition from primary tumor to metastasis. Also, this regulon was not enriched in a healthy control network, neither involved with a nonspecific proliferative signature, corroborating that it is related to neuroblastoma-specific pathways.

MAX expression was significantly altered in three independent gene expression studies. In the first, we have observed that neuroblastoma patients who did not present disease relapse in five years had significant higher levels of MAX expression than patients who either had or were already at relapse. Of important note is that only 3 of the 15 patients studied at relapse were from those included at the beginning of the study, meaning that we can treat this last group independently. This result suggests that higher MAX levels are associated to improved patient prognosis in patients with high risk, MYCN non-amplified metastatic neuroblastoma. Corroborating this data, we have analyzed a survival cohort of neuroblastoma patients and found that individuals with lower MAX expression had significantly decreased survival rates than the others. However, we have found the opposite results in a second neuroblastoma cohort. In this study, patients with higher MAX expression levels presented lower survival rates than the rest. It is possible that this incongruence is brought by the great variability in the two survival cohorts. Patients in these studies, albeit having mostly MYCN non-amplified tumors, were sorted across all risk stages and were subject to different types of therapy, thus making difficult our attempts at understanding this data. Nevertheless, we have also found positive association of MAX expression and improved prognosis in two independent cohorts of breast and lung cancer patients, corroborating its role in disease progression.

To address whether MAX expression could affect neuroblastoma by activating a differentiation pathway, we have analyzed a dataset of MYCN non-amplified neuroblastoma cells SH-SY5Y subjected to a differentiation protocol. Strikingly, we have found a peak of MAX expression between the third and fifth days of differentiation, lasting to the end of the experiment. This data makes us believe that MAX is one the late effectors SH-SY5Y differentiation, suggesting that its expression is implicated in favorable outcomes for MYCN non-amplified neuroblastoma patients by means of enhancing cell differentiation and/or impairing proliferation. These results are further corroborated by the increase of MAX protein levels we have observed in our SH-SY5Y cells differentiation protocol.

The MYC/MAX/MAD network is a key axis in the cell decision-making for differentiation and proliferation, and it is altered in several types of cancer [40–42]. Much has been said about the oncogenic roles of MYC family genes, but, other than being the obligatory heterodimer of MYC and MAD families of proteins, little is known about how variations in MAX expression can affect cell-cycle progression. In earlier cell differentiation studies, it was found that MAX expression remains constant throughout the process [43] (for a review, see [44] and references therein). One study demonstrates that
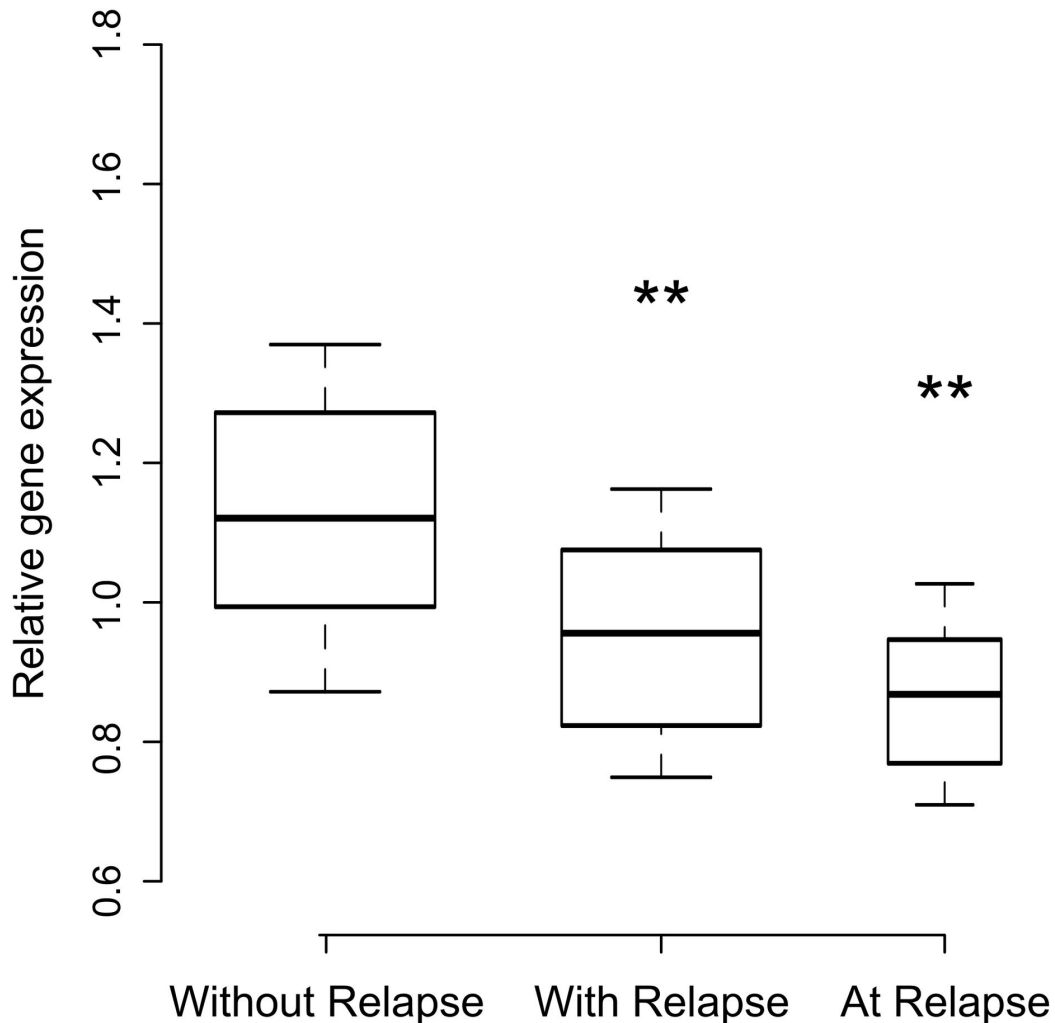
Figure 3: MAX expression in the GSE3446 patient cohort. Boxplots representations of patients who did not present relapse in five years, and patients that either had or were already at relapse, respectively. y-axis represents relative MAX expression (normalized sample value divided by MAX mean across all samples). Boxplot bars indicate lower and upper quartiles, central bar indicates mean, whiskers indicate one standard deviation of mean, and box widths are proportional to sample size. Double asterisks indicates significant differences from the first group (p<0.001, two-tailed pairwise t-tests with Bonferroni correction).

Figure 3. MAX expression in the GSE3446 patient cohort. Boxplots representations of patients who did not present relapse in five years, and patients that either had or were already at relapse, respectively. *y*-axis represents relative MAX expression (normalized sample value divided by MAX mean across all samples). Boxplot bars indicate lower and upper quartiles, central bar indicates mean, whiskers indicate one standard deviation of mean, and box widths are proportional to sample size. Double asterisks indicates significant differences from the first group (*p*<0.001, two-tailed pairwise *t*-tests with Bonferroni correction).
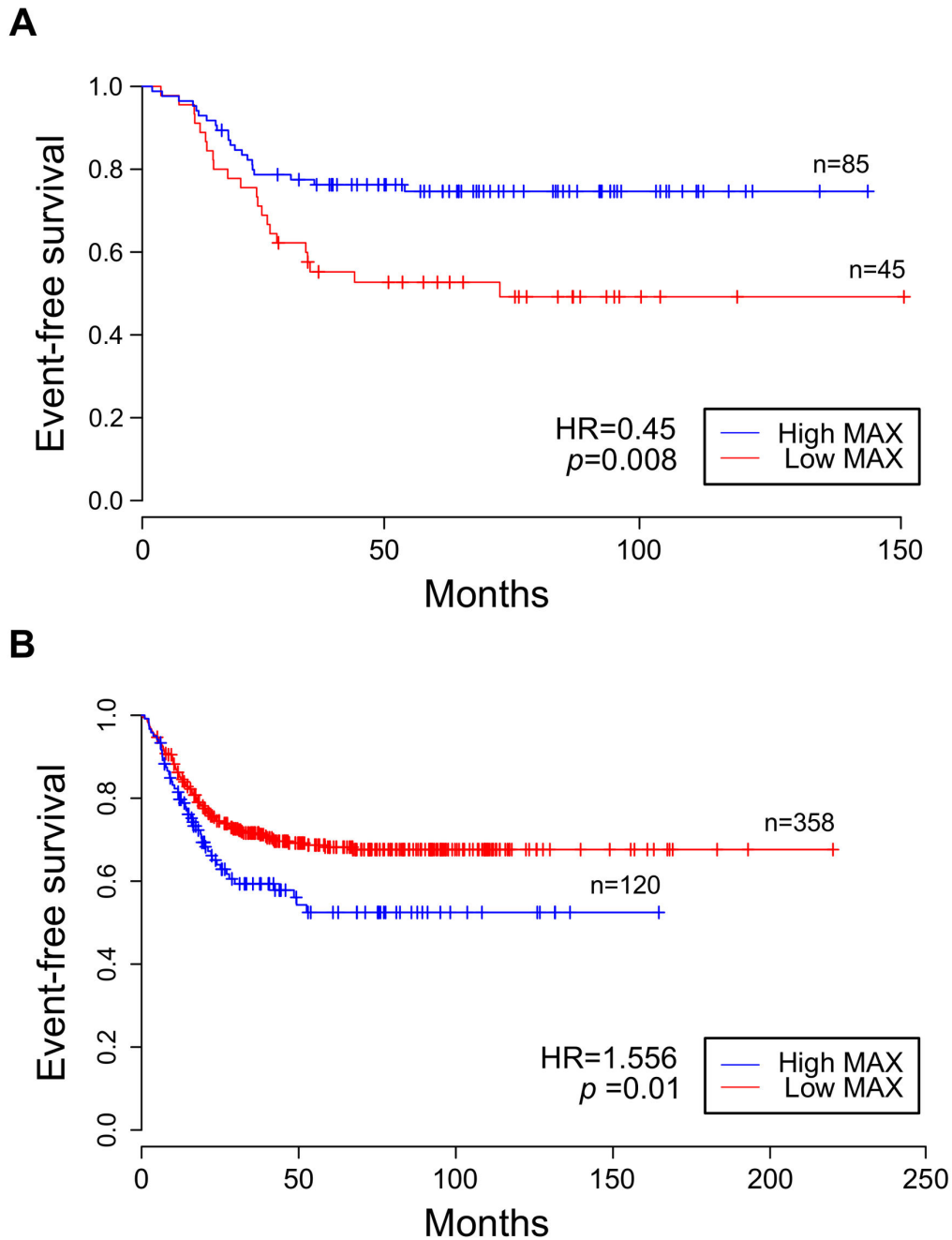
doi: 10.1371/journal.pone.0082457.g003

Figure 4: Kaplan-Meier plot of E-TABM-38 (A) and E-MTAB-179 (B) patient cohorts. x-axis indicates event-free survival time. y-axis represents the percentage of patients event-free survival. The red line represents patients with lower MAX expression, and the blue line, patients with higher MAX expression. Crosses mark censored data.

Figure 4. MAX expression and patient survival. Kaplan-Meier plots of E-TABM-38 (A) and E-MTAB-179 (B) patient cohorts. x-axis indicates event-free survival time. y-axis represents the percentage of patients event-free survival. The red line represents patients with lower MAX expression, and the blue line, patients with higher MAX expression. Crosses mark censored data.

doi: 10.1371/journal.pone.0082457.g004

Table 1. Survival statistics for MAX and TFEC in other type of cancers.

Gene	Cancer	Hazard Ratio	p value	n
MAX	Breast	0.75	1.E-06	2,878
	Lung	0.45	1.E-14	1,404
	Ovarian	ns	ns	1,171
TFEC	Breast	ns	ns	2,878
	Lung	0.57	1.E-09	1,404
	Ovarian	ns	ns	1,171

doi: 10.1371/journal.pone.0082457.t001

MAX $^{-/-}$ mice are subject to early embryonic lethality [45], indicating that this gene plays a vital role in development. These results suggest that this gene is indispensable to differentiation and proliferation processes, albeit only as an accessory part of its network. Lindeman and colleagues [46], however, have elegantly demonstrated that MAX has a more active role decreasing the size and frequency of tumors when overexpressed in lymphoma susceptible mice. In their study, authors have transfected mice lineages with two different MAX transgenes. They observed that both strains presented impaired lymphoproliferation and delayed tumor onset when co-expressed with a highly active MYC transgene. Interestingly, the authors noted that the MAX transgene with higher activity presented more pronounced tumor impairing capacity than the less active one, indicating that MAX has indeed tumor suppressing properties, and that the latter may be dose-dependent. These results are in accordance to related papers that link MAX overexpression with increased differentiation in other cell lines [47,48], and particularly, the work of Peverali and colleagues using neuroblastoma cells [49]. This author demonstrated that retinoic acid-treatment in SK-N-BE neuroblastoma cells overexpressing MAX induced differentiation more than twice as fast as with retinoic acid alone. Curiously, this cells are MYCN amplified, suggesting that MAX expression is sufficient to revert proliferation even in more aggressive neuroblastoma cell variants.

Because these studies were made using overexpression techniques, one may argue that they are not accurate reflections of biological processes occurring at more physiological conditions. The only currently known association with MAX and tumor biology in humans comes from studies of pheochromocytoma and paraganglioma patients. Several independent research centers have observed correlations regarding possible inactivating mutations in MAX exons and the appearance of these tumors [50–54]. These results further suggest the tumor suppressing properties of MAX. To our notice, the spontaneous increase in MAX expression during SH-SY5Y retinoic acid-induced differentiation was the first of this kind to be reported, and leads us to question whether this gene could have a more prominent role in neuroblastoma progression and neuronal cells progenitors development than previously described.

It is known that the MAX:MYC heterodimer induces cell proliferation by interacting with the TRRAP complex and histone acetyltransferases to transcriptionally activate cell cycle activators such as cyclins, and possibly repressing cell cycle inhibitors [55–57]. The MAX:MAD/MNT complex and the MAX homodimer, on the other hand, are proliferation repressors and inducers of terminal differentiation in various cell types [58–60]. The heterodimerization of MYC and MAX is preferential because the stabilization of their bHLH is thermodynamically more favorable than that of MAD and MNT with MAX [61]. Being the limiting factor of its network, one possible mechanism for how changes in MAX expression can interfere directly with proliferation/differentiation fate in cell is by regulating the amount of MYC that can activate gene expression. Lower quantities of MAX would bind preferentially with MYC and change the network balance towards proliferation. As MAX levels rise, all the available MYC binding sites would be occupied and more MAD:MAX, MNT:MAX and MAX:MAX dimers would be formed, switching the balance towards differentiation (Figure 7). In MYCN amplified tumors, a larger fraction of MAX would be used to form pro-proliferative MAX:MYCN heterodimers, explaining how these cells are more resistant to exit division [47]. In line with this, there have been studies with compounds that are able to disrupt the MYC:MAX complex in order to decrease malignancy [62,63]. Of particular relevance is the recent work of Montagne and colleagues [64], that used MAX bHLH as a protein transduction domain for decreasing MYC availability to interact with MAX, thus impairing proliferation in HeLa cells. These results demonstrate MAX capacity of directly interfering with MYC activity. Put together, this data bring MAX forward as a central player in its network, and urge us to dedicate more research in this intriguing subject.

There are questions made during this work that must be addressed in the near future. It would be interesting to compare whether MAX levels are different in stage 4s patients in order to clinically corroborate our findings with SH-SY5Y cells. There is also the putative role of ZNF101 and TFEC, which were pointed as master regulators but could not be implicated in this disease by means of their expression. We have found an indication that the latter is associated in lung cancer outcome, suggesting its importance may be greater than previously thought. The accurate roles of these genes in neuroblastoma progression are an open question and additional assays are needed to address this matter.

Conclusions

We have found evidences that MAX expression plays a role in neuroblastoma biology that has not been previously described, possibly as an additional regulator of the availability of MYC:MAX heterodimers and the balance of proliferation/differentiation. Our analyses also point that this gene may be used as a candidate predictor for positive clinical outcomes of MYCN non-amplified neuroblastoma patients. We have detected significant associations between higher MAX expression and improved survival rates for breast and lung cancer patients, suggesting that the clinical predictor potential can also be extended to other types of tumors. Lastly, we have observed that MAX expression was significantly altered during SH-SY5Y retinoic acid-induced differentiation, providing a
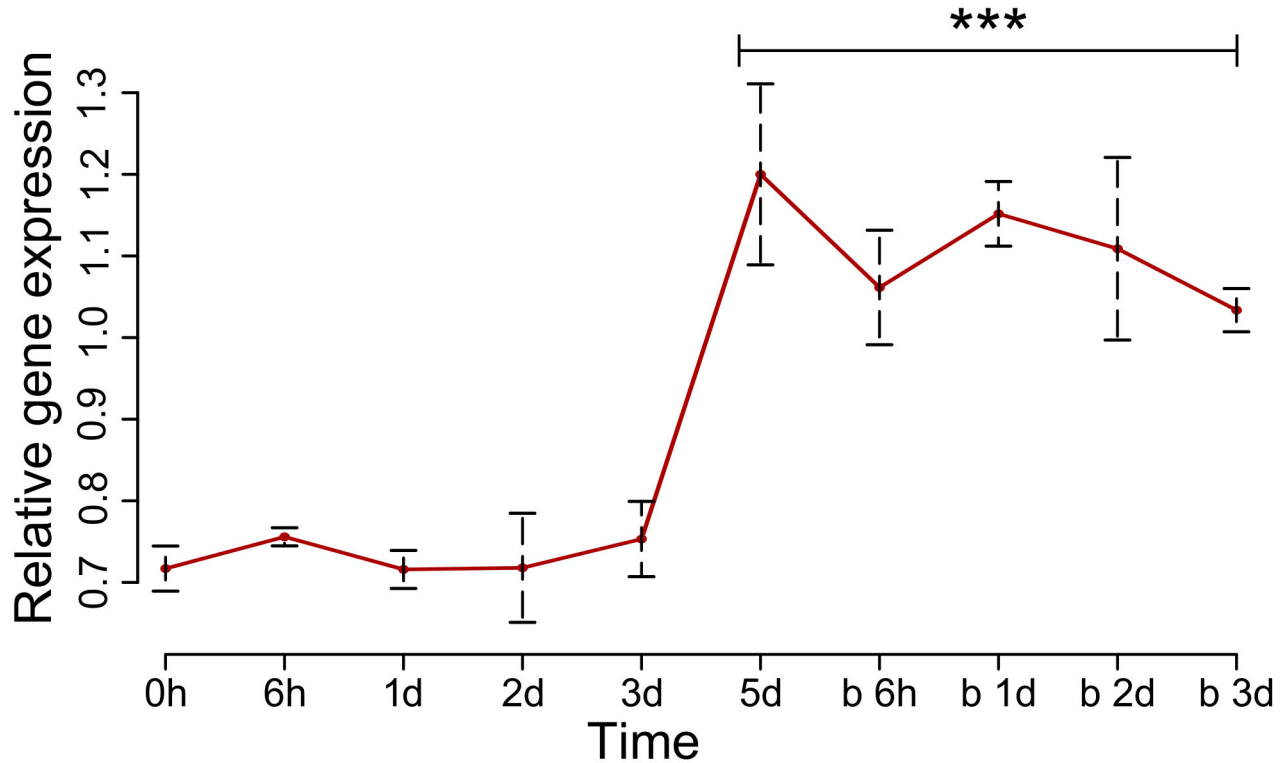
Figure 5: MAX expression in SH-SY5Y cells differentiation. x-axis indicates the time course of differentiation (times preceded by "b" indicate the addition of BDNF in the experimental protocol). y-axis indicates relative MAX expression (normalized sample value divided by MAX mean across all samples). Triple asterisks indicates statistically significant differences from the first three days (p<10-5, two-tailed pairwise t-tests with BH correction).

Figure 5. MAX expression in SH-SY5Y cells differentiation. *x*-axis indicates the time course of differentiation (times preceded by "b" indicate the addition of BDNF in the experimental protocol). *y*-axis indicates relative MAX expression (normalized sample value divided by MAX mean across all samples). Triple asterisks indicates statistically significant differences from the first three days ($p < 10^{-5}$, two-tailed pairwise *t*-tests with BH correction).

doi: 10.1371/journal.pone.0082457.g005

mechanism for which this gene could be linked to tumor progression.

Materials and Methods

Neuroblastoma regulatory network reconstruction and master regulator analysis

The neuroblastoma regulatory network was reconstructed and plotted using, respectively, the RTN [65] and RedeR [66]

packages for R Statistical Computing, available at Bioconductor [67]. This analysis uses the information-theoretic content (i.e. mutual information) of the gene expression profile for inferring relevant pairwise interactions among genes. Regulatory networks were reconstructed from the neuroblastoma biopsies datasets GSE16476 [68] and GSE3960 [69]. Human transcription factors were gathered from the Animal Transcription Factor Database [70]. Master regulator analyses were made using as query a genetic
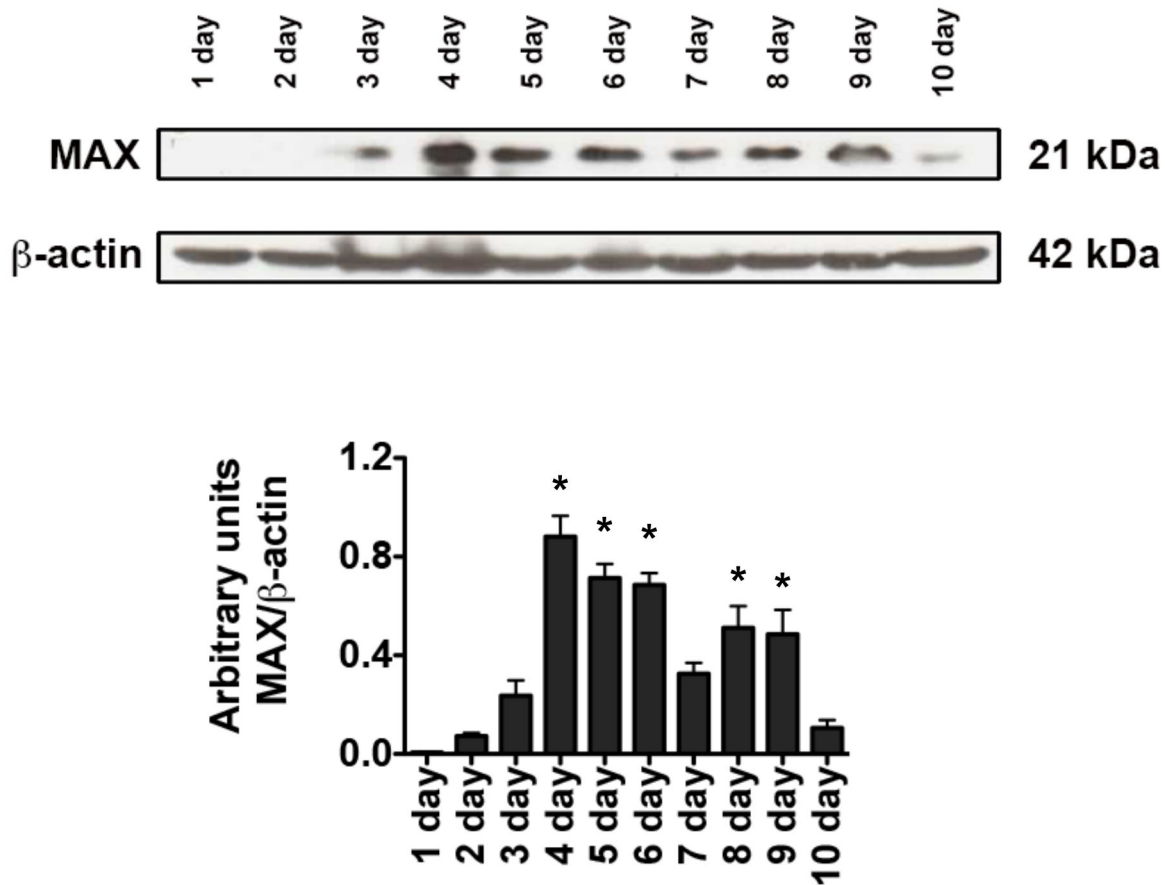
Figure 6: MAX immunocontent in differentiating SH-SY5Y neuroblastoma cells treated with retinoic acid. Asterisks indicate statistically significant (*p*<0.001) changes in the MAX/β-actin ratio in relation to the first day.

signature obtained from aggressive metastases. For assessing this metastatic signature, we have used Limma R package [71] on the GSE25623 dataset [72] in order to discover differentially expressed genes between neuroblastoma primary tumors and bone marrow metastases (for a schematic description of this workflow, we once again refer the reader to Figure 1). Master regulators DNA-binding motifs were searched using the FIMO tool for transcription factor binding site prediction [73]. We have considered target genes only those that had at least one motif occurrence in their flanking regions (either 2.5 Kb up and downstream from the gene start codon [74]). To assess whether regulons were significantly enriched with genes regulated by its transcription factor, we verified if our prediction was significantly higher than randomly sampling all human genes 100,000 times using regulon-sized samples. Statistical significance was calculated using one-tailed *z*-tests.

SH-SY5Y differentiation and Western blot analysis

Human neuroblastoma SH-SY5Y cells were obtained from the American Type Culture Collection (ATCC, Rockville, MD). Cells were cultivated using Dulbecco's Modified Eagle's Medium supplemented with 10% heat-inactivated fetal bovine serum, L-glutamine (2mM), and 0.28 mg/ml of gentamycin sulfate in a 5% CO_2 humidified incubator at 37°C. Differentiation protocol consisted in reducing fetal bovine serum concentration to 1% and adding retinoic acid (10μM) during 10 days. The culture medium was replaced every three days. Cells were grown in 6-well cluster dishes.
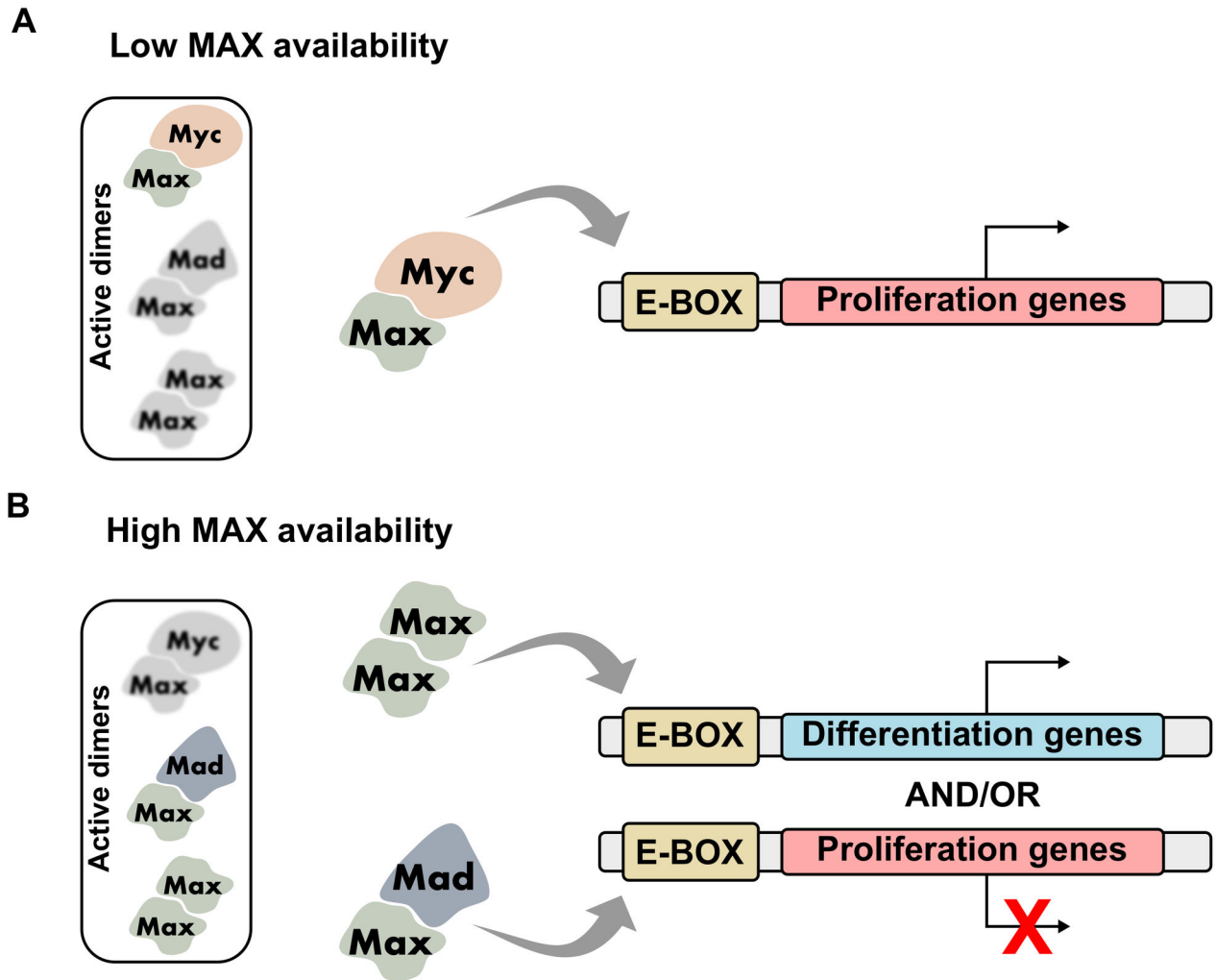
A

Low MAX availability

B

High MAX availability

Figure 7: Schematic representation of MAX expression in neuroblastoma cell behavior. Cells with lower MAX concentration preferentially form pro-proliferative MYC:MAX heterodimers, leading to increased proliferation and de-differentiation (A). With higher MAX levels, there is no more free MYC available and more MAD:MAX heterodimers and MAX homodimers are formed, shifting balance towards differentiation (B).

Figure 7. **Schematic representation of MAX expression in neuroblastoma cell behavior.** Cells with lower MAX concentration preferentially form pro-proliferative MYC:MAX heterodimers, leading to increased proliferation and de-differentiation (A). With higher MAX levels, there is no more free MYC available and more MAD:MAX heterodimers and MAX homodimers are formed, shifting balance towards differentiation (B).

Proteins (20µg) were separated using SDS-PAGE – 10% (w/v) acrylamide, 0.275% (w/v) bisacrylamide gels – and electrotransferred onto nitrocellulose membranes. Membranes were then incubated in Tris-buffered saline Tween-20 [TBS-T; 20 mM Tris–HCl, pH 7.5, 137 mM NaCl, 0.05% (v/v) Tween 20] for 1h at room temperature. Subsequently, membranes were incubated for 12h with polyclonal rabbit anti-MAX (1:1.000 dilution; Cell Signaling). After washing in TBST, blots were incubated with rabbit peroxidase-linked anti-immunoglobulin G (IgG) antibodies (1:10.000 dilution) for 1.5h at room temperature. Chemiluminescent bands were detected, and densitometric analysis was performed by Image-J® software. All analyses were performed in triplicate.

Statistical analyses

Unless stated otherwise, all statistical analyses were made using two-tailed t-tests. Pairwise t-tests were made with Bonferroni and BH p-value corrections when more than two groups were analyzed. The latter was used when testing for more than four groups. Survival curves for gene expression were drawn using Kaplan-Meier plot and tested with Cox proportional hazards model using the Survival [75] and Survcomp [76] packages for R statistical computing. The best cutoff for samples split was chosen by plotting all percentiles of gene expression between the upper and lower quartiles and selecting the best performing threshold. For breast, lung and ovarian cancers, statistics were made using the Kaplan-Meier Plotter web tool. Parameters used in all analyses were auto select split cutoff and only the JetSet best probes were selected [77]. All gene expression profiles used in this paper were obtained from Gene Expression Omnibus and ArrayExpress public databases.

Supporting Information

Figure S1. Kaplan-Meier plot of the 2012 breast cancer patient cohort available at the Kaplan-Meier Plotter web tool. x-axis indicates event-free survival time. y-axis represents the percentage of patients event-free survival. The black line represents patients with lower MAX expression, and the red line, patients with higher MAX expression. Crosses mark censored data.
(TIFF)

Figure S2. Kaplan-Meier plot of the unified lung cancer patient cohort available at the Kaplan-Meier Plotter web tool. x-axis indicates event-free survival time. y-axis represents the percentage of patients event-free survival. The black line represents patients with lower MAX expression, and the red line, patients with higher MAX expression. Crosses mark censored data.
(TIFF)

Figure S3. Kaplan-Meier plot of the 2013 ovarian cancer patient cohort available at the Kaplan-Meier Plotter web tool. x-axis indicates event-free survival time. y-axis represents the percentage of patients event-free survival. The black line represents patients with lower MAX expression, and the red

line, patients with higher MAX expression. Crosses mark censored data.
(TIFF)

Figure S4. Kaplan-Meier plot of the 2012 breast cancer patient cohort available at the Kaplan-Meier Plotter web tool. x-axis indicates event-free survival time. y-axis represents the percentage of patients event-free survival. The black line represents patients with lower TFEC expression, and the red line, patients with higher TFEC expression. Crosses mark censored data.
(TIFF)

Figure S5. Kaplan-Meier plot of the lung cancer patient cohort available at the Kaplan-Meier Plotter web tool. x-axis indicates event-free survival time. y-axis represents the percentage of patients event-free survival. The black line represents patients with lower TFEC expression, and the red line, patients with higher TFEC expression. Crosses mark censored data.
(TIFF)

Figure S6. Kaplan-Meier plot of the 2013 ovarian cancer patient cohort available at the Kaplan-Meier Plotter web tool. x-axis indicates event-free survival time. y-axis represents the percentage of patients event-free survival. The black line represents patients with lower TFEC expression, and the red line, patients with higher TFEC expression. Crosses mark censored data.
(TIFF)

Table S1. Overview of all Master Regulators found in our study and detailed composition of MAX, TFEC and ZNF101 regulons.
(XLSX)

Author Contributions

Conceived and designed the experiments: RD'OA RJSD MAAC JCFM. Performed the experiments: RD'OA MABP VMR. Analyzed the data: RD'OA RJSD MAAC JCFM. Contributed reagents/materials/analysis tools: MAAC DPG JCFM. Wrote the manuscript: RD'OA. Designed the software used in the analysis: MAAC.

References

1. Maris JM, Hogarty MD, Bagatell R, Cohn SL (2007) Neuroblastoma. Lancet 369: 2106–2120. doi:10.1016/S0140-6736(07)60983-0. PubMed: 17586306.
2. Howman-Giles R, Shaw PJ, Uren RF, Chung DKV (2007) Neuroblastoma and other neuroendocrine tumors. Semin Nucl Med 37: 286–302. doi:10.1053/j.semnuclmed.2007.02.009. PubMed: 17544628.
3. Spix C, Pastore G, Sankila R, Stiller CA, Steliarova-Foucher E (2006) Neuroblastoma incidence and survival in European children (1978-1997): report from the Automated Childhood Cancer Information System project. Eur J Cancer 42: 2081–2091. doi:10.1016/j.ejca.2006.05.008. PubMed: 16919772.
4. Parise IZS, Haddad BR, Cavalli LR, Pianovski MAD, Maggio EM et al. (2006) Neuroblastoma in southern Brazil: an 11-year study. J Pediatr

Hematol Oncol 28: 82–87. doi:10.1097/01.mph.0000199601.35010.52. PubMed: 16462579.
5. Mills SE (2002) Neuroectodermal neoplasms of the head and neck with emphasis on neuroendocrine carcinomas. Mod Pathol 15: 264–278. doi:10.1038/modpathol.3880522. PubMed: 11904342.
6. Modak S, Cheung N-KV (2010) Neuroblastoma: Therapeutic strategies for a clinical enigma. Cancer Treat Rev 36: 307–317. doi:10.1016/j.ctrv.2010.02.006. PubMed: 20227189.
7. Monclair T, Brodeur GM, Ambros PF, Brisse HJ, Cecchetto G et al. (2009) The International Neuroblastoma Risk Group (INRG) staging system: an INRG Task Force report. J Clin Oncol 27: 298–303. doi:10.1200/JCO.2008.16.6876. PubMed: 19047290.
8. Maris JM, Matthay KK (1999) Molecular biology of neuroblastoma. J Clin Oncol 17: 2264–2279. PubMed: 10561284.

9. Schwab M, Westermann F, Hero B, Berthold F (2003) Neuroblastoma: biology and molecular and chromosomal pathology. Lancet Oncol 4: 472–480. doi:10.1016/S1470-2045(03)01166-5. PubMed: 12901961.

10. Tonini GP, Boni L, Pession A, Rogers D, Iolascon A et al. (1997) MYCN oncogene amplification in neuroblastoma is associated with worse prognosis, except in stage 4s: the Italian experience with 295 children. J Clin Oncol 15: 85–93.

11. Westermann F, Schwab M (2002) Genetic parameters of neuroblastomas. Cancer Lett 184: 127–147. doi:10.1016/S0304-3835(02)00199-4. PubMed: 12127685.

12. Bagatell R, Beck-Popovic M, London WB, Zhang Y, Pearson ADJ et al. (2009) Significance of MYCN amplification in international neuroblastoma staging system stage 1 and 2 neuroblastoma: a report from the International Neuroblastoma Risk Group database. J Clin Oncol 27: 365–370. PubMed: 19047282.

13. Kushner BH, Cheung NK, LaQuaglia MP, Ambros PF, Ambros IM et al. (1996) International neuroblastoma staging system stage 1 neuroblastoma: a prospective study and literature review. J Clin Oncol 14: 2174–2180. PubMed: 8683252.

14. Bhatnagar SN, Sarin YK (2012) Neuroblastoma: a review of management and outcome. Indian J Pediatr 79: 787–792. doi:10.1007/s12098-012-0748-2. PubMed: 22528697.

15. Schmidt ML, Lukens JN, Seeger RC, Brodeur GM, Shimada H et al. (2000) Biologic factors determine prognosis in infants with stage IV neuroblastoma: A prospective Children's Cancer Group study. J Clin Oncol 18: 1260–1268. PubMed: 10715296.

16. Schmidt ML, Lal A, Seeger RC, Maris JM, Shimada H et al. (2005) Favorable prognosis for patients 12 to 18 months of age with stage 4 nonamplified MYCN neuroblastoma: a Children's Cancer Group Study. J Clin Oncol 23: 6474–6480. doi:10.1200/JCO.2005.05.183. PubMed: 16116154.

17. Asgharzadeh S, Pique-Regi R, Sposto R, Wang H, Yang Y et al. (2006) Prognostic significance of gene expression profiles of metastatic neuroblastomas lacking MYCN gene amplification. J Natl Cancer Inst 98: 1193–1203. doi:10.1093/jnci/djj330. PubMed: 16954472.

18. Babu MM, Lang B, Aravind L (2009) Methods to reconstruct and compare transcriptional regulatory networks. Methods Mol Biol 541: 163–180. doi:10.1007/978-1-59745-243-4_8. PubMed: 19381525.

19. Margolin A, Wang K, Lim WK, Kustagi M, Nemenman I et al. (2006) Reverse engineering cellular networks. Nat Protoc 1: 662–671.

20. Mehra S, Hu W-S, Karypis G (2004) A Boolean algorithm for reconstructing the structure of regulatory networks. Metab Eng 6: 326–339. doi:10.1016/j.ymben.2004.05.002. PubMed: 15491862.

21. Segal E, Shapira M, Regev A, Pe'er D, Botstein D et al. (2003) Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. Nat Genet 34: 166–176. doi:10.1038/ng1165. PubMed: 12740579.

22. Basso K, Margolin A, Stolovitzky G, Klein U, Dalla-Favera R et al. (2005) Reverse engineering of regulatory networks in human B cells. Nat Genet 37: 382–390.

23. Carro MS, Lim WK, Alvarez MJ, Bollo RJ, Zhao X et al. (2010) The transcriptional network for mesenchymal transformation of brain tumours. Nature 463: 318–325. doi:10.1038/nature08712. PubMed: 20032975.

24. Chen D-T, Nasir A, Culhane A, Venkataramu C, Fulp W et al. (2010) Proliferative genes dominate malignancy-risk gene signature in histologically-normal breast tissue. Breast Cancer Res Treat 119: 335–346. doi:10.1007/s10549-009-0344-y. PubMed: 19266279.

25. Venet D, Dumont JE, Detours V (2011) Most random gene expression signatures are significantly associated with breast cancer outcome. PLoS Comput Biol 7: e1002240. PubMed: 22028643.

26. Sandelin A, Alkema W, Engström P, Wasserman WW, Lenhard B (2004) JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res 32: D91–D94. doi:10.1093/nar/gnh092. PubMed: 14681366.

27. Portales-Casamar E, Thongjuea S, Kwon AT, Arenillas D, Zhao X et al. (2010) JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. Nucleic Acids Res 38: D105–D110. doi:10.1093/nar/gkp950. PubMed: 19906716.

28. Ledent V, Paquet O, Vervoort M (2002) Phylogenetic analysis of the human basic helix-loop-helix proteins. Genome Biol 3: RESEARCH0030. PubMed: 12093377.

29. Fernandez PC, Frank SR, Wang L, Schroeder M, Liu S et al. (2003) Genomic targets of the human c-Myc protein. Genes Dev 17: 1115–1129. doi:10.1101/gad.1067003. PubMed: 12695333.

30. Li Z, Van Calcar S, Qu C, Cavenee WK, Zhang MQ et al. (2003) A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells. Proc Natl Acad Sci U S A 100: 8164–8169. doi:10.1073/pnas.1332764100. PubMed: 12808131.

31. Yap C-S, Peterson AL, Castellani G, Sedivy JM, Neretti N (2011) Kinetic profiling of the c-Myc transcriptome and bioinformatic analysis of repressed gene promoters. Cell Cycle 10: 2184–2196. doi:10.4161/cc.10.13.16249. PubMed: 21623162.

32. Oberthuer A, Berthold F, Warnat P, Hero B, Kahlert Y et al. (2006) Customized oligonucleotide microarray gene expression-based classification of neuroblastoma patients outperforms current clinical risk stratification. J Clin Oncol 24: 5070–5078. doi:10.1200/JCO.2006.06.1879. PubMed: 17075126.

33. Oberthuer A, Juraeva D, Li L, Kahlert Y, Westermann F et al. (2010) Comparison of performance of one-color and two-color gene-expression analyses in predicting clinical endpoints of neuroblastoma patients. Pharmacogenomics J 10: 258–266. doi:10.1038/tpj.2010.53. PubMed: 20676065.

34. Gyorffy B, Lánczky A, Szállási Z (2012) Implementing an online tool for genome-wide validation of survival-associated biomarkers in ovarian-cancer using microarray data from 1287 patients. Endocr Relat Cancer 19: 197–208. doi:10.1530/ERC-11-0329. PubMed: 22277193.

35. Györffy B, Lanczky A, Eklund AC, Denkert C, Budczies J et al. (2010) An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. Breast Cancer Res Treat 123: 725–731. doi:10.1007/s10549-009-0674-9. PubMed: 20020197.

36. Nishida Y, Adati N, Ozawa R, Maeda A, Sakaki Y et al. (2008) Identification and classification of genes regulated by phosphatidylinositol 3-kinase- and TRKB-mediated signalling pathways during neuronal differentiation in two subtypes of the human neuroblastoma cell line SH-SY5Y. BMC Res Notes 1: 95. doi:10.1186/1756-0500-1-95. PubMed: 18957096.

37. D'Angio G, Evans A, Koop CE (1971) special pattern of widespread neuroblastoma with a favourable prognosis. Lancet 297: 1046–1049. doi:10.1016/S0140-6736(71)91606-0.

38. Bénard J, Raguénez G, Kauffmann A, Valent A, Ripoche H et al. (2008) MYCN-non-amplified metastatic neuroblastoma with good prognosis and spontaneous regression: a molecular portrait of stage 4S. Mol Oncol 2: 261–271. doi:10.1016/j.molonc.2008.07.002. PubMed: 19383347.

39. Garaventa A, Parodi S, De Bernardi B, Dau D, Manzitti C et al. (2009) Outcome of children with neuroblastoma after progression or relapse. A retrospective study of the Italian neuroblastoma registry. Eur J Cancer 45: 2835–2842. doi:10.1016/j.ejca.2009.06.010. PubMed: 19616426.

40. O'Hagan RC, Schreiber-Agus N, Chen K, David G, Engelman JA et al. (2000) Gene-target recognition among members of the myc superfamily and implications for oncogenesis. Nat Genet 24: 113–119. doi:10.1038/72761. PubMed: 10655054.

41. Hurlin PJ, Dezfouli S (2004) Functions of myc:max in the control of cell proliferation and tumorigenesis. Int Rev Cytol 238: 183–226. doi:10.1016/S0074-7696(04)38004-6. PubMed: 15364199.

42. Dang CV (2012) MYC on the path to cancer. Cell 149: 22–35. doi:10.1016/j.cell.2012.03.003. PubMed: 22464321.

43. Pulverer B, Sommer A, McArthur GA, Eisenman RN, Lüscher B (2000) Analysis of Myc/Max/Mad network members in adipogenesis: inhibition of the proliferative burst and differentiation by ectopically expressed Mad1. J Cell Physiol 183: 399–410. doi:10.1002/(SICI)1097-4652(200006)183:3. PubMed: 10797315.

44. Rottmann S, Lüscher B (2006) The Mad side of the Max network: antagonizing the function of Myc and more. Curr Top Microbiol Immunol 302: 63–122. doi:10.1007/3-540-32952-8_4. PubMed: 16620026.

45. Shen-Li H, O'Hagan RC, Hou H, Horner JW, Lee HW et al. (2000) Essential role for Max in early embryonic growth and development. Genes Dev 14: 17–22. PubMed: 10640271.

46. Lindeman GJ, Harris AW, Bath ML, Eisenman RN, Adams JM (1995) Overexpressed max is not oncogenic and attenuates myc-induced lymphoproliferation and lymphomagenesis in transgenic mice. Oncogene 10: 1013–1017. PubMed: 7898919.

47. Gu W, Cechova K, Tassi V, Dalla-Favera R (1993) Opposite regulation of gene transcription and cell proliferation by c-Myc and Max. Proc Natl Acad Sci U S A 90: 2935–2939. doi:10.1073/pnas.90.7.2935. PubMed: 8385351.

48. Cañelles M, Delgado MD, Hyland KM, Lerga A, Richard C et al. (1997) Max and inhibitory c-Myc mutants induce erythroid differentiation and resistance to apoptosis in human myeloid leukemia cells. Oncogene 14: 1315–1327. doi:10.1038/sj.onc.1200948. PubMed: 9178892.

49. Peverali FA, Orioli D, Tonon L, Ciana P, Bunone G et al. (1996) Retinoic acid-induced growth arrest and differentiation of neuroblastoma cells are counteracted by N-myc and enhanced by max overexpressions. Oncogene 12: 457–462. PubMed: 8570225.

50. Burnichon N, Cascón A, Schiavi F, Morales NP, Comino-Méndez I et al. (2012) MAX mutations cause hereditary and sporadic pheochromocytoma and paraganglioma. Clin Cancer Res 18: 2828–2837. doi:10.1158/1078-0432.CCR-12-0160. PubMed: 22452945.

51. Comino-Méndez I, Gracia-Aznárez FJ, Schiavi F, Landa I, Leandro-García LJ et al. (2011) Exome sequencing identifies MAX mutations as a cause of hereditary pheochromocytoma. Nat Genet 43: 663–667. doi:10.1038/ng.861. PubMed: 21685915.

52. Welander J, Söderkvist P, Gimm O (2011) Genetics and clinical characteristics of hereditary pheochromocytomas and paragangliomas. Endocr Relat Cancer 18: R253–R276. doi:10.1530/ERC-11-0170. PubMed: 22041710.

53. Kirmani S, Young WF (2008) Hereditary Paraganglioma-Pheochromocytoma Syndromes. University of Washington, Seattle.

54. Galan SR, Kann PH (2013) Genetics and molecular pathogenesis of pheochromocytoma and paraganglioma. Clin Endocrinol (Oxf) 78: 165–175. doi:10.1111/cen.12071. PubMed: 23061808.

55. McMahon SB, Wood MA, Cole MD (2000) The essential cofactor TRRAP recruits the histone acetyltransferase hGCN5 to c-Myc. Mol Cell Biol 20: 556–562. doi:10.1128/MCB.20.2.556-562.2000. PubMed: 10611234.

56. Park J (2001) The ATM-related domain of TRRAP is required for histone acetyltransferase recruitment and Myc-dependent oncogenesis. Genes Dev 15: 1619–1624. doi:10.1101/gad.900101. PubMed: 11445536.

57. Bouchard C, Dittrich O, Kiermaier A, Dohmann K, Menkel A et al. (2001) Regulation of cyclin D2 gene expression by the Myc/Max/Mad network: Myc-dependent TRRAP recruitment and histone acetylation at the cyclin D2 promoter. Genes Dev 15: 2042–2047. doi:10.1101/gad.907901. PubMed: 11511535.

58. Ayer DE, Lawrence QA, Eisenman RN (1995) Mad-max transcriptional repression is mediated by ternary complex formation with mammalian homologs of yeast repressor Sin3. Cell 80: 767–776. doi:10.1016/0092-8674(95)90355-0. PubMed: 7889570.

59. Hurlin PJ, Quéva C, Eisenman RN (1997) Mnt, a novel Max-interacting protein is coexpressed with Myc in proliferating cells and mediates repression at Myc binding sites. Genes Dev 11: 44–58. doi:10.1101/gad.11.1.44. PubMed: 9000049.

60. Ayer DE, Eisenman RN (1993) A switch from Myc:Max to Mad:Max heterocomplexes accompanies monocyte/macrophage differentiation. Genes Dev 7: 2110–2119. doi:10.1101/gad.7.11.2110. PubMed: 8224841.

61. Banerjee A, Hu J, Goss DJ (2006) Thermodynamics of protein-protein interactions of cMyc, Max, and Mad: effect of polyions on protein dimerization. Biochemistry 45: 2333–2338. doi:10.1021/bi0522551. PubMed: 16475822.

62. Wang H, Hammoudeh DI, Follis AV, Reese BE, Lazo JS et al. (2007) Improved low molecular weight Myc-Max inhibitors. Mol Cancer Ther 6: 2399–2408. doi:10.1158/1535-7163.MCT-07-0005. PubMed: 17876039.

63. Follis AV, Hammoudeh DI, Daab AT, Metallo SJ (2009) Small-molecule perturbation of competing interactions between c-Myc and Max. Bioorg Med Chem Lett 19: 807–810. doi:10.1016/j.bmcl.2008.12.025. PubMed: 19114306.

64. Montagne M, Beaudoin N, Fortin D, Lavoie CL, Klinck R et al. (2012) The Max b-HLH-LZ can transduce into cells and inhibit c-Myc transcriptional activities. PLOS ONE 7: e32172. doi:10.1371/journal.pone.0032172. PubMed: 22384171.

65. Fletcher MNC, Castro MAA, Wang X, de Santiago I, O'Reilly M et al. (2013) Master regulators of FGFR2 signalling and breast cancer risk. Nat Commun 4: 2464. PubMed: 24043118.

66. Castro MA, Wang X, Fletcher MN, Meyer KB, Markowetz F (2012) RedeR: R/Bioconductor package for representing modular structures, nested networks and multiple levels of hierarchical associations. Genome Biol 13: R29. doi:10.1186/gb-2012-13-4-r29. PubMed: 22531049.

67. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5: R80. doi:10.1186/gb-2004-5-10-r80. PubMed: 15461798.

68. Molenaar JJ, Koster J, Zwijnenburg Da, van Sluis P, Valentijn LJ et al. (2012) Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. Nature 483: 589–593. doi:10.1038/nature10910. PubMed: 22367537.

69. Wang Q, Diskin S, Rappaport E, Attiyeh E, Mosse Y et al. (2006) Integrative genomics identifies distinct molecular classes of neuroblastoma and shows that multiple genes are targeted by regional alterations in DNA copy number. Cancer Res 66: 6050–6062. doi:10.1158/0008-5472.CAN-05-4618. PubMed: 16778177.

70. Zhang HM, Chen H, Liu W, Liu H, Gong J et al. (2012) AnimalTFDB: a comprehensive animal transcription factor database. Nucleic Acids Res 40: D144–D149. doi:10.1093/nar/gkr965. PubMed: 22080564.

71. Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3: Article3

72. Morandi F, Scaruffi P, Gallo F, Stigliani S, Moretti S et al. (2012) Bone marrow-infiltrating human neuroblastoma cells express high levels of calprotectin and HLA-G proteins. PLOS ONE 7: e29922. doi:10.1371/journal.pone.0029922. PubMed: 22253825.

73. Grant CE, Bailey TL, Noble WS (2011) FIMO: scanning for occurrences of a given motif. Bioinformatics 27: 1017–1018. doi:10.1093/bioinformatics/btr064. PubMed: 21330290.

74. Kristiansson E, Thorsen M, Tamás MJ, Nerman O (2009) Evolutionary forces act on promoter length: identification of enriched cis-regulatory elements. Mol Biol Evol 26: 1299–1307. doi:10.1093/molbev/msp040. PubMed: 19258451.

75. Therneau T, Grambsch P (2000) Modeling survival data: extending the Cox model. Springer.

76. Haibe-Kains B, Desmedt C, Sotiriou C, Bontempi G (2008) A comparative study of survival models for breast cancer prognostication based on microarray data: does a single gene beat them all? Bioinformatics 24: 2200–2208. doi:10.1093/bioinformatics/btn374. PubMed: 18635567.

77. Li Q, Birkbak NJ, Gyorffy B, Szallasi Z, Eklund AC (2011) Jetset: selecting the optimal microarray probe set to represent a gene. BMC Bioinformatics 12: 474. doi:10.1186/1471-2105-12-474. PubMed: 22172014.