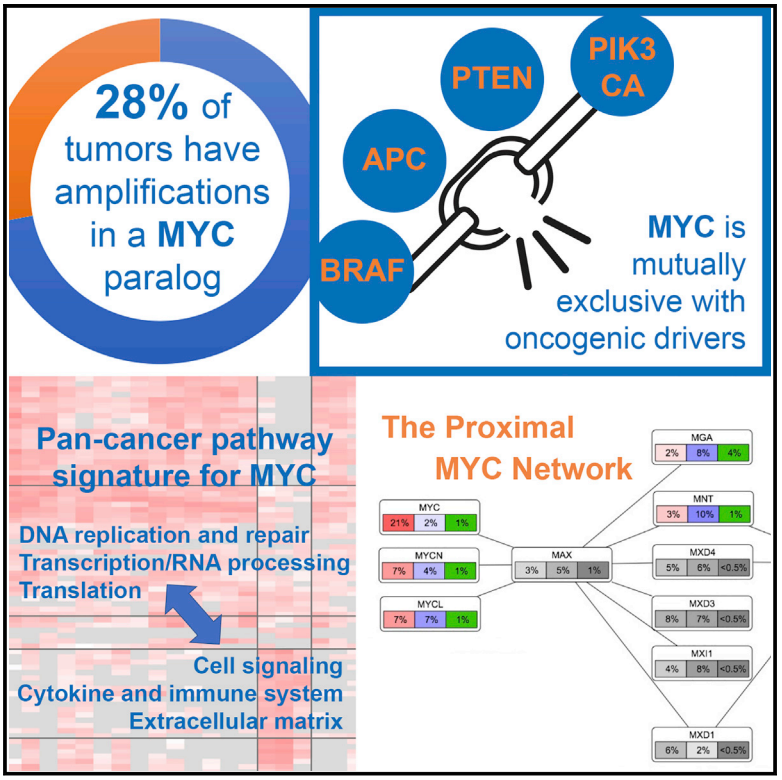
Article

Cell Systems

Pan-cancer Alterations of the MYC Oncogene and Its Proximal Network across the Cancer Genome Atlas

Graphical Abstract



Highlights

- MYC paralogs are significantly amplified (28% of all samples)
- MYC antagonists are mutated (MGA, 4% of samples) or deleted (MNT, 10% of samples)
- MYC alterations are mutually exclusive with *PIK3CA*, *PTEN*, *APC*, or *BRAF* alterations
- Expression analysis reveals pan-cancer and tumor-specific MYC-associated pathways

Authors

Franz X. Schaub, Varsha Dhankani, Ashton C. Berger, ..., Brady Bernard, Carla Grandori, The Cancer Genome Atlas Network

Correspondence

bbernard@systemsbiology.org (B.B.), carlagrandori@curefirst.org (C.G.)

In Brief

We present a computational study determining the frequency and extent of alterations of the MYC network across the 33 human cancers of TCGA. These data, together with MYC, positively correlated pathways as well as mutually exclusive cancer genes, will be a resource for understanding MYC-driven cancers and designing of therapeutics.

source: https://doi.org/10.7892/boris.126390 | downloaded: 4.12.2019

Schaub et al., 2018, Cell Systems 6, 282–300
March 28, 2018 © 2018 The Authors. Published by Elsevier Inc.
https://doi.org/10.1016/j.cels.2018.03.003

CellPress

OPEN ACCESS
CellPress

Cell Systems
Article

Pan-cancer Alterations of the MYC Oncogene and Its Proximal Network across the Cancer Genome Atlas

Franz X. Schaub,[1,2] Varsha Dhankani,[2] Ashton C. Berger,[3] Mihir Trivedi,[2] Anne B. Richardson,[2] Reid Shaw,[2] Wei Zhao,[4] Xiaoyang Zhang,[5] Andrea Ventura,[6] Yuexin Liu,[4] Donald E. Ayer,[7] Peter J. Hurlin,[8] Andrew D. Cherniack,[5] Robert N. Eisenman,[9] Brady Bernard,[10,11,*] Carla Grandori,[1,2,12,*] and The Cancer Genome Atlas Network

[1]Cure First, Seattle, WA, USA
[2]SEngine Precision Medicine, Seattle, WA, USA
[3]The Eli and Edythe L. Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA
[4]Department of Systems Biology, University of Texas MD Anderson Cancer Center, Houston, TX, USA
[5]Dana-Farber Cancer Institute, the Broad Institute of Harvard and MIT, and Harvard Medical School, Boston, MA, USA
[6]Cancer Biology and Genetics Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA
[7]Department of Oncological Sciences, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA
[8]Shriners Hospitals for Children Research Center, Department of Cell, Developmental and Cancer Biology, Oregon Health & Science University, Knight Cancer Institute, Oregon Health & Science University, Portland, OR, USA
[9]Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA
[10]Institute for Systems Biology, Seattle, WA, USA
[11]Providence Health and Services, Portland, OR, USA
[12]Lead Contact
*Correspondence: bbernard@systemsbiology.org (B.B.), carlagrandori@curefirst.org (C.G.)
https://doi.org/10.1016/j.cels.2018.03.003

SUMMARY

Although the *MYC* oncogene has been implicated in cancer, a systematic assessment of alterations of MYC, related transcription factors, and co-regulatory proteins, forming the proximal MYC network (PMN), across human cancers is lacking. Using computational approaches, we define genomic and proteomic features associated with MYC and the PMN across the 33 cancers of The Cancer Genome Atlas. Pan-cancer, 28% of all samples had at least one of the MYC paralogs amplified. In contrast, the MYC antagonists MGA and MNT were the most frequently mutated or deleted members, proposing a role as tumor suppressors. *MYC* alterations were mutually exclusive with *PIK3CA*, *PTEN*, *APC*, or *BRAF* alterations, suggesting that MYC is a distinct oncogenic driver. Expression analysis revealed MYC-associated pathways in tumor subtypes, such as immune response and growth factor signaling; chromatin, translation, and DNA replication/repair were conserved pan-cancer. This analysis reveals insights into MYC biology and is a reference for biomarkers and therapeutics for cancers with alterations of MYC or the PMN.

INTRODUCTION

The *MYC* gene was initially discovered as an oncogene (v-*MYC*) acquired from the host cell genome by a subgroup of avian leukemia viruses. Subsequently, the cellular *MYC* gene and its paralogs (*MYCN* and *MYCL*) were found to be subject to genetic alterations, such as amplification, chromosomal translocation, and viral integration, in a broad spectrum of cancers leading to tumorigenesis. In normal cells, expression of the endogenous *MYC* gene is upregulated in response to diverse mitogenic and developmental signals. The MYC protein functions as a transcription factor that responds to and integrates these signals into broad changes in gene expression, supporting cell growth and proliferation.

Many of the genetic alterations that occur in tumors act to uncouple *MYC* expression from its normal regulatory constraints, thereby resulting in high levels of MYC protein that are less sensitive to normal cellular and extracellular signals (for reviews see Dang and Eisenman, 2014). Such alterations include (1) point mutations in the *MYC* coding region that appear to increase MYC protein stability and activity as secondary events to translocations in lymphoma (Bahram et al., 2000; Hemann et al., 2005); (2) mutation or rare amplification of distal enhancers (Sur et al., 2012; Zhang et al., 2016); and (3) activating mutations in signal transduction pathways (e.g., Wnt, Notch) that augment *MYC* expression (Herranz et al., 2014; Muncan et al., 2006; Weng et al., 2006). Even relatively small constitutive changes in *MYC* expression level (>2-fold relative to normal) have been demonstrated to have biological consequences and influence tumorigenesis (Bazarov et al., 2001; Hofmann et al., 2015; Murphy et al., 2008). Earlier studies showed that multiple cancer types exhibit alterations at *MYC* family gene loci, usually associated with increased MYC mRNA and/or protein levels (Nesbit et al., 1999; Vita and Henriksson, 2006). Experiments in a number of tumor lines and in animal models of cancer indicated that, in many cases, MYC expression is required for tumor initiation, progression, or maintenance (for review see Gabay et al., 2014; Vita and

 © 2018 The Authors. Published by Elsevier Inc.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Henriksson, 2006). Therefore, it is reasonable to consider tumors with dysregulated MYC as "MYC-driven" or "MYC-addicted" tumors. However, the earlier meta-analyses indicated that within a given tumor type or subtype the fraction of tumors with *MYC* family gene rearrangements (amplification or translocations) can vary widely. For example, based on different published reports, *MYC* amplifications in breast cancer were found in 9%–48% of cases and 7%–78% in osteosarcoma (see Vita and Henriksson, 2006). These and other variations are likely due to different methodologies employed to detect rearrangements and to differences in sample sizes. More recently, a report broadly analyzing the landscape of focal amplifications in cancers found *MYC* amplification to be among the most frequent of all such events (Beroukhim et al., 2010).

The functional consequences of MYC de-regulated expression and its influence on gene expression programs and DNA replication or repair processes during normal and oncogenic proliferation have been a subject of intense research (Dominguez-Sola and Gautier, 2014; Sabo and Amati, 2014; Walz et al., 2014). MYC functions with its heterodimerization partner MAX, through recognition of specific DNA elements (Blackwell et al., 1993; Fernandez et al., 2003; Grandori et al., 1996; Guccione et al., 2006), and recruitment of transcriptional co-regulatory molecules linked to histone acetylation (Bouchard et al., 2001; Frank et al., 2001) to elevate expression of a broad, but selective, set of genes via the activation of specific chromatin marks. Upon dysregulation and overexpression, MYC binds to lower-affinity sites in promoters and enhancers in a dose-dependent manner resulting in ectopic regulation (activation or repression) of thousands of genes (between ~2,000 and 4,000 genes [de Pretis et al., 2017; Sabò et al., 2014]). In addition, a role for MYC in the global mRNA amplification observed during transition from quiescence to proliferation was proposed by examining gene expression changes in a B cell line, whose proliferative state was strictly dependent on a conditional MYC allele (Lin et al., 2012; Nie et al., 2012). Recent evidence, however, indicates that this may be largely an indirect effect of MYC as the amplification of the majority of mRNAs does not correlate with binding of MYC at their promoters (Kress et al., 2015).

While the effects of MYC as a transcription factor are often considered in isolation, it is important to consider its function in the context of a network of related transcription factors and interacting co-regulatory proteins that have the potential to influence MYC target gene binding and expression. We refer to this network as the proximal MYC network (PMN), which includes MAX, MGA, MXD1, MXD3, MXD4, MXI1, MNT, MLX, MLXIP, and MLXIPL. All of these proteins have related basic-helix-loop-helix zipper (bHLHZ) domains and can be considered members of the MYC bHLHZ superfamily. The different components of the network are connected through dimerization with MAX, MLX, or both. MAX, in addition to dimerizing with MYC paralogs, also forms heterodimers with the MXD family, comprised of MXI1, MNT, and MGA (Ayer et al., 1993; Hurlin et al., 1995, 1997, 1999; Meroni et al., 1997; Zervos et al., 1993) (Figure 1A). Although far less characterized than MYC proteins, these factors can compete with MYC for binding to MAX and for E-box sites in shared target genes. In contrast to the predominant transcriptional activation function of MYC, the MXD, MXI1, MNT, and MGA proteins repress transcription through the recruitment of corepressor complexes (for reviews see Conacci-Sorrell et al., 2014; Link and Hurlin, 2015). These opposing transcriptional activities, together with functional assays showing that MXDs, MNT, and MGA proteins can antagonize the transforming activity of MYC in cell culture assays, raised the possibility that they function as tumor suppressors. Supporting this hypothesis, recurrent deletions in *MXDs*, *MNT*, and *MGA* genes have been identified in some human tumors (Edelmann et al., 2017), and mouse studies suggest that at least MNT and MXI1 can behave as tumor suppressors (Dezfouli et al., 2006). By contrast, loss of MNT, such as dysregulated MYC, is pro-apoptotic, can exacerbate the apoptotic activity of MYC, and abrogate MYC-driven tumorigenesis (Link et al., 2012). A model that emerges from these studies is that a balance between the abundance and activity of MYC and MNT, and perhaps more generally between MYC, MXDs, MNT, and MGA, is needed to support oncogenesis (Diolaiti et al., 2015; Link and Hurlin, 2015). Moreover, the MAX-like protein MLX forms dimers with MLXIP and MLXIPL transcription factors (also known as MONDOA and CHREBP, respectively), which can either support or antagonize MYC function depending on cell context (Wilde and Ayer, 2015). Importantly, the nuclear localization and transcriptional activity of MLXIP and MLXIPL is highly dependent on nutrient flux, potentially connecting the functions of MYC and other PMN members to cellular metabolic state (Carroll et al., 2015; Diolaiti et al., 2015; Wilde and Ayer, 2015). Finally, different PMN members recruit other transcription factors, chromatin modifiers, and ubiquitin ligases that control their activity and abundance. Thus, changes in the copy-number, mutation, expression, and other alterations of PMN members and their interacting proteins may influence oncogenesis by increasing MYC expression, co-operating with or antagonizing MYC activity, and directly altering gene expression patterns independent of MYC, or a combination of these mechanisms (Diolaiti et al., 2015; Yang and Hurlin, 2017).

Here we performed a broad and unified analysis of genomic and expression data of the The Cancer Genome Atlas (TCGA) dataset with <9,000 samples covering 33 tumor types. We analyzed the frequency and extent of copy-number changes and mutations of *MYC* paralogs at the pan-cancer level. This was integrated with existing knowledge about MYC and the PMN to better understand the different roles that alterations of MYC and the PMN play on a pan-cancer level and in individual tumor types.

RESULTS

Pan-cancer Analysis of Copy-Number Alterations

MYC oncoproteins in solid tumors are mainly activated by copy gains, and it is well established that even small changes in MYC levels can drive ectopic proliferation of somatic cells and oncogenesis (Bazarov et al., 2001; Hofmann et al., 2015; Murphy et al., 2008). Therefore, we performed an in-depth analysis of copy-number alterations using purity- and ploidy-corrected focal copy-number data to provide the sensitivity necessary to detect low level copy-number gains expected to have a biological function. Focal copy-number events are defined by affecting less than 50% of the chromosome arms. At the pan-cancer level, *MYC* (*c-MYC*) is the most frequently amplified gene among the proximal network members across all cancer types, occurring

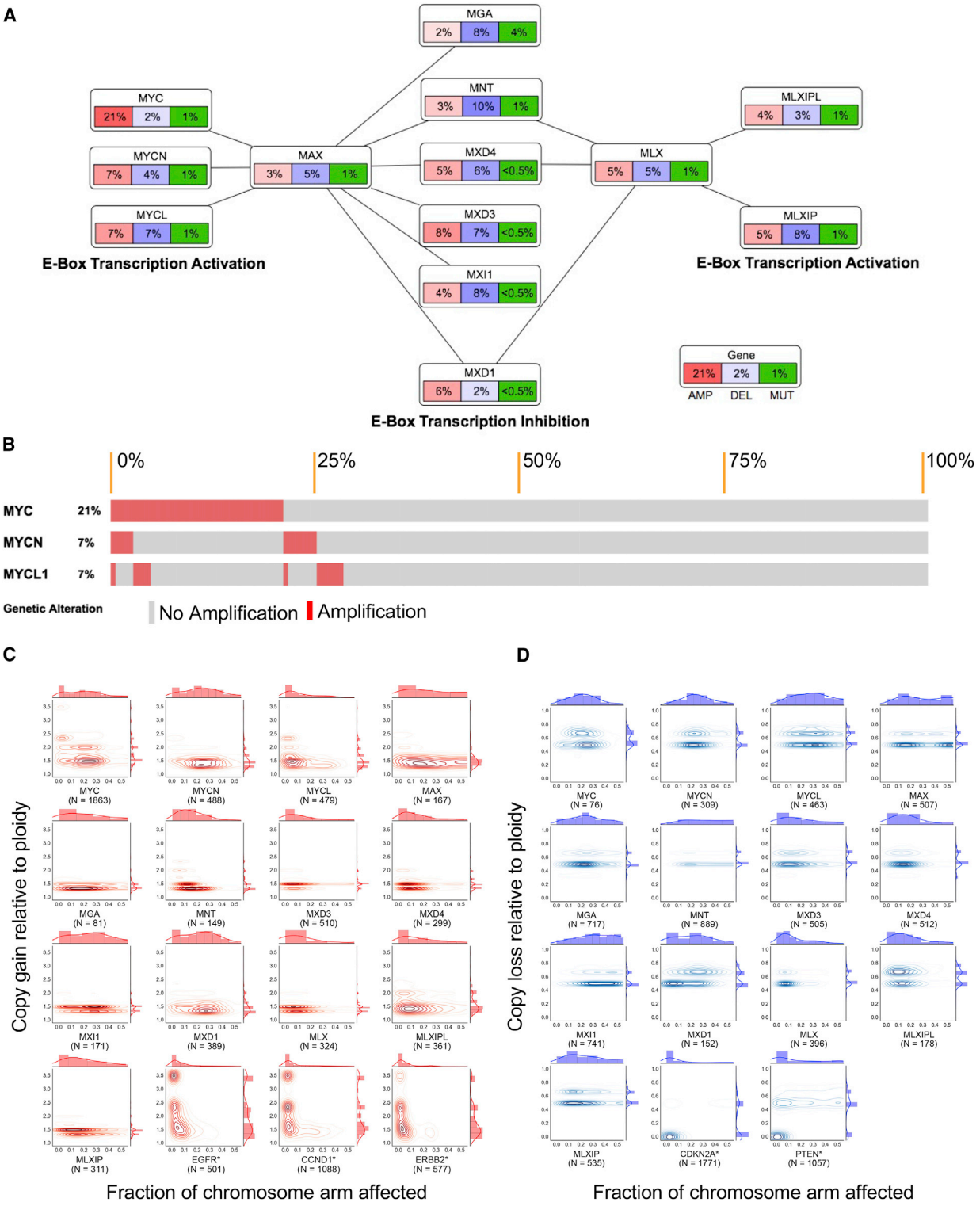


Figure 1. Proximal MYC Network

(A) Percentage of samples across 33 tumor types with focal range copy-number amplifications (leftmost box, red), focal range deletions (middle box, blue), and coding mutations (rightmost box, green) per gene of the MYC network.

(B–D) Oncoprint for focal amplifications of *MYC*, *MYCN*, and *MYCL* (B). Focal amplifications (C, red) and focal deletions (D, blue) across genes in the proximal MYC network visualized by the distribution of alteration size and amplitude. Oncogenes and tumor suppressors outside of the MYC network (denoted by *) were included for reference.

in 21% of samples (Figure 1A). *MYCN* and *MYCL* exhibit 7% focal amplification at a pan-cancer level. Overall, 28% of the analyzed samples in TCGA have at least one of the three MYC family members focally amplified (Figure 1B). The most frequent focal deletions in the PMN, including shallow deletions (only one copy of the gene lost), were seen in the transcriptional repressor and MYC antagonist, *MNT* (10%), aligning with its proposed role as a tumor suppressor in the MYC network (Yang and Hurlin, 2017) (Figure 1A). Deletions of the closely related *MXD3* (7%), and *MXD4* (6%) genes, as well as of *MGA* (8%), were also observed at the pan-cancer level. Shallow deletions of *MAX* were found in 5% of all TCGA samples. Even though the association of MAX with MYC is critical for its transforming potential in cell culture assays, recurrent shallow deletions of *MAX* were observed in some cancers, particularly pheochromocytomas and gastrointestinal stromal tumors (Burnichon et al., 2012; Comino-Méndez et al., 2011; Pantaleo et al., 2017; Schaefer et al., 2017). These findings raise the possibility that the loss of MAX dimerization with MXD, MNT, and MGA, all transcriptional MYC antagonists, favors tumorigenesis for some cell types. Also, for *MLX*, shallow deletions were observed in 5% of tumors (Figure 1A). MLX is an obligatory dimerization partner with the MONDO family (MLXIP or MLXIPL), but can bind MNT and a subset of MXD proteins (Figure 1A). Similar to MAX, shallow deletion of MLX could interfere with the repressor activity of MXD and MNT, as well as alter MYC-induced metabolic reprogramming by disabling of MLXIP and MLXIPL (O'Shea and Ayer, 2013).

The same analysis was performed considering broad deletions and amplifications using a cutoff of ±0.5 ploidy for the relative copy-number values (pan-cancer distribution shown in Figure S1B). This cutoff enabled the detection of copy-number gain or loss of a single MYC allele present in 100% of cells in the sample. Unlike focal events, the broad copy-number alterations can involve whole chromosome arms. In this case, *MYC* amplification frequency increases to 30% and the frequency of *MNT* deletions doubles (21%) (Figure S1C). The latter may be due to *MNT* and *TP53* being on the same distal arm of chromosome 17. For the rest of the PMN, both broad and focal copy-number events occur at similar frequencies.

Distinct Subgroups Based on Chromosome Fraction and Copy Number

To learn more about the characteristics of focal copy-number alterations, we plotted the amplitude of copy gain or copy loss against the fraction of the chromosome affected (Figures 1C and 1D). Only samples with focal GISTIC +1 and −1 were considered. Even though focal deletions were defined as less than 50% of the chromosome arm, this graphical representation highlights that the majority of events affect about 20% of the chromosome arm. Furthermore, distinct groups emerge based on amplitude and size. *MYC* amplifications (Figure 1B), for instance, shows three distinct groups: low copy increase affecting everything from small amplicons up to large (40% of the chromosome arm), an intermediate copy gain group with mostly large events, and a small subset with high copy gains affecting a small proportion of the chromosome. These patterns are similar to clinically significant drivers such as *EGFR*, *CCND1*, and *ERBB2*, with the only difference of *MYC* having also larger amplifications. *MYCN* only shows one group with low copy gain affecting a large proportion of the chromosome arm (Figure 1B).

Deletions in the PMN are mostly shallow, only affecting one copy of the gene, in contrast to known tumor suppressors such as *CDKN2A* and *PTEN* (Figure 1D). *MNT*, the most frequently deleted member of the PMN, shows a wide distribution in the size of the deletions.

Pan-cancer Frequency of Mutations in PMN

Mutations in *MYC* that may affect MYC protein stability and/or activity have been described to occur in Burkitt's lymphoma and diffuse large B cell lymphoma (Adhikary et al., 2005; Salghetti et al., 1999; Zhang et al., 2013). In this analysis, recurrent mutations altering two particular amino acids (P74 and S161) were observed at a frequency of ~0.07% (Figures 1A and S1D). P74 is within the MYC box 1 phosphodegron (sequence conserved for all FBXW7 degrons) and is strongly predicted to inactivate the degron and lead to increased stability of MYC (O'Neil et al., 2007). S161 is proximal to the 3′ end of MYC box II, and since MYC box II is the major binding site for MYC co-activator complexes, it is possible that phosphorylation of S161 may influence transcriptional activity. The serine in this position is not conserved, while the core of MYC box II is highly conserved in MYC paralogs.

In addition, MGA was mutated at 4% across all cancer types, with 30% of the 523 mutations identified in MGA predicted to truncate the protein, thereby eliminating the bHLHZ domain and interaction with MAX (Washkowitz et al., 2015) (Figure S1E). Such loss of function mutations in MGA, an essential gene encoding a >3,000 amino acid protein that contains both T-box and bHLHZ DNA binding domains, were recently reported to occur in chronic lymphocytic leukemia (De Paoli et al., 2013) and lung adenocarcinomas (Cancer Genome Atlas Research Network, 2014a).

Genetic Alterations among Individual Cancer Types

In addition to comprehensive characterization of MYC and PMN alterations across all TCGA samples, genetic alterations in individual cancer types were evaluated (for cancer-type-specific abbreviations, see Table S1). The data are presented as hierarchically clustered heatmaps highlighting tumor types with similar patterns. Focal amplifications involving *MYC* (Figure 2A) occur most frequently in OV (64.8%), followed by ESCA (45.3%) and LUSC (37.2%). These tumor types group together with UCS and BLCA for a clear subgroup highlighted by frequent amplifications of all MYC paralogs. A second group with high *MYC* amplification, but lower percentages for *MYCN* and *MYCL*, includes STAD, LUAD, BRCA, and LIHC. The exception among the 33 cancer types are THYM, THCA, KICH, LAML, KIRP, and PCPG, which have infrequent amplifications of MYC paralogs and the PMN (between 5 and 0.4%). The same group of tumor types has also very few deletions among members of the PMN. In general, 23/33 cancer types have at least 10% of samples with *MYC* focally amplified, whereas *MYCL* and *MYCN* are amplified less frequently than *MYC*. Data for broad amplification are shown in Figure S3A, indicating that in UVM and KIRP the majority of amplifications are not focal.

Among the PMN, *MNT* was the most focally deleted gene occurring in more than 20% of LIHC, LUAD, SARC, and UCS

samples. However, other PMN members involved in E-box transcription repression, such as *MGA* and *MXI1*, are also frequently deleted.

Because multiple members of the PMN act co-operatively or antagonistically, and show cross-regulation, the network can be considered as a single transcriptional module, we combined focal deletions in suppressors (e.g., *MNT*, *MGA*), focal amplification in drivers (e.g., *MYC*) and mutations (e.g., *MGA*) of all PMN members. Based on this analysis, almost every cancer type has at least one member of the PMN affected in at least 10% of the samples (Figure 2D), and 24/33 tumor types exhibit alteration in at least 50% of the samples. The percentage of alterations varies widely among tumor types, with OV cancer showing almost 100% samples with alterations in the MYC network, whereas THCA exhibits less than 5% of samples with MYC/PMN alterations (Figure 2D).

Tumor-Type-Specific Amplification Size of MYC

The size of amplifications involving *MYC* varies among tumor types. We defined an arbitrary threshold of 0.1 to separate the amplifications into two groups based on the chromosome fraction affected. On a pan-cancer level, 75% of samples with *MYC* focal amplification have events affecting more than a 0.1 fraction of the chromosome arm, and for 25% the fraction of the chromosome is less than 0.1, similar to the range observed for other drivers such as *EGFR* and *ERBB2* (Figures 1B and 2E). Individual tumor types, such as UCEC, STAD, ESCA, UCS, and SARC, have significantly more amplification events, which affect less than a 0.1 fraction of the chromosome arm. For LUSC, HNSC, LIHC, LGG, UVM, and SKCM on the other side, larger amplicons are more frequent (Figures 2F and S3C). Among tumor subtypes, a significant difference was detected only in ESCA (Figure S3D), where small amplification of *MYC* are more frequent in the chromosomal instability subtype.

Pan-cancer Analysis Reveals Mutual-Exclusivity between MYC Alterations and Common Oncogenic Drivers

The large number of samples in the TCGA (9,125 samples), and the fact that *MYC* is frequently altered across many tumor types, prompted us to look for genes with alterations that are either mutually exclusive or co-occurring with *MYC*. However, the analysis for co-occurrence with all genetic alterations (copy number and mutations) typically returned genes located within the same chromosome arm as *MYC* (Figure S3A). We therefore focused on mutual exclusivity allowing for the discovery of alterations (copy number and mutation), which occur more frequently than expected by chance without *MYC*. We used the DISCOVER method (Canisius et al., 2016) to calculate the significance that alterations in a given gene are less likely to co-occur with *MYC*. With a false discovery rate of 1%, 370 genes were found to be mutually exclusive with *MYC*. Most strikingly, the top four most significant genes were all known oncogenic drivers: *PTEN*, *BRAF*, *APC*, and *PIK3CA* (Figures 3A and 3B; Table S2). Additional known oncogenes identified had higher but still significant q values, such as *KRAS*, *NRAS*, *IDH1*, and *MTOR*. On a pan-cancer level we did not observe any mutual exclusivity between *MYC* and other members of the PMN.

Oncogenic drivers could simply be inducing *MYC* expression and might therefore appear to be mutually exclusive. We therefore compared *MYC* expression between gene of interest altered, *MYC* altered, both altered, or none of them (Figure 3C). In the case of *PTEN* and *BRAF*, *MYC* expression levels were unchanged in the presence of alterations in these genes, suggesting that *MYC* was not induced in these tumors. In the case of *APC* and *PIK3CA* alterations, *MYC* expression was elevated suggesting that these alterations are sufficient to activate *MYC* expression, without the need of gene amplification, consistent with previous reports (Ilic et al., 2011; Muncan et al., 2006).

The same analysis was performed for individual tumor types (Table S2; Figures 4A–4C). Only 3 out of the 33 tumor types had significant results: BRCA (33.4% *MYC* alterations), UCEC (21.5% *MYC* alterations), and LGG (12.3% *MYC* alterations). In BRCA, *PIK3CA* alterations are mutually exclusive with *MYC* alterations (Figure 4A; Table S10). UCEC had ten significantly mutually exclusive genes, of which eight are shown. The list includes *PTEN* and *KRAS* as common oncogenic drivers (Figure 4B; Table S10). In LGG cancers, mutations of the CIC and *FUBP1* genes were mutually exclusive with *MYC*. FUBP1 is known to bind to regulatory sequences of *MYC* and in the absence of this transcription factor, endogenous expression of MYC is blocked (He et al., 2000), indicating the functional dependence of MYC from FUBP1.

Elevated MYC Expression across Multiple Cancer Types

As described in the Introduction, genetic alterations alone are not sufficient to characterize all MYC-driven tumors, and we therefore proceeded to analyze *MYC* expression levels across the 33 cancer types. Overall, *MYC* expression was significantly (Hedges' g effect size = 0.71) increased in the samples with somatic *MYC* alterations, but not in samples where alteration only occurred in a PMN member (Hedges' g effect size = 0.17; Figure 5A). Similar observations were made for *MYCN* and *MYCL* (Figures 5B and 5C). The lack of effect of PMN alterations on *MYC* expression is consistent with reports that PMN members, such as MNT and MGA (Hurlin et al., 1997, 1999) are antagonistic of MYC transcriptional targets but not of the *MYC* gene itself.

MYC levels can also be altered through upstream signal transduction pathways, epigenetic changes, and regulation of mRNA and protein stability. *MYC* mRNA levels were elevated in most tumor types, with highest levels detected in COAD, HNSC, ESCA, READ, and UVM, followed by OV, LUSC, MESO, SKCM, and STAD (Figure 5D). Interestingly, cancers with infrequent amplification of *MYC* also had the lowest expression levels: KICH, PCPG, and THCA. In summary, cancers with the highest frequency of *MYC* copy-number gains showed elevated average expression at the mRNA level.

In addition, because *MYC* has been well established to be post-transcriptionally regulated by multiple ubiquitin/proteasomal degradation pathways, we analyzed protein expression data which were quantified using the reverse phase protein array (RPPA) platform. These data are available for MYC, but not for MYCN or MYCL. MYC protein levels are the highest in OV and KICH, which is in concordance with the high frequency of copy-number increases of *MYC* in OV cancer (Figure 5E). *MYC*-low mRNA expression and infrequent gene amplification in KICH suggests that MYC is stabilized either by a post-translational

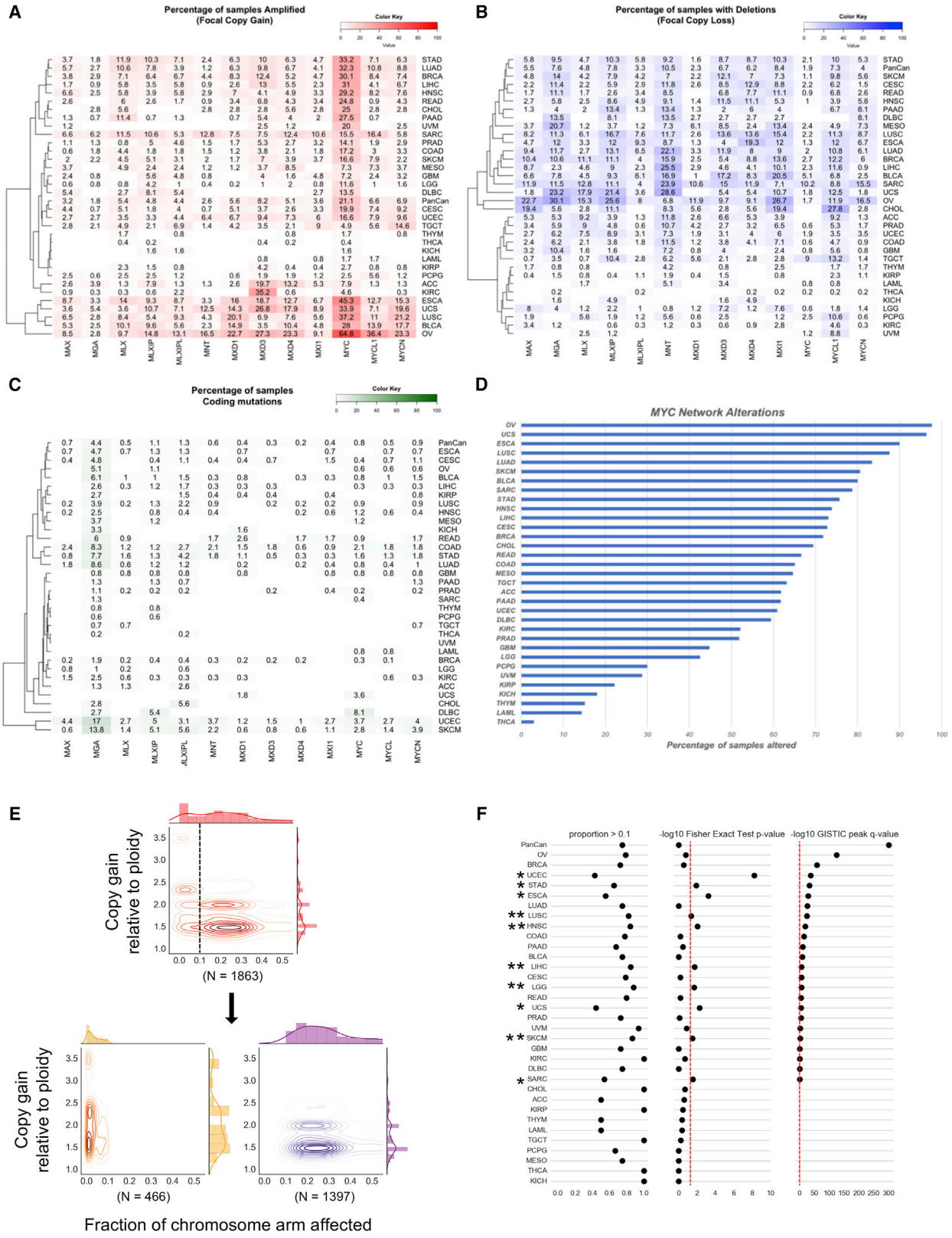


Figure 2. Tumor-Type-Specific Alterations of PMN Members

(A) Percentage of samples with focal amplifications per gene per tumor type.

(B) Percentage of samples with focal deletions respectively per gene per tumor type.

(legend continued on next page)

mechanism, or translation itself could be regulated in this cancer. Colorectal cancer was the third highest cancer with MYC protein level, possibly explained by the activation of the Wnt/β-catenin signaling pathway occurring in the majority of colorectal cancer samples (Ciriello et al., 2013). Brain tumors, LGG, and GBM, had the lowest levels of MYC protein. It has been shown in pediatric GBM that MYCN is mostly elevated (Huang and Weiss, 2013), but our data show elevated mRNA levels for both MYCN and MYC. LGG, on the other hand, have higher *MYCN* mRNA levels compared with *MYC*. Lower MYC protein was also detected using RPPA data in DLBC, with only a small subset (5%–10%) exhibiting high MYC protein expression. Generally, *MYC* mRNA correlates well with MYC protein levels in most tumor types, with UVM, CHOL, ACC, and KICH being notable exceptions (Figure 5F).

In contrast to *MYC*, *MYCN* mRNA expression was overall lower at the pan-cancer level, with the most prominent increase detected in a small subset of cancer types, namely: LGG, OV, PCPG, TGCT, and GBM. This highlights a more disease/tissue-specific expression pattern for *MYCN* compared with *MYC* (Figure 5D). However, *MYCN* gene amplification status across a variety of cancer types (Figure 1B) was not limited to these tissues, indicating that MYCN alterations are widespread across different tissues.

Unique Gene Set Enrichment Patterns for MYC

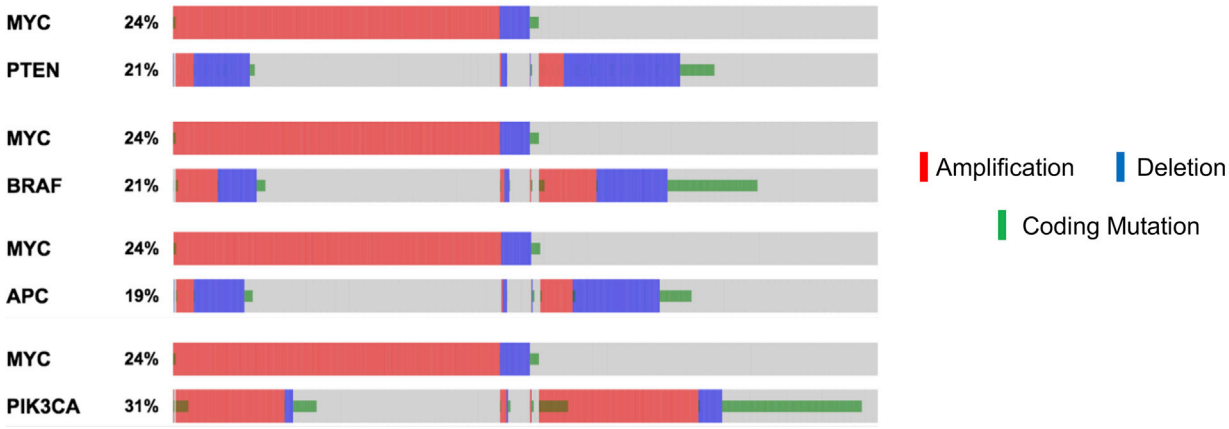
To find gene sets which are associated with activation of the MYC pathway, *MYC* and *MYCN* expression were utilized to identify sets of genes with both significant positive and negative Spearman rank correlation coefficients in a pan-cancer level. The coefficients were then used to generate a weighted genome-wide gene list to be used for gene set enrichment analysis. For this analysis, gene ontology molecular function category comprising 901 gene sets was corrected for redundancy and reduced to 396 gene sets (using the reduce_overlap function of the R GOplot package, see Tables S4 and S5). From the positively correlated gene sets a commonly shared pan-cancer enriched gene set emerged, with the exception of THCA, PCPG, and TGCT. Heatmaps derived from the enrichment scores of the top 100 pathways are shown for *MYC* and *MYCN* (Figures 6A and 6B; Table S6). Hierarchical clustering of cancer types and positively correlated pathways defines three groups, mainly distinguished by both the strength of correlation with *MYC* expression as well as the specific enrichment of certain gene sets (designated groups I, II, and III in Figure 6A). Groups I and II, together comprising 30 cancer types, showed the highest similarity, with enrichment of pathways that is in line with previous knowledge and are therefore referred as ''canonical'' MYC gene sets, such as transcription and RNA processing, chromatin remodeling (Lüscher and Vervoorts, 2012; Nakagawara et al., 1987), and translational processes (Cowling and Cole, 2007), including gene sets corresponding to ribosomal structural components and rRNA synthetic apparatus (Grandori et al., 2005; Grewal et al., 2005), as well as DNA replication and repair (Dominguez-Sola and Gautier, 2014; Rohban and Campaner, 2015). The canonical signature was the predominant feature of the 14 tumor types of group I, exhibiting *MYC* amplification in >20% of samples, indicating that these expression changes are linked to de-regulated *MYC* expression from gene copy-number alterations (Figure 6A).

Interestingly, this broad and unbiased pan-cancer analysis also revealed a new aspect of MYC-associated broad transcriptional changes in groups II and III. These groups are characterized by gene sets enriched for cytokines, immune response, and extracellular matrix components. In addition, groups II and III shared enrichment for growth factor signaling pathways, consistent with the notion that MYC can be induced in response to several extracellular stimuli, thus linking activation of the MYC pathway to microenvironment cues in a subset of cancers. For simplicity, we will refer to the chemokine/immune-response and signaling signature to as ''non-canonical'' MYC signatures (see Figure 6A, highlighted in orange). Among the signaling/growth factor pathways were transforming growth factor β, epidermal growth factor receptor (EGFR), insulin receptors, hormone receptors, and G-protein-coupled receptors, indicating that a diversity of signaling pathways may be able to activate MYC or vice versa, and that MYC could affect their expression. WNT signaling was found to be enriched across all three groups, with the highest enrichment score in cancers of group III. The notion that elevated MYC expression in certain cancers occurs irrespective of gene amplification events is exemplified by the activity of WNT in COAD and READ. These cancer types exhibit the highest levels of *MYC* mRNA expression, yet not the highest percentage of *MYC* gene amplification; also KIRP and KIRC both exhibit elevated *MYC* mRNA expression (see Figure 5D), despite low percentage of *MYC* amplification (2%–3%). Altogether, these results underline the importance of signaling pathways contributing to *MYC* elevated expression in many cancers. Strikingly, group III, which consists of only three cancer types, TCGT, THCA, and PCPG, exclusively exhibited high correlation with the ''non-canonical MYC signature,'' comprising cytokine, immune response, and signaling pathways, but lacking hallmarks of the canonical MYC signature. Interestingly, these three cancers exhibit low *MYC* expression (Figure 5D) and low percentage of *MYC* copy-number alterations (see Figure 2B).

Overall, the definition of MYC-associated gene expression at the pan-cancer level confirmed known hallmarks of MYC pathway activity across 30 of the 33 cancer types, while also

(C) Percentage of samples with protein coding mutations per gene per tumor type.
(D) Percentage of samples showing any alterations in at least one of the PMN members per tumor type.
(E) Focal amplification of *MYC* visualized by the distribution of alteration size and amplitude. An arbitrary threshold of 0.1 was used to define two groups with either amplification larger than a 0.1 fraction of the chromosome arm, or less than 0.1.
(F) Diagram of various metrics calculated for focal copy-number amplifications targeting *MYC*. Proportion > 0.1 (left) demonstrates the amount of samples for a given tumor type with focal amplifications that span greater than a 0.1 fraction of the chromosome arm. Fisher's exact test p value metric (middle) resulted from Fisher's exact tests comparing the fractions of samples on either side of the 0.1 cutoff for each tumor type with the rest of the tumor types, with the red line representing the equivalent of p = 0.05. The GISTIC peak q value (right) is only available for tumor types in which GISTIC identified significant regions of focal copy-number amplification affecting *MYC*, with the red line representing the equivalent of q = 1.00. Tumor types with significantly smaller amplification are labeled with one asterisk (*), whereas tumor types which have significantly larger amplifications are labeled with two asterisks (**).

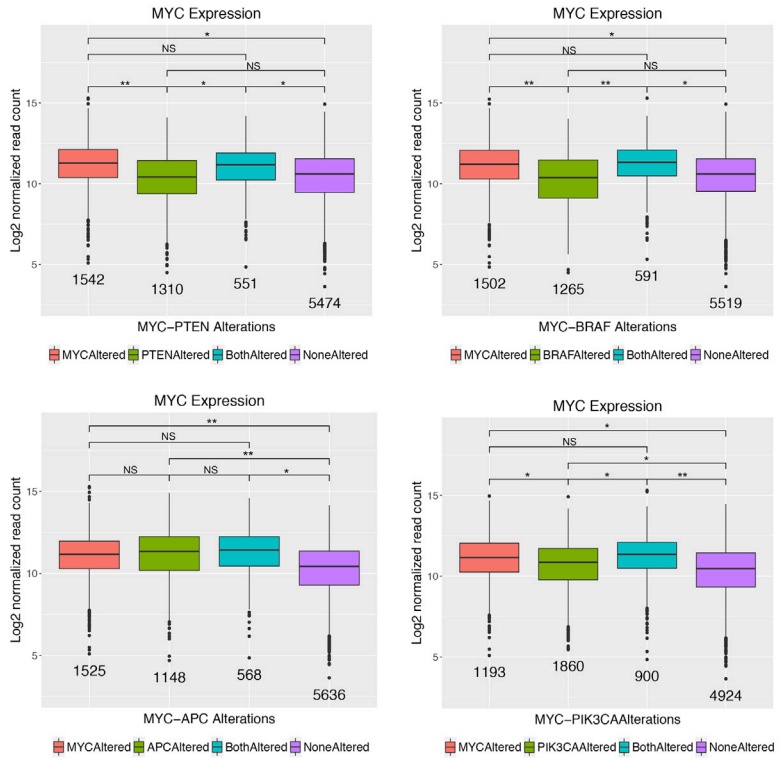
A



B

Gene	MYC wt Gene wt	MYC wt Gene alt	MYC alt Gene wt	MYC alt Gene alt	p-Value	q-Value
PTEN	5478	1313	1542	551	1.55E-20	2.87E-16
BRAF	5524	1267	1502	591	1.13E-13	1.11E-09
APC	5642	1149	1525	568	1.74E-10	1.18E-06
PIK3CA	4928	1863	1193	900	2.85E-10	1.45E-06

C



(legend on next page)

revealing novel associations for example with immune response and cytokine signaling for future studies.

Unique Gene Set Enrichment Patterns for MYCN

While *MYCN* was less prominently amplified at the pan-cancer level than *MYC* (7% versus 21%), several cancer types such as UCS, TGCT, OV, LUSC, ESCA, and BLCA had elevated copy-number changes (>10%), which was also reflected in higher *MYCN* expression levels. To explore potential differences and/or similarities with respect to global gene expression signatures between these two MYC paralogs, a genome-wide correlation with *MYCN* across the 33 tumor types was also carried out. This analysis revealed common features with MYC-associated pathways as well as MYCN-specific gene sets (Figure 6B). Among the MYCN-associated signature, across all cancers with the exception of LAML, LGG, and TGCT, was the enrichment for cell-signaling and developmental pathways including WNT, NOTCH, and Ephrin receptors. Interestingly, gene sets enriched for WNT signaling appeared more significantly in MYCN than MYC, and altogether developmental pathways were highly represented in the *MYCN*-associated gene expression signature. Similar to MYC, the MYCN pan-cancer signature was enriched for epigenetic pathways linked to histone acetylation and chromatin modifications. Metabolism linked genes were noted as MYCN-associated across all cancer types, but only detected on a cancer-type basis for MYC.

Unsupervised clustering of the positively correlated gene set distinguished three major groups (Figure 6B). Groups II and III were most similar with respect to cell-signaling, and also characterized by enrichment for neuronal function gene sets, including genes related to glutamate receptor function, ligand-gated ion channels, calcium ion transport, and acetylcholine binding. This neuronal-like signature (Figure 6B, light orange) is another distinguishing feature between MYCN and MYC paralogs. Group III and, to a lesser extent group II, also contained the ''non-canonical'' MYC signature, including cytokine, immune system, and extracellular matrix genes. However, the cancer types exhibiting the non-canonical signature did not overlap with MYC cancer types with such a signature, except for all the kidney cancers: KIRP, KIRC, and KICH (Figure 6B).

Finally, group 1, comprising LAML, LGG, TGCT, READ, PAAD, BLCA, THCA, PCPG and GBM, OV and COAD, correlated with MYC canonical signature with respect to DNA replication/repair and chromatin. Interestingly this group contained cancer types with the highest *MYCN* expression (see Figure 5D) such LGG, GBM and TGCT, and OV. This observation is consistent with the high threshold level of MYCN reached in these cancers, to potentially drive the canonical signature. In contrast, group I, altogether lacked the enrichment for cytokine/immune system as well as gene sets related to cell signaling (Figure 6B). In summary, while both MYC- and MYCN-enriched pathways commonly exhibited hallmarks of the canonical MYC signature, MYCN was unique in its association with genes related to neuronal function and developmental pathways.

A Pan-cancer MicroRNA Signature Associated with MYC

MicroRNAs (miRNAs) are key regulators of gene expression and important players in the pathogenesis of human cancers (Bartel, 2004; Ventura and Jacks, 2009). To gain insights into the interaction between miRNAs and MYC in human cancers, we calculated the Spearman rank correlation coefficients between 662 human miRNAs and *MYC* mRNA levels across each of the 33 cancer types. This analysis revealed a subset of miRNAs whose expression correlates with *MYC* levels across multiple cancer types. Eighty-two miRNAs showed an absolute Spearman rho correlation greater than 0.35 in at least three of the 33 cancer types examined. Based on their correlation to *MYC* levels, these miRNAs can be divided in three major groups (Figure 7).

The first include miRNAs whose expression positively correlates with *MYC* expression across the vast majority of studies. This group is largely composed by members of the miR-17–92 cluster and of its paralog, miR-107b–25. miR-17–92 is a polycistronic miRNA locus encoding six distinct miRNAs (miR-17, miR-18a, miR-19a, miR-19b-1, miR-20a, and miR-92a-1). Studies in mouse models of human cancers have demonstrated that this cluster, also known as Oncomir-1, is a bona fide oncogene and a direct MYC transcriptional target (He et al., 2005; O'Donnell et al., 2005). Furthermore, induction of miR-17–92 by MYC has been reported to be crucial for tumor cell survival and for tumor progression in multiple cancer types (Han et al., 2015; Li et al., 2014; Mu et al., 2009; Olive et al., 2009). miR-21, another bona fide human oncogene (Medina et al., 2010), and miR-27a, also are also generally positively correlating with *MYC* expression.

The second group consists of miRNAs, whose expression is consistently negatively correlated with *MYC*. Prominent examples are members of the miR-29, miR-30, miR-125a, and let-7 families of miRNAs. These miRNAs have been previously shown to be directly repressed by MYC (Chang et al., 2007), and are suspected to act as potential tumor suppressors. Other prominent miRNAs belonging to this second group are miR-200a and miR-200b. These two related miRNAs have been extensively studied and linked due their ability to modulate epithelial-mesenchymal transition (EMT) by forming a negative feedback loop with the master transcription factors ZEB1 and ZEB2 (Burk et al., 2008; Gregory et al., 2008; Korpal et al., 2008; Park

Figure 3. Pan-cancer Mutual Exclusivity with Focal Copy-Number Events and Mutations of *MYC*

(A) Oncoprint of MYC with the four most mutually exclusively altered genes *PTEN*, *BRAF*, *APC*, and *PIK3CA*.

(B) The table lists the top four genes most mutually exclusive with respect to *MYC*. Columns 2–5 show counts of samples with no alterations in *MYC* or the gene, with just the gene altered, with just *MYC* altered, and with both *MYC* and the gene altered. Columns 6 and 7 show p and q values as computed by the DISCOVER method.

(C) Boxplots compare *MYC* expression between groups of samples defined by pairwise alteration status of *PTEN*, *BRAF*, *APC*, and *PIK3CA*, respectively, with *MYC*. Hedges' g effect sizes are indicated for each pair of boxplots. NS indicates an effect size magnitude <0.2 (negligible), * indicates an effect size magnitude <0.5 (small) and ** indicates an effect size magnitude <0.8 (medium).

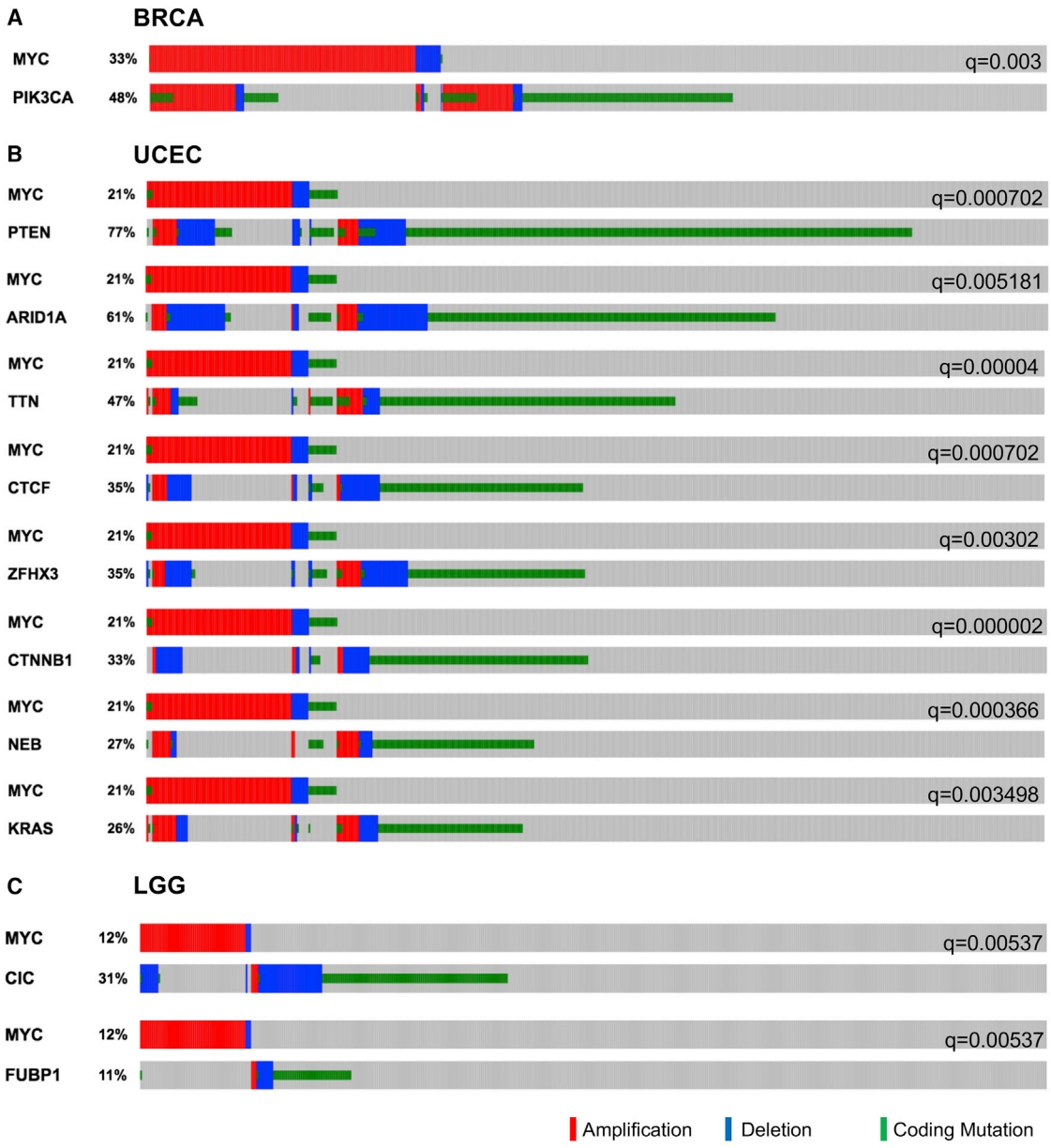


Figure 4. Cancer-Type-Specific Mutual Exclusivity with *MYC*

(A–C) Oncoprints for genes significantly mutually exclusive with *MYC* within BRCA, UCEC (top 8 genes are shown), and LGG, respectively.

et al., 2008). Consistent with this model, downregulation of these two miRNAs has been shown to be sufficient to promote EMT, while their overexpression can prevent EMT and inhibit cancer cell migration (Korpal et al., 2008).

Finally, a third and larger group of miRNAs was identified whose expression correlates with *MYC* in only a more restricted subset of cancer types, or that show positive correlation in some studies, but negative correlation in others. Several notable putative oncogenic and tumor-suppressive miRNAs belong in this group. For example, expression of miR-221 and miR-222, which have been proposed to promote cell proliferation and tumorigenesis, at least in part via p27-Kip repression (Kedde et al., 2010; Kim et al., 2009; Korpal et al., 2008; le Sage et al., 2007), positively correlated with *MYC* levels in PCPG, BLCA, CESC, and TGCT, but negatively correlated in THYM, LGG, and SKCM. A similar behavior is observed for miR-150, miR-155, miR-223, and miR146b.

Overall, these data show that there is clear correlation between *MYC* expression and several key oncogenic miRNA expression across multiple human cancers.

DISCUSSION

MYC was discovered nearly four decades ago, and since that time has been the subject of over 26,000 publications resulting in a substantial amount of information about the normal and

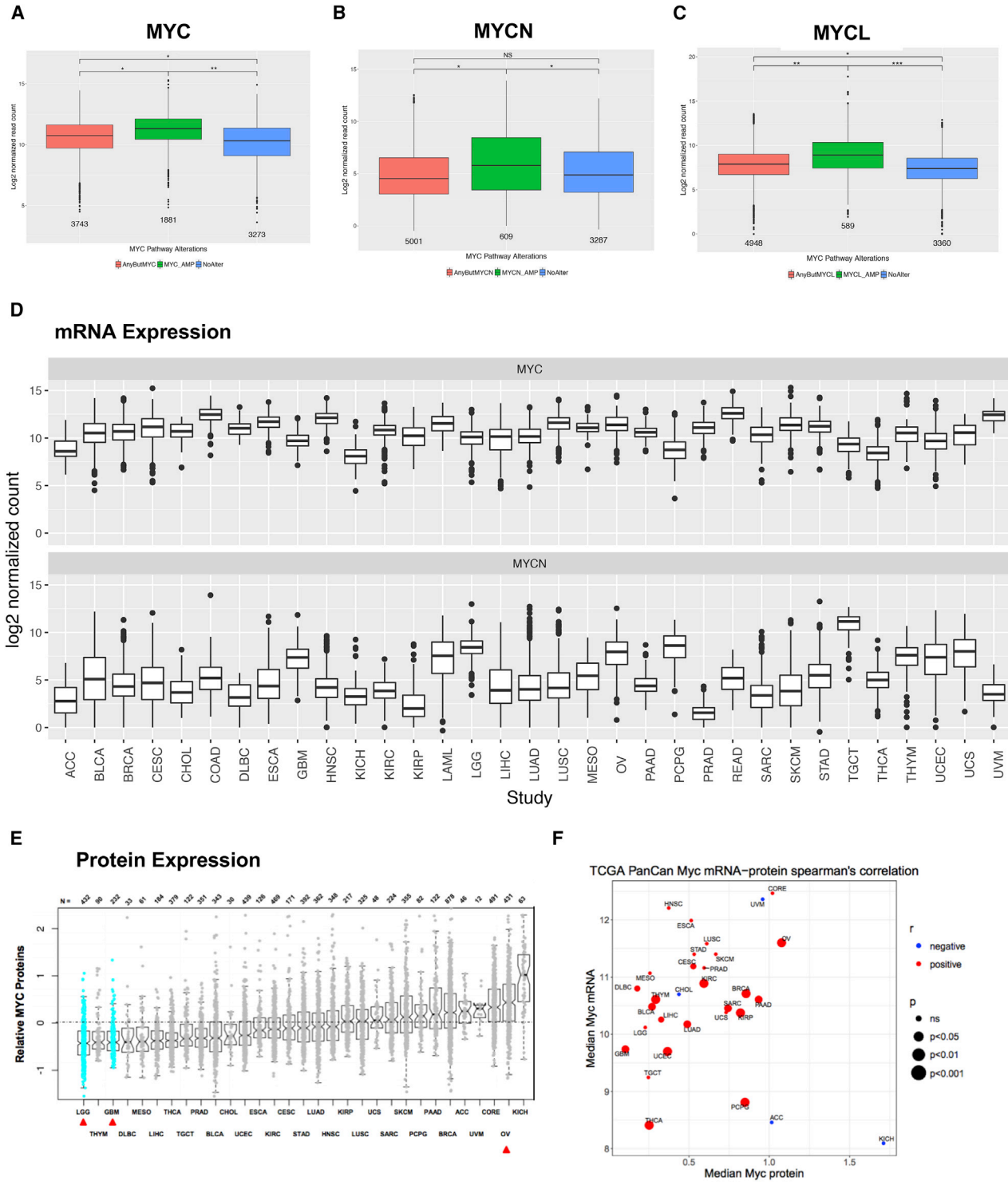


Figure 5. MYC Expression

(A–C) Comparison of *MYC*, *MYCN*, and *MYCL* expressions, respectively, between cohorts defined by focal amplification state of the genes themselves. For each of the three genes, from left to right, the boxplots represent: a cohort with any PMN gene other than *MYC*, *MYCN*, or *MYCL* altered, respectively, a cohort with *MYC*, *MYCN*, or *MYCL* amplified, respectively, and a cohort with no PMN genes altered. NS indicates an effect size magnitude <0.2 (negligible), * indicates an effect size magnitude <0.5 (small), ** indicates an effect size magnitude <0.8 (medium), and *** indicates an effect size magnitude ≥0.8 (large).

(D) Distribution of *MYC* and *MYCN* gene expression per tumor type.

(F) Distribution of MYC protein expression per tumor type.

(E) Correlation between median *MYC* mRNA expression and median MYC protein expression per tumor type.

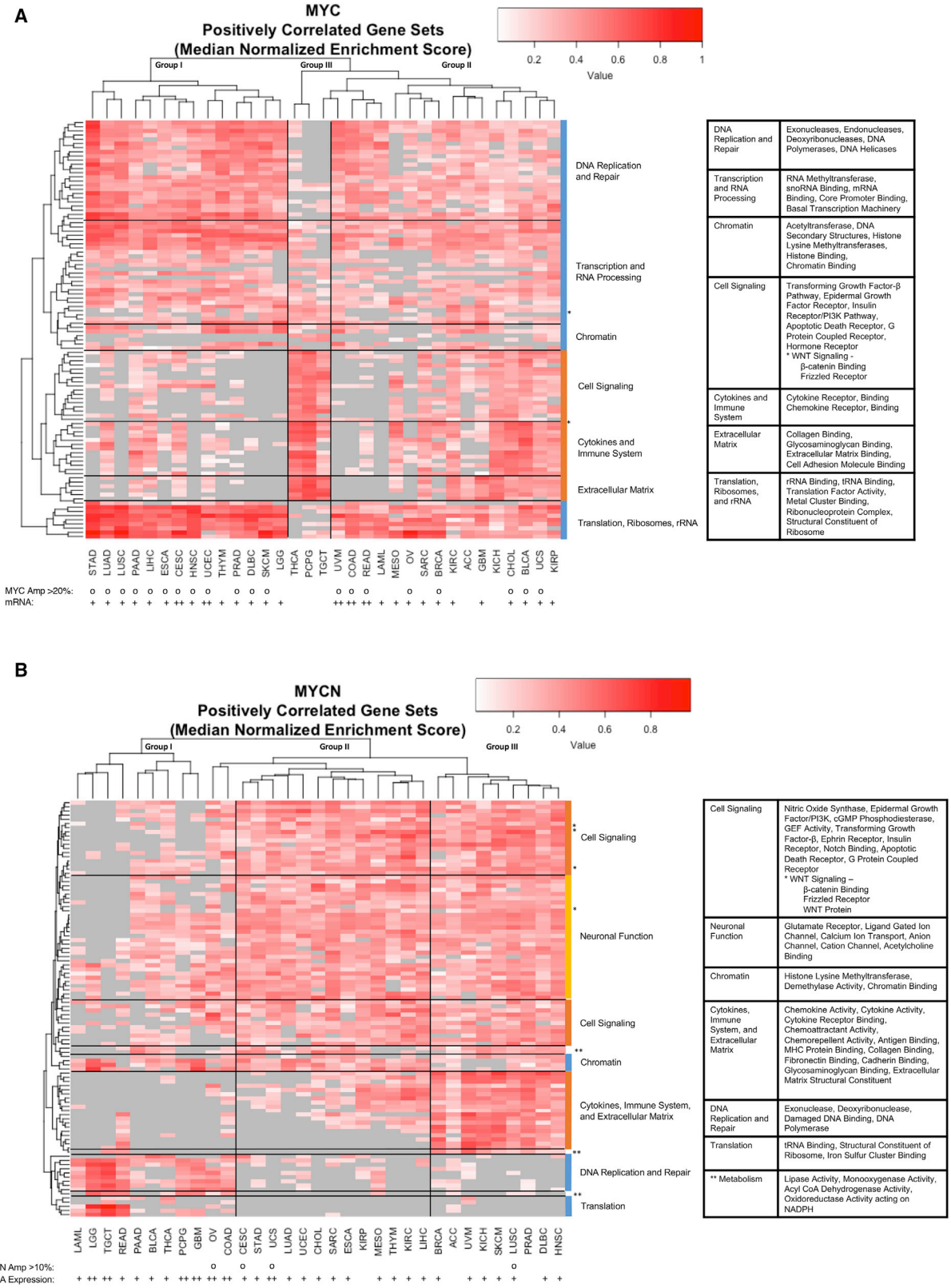


Figure 6. Pan-cancer Gene Set Enrichment Patterns

(A and B) Heatmap shows clustering of tumor types based on top 100 most positively correlated gene sets from gene ontology molecular function category for *MYC* (A) and *MYCN* (B). Each cell of the heatmap is colored by the normalized enrichment score of a gene set for a tumor type. Gray cells indicate lack of

(legend continued on next page)

oncogenic functions of MYC. The involvement of MYC as an oncogenic driver was described in a wide spectrum of tumor types as shown by early reports in leukemia (Dalla-Favera et al., 1983; Nowell et al., 1983), B cell lymphomas (Hayward et al., 1981), lung cancer (Little et al., 1983), and neuroblastoma (Schwab et al., 1983). Many of the analyses of MYC alterations in different tumor types have been scattered among many publications and meta-analyses. The TCGA permits for the first time a pan-cancer analysis of the MYC network using a uniform dataset and brings the field a step closer to classify tumors with and without MYC alterations. Gene amplifications of MYC paralogs occur in 28% of all cancer samples analyzed herein, suggesting that MYC paralogs have an important function in tumorigenesis with potential therapeutic implications for many of the 33 major cancer types described by TCGA.

Because MYC and its proximal network are altered in multiple ways, characterization by MYC copy number alone is insufficient to reveal the extent of MYC involvement in tumorigenesis. Therefore, the present analysis also includes genetic alterations, mRNA, and protein expression analysis of MYC paralogs (MYC, MYCN, and MYCL), and alterations of genes comprising the PMN, a group of MYC-related transcription factors that have been implicated in MYC pathway activity.

PMN Balance

MYC is the only member of the PMN which shows mainly amplifications, while *MAX*, *MNT*, and *MGA* have predominantly shallow deletions. All other members of the PMN display a mixture of amplifications and deletions raising the possibility that cells have to maintain a specific balance in the network. The partial loss of MAX, an essential dimerization partner of MYC, seems to be counterintuitive, but loss of MAX dimerization with MXD, MNT, and MGA, all transcriptional MYC antagonists, may play a critical role in tumorigenesis for some cell types by reducing negative control of the PMN (Diolaiti et al., 2015; Nilsson et al., 2004; Yang and Hurlin, 2017). In addition, MYC activity can also be enhanced by loss of MNT (Hurlin et al., 1997, 1999), which is focally deleted in 10% of all samples, or by inactivation of MGA, which occurs by focal deletions and by truncating mutations (with 9% and 4% frequency, respectively, Figures 1 and S2). Since mutual exclusivity between *MYC* amplification with deletions or inactivation of other members of the PMN was not observed, the oncogenic activity of MYC is most likely activated by a combination of alterations in different PMN members. In 80% of the samples with *MYC* amplification, *MYC* is altered in conjunction with other PMN members. This further highlights the possibility that activation or loss of any of the negative regulators may balance MYC activity in complex patterns, potentially as a result of stochastic events (Carroll et al., 2015; Link et al., 2012). Besides MYC, sole alterations in any of the remaining PMN members are rarely observed (Figure S4B). However, a clear quantification of the PMN balance is currently not possible and future studies are needed to provide evidences on the roles of the different alterations in the PMN for cancer development.

Low-Level Copy-Number Changes Are a Feature of MYC in Solid Tumors

Previous observations in hematopoietic cell lines and in neuroblastoma reported copy-number changes on the order of 10–100 for *MYC* or *MYCN*; however, the present analysis and work from others indicate that, while the frequency of *MYC* amplification across solid tumors is highly significant (TCGA paper in press), the fold copy-number changes are small (mainly between 0.5 and 2.5). For this reason, copy-number data were corrected for ploidy enabling a low threshold to identify single-copy gains of *MYC* and PMN members. Even using these inclusive criteria to call samples with copy-number alterations, we potentially underestimate the extent of their involvement, as shown in the violin plots representing copy-number distribution (Figure S1B). In consideration of the potential heterogeneity and infiltration of the tumor specimens by non-cancer stromal components, further analysis will be required to verify copy-number frequencies with lower threshold settings.

Analyzing both amplitude and size of copy-number alterations of *MYC* amplifications unveils three distinct groups. We see statistically significant differences in *MYC* amplification size between tumor types (Figure 2F), but overall each tumor type has samples fitting into each of the three groups characterized by either low- and intermediate-amplitude amplifications affecting a wide range of the chromosome arm, and high amplitude amplifications affecting only a small subset of the chromosome arm. For deletions, we rarely see loss of more than one copy, suggesting that most of the PMN might be haploinsufficient. This pattern has been observed in classical tumor suppressors as described for *TP53* and *P27KIP1* (Fero et al., 1998; Payne and Kemp, 2005), and is shown in this dataset and a recent publication (Vidotto et al., 2018) for *PTEN*, which exhibits a group with shallow deletions (one copy loss) and a wide range of the chromosome affected.

Mutually Exclusive Oncogenic Drivers and MYC Alterations

The frequent copy-number changes of *MYC* across many tumor types and the large body of experimental evidence defines MYC as a clear oncogenic driver. We used a genome-wide mutual exclusivity analysis with *MYC* to find genes that are less likely to co-occur with *MYC*. Pan-cancer mutual-exclusivity analysis resulted in a higher count of statistically significant results when compared with tumor-type-specific analysis. This was expected due to smaller sample sizes in individual tumor types resulting in loss of statistical power. Strikingly, the top mutually exclusive genes (*PTEN*, *BRAF*, *PIK3CA*, and *APC*) are all known oncogenic drivers linked to cell signaling (Figures 3 and 4; Table S10). This observation could be linked to unique and a possibly stochastic accumulation of genetic alterations in a given cancer, which may be sufficient to drive tumor formation and therefore appear mutually exclusive. *PTEN*, for example, is most frequently mutated in UCEC, where *MYC* is also altered at high frequency

enrichment. Dots below tumor type denote high MYC amplification, while plus signs denote high mRNA expression. Blue lines on the heatmaps mark gene sets corresponding to the canonical MYC signature, orange lines correspond to the non-canonical MYC signature, and yellow lines correspond to neuronal function, found in MYCN only. Tables contain main gene sets found in each cluster category. One asterisk marks a WNT signaling gene set, and two asterisks mark a metabolic gene set.

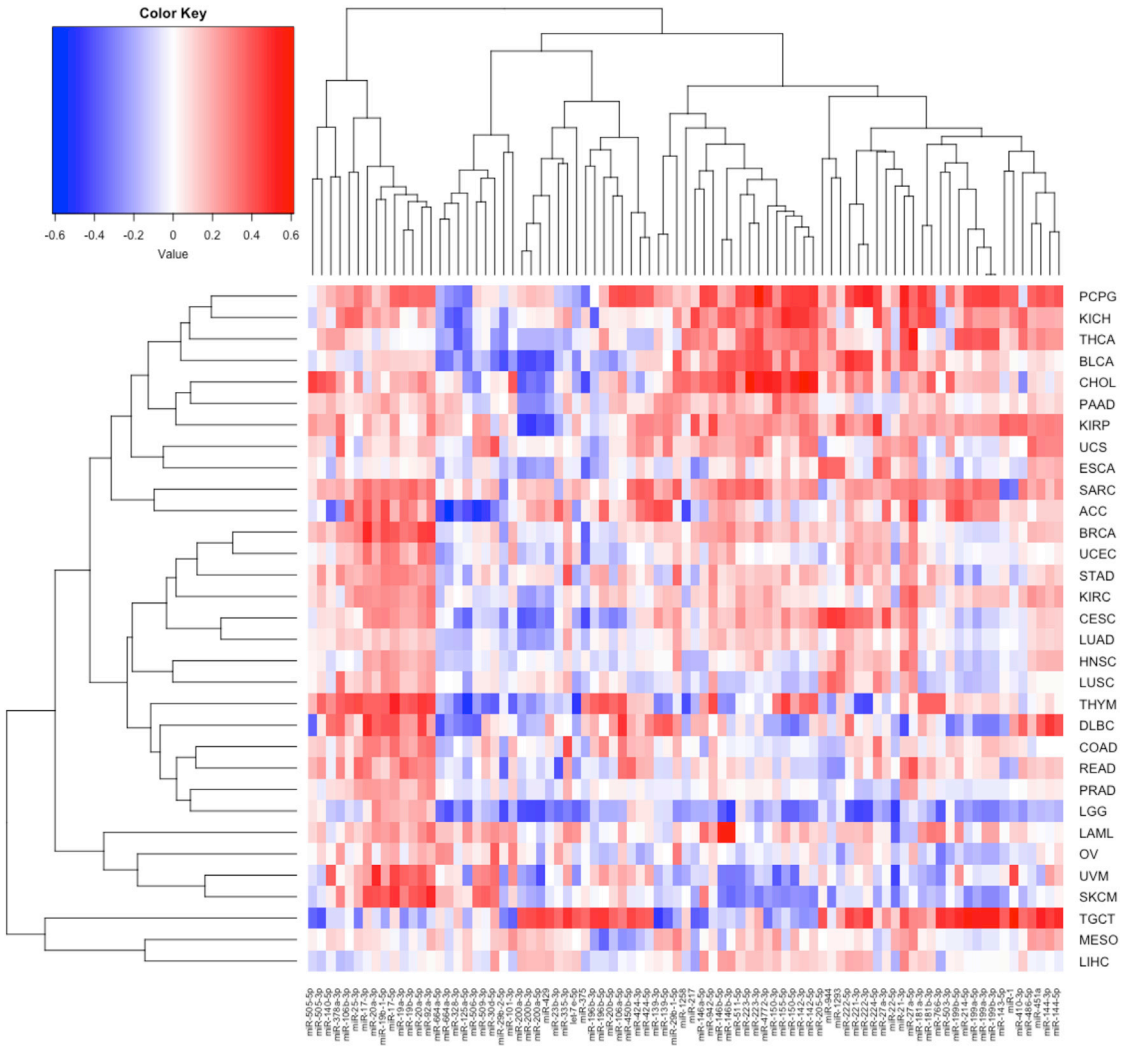


Figure 7. Heatmap of Correlation (Spearman Rho Value) between miRNA and *MYC* Expression across the 33 Cancer Studies

Only miRNAs that showed an absolute correlation value equal or greater than 0.35 in at least three studies were included (the complete list is provided in Table S9). Red indicates positive correlation, blue negative correlation.

(21%; Figure 4B; Table S10), suggesting that either pathway can be a driver of the same cancer type. Alternatively, mutual exclusivity between *MYC* and other recurrently altered genes may be a reflection that mutations in these oncogenes and tumor suppressors may not increase their fitness and potentially exhibit synthetic lethality.

Pan-cancer Canonical and Non-canonical MYC-Specific Gene Set Enrichment

Alteration of *MYC* expression leads to broad transcriptional changes through both direct and indirect effects, as observed in many experimental systems (as reviewed in Kress et al., 2015). Here, in order to gain an overview at the pan-cancer level of pathways correlating with *MYC* expression, we investigated both positively and negatively correlated genes. Gene set enrichment analysis was significant only among positively correlated genes, which revealed conserved pathways across 30 of the 33 tumor types. These pathways comprised DNA replication and repair processes, including genes such as DNA helicases, exonucleases, polymerases, and telomerase. Chromatin binding and remodeling processes were prominent, including acetylation and methylation, as well as multiple components of the basic transcription machinery. Altogether, we referred to these gene sets as canonical pathways, as previously observed in experimental systems upon manipulation of MYC levels (Dominguez-Sola et al., 2007; Gomez-Roman et al., 2006; Grandori et al., 2003, 2005; Grewal et al., 2005; Hnisz et al., 2013; Johnston et al., 1999; Moser et al., 2012; Robinson et al., 2009; for a review see Ruggero, 2009). Finally, since transcriptional and chromatin remodeling processes are at the core of MYC and PMN function, the enrichment in these pathways reflects biological function of the MYC network, consistent with the strongest correlation in group I cancers harboring frequent *MYC* alterations and with high expression (Figure 6A). Interestingly, MYC-synthetic lethal genes are enriched for similar pathways to the canonical MYC signature (Cermelli et al., 2014).

The three cancer outliers were THCA, PCPG, and TGCT, which also have the lowest level of *MYC* copy-number and expression changes. THCA, for instances, has a high frequency of BRAF mutation (Cancer Genome Atlas Research Network, 2014b), which is mutually exclusive to MYC based on our analysis. This is another indication, that these tumor types are largely not driven by genetic alterations of *MYC*.

In addition, a previously unrecognized non-canonical signature emerged, detectable in groups II and III, comprised of 11 tumor types (Figure 6A), indicating an association of *MYC* expression with extracellular signaling, the immune system (growth factors and cytokines), and extracellular matrix (Figure 6A). It is conceivable that upregulation of MYC in these cancers might occur in response to alterations of signaling pathways, as documented for WNT signaling in COAD. It is currently unclear how MYC expression could influence the immune response in a subset of cancers, including KIRP and KICH, UCS, BLCA, and CHOL. How MYC may influence the tumor immune microenvironment is of great interest, given the recent study showing that MYC is able to control PD-L1 and CD47 expression (Casey et al., 2016). Prominent signaling pathways and immune response were also observed in group III, with infrequent *MYC* copy-number changes (THCA, PCPG, and TGCT). Finally, each cancer type had a unique ranking of top associated pathways, indicating potential cancer-type-specific features (Table S7 for Molecular Processes and Table S8 for Biological Processes). Future in-depth analysis of these cancer-type-specific pathways has the potential to give insights about the biology and potential novel therapeutic vulnerabilities of these tumor types.

Pan-cancer MYCN Alterations and Associated Pathways

To date, MYCN has been largely considered a driver of selected pediatric cancers, such as neuroblastoma, where 25% of cases exhibit high levels of *MYCN* copy-number gains with clear prognostic significance (Nakagawara et al., 1987). The current analysis reveals that *MYCN* alterations, albeit less frequent than *MYC*, are widespread across several cancer types (Figure 2B), pointing to *MYCN* as a prominent oncogene in adult cancers. Indeed, similarly to neuroblastoma, *MYCN* amplification reaches 20% in TCGA cancer samples such as OV, UCS, and LUSC. Detection of an MYCN-associated gene expression signature allowed for clustering of the 33 cancer types into three groups, according to MYCN-specific enriched gene sets. Pathways unique to MYCN involved neuronal function and development spanning the majority of cancer types. These MYCN-specific hallmarks may enable derivation of biomarkers to identify MYCN-driven cancers that are caused by alterations other than copy-number changes (Figure 6B). Interestingly, among the gene expression signatures shared with MYC were aspects of both the canonical and non-canonical signatures (Figure 6B). Thus, at the pan-cancer level, MYC and MYCN appear distinct with respect to tumor types (with the exception of OV, where both are implicated at high frequency) and with respect to differences in associated pathways, which could provide surrogate biomarkers specific for MYC or MYCN. These results support and broaden, at a pan-cancer level, previous studies demonstrating a distinction between MYC and MYCN overexpression on determining specific cell lineages in genetic models of medulloblastoma (Roussel and Robinson, 2013; Vo et al., 2016).

Conclusions

The data and analyses presented herein are the first comprehensive assessment of MYC alterations across diverse tumor types. Over the years, a large body of evidence for the importance of MYC in tumorigenesis has been developed and a multitude of biological models created to uncover the complex mechanisms of MYC function. This study, by integrating this information with the comprehensive TCGA dataset, provides a resource for further understanding MYC-driven cancers, spurring new research areas and the basis for novel biomarkers and therapeutic approaches.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- METHODS DETAILS
  - Gene Expression
  - *Protein Expression by RPPA*: TCGA-RPPA-pancan-clean-v2.txt
  - miRNA Expression
  - Somatic Mutations
  - GSEA
  - DISCOVER
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Gene Alterations
  - Hedges' g Effect Size
  - Mutual Exclusivity with MYC
  - MYC-associated Gene Expression Signature and Gene Set Enrichment Analysis
- DATA AND SOFTWARE AVAILABILITY
  - R Code

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and ten tables and can be found with this article online at https://doi.org/10.1016/j.cels.2018.03.003.

ACKNOWLEDGMENTS

The MYC team would like to acknowledge Dr. Theo Knijnenburg for guidance and discussions regarding the statistical analysis and interpretation conducted as part of this research. We would also like to thank Dr. Bruno Amati, Dr. Christopher Kemp, and Dr. Goldie Lui for helpful input on the manuscript. This work was supported in part with grants from NIH (4U01CA176303-04 to C.G., U24CA143867 to A.C.B. and A.D.C., RO1CA57138 to R.N.E., and 5R01CA149707 to A.V.) and research funding from SEngine Precision Medicine and Cure First to C.G., F.X.S., V.D., and R.S. P.J.H. is supported by grants from Shriners Hospitals for Children and Leukemia and Lymphoma Society.

The TCGA Research network was funded by the following grants: U54 HG003273 to R.A. Gibbs, U54 HG003067 to S. Gabriel and E.S. Lander, U54 HG003079 to R.K. Wilson, U24 CA143799 to T.P. Speed and P.T. Spellman, U24 CA143835 to I. Shmulevich, U24 CA143840 to M. Ladanyi and C. Sander, U24 CA143843 to R.A. Gibbs and D.A. Wheeler, U24 CA143845 to L. Chin and G. Getz, U24 CA143848 to D.N. Hayes and C.M. Perou, U24 CA143858 to J. Stuart, C. Benz, and D.H. Haussler, U24 CA143866 to M.A. Marra, U24 CA143867 to S. Gabriel and M.L. Meyerson, U24 CA143882 to S.B. Baylin and P.W. Laird, U24 CA143883 to G.B. Mills, J.N. Weinstein, R. Akbani, and W.K.A. Yung, U24 CA144025 to R.S. Kucherlapati, and P30 CA016672 to G.B. Mills.

AUTHOR CONTRIBUTIONS

F.X.S. contributed to conceptualization, investigation, supervision, data visualization, and wrote the manuscript. V.D., A.C.B., and W.Z. developed computational methodologies, while M.T., X.Z., and R.S. carried out initial bioinformatics methodologies. A.V. was responsible for methodologies and writing on miRNAs. D.E.A., P.J.H., and R.N.E were responsible for MYC field expert review, data interpretation, and writing of the original manuscript. A.B.R. contributed to data curation. A.D.C. and Y.L. contributed to supervision of computational methodologies. B.B. contributed to co-ordination and supervision of computational methodologies, reviewed the manuscript, and communicated with TCGA. C.G. contributed to the conceptualization, supervision, data interpretation, writing of the manuscript, and provided funding.

DECLARATION OF INTERESTS

Michael Seiler, Peter G. Smith, Ping Zhu, Silvia Buonamici, and Lihua Yu are employees of H3 Biomedicine. Parts of this work are the subject of a patent application: WO2017040526 titled "Splice variants associated with neomorphic sf3b1 mutants." Shouyoung Peng, Anant A. Agrawal, James Palacino, and Teng Teng are employees of H3 Biomedicine. Andrew D. Cherniack, Ashton C. Berger, and Galen F. Gao receive research support from Bayer Pharmaceuticals. Gordon B. Mills serves on the External Scientific Review Board of Astrazeneca. Anil Sood is on the Scientific Advisory Board for Kiyatec and is a shareholder in BioPath. Jonathan S. Serody receives funding from Merck. Kyle R. Covington is an employee of Castle Biosciences. Preethi H. Gunaratne is founder, CSO, and shareholder of NextmiRNA Therapeutics. Christina Yau is a part-time employee/consultant at NantOmics. Franz X. Schaub is an employee and shareholder of SEngine Precision Medicine. Carla Grandori is an employee, founder, and shareholder of SEngine Precision Medicine. Robert N. Eisenman is a member of the Scientific Advisory Boards and shareholder of Shenogen Pharma and Kronos Bio. Daniel J. Weisenberger is a consultant for Zymo Research Corporation. Joshua M. Stuart is the founder of Five3 Genomics and shareholder of NantOmics. Marc T. Goodman receives research support from Merck. Andrew J. Gentles is a consultant for Cibermed. Charles M. Perou is an equity stock holder, consultant, and Board of Directors member of BioClassifier and GeneCentric Diagnostics and is also listed as an inventor on patent applications on the Breast PAM50 and Lung Cancer Subtyping assays. Matthew Meyerson receives research support from Bayer Pharmaceuticals; is an equity holder in, consultant for, and Scientific Advisory Board chair for OrigiMed; and is an inventor of a patent for EGFR mutation diagnosis in lung cancer, licensed to LabCorp. Eduard Porta-Pardo is an inventor of a patent for domainXplorer. Han Liang is a shareholder and scientific advisor of Precision Scientific and Eagle Nebula. Da Yang is an inventor on a pending patent application describing the use of antisense oligonucleotides against specific lncRNA sequence as diagnostic and therapeutic tools. Yonghong Xiao was an employee and shareholder of TESARO. Bin Feng is an employee and shareholder of TESARO. Carter Van Waes received research funding for the study of IAP inhibitor ASTX660 through a Cooperative Agreement between NIDCD, NIH, and Astex Pharmaceuticals. Raunaq Malhotra is an employee and shareholder of Seven Bridges. Peter W. Laird serves on the Scientific Advisory Board for AnchorDx. Joel Tepper is a consultant at EMD Serono. Kenneth Wang serves on the Advisory Board for Boston Scientific, Microtech, and Olympus. Andrea Califano is a founder, shareholder, and advisory board member of DarwinHealth. and a shareholder and advisory board member of Tempus. Toni K. Choueiri serves as needed on advisory boards for Bristol-Myers Squibb, Merck, and Roche. Lawrence Kwong receives research support from Array BioPharma. Sharon E. Plon is a member of the Scientific Advisory Board for Baylor Genetics Laboratory. Beth Y. Karlan serves on the Advisory Board of Invitae.

Received: July 26, 2017
Revised: February 6, 2018
Accepted: March 2, 2018
Published: March 28, 2018

REFERENCES

Adhikary, S., Marinoni, F., Hock, A., Hulleman, E., Popov, N., Beier, R., Bernard, S., Quarto, M., Capra, M., Goettig, S., et al. (2005). The ubiquitin ligase HectH9 regulates transcriptional activation by Myc and is essential for tumor cell proliferation. Cell *123*, 409–421.

Ayer, D.E., Kretzner, L., and Eisenman, R.N. (1993). Mad: a heterodimeric partner for Max that antagonizes Myc transcriptional activity. Cell *72*, 211–222.

Bahram, F., von der Lehr, N., Cetinkaya, C., and Larsson, L.G. (2000). c-Myc hot spot mutations in lymphomas result in inefficient ubiquitination and decreased proteasome-mediated turnover. Blood *95*, 2104–2110.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281–297.

Bazarov, A.V., Adachi, S., Li, S.F., Mateyak, M.K., Wei, S., and Sedivy, J.M. (2001). A modest reduction in c-myc expression has minimal effects on cell growth and apoptosis but dramatically reduces susceptibility to Ras and Raf transformation. Cancer Res. *61*, 1178–1186.

Beroukhim, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M., et al. (2010). The landscape of somatic copy-number alteration across human cancers. Nature *463*, 899–905.

Blackwell, T.K., Huang, J., Ma, A., Kretzner, L., Alt, F.W., Eisenman, R.N., and Weintraub, H. (1993). Binding of myc proteins to canonical and noncanonical DNA sequences. Mol. Cell. Biol. *13*, 5216–5224.

Bouchard, C., Dittrich, O., Kiermaier, A., Dohmann, K., Menkel, A., Eilers, M., and Lüscher, B. (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.

Burk, U., Schubert, J., Wellner, U., Schmalhofer, O., Vincan, E., Spaderna, S., and Brabletz, T. (2008). A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. EMBO Rep. *9*, 582–589.

Burnichon, N., Cascón, A., Schiavi, F., Morales, N.P., Comino-Méndez, I., Abermil, N., Inglada-Pérez, L., de Cubas, A.A., Amar, L., Barontini, M., et al. (2012). MAX mutations cause hereditary and sporadic pheochromocytoma and paraganglioma. Clin. Cancer Res. *18*, 2828–2837.

Cancer Genome Atlas Research Network. (2014a). Comprehensive molecular profiling of lung adenocarcinoma. Nature *511*, 543–550.

Cancer Genome Atlas Research Network. (2014b). Integrated genomic characterization of papillary thyroid carcinoma. Cell *159*, 676–690.

Canisius, S., Martens, J.W., and Wessels, L.F. (2016). A novel independence test for somatic alterations in cancer shows that biology drives mutual exclusivity but chance explains most co-occurrence. Genome Biol. *17*, 261.

Carroll, P.A., Diolaiti, D., McFerrin, L., Gu, H., Djukovic, D., Du, J., Cheng, P.F., Anderson, S., Ulrich, M., Hurley, J.B., et al. (2015). Deregulated Myc requires MondoA/Mlx for metabolic reprogramming and tumorigenesis. Cancer Cell *27*, 271–285.

Carter, S.L., Cibulskis, K., Helman, E., McKenna, A., Shen, H., Zack, T., Laird, P.W., Onofrio, R.C., Winckler, W., Weir, B.A., et al. (2012). Absolute quantification of somatic DNA alterations in human cancer. Nat. Biotechnol. *30*, 413–421.

Casey, S.C., Tong, L., Li, Y., Do, R., Walz, S., Fitzgerald, K.N., Gouw, A.M., Baylot, V., Gütgemann, I., Eilers, M., et al. (2016). MYC regulates the antitumor immune response through CD47 and PD-L1. Science *352*, 227–231.

Cermelli, S., Jang, I.S., Bernard, B., and Grandori, C. (2014). Synthetic lethal screens as a means to understand and treat MYC-driven cancers. Cold Spring Harb. Perspect. Med. *4*, https://doi.org/10.1101/cshperspect.a014209.

Chang, T.-C., Wentzel, E.A., Kent, O.A., Ramachandran, K., Mullendore, M., Lee, K.H., Feldmann, G., Yamakuchi, M., Ferlito, M., Lowenstein, C.J., et al. (2007). Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. Mol. Cell *26*, 745–752.

Ciriello, G., Miller, M.L., Aksoy, B.A., Senbabaoglu, Y., Schultz, N., and Sander, C. (2013). Emerging landscape of oncogenic signatures across human cancers. Nat. Genet. *45*, 1127–1133.

Comino-Méndez, I., Gracia-Aznárez, F.J., Schiavi, F., Landa, I., Leandro-García, L.J., Letón, R., Honrado, E., Ramos-Medina, R., Caronia, D., Pita, G., et al. (2011). Exome sequencing identifies MAX mutations as a cause of hereditary pheochromocytoma. Nat. Genet. *43*, 663–667.

Conacci-Sorrell, M., McFerrin, L., and Eisenman, R.N. (2014). An overview of MYC and its interactome. Cold Spring Harb. Perspect. Med. *4*, a014357.

Cowling, V.H., and Cole, M.D. (2007). The Myc transactivation domain promotes global phosphorylation of the RNA polymerase II carboxy-terminal domain independently of direct DNA binding. Mol. Cell. Biol. *27*, 2059–2073.

Dalla-Favera, R., Westin, E., Gelmann, E.P., Martinotti, S., Bregni, M., Wong-Staal, F., and Gallo, R.C. (1983). The human onc gene c-myc: structure, expression, and amplification in the human promyelocytic leukemia cell line HL-60. Haematol. Blood Transfus. *28*, 247–254.

Dang, C.V., and Eisenman, R.N. (2014). Myc and the Pathway to Cancer (Cold Spring Harbor).

De Paoli, L., Cerri, M., Monti, S., Rasi, S., Spina, V., Bruscaggin, A., Greco, M., Ciardullo, C., Famà, R., Cresta, S., et al. (2013). MGA, a suppressor of MYC, is recurrently inactivated in high risk chronic lymphocytic leukemia. Leuk. Lymphoma *54*, 1087–1090.

de Pretis, S., Kress, T.R., Morelli, M.J., Sabo, A., Locarno, C., Verrecchia, A., Doni, M., Campaner, S., Amati, B., and Pelizzola, M. (2017). Integrative analysis of RNA polymerase II and transcriptional dynamics upon MYC activation. Genome Res. *27*, 1658–1664.

Dezfouli, S., Bakke, A., Huang, J., Wynshaw-Boris, A., and Hurlin, P.J. (2006). Inflammatory disease and lymphomagenesis caused by deletion of the Myc antagonist Mnt in T cells. Mol. Cell. Biol. *26*, 2080–2092.

Diolaiti, D., McFerrin, L., Carroll, P.A., and Eisenman, R.N. (2015). Functional interactions among members of the MAX and MLX transcriptional network during oncogenesis. Biochim. Biophys. Acta *1849*, 484–500.

Dominguez-Sola, D., and Gautier, J. (2014). MYC and the control of DNA replication. Cold Spring Harb. Perspect. Med. *4*, https://doi.org/10.1101/cshperspect.a014423.

Dominguez-Sola, D., Ying, C.Y., Grandori, C., Ruggiero, L., Chen, B., Li, M., Galloway, D.A., Gu, W., Gautier, J., and Dalla-Favera, R. (2007). Non-transcriptional control of DNA replication by c-Myc. Nature *448*, 445–451.

Edelmann, J., Tausch, E., Landau, D.A., Robrecht, S., Bahlo, J., Fischer, K., Fink, A.M., Bloehdorn, J., Holzmann, K., Böttcher, S., et al. (2017). Frequent evolution of copy number alterations in CLL following first-line treatment with FC(R) is enriched with TP53 alterations: results from the CLL8 trial. Leukemia *31*, 734–738.

Fernandez, P.C., Frank, S.R., Wang, L., Schroeder, M., Liu, S., Greene, J., Cocito, A., and Amati, B. (2003). Genomic targets of the human c-Myc protein. Genes Dev. *17*, 1115–1129.

Fero, M.L., Randel, E., Gurley, K.E., Roberts, J.M., and Kemp, C.J. (1998). The murine gene p27Kip1 is haplo-insufficient for tumour suppression. Nature *396*, 177–180.

Frank, S.R., Schroeder, M., Fernandez, P., Taubert, S., and Amati, B. (2001). Binding of c-Myc to chromatin mediates mitogen-induced acetylation of histone H4 and gene activation. Genes Dev. *15*, 2069–2082.

Gabay, M., Li, Y., and Felsher, D.W. (2014). MYC activation is a hallmark of cancer initiation and maintenance. Cold Spring Harb. Perspect. Med. *4*.

Gomez-Roman, N., Felton-Edkins, Z.A., Kenneth, N.S., Goodfellow, S.J., Athineos, D., Zhang, J., Ramsbottom, B.A., Innes, F., Kantidakis, T., Kerr, E.R., et al. (2006). Activation by c-Myc of transcription by RNA polymerases I, II and III. Biochem. Soc. Symp. 141–154.

Grandori, C., Gomez-Roman, N., Felton-Edkins, Z.A., Ngouenet, C., Galloway, D.A., Eisenman, R.N., and White, R.J. (2005). c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I. Nat. Cell Biol. *7*, 311–318.

Grandori, C., Mac, J., Siebelt, F., Ayer, D.E., and Eisenman, R.N. (1996). Myc-Max heterodimers activate a DEAD box gene and interact with multiple E box-related sites in vivo. EMBO J. *15*, 4344–4357.

Grandori, C., Wu, K.-J., Fernandez, P., Ngouenet, C., Grim, J., Clurman, B.E., Moser, M.J., Oshima, J., Russell, D.W., Swisshelm, K., et al. (2003). Werner syndrome protein limits MYC-induced cellular senescence. Genes Dev. *17*, 1569–1574.

Gregory, P.A., Bert, A.G., Paterson, E.L., Barry, S.C., Tsykin, A., Farshid, G., Vadas, M.A., Khew-Goodall, Y., and Goodall, G.J. (2008). The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. Nat. Cell Biol. *10*, 593–601.

Grewal, S.S., Li, L., Orian, A., Eisenman, R.N., and Edgar, B.A. (2005). Myc-dependent regulation of ribosomal RNA synthesis during *Drosophila* development. Nat. Cell Biol. *7*, 295–302.

Guccione, E., Martinato, F., Finocchiaro, G., Luzi, L., Tizzoni, L., Dall' Olio, V., Zardo, G., Nervi, C., Bernard, L., and Amati, B. (2006). Myc-binding-site recognition in the human genome is determined by chromatin context. Nat. Cell Biol. *8*, 764–770.

Han, Y.C., Vidigal, J.A., Mu, P., Yao, E., Singh, I., González, A.J., Concepcion, C.P., Bonetti, C., Ogrodowski, P., Carver, B., et al. (2015). An allelic series of miR-17∼92-mutant mice uncovers functional specialization and cooperation among members of a microRNA polycistron. Nat. Genet. *47*, 766–775.

Hayward, W.S., Neel, B.G., and Astrin, S.M. (1981). Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukosis. Nature *290*, 475–480.

He, L., Liu, J., Collins, I., Sanford, S., O'Connell, B., Benham, C.J., and Levens, D. (2000). Loss of FBP function arrests cellular proliferation and extinguishes c-myc expression. EMBO J. *19*, 1034–1044.

He, L., Thomson, J.M., Hemann, M.T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S.W., Hannon, G.J., et al. (2005). A microRNA polycistron as a potential human oncogene. Nature *435*, 828–833.

Hedges, L.V. (1981). Distribution theory for glass's estimator of effect size and related estimators. J. Educ. Behav. Stat. *6*, 107–128.

Hemann, M.T., Bric, A., Teruya-Feldstein, J., Herbst, A., Nilsson, J.A., Cordon-Cardo, C., Cleveland, J.L., Tansey, W.P., and Lowe, S.W. (2005). Evasion of the p53 tumour surveillance network by tumour-derived MYC mutants. Nature *436*, 807–811.

Herranz, D., Ambesi-Impiombato, A., Palomero, T., Schnell, S.A., Belver, L., Wendorff, A.A., Xu, L., Castillo-Martin, M., Llobet-Navás, D., Cordon-Cardo, C., et al. (2014). A NOTCH1-driven MYC enhancer promotes T cell development, transformation and acute lymphoblastic leukemia. Nat. Med. *20*, 1130–1137.

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. Cell *155*, 934–947.

Hofmann, J.W., Zhao, X., De Cecco, M., Peterson, A.L., Pagliaroli, L., Manivannan, J., Hubbard, G.B., Ikeno, Y., Zhang, Y., Feng, B., et al. (2015). Reduced expression of MYC increases longevity and enhances healthspan. Cell *160*, 477–488.

Huang, M., and Weiss, W.A. (2013). G34, another connection between MYCN and a pediatric tumor. Cancer Discov. *3*, 484–486.

Hurlin, P.J., Quéva, C., and Eisenman, R.N. (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.

Hurlin, P.J., Quéva, C., Koskinen, P.J., Steingrímsson, E., Ayer, D.E., Copeland, N.G., Jenkins, N.A., and Eisenman, R.N. (1995). Mad3 and Mad4: novel Max-interacting transcriptional repressors that suppress c-myc dependent transformation and are expressed during neural and epidermal differentiation. EMBO J. *14*, 5646–5659.

Hurlin, P.J., Steingrìmsson, E., Copeland, N.G., Jenkins, N.A., and Eisenman, R.N. (1999). Mga, a dual-specificity transcription factor that interacts with Max and contains a T-domain DNA-binding motif. EMBO J. *18*, 7019–7028.

Ilic, N., Utermark, T., Widlund, H.R., and Roberts, T.M. (2011). PI3K-targeted therapy can be evaded by gene amplification along the MYC-eukaryotic translation initiation factor 4E (eIF4E) axis. Proc. Natl. Acad. Sci. USA *108*, E699–E708.

Johnston, L.A., Prober, D.A., Edgar, B.A., Eisenman, R.N., and Gallant, P. (1999). *Drosophila* myc regulates cellular growth during development. Cell *98*, 779–790.

Kedde, M., van Kouwenhove, M., Zwart, W., Oude Vrielink, J.A.F., Elkon, R., and Agami, R. (2010). A Pumilio-induced RNA structure switch in p27-3' UTR controls miR-221 and miR-222 accessibility. Nat. Cell Biol. *12*, 1014–1020.

Kim, Y.K., Yu, J., Han, T.S., Park, S.Y., Namkoong, B., Kim, D.H., Hur, K., Yoo, M.W., Lee, H.J., Yang, H.K., et al. (2009). Functional links between clustered microRNAs: suppression of cell-cycle inhibitors by microRNA clusters in gastric cancer. Nucleic Acids Res. *37*, 1672–1681.

Korpal, M., Lee, E.S., Hu, G., and Kang, Y. (2008). The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. J. Biol. Chem. *283*, 14910–14914.

Kress, T.R., Sabò, A., and Amati, B. (2015). MYC: connecting selective transcriptional control to global RNA production. Nat. Rev. Cancer *15*, 593–607.

le Sage, C., Nagel, R., Egan, D.A., Schrier, M., Mesman, E., Mangiola, A., Anile, C., Maira, G., Mercatelli, N., Ciafrè, S.A., et al. (2007). Regulation of the p27(Kip1) tumor suppressor by miR-221 and miR-222 promotes cancer cell proliferation. EMBO J. *26*, 3699–3708.

Li, Y., Choi, P.S., Casey, S.C., Dill, D.L., and Felsher, D.W. (2014). MYC through miR-17-92 suppresses specific target genes to maintain survival, autonomous proliferation, and a neoplastic state. Cancer Cell *26*, 262–272.

Lin, C.Y., Lovén, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., and Young, R.A. (2012). Transcriptional amplification in tumor cells with elevated c-Myc. Cell *151*, 56–67.

Link, J.M., and Hurlin, P.J. (2015). The activities of MYC, MNT and the MAX-interactome in lymphocyte proliferation and oncogenesis. Biochim. Biophys. Acta *1849*, 554–562.

Link, J.M., Ota, S., Zhou, Z.Q., Daniel, C.J., Sears, R.C., and Hurlin, P.J. (2012). A critical role for Mnt in Myc-driven T-cell proliferation and oncogenesis. Proc. Natl. Acad. Sci. USA *109*, 19685–19690.

Little, C.D., Nau, M.M., Carney, D.N., Gazdar, A.F., and Minna, J.D. (1983). Amplification and expression of the c-myc oncogene in human lung cancer cell lines. Nature *306*, 194–196.

Lüscher, B., and Vervoorts, J. (2012). Regulation of gene transcription by the oncoprotein MYC. Gene *494*, 145–160.

Medina, P.P., Nolde, M., and Slack, F.J. (2010). OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma. Nature *467*, 86–90.

Mermel, C.H., Schumacher, S.E., Hill, B., Meyerson, M.L., Beroukhim, R., and Getz, G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. *12*, R41.

Meroni, G., Reymond, A., Alcalay, M., Borsani, G., Tanigami, A., Tonlorenzi, R., Lo Nigro, C., Messali, S., Zollo, M., Ledbetter, D.H., et al. (1997). Rox, a novel bHLHZip protein expressed in quiescent cells that heterodimerizes with Max, binds a non-canonical E box and acts as a transcriptional repressor. EMBO J. *16*, 2892–2906.

Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat. Genet. *34*, 267–273.

Moser, R., Toyoshima, M., Robinson, K., Gurley, K.E., Howie, H.L., Davison, J., Morgan, M., Kemp, C.J., and Grandori, C. (2012). MYC-driven tumorigenesis is inhibited by WRN syndrome gene deficiency. Mol. Cancer Res. *10*, 535–545.

Mu, P., Han, Y.C., Betel, D., Yao, E., Squatrito, M., Ogrodowski, P., de Stanchina, E., D'Andrea, A., Sander, C., and Ventura, A. (2009). Genetic dissection of the miR-17 92 cluster of microRNAs in Myc-induced B-cell lymphomas. Genes Dev. *23*, 2806–2811.

Muncan, V., Sansom, O.J., Tertoolen, L., Phesse, T.J., Begthel, H., Sancho, E., Cole, A.M., Gregorieff, A., de Alboran, I.M., Clevers, H., et al. (2006). Rapid loss of intestinal crypts upon conditional deletion of the Wnt/Tcf-4 target gene c-Myc. Mol. Cell. Biol. *26*, 8418–8426.

Murphy, D.J., Junttila, M.R., Pouyet, L., Karnezis, A., Shchors, K., Bui, D.A., Brown-Swigart, L., Johnson, L., and Evan, G.I. (2008). Distinct thresholds govern Myc's biological output in vivo. Cancer Cell *14*, 447–457.

Nakagawara, A., Ikeda, K., Tsuda, T., and Higashi, K. (1987). N-myc oncogene amplification and prognostic factors of neuroblastoma in children. J. Pediatr. Surg. *22*, 895–898.

Nesbit, C.E., Tersak, J.M., and Prochownik, E.V. (1999). MYC oncogenes and human neoplastic disease. Oncogene *18*, 3004–3016.

Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D.R., Tessarollo, L., Casellas, R., et al. (2012). c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. Cell *151*, 68–79.

Nilsson, J.A., Maclean, K.H., Keller, U.B., Pendeville, H., Baudino, T.A., and Cleveland, J.L. (2004). Mnt loss triggers Myc transcription targets, proliferation, apoptosis, and transformation. Mol. Cell. Biol. *24*, 1560–1569.

Nowell, P., Finan, J., Dalla-Favera, R., Gallo, R.C., ar-Rushdi, A., Romanczuk, H., Selden, J.R., Emanuel, B.S., Rovera, G., and Croce, C.M. (1983). Association of amplified oncogene c-myc with an abnormally banded chromosome 8 in a human leukaemia cell line. Nature *306*, 494–497.

O'Donnell, K.A., Wentzel, E.A., Zeller, K.I., Dang, C.V., and Mendell, J.T. (2005). c-Myc-regulated microRNAs modulate E2F1 expression. Nature *435*, 839–843.

O'Neil, J., Grim, J., Strack, P., Rao, S., Tibbitts, D., Winter, C., Hardwick, J., Welcker, M., Meijerink, J.P., Pieters, R., et al. (2007). FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to gamma-secretase inhibitors. J. Exp. Med. *204*, 1813–1824.

O'Shea, J.M., and Ayer, D.E. (2013). Coordination of nutrient availability and utilization by MAX- and MLX-centered transcription networks. Cold Spring Harb. Perspect. Med. *3*, a014258.

Olive, V., Bennett, M.J., Walker, J.C., Ma, C., Jiang, I., Cordon-Cardo, C., Li, Q.-J., Lowe, S.W., Hannon, G.J., and He, L. (2009). miR-19 is a key oncogenic component of mir-17-92. Genes Dev. *23*, 2839–2849.

Pantaleo, M.A., Urbini, M., Indio, V., Ravegnini, G., Nannini, M., De Luca, M., Tarantino, G., Angelini, S., Gronchi, A., Vincenzi, B., et al. (2017). Genome-wide analysis identifies MEN1 and MAX mutations and a neuroendocrine-like molecular heterogeneity in quadruple WT GIST. Mol. Cancer Res. *15*, 553–562.

Park, S.-M., Gaur, A.B., Lengyel, E., and Peter, M.E. (2008). The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. Genes Dev. *22*, 894–907.

Payne, S.R., and Kemp, C.J. (2005). Tumor suppressor genetics. Carcinogenesis *26*, 2031–2045.

Robinson, K., Asawachaicharn, N., Galloway, D.A., and Grandori, C. (2009). c-Myc accelerates S-phase and requires WRN to avoid replication stress. PLoS One *4*, e5951.

Rohban, S., and Campaner, S. (2015). Myc induced replicative stress response: how to cope with it and exploit it. Biochim. Biophys. Acta *1849*, 517–524.

Roussel, M.F., and Robinson, G.W. (2013). Role of MYC in medulloblastoma. Cold Spring Harb. Perspect. Med. *3*.

Ruggero, D. (2009). The role of Myc-induced protein synthesis in cancer. Cancer Res. *69*, 8839–8843.

Sabo, A., and Amati, B. (2014). Genome recognition by MYC. Cold Spring Harb. Perspect. Med. *4*.

Sabò, A., Kress, T.R., Pelizzola, M., de Pretis, S., Gorski, M.M., Tesi, A., Morelli, M.J., Bora, P., Doni, M., Verrecchia, A., et al. (2014). Selective transcriptional regulation by Myc in cellular growth control and lymphomagenesis. Nature *511*, 488–492.

Salghetti, S.E., Kim, S.Y., and Tansey, W.P. (1999). Destruction of Myc by ubiquitin-mediated proteolysis: cancer-associated and transforming mutations stabilize Myc. EMBO J. *18*, 717–726.

Schaefer, I.M., Wang, Y., Liang, C.W., Bahri, N., Quattrone, A., Doyle, L., Mariño-Enríquez, A., Lauria, A., Zhu, M., Debiec-Rychter, M., et al. (2017). MAX inactivation is an early event in GIST development that regulates p16 and cell proliferation. Nat. Commun. *8*, 14674.

Schwab, M., Alitalo, K., Klempnauer, K.H., Varmus, H.E., Bishop, J.M., Gilbert, F., Brodeur, G., Goldstein, M., and Trent, J. (1983). Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumour. Nature *305*, 245–248.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Sur, I.K., Hallikas, O., Vähärautio, A., Yan, J., Turunen, M., Enge, M., Taipale, M., Karhu, A., Aaltonen, L.A., and Taipale, J. (2012). Mice lacking a Myc enhancer that includes human SNP rs6983267 are resistant to intestinal tumors. Science *338*, 1360–1363.

Ventura, A., and Jacks, T. (2009). MicroRNAs and cancer: short RNAs go a long way. Cell *136*, 586–591.

Vidotto, T., Tiezzi, D.G., and Squire, J.A. (2018). Distinct subtypes of genomic PTEN deletion size influence the landscape of aneuploidy and outcome in prostate cancer. Mol. Cytogenet. *11*, 1.

Vita, M., and Henriksson, M. (2006). The Myc oncoprotein as a therapeutic target for human cancer. Semin. Cancer Biol. *16*, 318–330.

Vo, B.T., Wolf, E., Kawauchi, D., Gebhardt, A., Rehg, J.E., Finkelstein, D., Walz, S., Murphy, B.L., Youn, Y.H., Han, Y.-G., et al. (2016). The interaction of myc with Miz1 defines medulloblastoma subgroup identity. Cancer Cell *29*, 5–16.

Walter, W., Sánchez-Cabo, F., and Ricote, M. (2015). GOplot: an R package for visually combining expression data with functional analysis. Bioinformatics *31*, 2912–2914.

Walz, S., Lorenzin, F., Morton, J., Wiese, K.E., von Eyss, B., Herold, S., Rycak, L., Dumay-Odelot, H., Karim, S., Bartkuhn, M., et al. (2014). Activation and repression by oncogenic MYC shape tumour-specific gene expression profiles. Nature *511*, 483–487.

Washkowitz, A.J., Schall, C., Zhang, K., Wurst, W., Floss, T., Mager, J., and Papaioannou, V.E. (2015). Mga is essential for the survival of pluripotent cells during peri-implantation development. Development *142*, 31–40.

Weng, A.P., Millholland, J.M., Yashiro-Ohtani, Y., Arcangeli, M.L., Lau, A., Wai, C., Del Bianco, C., Rodriguez, C.G., Sai, H., Tobias, J., et al. (2006). c-Myc is an important direct target of Notch1 in T-cell acute lymphoblastic leukemia/lymphoma. Genes Dev. *20*, 2096–2109.

Wilde, B.R., and Ayer, D.E. (2015). Interactions between Myc and MondoA transcription factors in metabolism and tumourigenesis. Br. J. Cancer *113*, 1529–1533.

Yang, G., and Hurlin, P.J. (2017). MNT and emerging concepts of MNT-MYC antagonism. Genes (Basel) *8*, https://doi.org/10.3390/genes8020083.

Zack, T.I., Schumacher, S.E., Carter, S.L., Cherniack, A.D., Saksena, G., Tabak, B., Lawrence, M.S., Zhsng, C.Z., Wala, J., and Mermel, C.H. (2013). Pan-cancer patterns of somatic copy number alteration. Nat. Genet. *45*, 1134–1140.

Zervos, A.S., Gyuris, J., and Brent, R. (1993). Mxi1, a protein that specifically interacts with Max to bind Myc-Max recognition sites. Cell *72*, 223–232.

Zhang, Q., Spears, E., Boone, D.N., Li, Z., Gregory, M.A., and Hann, S.R. (2013). Domain-specific c-Myc ubiquitylation controls c-Myc transcriptional and apoptotic activity. Proc. Natl. Acad. Sci. USA *110*, 978–983.

Zhang, X., Choi, P.S., Francis, J.M., Imielinski, M., Watanabe, H., Cherniack, A.D., and Meyerson, M. (2016). Identification of focally amplified lineage-specific super-enhancers in human epithelial cancers. Nat. Genet. *48*, 176–182.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
EB++AdjustPANCAN_IlluminaHiSeq_ RNASeqV2.geneExp.tsv	The Cancer Genome Atlas Research Network, unpublished data	https://gdc.cancer.gov/about-data/publications/pancanatlas
TCGA-RPPA-pancan-clean-v2.txt	The Cancer Genome Atlas Research Network, unpublished data	https://gdc.cancer.gov/about-data/publications/pancanatlas
GISTIC.focal_data_by_genes.conf_95.tsv	The Cancer Genome Atlas Research Network, unpublished data	https://gdc.cancer.gov/about-data/publications/pancanatlas
pancanMiRs_EBadjOnProtocolPlatformWithout RepsWithUnCorrectMiRs_08_04_16.csv	The Cancer Genome Atlas Research Network, unpublished data	https://gdc.cancer.gov/about-data/publications/pancanatlas
pancan.merged.v0.2.5.filtered.maf.gz	The Cancer Genome Atlas Research Network, unpublished data	https://gdc.cancer.gov/about-data/publications/pancanatlas
Software and Algorithms		
GSEA	The Broad Institute	https://software.broadinstitute.org/gsea/index.jsp
DISCOVER	The Netherlands Cancer Institute	http://ccb.nki.nl/software/discover/
R	The R Foundation	https://www.r-project.org/

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact Carla Grandori (carlagrandori@curefirst.org)

METHODS DETAILS

We analyzed TCGA pan-cancer Atlas cohort defined by the whitelist commonly agreed upon by TCGA AWGs for all analyses. The cohort consisted of 9,125 samples of 33 different histopathologic cancer types representing most major classes of human adult cancer.

Gene Expression

Gene expression data were available for 20502 genes and 9118 samples across 33 tumor types. (File: EB++AdjustPANCAN_IlluminaHiSeq_RNASeqV2.geneExp.tsv)

Protein Expression by RPPA: TCGA-RPPA-pancan-clean-v2.txt

Copy Number Variation

The PanCanAtlas Aneuploidy group (Taylor et al. submitted) produced tumor cancer cell ploidy, heterogeneity, and allelic copy number estimates by running ABSOLUTE (Carter et al., 2012) on segmented Affymetrix SNP6.0 array copy-number data. To evaluate gene amplification and deletions, the allelic copy number estimates were normalized using thresholded cancer cell ploidy to obtain new copy ratio estimates adjusted for tumor purity and ploidy. The purity and ploidy values were also used to perform ISAR correction (Zack et al., 2013) on segmented marker-generated copy ratios, which were then used with GISTIC2.0 (Mermel et al., 2011) to compute relative linear copy number values representing both broad arm-level events and focal events (less than 50% of the chromosome arm), providing metrics to assess the focality of the somatic copy number aberrations.

Overall, we had focal copy number data for 8884 tumor samples, and broad range copy number data for 8785 tumor samples across 33 tumor types and 24203 genes. (Files: GISTIC.focal_data_by_genes.conf_95.tsv, ABSOLUTE.relative_gene_scores)

miRNA Expression

We analysed miRNA expression data for 662 miRNAs and 10824 samples across 33 tumor types. (File: pancanMiRs_EBadjOnProtocolPlatformWithoutRepsWithUnCorrectMiRs_08_04_16.csv)

Somatic Mutations

Only protein coding mutations were retained for downstream analyses (Variant_Classification one of Frame_Shift_Del,Frame_Shift_Ins,In_Frame_Del,In_Frame_Ins,Missense_Mutation, Nonsense_Mutation, Nonstop_Mutation, Splice_Site, and Translation_Start_Site). Further mutations calls were required to be made by two or more mutations callers (NCALLERS>1). Overall, mutation calls were available for 19684 genes for 10133 tumor samples. (File: MC3 Mutation Annotation File pancan.merged.v0.2.5.filtered.maf.gz)

GSEA

The Gene Set Enrichment Analysis tool was used to discover gene sets that were enriched in gene lists ranked by correlation with MYC and MYCN expression individually. The desktop version for this software is available for download at:
http://software.broadinstitute.org/gsea/index.jsp

DISCOVER

The DISCOVER method developed by Canisius et. al. was used to evaluate mutual exclusivity between MYC and 24202 genes across 8884 tumor samples encompassing 33 tumor-types. The same method was also used to evaluate mutual exclusivity with MYC within individual tumor types. The method is documented on Github at:
https://github.com/NKI-CCB/DISCOVER

QUANTIFICATION AND STATISTICAL ANALYSIS

Gene Alterations

Broad and focal copy number variation data was corrected for tumor ploidy and purity (Carter et al., 2012). For broad range copy number variation, a corrected value >= 1.5 was considered amplification and a corrected value <= 0.5 was considered deletion. For focal copy number variation, a corrected value > 0 was considered amplification and a corrected value < 0 was considered deletion.

For mutual exclusivity analysis as well as for correlation analyses between alterations in MYC network genes and expressions of various pathways, a gene was considered altered if it had a focal copy number loss or gain or if it had a coding mutation as defined below in the Data and Software Availability section.

Hedges' g Effect Size

For Figures 3C and 5A, Hedges' g (Hedges, 1981) effect sizes were computed to quantify differences between pairs of groups defined by combined alteration state of MYC with PTEN, BRAF, APC, or PIK3CA respectively. An effect size of magnitude $|g|<0.2$ was considered "negligible", $0.2<=|g|<0.5$ was considered "small", $0.5<=|g|<0.8$ was considered "medium", and $|g|>=0.8$ was considered "large".

Mutual Exclusivity with MYC

For each of the 24202 protein coding genes across the whole genome, mutual exclusivity with respect to MYC was evaluated using the DISCOVER method (Canisius et al., 2016). This evaluation was done across all 33 tumor types collectively, as well as for each tumor type individually. A false discovery rate of 1% was used to indicate statistical significance.

MYC-associated Gene Expression Signature and Gene Set Enrichment Analysis

We computed pairwise Spearman correlation coefficients between each individual MYC gene (c-MYC, MYCN and MYCL) expression and expressions of 20502 genes (whole genome). The ranked gene list based on decreasing order of Spearman coefficients was used as input to GSEA (Subramanian et al., 2005; Mootha et al., 2003) to evaluate enrichment of gene sets in the Gene Ontology Molecular Function category (c5.mf.v5.2.symbols.gmt). Both the correlation analysis and the gene set enrichment analysis were done separately for each of the 33 different tumor types. The output of GSEA consists of a normalized enrichment score for each gene set for each tumor type. The enrichment score is normalized for size of the gene set (number of genes in the gene set) and for correlation between the gene sets and the ranked gene lists. Since gene sets inherently have overlaps in terms of the genes they contain, we used the *reduce_overlap* function in the R *GOplot* package (Walter, Sánchez-Cabo, and Ricote, 2015) to retain only those gene sets in the GSEA results that had less than 75% overlap with other gene sets. In other words, if 2 gene sets had more than 75% overlap, only one of them was retained in the result set for downstream analyses and visualizations. The heatmaps in Figure 6 show the top 100 gene sets with highest median normalized enrichment score (NES) across all the 33 tumor types.

DATA AND SOFTWARE AVAILABILITY

The raw data, processed data and clinical data can be found at the legacy archive of the GDC (https://portal.gdc.cancer.gov/legacy-archive/search/f) and the PancanAtlas publication page (https://gdc.cancer.gov/about-data/publications/pancanatlas). The mutation data can be found here (https://gdc.cancer.gov/about-data/publications/mc3-2017). TCGA data can also be explored through the Broad Institute FireBrowse portal (http://gdac.broadinstitute.org) and the Memorial Sloan Kettering Cancer Center cBioPortal (http://www.cbioportal.org). Details for software availability are in the Key Resource Table.

R Code

Analysis scripts are publicly available at https://github.com/CureFirstResearch/MYC.