Premetazoan Ancestry of the Myc–Max Network

Susan L. Young †1,2, Daniel Diolaiti, †3, Maralice Conacci-Sorrell, 3 Iñaki Ruiz-Trillo, 4,5 Robert N. Eisenman, *,3 and Nicole King, *,1,2

1 Department of Molecular and Cell Biology, Division of Genetics, Genomics and Development, University of California, Berkeley
2 Center for Integrative Genomics, University of California, Berkeley
3 Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington
4 Departament de Genètica & Institut de Recerca en Biodiversitat (Irbio), Universitat de Barcelona, Barcelona, Spain
5 Institució Catalana per a la Recerca i Estudis Avançats (ICREA), Barcelona, Spain

† These authors contributed equally to this study.
*Corresponding authors: E-mail: eisenman@fhcrc.org; nking@berkeley.edu.
Associate editor: Billie Swalla

Abstract

The origin of metazoans required the evolution of mechanisms for maintaining differentiated cell types within a multicellular individual, in part through spatially differentiated patterns of gene transcription. The unicellular ancestor of metazoans was presumably capable of regulating gene expression temporally in response to changing environmental conditions, and spatial cell differentiation in metazoans may represent a co-option of preexisting regulatory mechanisms. Myc is a critical regulator of cell growth, proliferation, and death that is found in all metazoans but absent in other unicellular lineages, including fungi and plants. Homologs of Myc and its binding partner, Max, exist in two of the closest living relatives of animals, the choanoflagellate Monosiga brevicollis (Mb) and Capsaspora owczarzaki, a unicellular opisthokont that is closely related to metazoans and choanoflagellates. We find that Myc and Max from M. brevicollis heterodimerize and bind to both canonical and noncanonical E-boxes, the DNA-binding sites through which metazoan Myc proteins act. Moreover, in M. brevicollis, MbMyc protein can be detected in nuclear and flagellar regions. Like metazoan Max proteins, MbMax can form homodimers that bind to E-boxes. However, cross-species dimerization between Mb and human Myc and Max proteins was not observed, suggesting that the binding interface has diverged. Our results reveal that the Myc/Max network arose before the divergence of the choanoflagellate and metazoan lineages. Furthermore, core features of metazoan Myc function, including heterodimerization with Max, binding to E-box sequences in DNA, and localization to the nucleus, predate the origin of metazoans.

Key words: Myc, Max, multicellularity, choanoflagellates, Monosiga brevicollis, Capsaspora owczarzaki, evolution.

Introduction

The ability to tightly regulate cell growth and proliferation is crucial to the development and maintenance of multicellularity in metazoans. The transcription factor Myc regulates cell growth, proliferation, differentiation, and apoptosis in response to intercellular signaling and environmental cues (Grandori et al. 2000; Liu and Levens 2006). In addition, Myc homologs have been identified in all major metazoan lineages (Gallant et al. 1996; Simionato et al. 2007; Hartl et al. 2010). Myc has previously been identified in only two nonmetazoan lineages—the choanoflagellates (Brown et al. 2008; King et al. 2008) and Capsaspora owczarzaki, a unicellular relative of choanoflagellates and metazoans (Ruiz-Trillo et al. 2008; Sebe-Pedros et al. 2011). The discovery of a Myc homolog in the genome of the unicellular choanoflagellate Monosiga brevicollis revealed that Myc evolved before the origin and diversification of metazoans. Because choanoflagellates are the closest living relatives of metazoans, yet lack complex multicellularity, investigating the function of Myc in choanoflagellates promises to illuminate the role of Myc before the origin of metazoans and may reveal how the strict regulation of cell life and death in metazoans arose from a unicellular context. Furthermore, comparative studies of transcription factor function in metazoans and their unicellular relatives may reveal how ancestral modes of regulating gene expression in response to transient environmental cues might have been co-opted into the spatiotemporal regulation of cell differentiation in metazoans.

Myc is a basic-helix-loop-helix leucine zipper (bHLH-LZ) transcription factor and its function is regulated by interactions with other bHLH-LZ proteins. Heterodimerization of Myc with the bHLH-LZ protein Max (Myc-associated factor X) is required for binding to the canonical Myc target sequence, the E-box (Blackwood and Eisenman 1991; Blackwood et al. 1992). Conversely, Myc activity can be antagonized indirectly by the dimerization of Max with other bHLH-LZ transcription factors such as Mxd, Mnt, or itself, thereby reducing the availability of Max for dimerization with Myc (Ayer et al. 1993; Grandori et al. 2000).

Myc function is also shaped by its subcellular localization. In proliferative cells, Myc is localized to the nucleus...
where it promotes the transcription of genes required for ribosome biogenesis, cell cycle regulation, and nucleotide and amino acid biosynthesis (for review, see Eilers and Eisenman 2008). As metazoan cells differentiate and cease to proliferate, Myc expression is repressed and at least in certain biological contexts Myc protein is restricted to the cytoplasm, where it is inactive as a transcription factor (Craig et al. 1993; Wakamatsu et al. 1993; Wang et al. 1997; Conacci-Sorrell et al. 2010).

Metazoan Myc proteins are characterized by a C-terminal bHLH-LZ domain and a conserved set of N-terminal “Myc homology box” motifs (MBI–IV) that can be used to distinguish Myc homologs from other bHLH-LZ transcription factors (Stone et al. 1987; Kato et al. 1990; Spotts et al. 1997; McMahon et al. 1998; McMahon et al. 2000; Herbst et al. 2005; Cowling and Cole 2006; Kurland and Tansey 2008). MBI and MBII are the most highly conserved of the Myc homology boxes and are found in vertebrate, insect, and cnidarian Myc proteins (Brough et al. 1995; Gallant et al. 1996; Gallant 2006; Hartl et al. 2010). MBI and MBIV are not as well conserved. MBIV appears to be a vertebrate-specific motif (Cowling et al. 2006), although MBI has been identified in vertebrate and cnidarian Myc proteins, it is highly diverged in Drosophila melanogaster dMyc (Gallant et al. 1996; Hartl et al. 2010). Human and mouse c-Myc and D. melanogaster dMyc are the best studied of the Myc homologs, and the conservation of function between these homologs, as demonstrated by similarities in their regulation of growth control genes and the ability of dMyc to partially substitute for c-Myc, corroborates the structural conservation observed in the domain architecture of Myc homologs from diverse metazoan phyla (Gallant et al. 1996; Schreiber-Agus et al. 1997; Trumpf et al. 2001).

We previously reported a candidate homolog of Myc in the genome of M. brevicollis, and others have reported a M. brevicollis homolog of Max (Brown et al. 2008; King et al. 2008). Here, we report that MbMyc and C. owczarzaki Myc contain conserved sequence motifs that predict their ability to interact with Max. In addition, we show that MbMyc performs core functions of metazoan Myc proteins, namely heterodimerization with Max and localization to the nucleus and flagellum in cells from proliferating cultures. Notably, MbMyc/MbMax heterodimers form and bind to E-boxes in a species-specific manner. These results extend the origin of the Myc–Max network to the unicellular progenitors of metazoans and demonstrate that functional characteristics of Myc proteins predate the divergence of the choanoflagellate and metazoan lineages.

Materials and Methods

Nematostella vectensis and Trichoplax adhaerens Mxd/Mnt homologs were identified by tblastn searches in their respective JGI genome browsers (http://genome.jgi-psf.org/Nemve1 and http://genome.jgi-psf.org/Triad1) using the Mxd/Mnt protein sequences listed in the supplementary materials. The sequence data sets from two closely related homoscleromorph sponges, Oscarella carmela and Oscarella sp., are each incomplete, so we used a merger of data from these sister groups to represent sponges in our analysis. Myc and Mxd homologs in expressed sequence tags (ESTs) from the sponge O. carmela have been previously reported as unpublished observations (Simionato et al. 2007) but have not been explicitly described. Using the Myc, Max, and Mxd/Mnt protein sequences listed in the supplementary materials, Oscarella homologs were identified by tblastn searches in a draft genome assembly as well as EST and RNA-Seq data sets (Nichols et al. 2006 and supplementary materials, Supplementary Material online). The homologs described herein are O. carmela Myc and Oscarella sp. Max and Mnt like. A partial Oscarella sp. Myc protein was identified by tblastn, but the draft genome assembly does not include the N-terminal region. Reciprocal best blast searches were used to validate the best hits. Protein and nucleotide sequences are available in the supplementary materials (Supplementary Material online).

Identification of N-Terminal Myc Homology Domains

To investigate whether MbMyc and C. owczarzaki Myc contain any of the previously characterized Myc homology boxes (MBI–MBIV) that lie within the N-terminal transcription domain of metazoan Myc transcription factors, alignments were made in ClustalW with iteratively smaller segments of the N-terminal region to identify the best matches for these short motifs that do not have a conserved position within Myc proteins.

Modeling of 3D Organization of MbMyc and MbMax Conservation

Coordinates for the human Myc–Max DNA co-complex crystal structure (1NKP) were obtained from the RCSB Protein Data Bank (www.pdb.org; Nair and Burley 2003). Residues 353–434 of human c-Myc (accession P01106) were aligned with residues 279–383 of MbMyc in ClustalW (Thompson et al. 1994). Residues that were conserved or similar as defined by the Gonnet matrix (Gonnet et al. 1992) were highlighted in pink on the 1NKP structure in MacPyMol (http://www.pymol.org). Residues 22–103 of human Max (accession NP_002373) were aligned with residues 35–119 of MbMax. Conserved and similar residues were highlighted in yellow on the 1NKP structure as above.

Cloning and Expression of MbMyc and MbMax

Full-length MbMyc and MbMax coding sequences were cloned into expression vectors by polymerase chain reaction amplifying the entire coding region of MbMyc and MbMax from M. brevicollis cDNA using the following primers:

1. Myc antibody forward: 5’-CCGGAATTCATGAGCTCGTGTGCAAGATCGCTGTTG-3’, which contains an EcoRI restriction site (underlined).
2. Myc antibody pMal: reverse, 5’-CCGCGATCCAGTTAC-GAAGTGGC-3’, which contains a BamHI restriction site (underlined) as well as a stop codon (in bold).
3. Myc antibody pGex reverse: 5’-CCGGTCCAGGTAC-GAAGTGGC-3’, which contains a Sall restriction site (underlined) as well as a stop codon (in bold).
4. Max pGex forward: 5’-CGAGAGGCCGCGCTTTATGG-TGCAGTCAGGCCT-3’, which contains a BamHI restriction site (underlined).
5. Max pGex reverse: 5’-CGAGAGGCCGCGCTTTATGG-TGCAGTCAGGCCT-3’, which contains a NotI restriction site (underlined) as well as a stop codon (in bold).

We cloned the amplified cDNA into pGEX-6P-1 (Amer- sham Biosciences, Piscataway, NJ) and pMAL-c2X (New England Biolabs, Ipswich, MA) expression vectors using the engineered restriction sites to generate three fusion proteins: MbMyc N-terminally tagged with glutathione-S-transferase (GST), MbMyc N-terminally tagged with maltose-binding protein (MBP), and MbMax N-terminally tagged with GST. We then transformed Escherichia coli BL21 cells to express each construct as well as the empty vector to express each tag on its own and purified the fusion proteins and their respective tags using affinity chromatography. Amylose columns were used to purify MBP-MbMyc, and GST and glutathione–sepharose columns were used to purify GST-MbMyc, GST-MbMax, and GST.

In Vitro Protein–Protein Interaction Assay
GST-MbMax and GST-tag alone were expressed in E. coli (BL21) and purified over glutathione–sepharose columns without elution. Purified MBP-MbMyc was incubated with the GST-MbMax– or GST-coated glutathione columns in 10 mM Tris–HCl, 150 mM KCl, 0.5 mM ethylenediaminetetraacetic acid (EDTA), 0.1% Triton-X 100, 12.5% glycerol, and 1 mM dithiothreitol (DTT). The bound proteins were eluted by boiling in 1× Laemmli buffer. The eluates were run on sodium dodecyl sulfate–polyacrylamide gels and subjected to western blotting by transferring the separated proteins onto a nitrocellulose membrane and probing with a 1:500 dilution of the anti-MbMyc antibody in 5% powdered milk/Tris-buffered saline, 0.1% Tween-20. For cross-species heterodimerization analysis, 20 μl of wheat germ extract containing the indicated [35S]methionine-labeled proteins by allowing the sera to flow over E. coli lysate–coated beads. The flow through was collected and then allowed to bind to GST-MbMyc bound to a glutathione–sepharose column. After extensive washing, the affinity-purified antibodies were eluted and stored in 0.02% sodium azide, 2.5% normal goat serum in phosphate-buffered saline. The specificity of the affinity-purified antibody was confirmed by competition experiments in which preincubation of the purified antibody with recombinant Myc before use in immunofluorescence staining was found to prevent all detection of endogenous Myc (supplementary fig. S4, Supplementary Material online).

In Vitro Transcription and Translation
MbMyc and MbMax cDNA were subcloned in the pCS2+ vector to allow SP6 RNA polymerase–dependent transcription. pCS2-HsMyc, pRK7-HsHMax, pCS2-MbMyc, and pCS2-MbMax were in vitro transcribed and translated alone or in combination by using the TNT-coupled wheat germ extract system (Promega) following the manufacturer’s protocol. Either unlabeled methionine or [35S]methionine was added to the reaction.

Electrophoretic Mobility Shift Assays
Oligonucleotide probes used for electrophoretic mobility shift assays (canonical E-box sequences indicated in bold). Sense oligonucleotide sequence: 5’-CGCGCGCGGC-CACGCGGCAAACTA-3’. Antisense oligonucleotide sequence: 5’-AGGGTGGTCGCTCAGGCTT-3’. Mutated E-box, noncanonical E-box-1, and noncanonical E-box-2 sense and antisense oligonucleotides were identical to the canonical E-box oligonucleotides but contain respectively the following E-box sequences: CTGAG, CATGTG (1), and CACGCG (2) (Blackwell et al. 1993; Perini et al. 2005). Complementary sense and antisense oligonucleotides were annealed, and 50 ng of annealed probe was labeled with [α-32P]dCTP by performing a fill-in reaction with the entire Reaction of DNA polymerase I. Labeled probes were purified using Microcon Ultracel YM-10 filter devices. Wheat germ extracts containing the indicated in vitro transcribed and translated (IVT) proteins were incubated with approximately 0.5 ng of probe (2 × 105 to 1 × 106 cpm) at 4°C in electrophoretic mobility shift assay-binding buffer (20 mM Tris–HCl, pH 8, 100 mM NaCl, 0.5% NP-40, 1 mM DTT, 1 mM EDTA, 1 mM phenylmethyl-sulfonyl fluoride [PMSF], Roche complete EDTA–free protease inhibitor). After 2-h incubation at 4°C, beads were extensively washed with GST-binding buffer and proteins were eluted by boiling the samples in 1× Laemmli buffer.

Anti-MbMyc Antibody Development
Antibodies against MBP-MbMyc were raised in rabbits (Covance Research Products Inc., Princeton, NJ). To isolate MbMyc-specific antibodies from immunized rabbit sera, we first removed antibodies cross-reactive to E. coli proteins by allowing the sera to flow over E. coli lysate–coated beads. The flow through was collected and then allowed to bind to GST-MbMyc bound to a glutathione–sepharose column. After extensive washing, the affinity-purified antibodies were eluted and stored in 0.02% sodium azide, 2.5% normal goat serum in phosphate-buffered saline. The specificity of the affinity-purified antibody was confirmed by competition experiments in which preincubation of the purified antibody with recombinant Myc before use in immunofluorescence staining was found to prevent all detection of endogenous Myc (supplementary fig. S4, Supplementary Material online).

MbMyc Immunofluorescence Staining and Imaging
For the study of Myc subcellular localization, M. brevicollis cultured with Flavobacterium sp. NK107 (ATCC strain number PRA-258) was split 1:5 into fresh growth media and grown to 1.7 × 106 cells/ml (ca. 1 day). Cells were...
pelleted by spinning for 10 min at 500 \times g and then resuspended in a small volume of artificial seawater from which 1-mL aliquots were dispensed across 3 \times 1.5-ML tubes. Formaldehyde was added to a final concentration of 4%, and cells were pipetted onto poly-L-lysine–coated glass coverslips to fix as they settled on the slip for 20 min. The 4% formaldehyde fixative was removed, and the cells were fixed in 6% acetic acid in artificial seawater for 5 min. The coverslips were then washed with PEM (100 mM 1,4-piperazinediethane-sulfonic acid, pH 6.9, 1 mM ethyleneglycol-bis(aminooxyethylether)-tetraacetic acid, 0.1 mM MgSO\(_4\)) and blocked in PEM/1% BSA/0.3% Triton-X 100 for 30 min. The block was removed and replaced with a 1:100 dilution of MbMyc antibody in blocking solution and incubated at 4\(^\circ\)C overnight. The coverslips were then washed three times with blocking solution, and a 1:400 dilution of Alexa Fluor 488 goat anti-rabbit IgG (H+L) secondary antibody (Molecular Probes, Carlsbad, CA) was added and incubated for 1 h at room temperature. The coverslips were then washed three times with blocking solution and then incubated with a 5 U/100 \muL rhodamine-phalloidin (Invitrogen) and 10 ng/\muL 4',6-diamidino-2-phenylindole (DAPI) in PEM solution for 20 min. Finally, the coverslips were washed twice with PEM and then were mounted to glass slides in Prolong Gold with DAPI (Invitrogen) and sealed with clear nail polish.

Images were taken at 16-bit depth with a 100 \times oil immersion objective on an inverted Leica microscope as Z-stacks. Images shown are Z-projections created with the Max Intensity ZProjection feature of ImageJ (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, MA, http://rsb.info.nih.gov/ij/, 1997–2009). The highest resolution slices from each channel’s stack were used so that the nuclei of multiple cells within a single field could be in focus. The projections were converted to 8-bit image depth in ImageJ. Finally, the images were contrast adjusted and then color and differential interference contrast optics overlays were cropped to regions of interest in Adobe Photoshop CS (Adobe Systems Inc., San Jose, CA).

**Results**

**Conservation of Myc and Max in Unicellular Relatives of Metazoans**

To investigate the extent to which Myc, Max, and Mxd/Mnt genes are conserved in nonmetazoans, we reanalyzed the genomes of *M. brevicollis* and *C. owczarzaki*. Consistent with prior findings, *M. brevicollis* contains homologs of Myc and Max but seemingly lacks the Max-binding partners Mxd or Mnt (Brown et al. 2008; King et al. 2008). In contrast, the genome of *C. owczarzaki* contains homologs of Myc, Max, and Mnt (figs. 1 and 2) (Sebe-Pedros et al. 2011).

Key features of Myc and Max proteins, including overall protein domain architecture and specific residues in the basic region important for recognition of the E-box, are conserved in *M. brevicollis* and *C. owczarzaki* (fig. 1; supplementary fig. S1, Supplementary Material online; Landschulz et al. 1988; Ferre-D’Amare et al. 1993; Nair and Burley 2003). The strongest conservation among metazoan Myc and Max protein sequences can be found in their bHLH-LZ regions. The bHLH-LZ regions of the *M. brevicollis* proteins MbMyc and MbMax share 67% and 84% amino acid similarity, respectively, with the bHLH-LZ regions from human Myc and Max. The bHLH-LZ regions of Myc and Max homologs from *M. brevicollis* and *C. owczarzaki* share more similarity with metazoan Myc and Max proteins than with non-Myc/Max bHLH proteins, respectively, represented by human MyoD (fig. 1B, see also Brown et al. 2008). In addition, diagnostic residues that distinguish Myc and Max from other bHLH transcription factor proteins (Atchley and Fernandes 2005) are found in MbMyc and MbMax, providing further support for their homology (supplementary fig. S1, Supplementary Material online). Notably, the leucine zipper domains of MbMax and *C. owczarzaki* Max contain residues (QN) considered to promote heterodimerization and disfavor homodimerization (fig. 1D; Nair and Burley 2003). However, the leucine zipper domains of MbMyc and *C. owczarzaki* Myc lack a pair of residues (RR) previously found to determine the heterodimerization of human c-Myc with Max (Nair and Burley 2003).

To explore whether MbMyc and *C. owczarzaki* Myc contain any of the previously characterized Myc homology boxes (MBI–MBIV) that lie within the N-terminal transactivation domain of metazoan Myc transcription factors, alignments were made in ClustalW with iteratively smaller segments of the N-terminal region to identify the best matches for these small motifs within the relatively large Myc proteins (fig. 1A, E, and F). MBI and MBIV were not detected in either MbMyc or *C. owczarzaki* Myc. In contrast, conservation of the MBII domain extends from vertebrates to sponges to and to a lesser degree to *M. brevicollis* and *C. owczarzaki* (fig. 1F). Like MBII, the MBIII domain is conserved from vertebrates to sponges to *C. owczarzaki* (fig. 1F). Although MbMyc lacks the clear conservation of the MBIII domain seen in *C. owczarzaki* (fig. 1F), the fact that this domain is conserved between diverse metazoans and *C. owczarzaki* suggests that it was an ancestral feature that was not conserved in the *M. brevicollis* lineage.

**MbMyc/MbMax Heterodimers Bind to Canonical and Noncanonical E-boxes**

To determine the extent to which MbMyc and MbMax behave like metazoan Myc and Max proteins, we first tested their ability to heterodimerize and bind E-boxes, the best characterized targets of Myc activity. The amino acid sequence conservation of MbMyc and MbMax mapped onto a Myc–Max DNA co-complex structure reveals that conserved residues are concentrated at the predicted interfaces between Myc and Max and between the two proteins and DNA (fig. 3A). This suggests that MbMyc and MbMax might have the capacity to form heterodimers on DNA (Burley et al. 1993). The prediction that MbMyc and MbMax heterodimerization is supported by in vitro protein interaction assays performed with bacterially expressed proteins (fig. 3B; supplementary fig. S2, Supplementary Material online) as well as in vitro translated proteins (see fig. 5). Therefore,
structure/function predictions and in vitro biochemical assays indicate that MbMyc and MbMax can heterodimerize. In animals, Myc must form a heterodimer with Max to bind to E-boxes. To test binding of MbMyc–MbMax heterodimers to DNA, we incubated increasing amounts of IVT MbMyc and MbMax with a radioactively labeled oligonucleotide probe containing the canonical vertebrate (and *D. melanogaster*) Myc–Max E-box–binding sequence (CACGTG) (fig. 4 A, lanes 6–9). As a positive control, we also incubated the same probe with increasing amounts of IVT extracts containing HsMyc and HsMax (fig. 4A, lanes 2–5). Figure 4A shows that both HsMyc–HsMax and MbMyc–MbMax heterodimers bind to the E-box sequence. That the observed electrophoretic mobility shift is specifically due to the binding of the HsMyc–HsMax or MbMyc–MbMax heterodimers to the CACGTG consensus sequence is consistent with our finding that no probe shift is detected when the same samples are incubated with an identical oligonucleotide probe containing a mutated E-box (CTCGAG) (fig. 4A, lanes 11–18). We also detected a strong mobility shift due to the binding of a lower–molecular weight complex in the MbMyc samples, which we attribute to formation of MbMax homodimers (see below).

To further examine the identity and properties of the MbMyc–MbMax DNA complex, we performed supershift...
and competition experiments (fig. 4B). We added to the binding reaction either a nonspecific rabbit IgG or an anti-MbMyc antibody (fig. 4B, lanes 4 and 5) and found that the anti-MbMyc antibody specifically inhibits the formation of the high–molecular weight complex. Interestingly, the anti-MbMyc antibody does not inhibit the formation of the faster migrating protein/DNA complex (labeled MbMax–MbMax; fig. 4B; cf. lanes 3–5). To test the specificity of the binding, we also performed a competition experiment in which we added to the binding reaction increasing amounts of unlabeled competitor oligonucleotide containing CACGTG or a mutated CTCGAG sequence. As shown in figure 4B (lanes 9–11), the unlabeled canonical CACGTG oligonucleotide efficiently competes for the binding of the MbMyc–MbMax heterodimers, whereas a CTCGAG-containing probe fails to compete for binding to MbMyc–MbMax complex. In addition, cold probes containing noncanonical CACGTG sequences previously shown to be recognized by the human Myc–Max complex compete, although less efficiently, for the binding of the MbMyc–MbMax complex (fig. 4B, lanes 15–17 and 18–20). Consistent with the notion that the low–molecular weight complex is composed of MbMax homodimers, we found that this complex is unaffected by addition of the specific anti-MbMyc antibody to the reaction, whereas it does specifically bind to both canonical and noncanonical E-boxes as shown by the competition experiment.

To determine whether the higher–molecular weight complex is composed of MbMyc homodimers or MbMyc–MbMax heterodimers, we incubated a probe containing a canonical E-box sequence with IVT samples in which each protein had been in vitro translated in separate reactions. The results shown in figure 4C clearly indicate that only MbMax is able to homodimerize and bind DNA efficiently. Using HsMax, a specific band is detected only following relatively long exposure times, confirming previous reports showing that HsMax can also homodimerize in vitro. No signals were detected in the samples containing either HsMyc or MbMyc alone, even after longer exposures. Taken together, these results allow us to conclude that the upper and the lower bands detected in the IVT MbMyc + MbMax samples correspond, respectively, to MbMyc–MbMax heterodimer and MbMax homodimer–containing complexes.

**FIG. 2.** Phylogenetic distribution of Myc, Max, and Mxd/Mnt family members. Myc, Max, and Mxd/Mnt proteins are found in metazoans and some of their unicellular relatives but are not found in yeast or non-opisthokont lineages. The presence in each genome of at least one representative of the Myc, Max, and Mxd/Mnt protein families is indicated by a colored box. The presence or absence of Myc, Max, and Mxd/Mnt in the species shown was determined by literature searches when data were available or reciprocal best Blast analysis when homology had not been previously assessed (see Materials and Methods). "Choanos": choanoflagellates. "Opisthokonts": a clade containing metazoans, choanoflagellates, *Capaspora owczarzaki*, fungi, and diverse unicellular eukaryotes, most of which are characterized by the presence of a single posterior flagellum. Full names of species represented by genus name in figure: *Saccharomyces cerevisiae* (yeast), *C. owczarzaki*, *Monosiga brevicollis* (choanoflagellate), *Oscarella carmela* and *Oscarella* sp. (sponge), *Trichoplax adhaerens* (placzoan), *Nematostella vectensis* (sea anemone), *Caenorhabditis elegans* (nematode), *Drosophila melanogaster* (fly), *Mus musculus* (mouse), and *Homo sapiens* (human). The cladogram shown represents a consensus of recent molecular phylogenies, including Ruiz-Trillo et al. (2008) and Srivastava et al. (2010).

**FIG. 3.** MbMyc heterodimerizes with MbMax: (A) Mapping conserved residues from MbMyc and MbMax onto the crystal structure of a human Myc–Max DNA complex predicts that the *Monosiga brevicollis* homologs are capable of binding to form heterodimers. Residues that are conserved or similar in MbMyc and human c-Myc (pink) and those that are conserved or similar in MbMax and human Max (yellow) are highlighted on the crystal structure of the bHLH-LZ domains of a human Myc–Max heterodimer bound to an E-box–containing DNA oligo. (B) Purified MBP-MbMyc (arrow) binds to the GST-MbMax column but not to a column containing GST alone.
Species Specificity of MbMyc/MbMax Heterodimerization

Mammalian Myc and Max proteins have been shown to readily form heterodimers with evolutionarily distant Max and Myc proteins from *Xenopus laevis* (King et al. 1993). Furthermore, *D. melanogaster* Myc was first identified using mammalian Max in a two-hybrid screen, whereas dMyc and mammalian Myc were found to be active in mammalian and *D. melanogaster* cells, respectively, implying cross-species interaction with endogenous Max (Gallant et al. 1996; Schreiber-Agus et al. 1997; Benassayag et al. 2005). To directly examine the ability of MbMyc and MbMax to heterodimerize with the reciprocal proteins from humans, we performed GST pull-down experiments in which extracts containing IVT HsMyc or IVT HsMax (fig. 5A and B) were incubated either with GST alone or with GST-MbMax or GST-MbMyc fusion proteins immobilized on glutathione–sepharose beads. As shown in figure 5A and B, neither HsMyc or HsMax efficiently bind to their reciprocal *M. brevicollis* heterodimerization partners. Finally, to test whether MbMax can homodimerize with HsMax, IVT HsMax was incubated with beads containing GST or GST-MbMax (fig. 5C). Consistent with our DNA-binding data (fig. 4), the results show that IVT MbMax homodimerizes with GST-MbMax, whereas IVT HsMax does not appear to efficiently interact with GST-MbMax.

To test for cross-species heterodimerization on DNA, we next cotranslated either HsMyc + MbMax or HsMax + MbMyc proteins in vitro and incubated the IVT reactions with a labeled oligonucleotide probe containing a canonical E-box sequence. As shown in figure 4D (lanes 6–13), we were unable to detect any high–molecular weight complexes in the samples where human and choanoflagellate proteins were reciprocally matched, although MbMax homodimers were observed. Under the same conditions, MbMax and MbMyc formed heterodimers and MbMax homodimers (fig. 4D, lanes 2–5). These results suggest that although human and choanoflagellate proteins share a high degree of sequence similarity, they are not able to reciprocally heterodimerize.

MbMyc Localizes to the Nucleus

To investigate whether MbMyc might function as a transcriptional regulator in vivo, we determined its localization in *M. brevicollis* cells collected from a proliferating culture. MbMyc in most cells was concentrated in the nucleus, whereas also being distributed diffusely throughout the cell body, including the flagellum (fig. 6; supplementary fig. S5).

**Fig. 4.** Electrophoretic mobility shift assays to examine MbMyc and MbMax dimerization and DNA-binding specificity. (A) HsMyc + HsMax or MbMyc + MbMax proteins were cotranslated in vitro using a wheat germ translation system (see Materials and Methods). Increasing amounts of each reaction (1, 2.5, 5, and 10 μl) were incubated with a 32p-labeled oligonucleotide probe containing either a canonical (lanes 1–9) or a mutated E-box sequence (lanes 10–18). (B) Supershift and competition experiments. Anti-MbMyc or nonimmune IgG was added to the MbMyc + MbMax + oligonucleotide probe reactions as indicated (lanes 1–8). Unlabeled probes containing canonical or noncanonical E-box sequences were used in competition reactions with MbMyc + MbMax and labeled probe as indicated (lanes 9–20). HsMyc–HsMax extracts were used as a positive control (lanes 2 and 7). (C) Human and choanoflagellate Myc and Max proteins were in vitro translated separately and increasing amounts of each extract (1, 2.5, 5, and 10 μl) incubated with a labeled probe containing a canonical E-box sequence.
Supplementary Material online). A subset of cells contained primarily cytoplasmic MbMyc and little detectable nuclear localization (supplementary fig. S5, Supplementary Material online). It is not currently known what regulates whether MbMyc localizes to the nucleus or cytoplasm. Nonetheless, the observation that MbMyc is enriched in the nucleus of most cells is consistent with the hypothesis that, like its metazoan homologs, MbMyc can act as a transcription factor.

Discussion

Our results demonstrate that the choanoflagellate M. brevicollis possesses homologs of metazoan Myc and Max transcription factors, both in terms of sequence conservation and protein function. Considering the ability of MbMyc/MbMax heterodimers and MbMax homodimers to bind to E-boxes and the localization of MbMyc to the nucleus in proliferating cells, two properties of Myc function that are well conserved in metazoans, one can reasonably infer that these aspects of Myc function were present in the last common ancestor of choanoflagellates and metazoans. Furthermore, because representatives of the Myc, Max, and Mxd/Mnt protein families exist in the unicellular eukaryote C. owczarzaki (Sebe-Pedros et al. 2011), it is likely that the regulatory network evolved prior to the diversification of metazoans and choanoflagellates. That the MBII and MBIII motifs are more readily identified in the N-terminus of C. owczarzaki Myc than in MbMyc argues that while these are ancestral features of all Myc proteins, they have since diverged in the choanoflagellate lineage. For MBII, this is perhaps unsurprising given that this motif mediates interactions with transcriptional cofactors, some of which are not present or highly diverged in M. brevicollis (McMahon et al. 1998; King et al. 2008). Consistent with divergence in transcriptional cofactor binding, we have been unable to detect any biological or transcriptional activity of MbMyc or MbMax in mammalian cells (data not shown).
Another potentially significant feature of Myc proteins from *M. brevicollis* and *C. owczarzaki* is the absence of the two arginine residues (Arg423–Arg424) that were reported to mediate specificity and affinity of binding between human Myc and Max proteins (Nair and Burley 2003). Although the bHLH-LZ domains of human and choanoflagellate Myc proteins are highly conserved, MbMyc differs from HsMyc in that at the positions equivalent to Arg423–Arg424, MbMyc contains isoleucine and asparagine (I372–N373). Thus, in MbMyc, two polar positively charged residues are replaced by two neutral residues, most likely lowering the affinity between MbMyc and HsMax and providing a possible explanation for our inability to demonstrate their interaction (figs. 4 and 5). On the other hand, when we compared HsMax and MbMax, they both encode a glutamine and an asparagine in the equivalent positions (HsMax Q91–N92 and MbMax Q108–N109). Based on this observation, we would expect HsMax to interact with both HsMyc and MbMyc. However, at least under our experimental conditions, this is not the case, suggesting that other residues may be important determinants of affinity and specificity of heterodimerization between MbMyc and MbMax. Further experiments, including crystallography of the MbMyc–MbMax/DNA complex and site-specific mutagenesis of the bHLH-LZ domains, would be required to establish the structural determinants of MbMyc/MbMax heterodimerization.

Bilaterian Myc proteins are powerful regulators of potentially destructive cellular processes (e.g., apoptosis, metabolism, proliferation). These functions are thought to be kept in check by Mxd/Mnt, bHLH-LZ family proteins that heterodimerize with Max, bind E-boxes and antagonize Myc transcriptional activity. Our study suggests that this negative regulation may be an ancestral feature of the Myc network that was lost in the lineage leading to *M. brevicollis*. In turn, this could suggest that the regulation of critical cellular processes like apoptosis, growth, and proliferation that are under the auspices of bilaterian Myc proteins are not part of MbMyc function or perhaps that the unicellular lifestyle of *M. brevicollis* presents a more relaxed environment in which MbMyc can operate. Alternatively, MbMyc may be controlled by as-of-yet unidentified negative regulators. In this regard, it is intriguing that MbMax appears to homodimerize and bind E-box DNA more efficiently than does HsMax (fig. 4). Earlier work had shown that highly overexpressed mammalian Max is capable of abrogating Myc transcriptional activity (Kretzner et al. 1992). Perhaps MbMax homodimerization may function as a primitive Mxd analog by competing with MbMyc–MbMax heterodimers for binding to E-box DNA. Like metazoan Myc proteins, MbMyc may also regulate nontranscriptional processes, such as DNA replication or α-tubulin acetylation (Domínguez-Sola et al. 2007; Conacci-Sorrell et al. 2010). Interestingly, the presence of MbMyc in the choanoflagellate flagellum, a microtubule-based structure, raises the possibility that MbMyc may interact with tubulin monomers and/or microtubules as does c-Myc (Alexandrova et al. 1995; Conacci-Sorrell et al. 2010).

Myc proteins in bilaterians regulate a remarkable diversity of cellular processes, many of which are fundamental to cellular physiology. A thorough investigation of the function of Myc proteins in early metazoans and their unicellular relatives will be necessary to determine whether the breadth of early Myc functions was similarly grand or perhaps more limited in scope. The large repertoire of bilaterian genes regulated by Myc may have arisen from a simpler form of the network in unicellular organisms. A previous survey of the *M. brevicollis* genome predicted the presence of Myc-binding sites upstream of ribosome biogenesis genes, suggesting that MbMyc might regulate their transcription (Brown et al. 2008). Understanding how the Myc regulatory network influences the biology of *M. brevicollis* and *C. owczarzaki* will require a complete characterization of their protein-binding partners and DNA sequence targets. In addition to defining the ancestral functions of this critical regulatory network, detailed functional analyses of MbMyc may help illuminate the largely unexplored mechanisms of gene regulation in the unicellular progenitors of metazoans.

**Supplementary Material**

Supplementary materials and figures S1–S5 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

**Acknowledgments**

We are grateful to Scott Nichols, Michelle Ulrich, and Celaine Ngouenet for technical assistance and Jody Westbrook for helpful comments on the manuscript. This work was...
supported by funds from the American Cancer Society, the Gordon and Betty Moore Foundation Marine Microbiology Initiative, and the Pew Scholars Program to N.K. and NIH/NCI grant R37CA57138 to R.N.E. N.K. is a Scholar in the Canadian Institute for Advanced Research.

References


Spotts GDV, Patel SV, Xiao Q, Hann SR. 1997. Identification of downstream-initiated c-Myc proteins which are dominant...