© 2017. Published by The Company of Biologists Ltd | Development (2017) 144, 292-304 doi:10.1242/dev.139220

RESEARCH ARTICLE

High-throughput analysis reveals novel maternal germline RNAs crucial for primordial germ cell preservation and proper migration

Dawn A. Owens[1], Amanda M. Butler[1], Tristan H. Aguero[1], Karen M. Newman[1], Derek Van Booven[2] and Mary Lou King[1,*]

ABSTRACT

During oogenesis, hundreds of maternal RNAs are selectively localized to the animal or vegetal pole, including determinants of somatic and germline fates. Although microarray analysis has identified localized determinants, it is not comprehensive and is limited to known transcripts. Here, we utilized high-throughput RNA-sequencing analysis to comprehensively interrogate animal and vegetal pole RNAs in the fully grown *Xenopus laevis* oocyte. We identified 411 (198 annotated) and 27 (15 annotated) enriched mRNAs at the vegetal and animal pole, respectively. Ninety were novel mRNAs over 4-fold enriched at the vegetal pole and six were over 10-fold enriched at the animal pole. Unlike mRNAs, microRNAs were not asymmetrically distributed. Whole-mount *in situ* hybridization confirmed that all 17 selected mRNAs were localized. Biological function and network analysis of vegetally enriched transcripts identified protein-modifying enzymes, receptors, ligands, RNA-binding proteins, transcription factors and co-factors with five defining hubs linking 47 genes in a network. Initial functional studies of maternal vegetally localized mRNAs show that *sox7* plays a novel and important role in primordial germ cell (PGC) development and that *ephrinB1* (*efnb1*) is required for proper PGC migration. We propose potential pathways operating at the vegetal pole that highlight where future investigations might be most fruitful.

KEY WORDS: *Xenopus*, RNA-seq, PGCs, Localized RNAs, Gene pathways, Scaffold

INTRODUCTION

For many organisms, oogenesis is a protracted affair during which RNAs and proteins are synthesized and stored for later use during embryogenesis. Patterning of the early *Xenopus* embryo is determined by these components and includes specification of the three axes (animal/vegetal, dorsal/ventral, left/right), the three primary germ layers, and the germ cell lineage (King, 2014). The animal/vegetal (A/V) axis is the first to be established and specifies where the three primary germ layers will arise in the embryo. Visible signs of A/V polarity are obvious in the stage I oocyte as the Balbiani body (BB) (or mitochondrial cloud) forms in close association with the nucleus and faces the future vegetal pole. The BB contains the maternal stockpile of mitochondria as well as the germline determinants embedded within germ plasm. Later in oogenesis, the BB components accumulate at the vegetal pole, becoming tightly associated with the subcortical region.

The identity of the maternal RNAs and proteins that participate in embryonic patterning, and thus normal development, are of great interest. Initial screens selecting mRNAs enriched at either pole identified both somatic determinants, such as *vg1*, *vegT* and *wnt11*, and germ cell determinants including *nanos1*, *deadsouth*, *xdazl* and *xpat* (Mowry, 1996; King, 2014; Aguero et al., 2016). These mRNAs defined two patterns of RNA localization during oogenesis that appeared to align with their embryonic functions: BB-localized RNAs that function in germline identity (early pathway); and RNAs that are uniformly distributed in stage I but vegetally localized during stages II-IV (late pathway) and function in somatic patterning. However, as the number of known localized RNAs increased, some were found that used both the early and late pathways (*hermes*, *fatvg*), or used the late pathway but, after fertilization, were found only in the germ plasm (*dead-end*). Loss-of-function studies revealed that these RNAs indeed have multiple functions important to both somatic and germ cell lineages (Houston, 2013).

How the asymmetric distribution of maternal RNA controls embryonic patterning represents a key area of research in developmental biology. Recent microarray data using cortical RNAs as probes have identified several hundred transcripts at the vegetal pole and many fewer localized at the animal pole (Cuykendall and Houston, 2010). Although microarrays have identified transcripts localized to the vegetal cortex and to germ plasm, this type of analysis is limited in sensitivity and to known transcripts. A comprehensive analysis identifying the RNAs, both coding and non-coding, that are significantly enriched at either the animal or vegetal pole is an important first step towards understanding the maternal contribution to embryonic patterning.

In the present study, we utilized high-throughput RNA-sequencing (RNA-seq) analysis to interrogate both animal and vegetal pole localized RNAs in the fully grown oocyte. We identified 411 vegetally localized mRNAs and, of those, 198 are previously identified genes currently in the Xenbase database (Karpinka et al., 2015). Analysis of vegetally enriched transcripts identified receptors, ligands, RNA binding proteins, protein modifying enzymes and transcription factors, as well as defined gene hubs. Functional analysis of key genes confirmed their roles in primordial germ cell (PGC) development. We also identified eight microRNAs (miRNAs), all uniformly distributed, suggesting that early embryonic patterning is not regulated by localized maternal miRNAs but rather their localized mRNA targets. Analysis of non-coding RNAs must await further annotation of the *Xenopus tropicalis* or *laevis* genome. Here, we present a comprehensive analysis of identified RNAs found enriched at either the animal or vegetal pole. Our findings strongly support the vegetal pole as a major signaling center that patterns the early embryo.

[1]Department of Cell Biology, University of Miami Miller School of Medicine, 1011 NW 15th St, Miami, FL 33136, USA. [2]The Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, 1011 NW 15th St, Miami, FL 33136, USA.

*Author for correspondence (mking@med.miami.edu)

M.L.K., 0000-0002-1923-4283

Received 28 April 2016; Accepted 25 November 2016

Development (2017) 144, 292-304 doi:10.1242/dev.139220

RESULTS

RNA-seq analysis of vegetal and animal poles

To identify transcripts localized at either the vegetal or animal pole, RNA was isolated from the respective poles (each comprising ~10-20% of total oocyte) of stage VI *X. laevis* oocytes and subjected to RNA-seq analysis. A total of six samples, comprising three vegetal and three oocyte-matched animal poles, were included in the analysis as described in Materials and Methods. The total number of reads for all three samples of vegetal and matched animal poles were virtually identical, revealing sample precision (Fig. 1A). The reads were aligned to the version 7.1 *X. laevis* genome (Xenbase.org) and principal component analysis (PCA) was performed on the normalized results. Two-dimensional PCA showed that transcripts identified in the vegetal pole samples cluster together and away from the animal pole samples, which also cluster together (Fig. 1B). The identified transcripts having an FDR<0.05 and an FPKM≥5 were used to generate a scatter plot (Fig. 1C). The data support and extend previous analyses that show both a greater complexity and fold enrichment of RNAs at the vegetal pole in comparison to the animal pole (Cuykendall and Houston, 2010; De Domenico et al., 2015). As expected, mRNAs known to be localized at the vegetal pole (*nanos1*, *dazl*, *ddx25*/*deadsouth*) and animal pole [*dand5* (*coco*) and *slc18a2* (*vmat2*)] were identified as well as novel mRNAs. Over five thousand transcripts (5717) were found differentially expressed between the animal and vegetal poles based on a minimum q-value of 0.05 and 2-fold change.

Transcripts were considered significantly enriched at the vegetal pole if they had at least a 4-fold increase compared with the animal pole. The stringent criteria set yielded a total of 411 vegetally enriched transcripts, 198 of which were annotated (Fig. 1D, Table S1). Of the 198 transcripts identified, 38 have been shown to be vegetally localized, with 23 of them being specifically associated with germ plasm (Cuykendall and Houston, 2010; Claussen et al., 2015; De Domenico et al., 2015). Transcripts were considered localized at the animal pole if they were at least 10-fold enriched compared with the vegetal pole. Under these conditions, 27, including 15 annotated, mRNAs were enriched at the animal pole (Fig. 1D). All annotated and unannotated transcripts can be found in Tables S2 and S3, respectively.

Biological process and network analysis of vegetally enriched transcripts

To identify possible gene functions, the 198 annotated vegetal mRNAs were manually data mined using GeneCards (www.genecards.org). Table 1 shows the 40 most enriched transcripts in the vegetal pole. Ten categories were established based on function (Fig. 2A). The top six categories were: signal transduction (26%), transport (13%), transcription (7%), cytoskeletal related (8%), the ubiquitin pathway (7%) and cell cycle (7%). Enzymes are often key players in regulating gene pathways; therefore, we also identified and categorized the enzymes represented in our vegetally enriched data set. Enzymes represent 60/198 (30%) of localized transcripts. Nine categories of enzymes were identified, with kinases (18%),

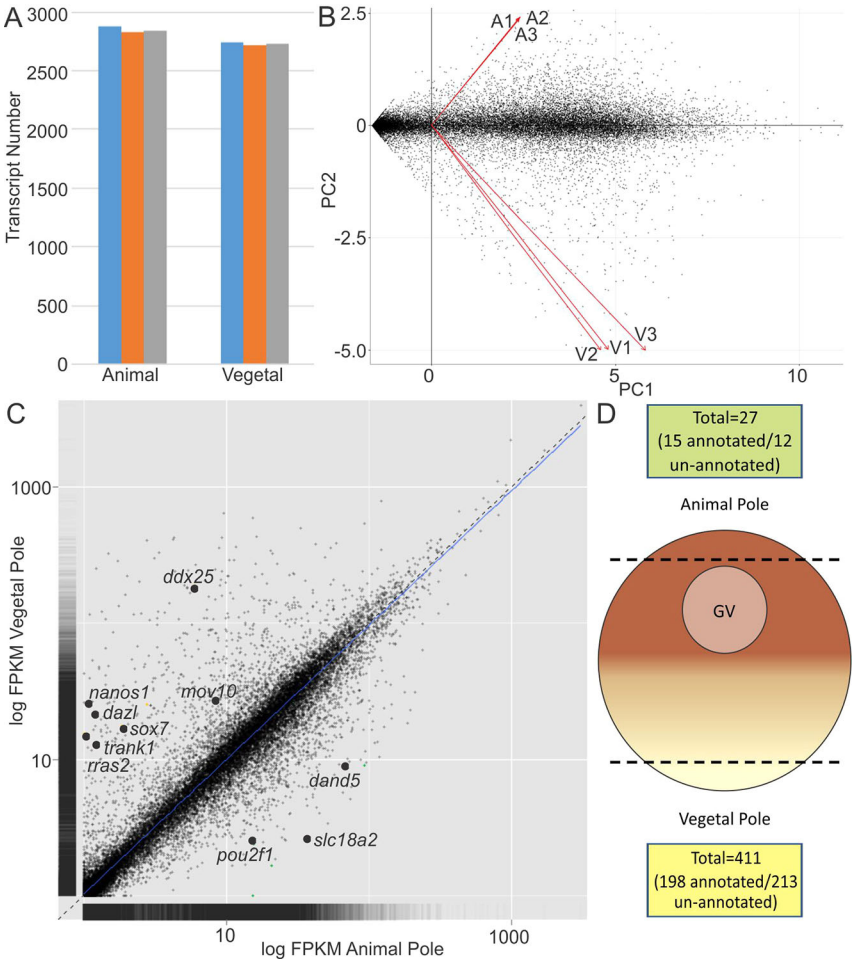

Fig. 1. RNA-seq analysis of vegetal versus animal pole transcripts in stage VI *X. laevis* oocytes. (A) Total read counts for oocyte-paired vegetal and animal pole samples. Bars of the same color represent vegetal (n=3) and animal (n=3) pole samples extracted from the same oocytes. (B) Two-dimensional principal component analysis of vegetal and animal pole transcripts. Vectors V1, V2, V3 and A1, A2, A3 represent vegetal and animal pole samples, respectively. (C) Scatter plot comparing vegetal and animal pole transcripts. (D) Differential expression analysis of vegetal versus animal pole transcripts with FDR≤0.05 and FPKM≥5. GV, germinal vesicle.

Development (2017) 144, 292-304 doi:10.1242/dev.139220

metabolism related (17%), ubiquitin pathway (13%) and ATPases/GTPases (12%) making up the majority (Fig. 2B).

Vegetally localized mRNAs were subject to gene pathway and network analysis by GeneGo. Significantly related gene ontology (GO) processes were grouped into seven categories based on gene expression (Table 2). Consistent with a role in embryonic patterning, these categories included: developmental processes, signaling regulation, localization, phosphate metabolic processes, cellular protein metabolic processes, cell cycle, and gamete generation. Interestingly, genes involved in neurogenic processes such as neuroblast proliferation, including *fgfr2*, *frizzled1* and *ephrinB1* (*efnb1*), were well represented in our data set, composing 12% of annotated genes (Table 2).

We next investigated potential gene networks present within the 198 vegetally enriched transcripts. Using MetaCore analysis (GeneGo), we identified 47 genes that form a direct interaction network (Fig. 2C). These genes encode protein-modifying enzymes [*caspase 3* (*casp3*), *cathepsin C* (*ctsc*), *pcsk6/pace4*, *tesk1/2*, *senp1*], receptors [*fgfr2*, *frizzled1* (*fzd1*), *a2mr/lrp1*], ligands (*wnt11*), and five key transcription factors or co-factors (*e2f1*, *irf8*, *err1/esrra*, *p300/ep300* and *sox7*), the first four of which represent network hubs. These mRNAs were validated as vegetally localized by either RT-qPCR (Fig. 2D) or WISH (Fig. 3). Published studies on these factors suggest their involvement in regulating the cell cycle (*e2f1*), endoderm specification (*sox7*), metabolic pathways (*err1*) and lineage commitment (*irf8*) (Costa et al., 2013; Johansen et al., 2016; Minderman et al., 2016; Stovall et al., 2014).

miRNA target RNAs are localized at the vegetal pole

Germ plasm RNAs must be post-transcriptionally regulated for germline survival (Lai et al., 2012; reviewed by Lai and King, 2013). RNA degradation within germ plasm may be regulated by miRNAs (Bartel, 2004; Yamaguchi et al., 2014). Therefore, we mined our data to identify vegetally localized miRNAs. Our analysis identified only eight miRNAs that were expressed in both the vegetal and animal poles: *15c*, *18a*, *19b*, *20a*, *92a*, *363*, *427* and *429*. Surprisingly, none was significantly enriched at either pole (data not shown).

We next determined if the predicted mRNA targets of the eight identified miRNAs were vegetally localized. Interestingly, predicted targets of 7/8 miRNAs are enriched in the vegetal pole (Table 3). In total, 13 vegetally localized target mRNAs (listed in Table 3) were identified that contain at least one recognition sequence conserved between *X. tropicalis* and human for their respective miRNAs. These results suggest that if early embryonic patterning is regulated by miRNA activity, it is not by localizing miRNAs to the vegetal pole but rather by targeting specific vegetally localized RNAs.

Expression of vegetally localized RNAs during development

We chose 17 transcripts (*xpat*, *efnb1*, *rras2*, *mov10*, *otx1*, *sox7*, *spire1*, *wnk2*, *e2f1*, *sybu*, *atrx*, *hook2*, *tob2*, *rnf38*, *trank1*, *wwtr1* and *parn*) for WISH analysis to determine their expression pattern during embryogenesis (Fig. 3, Fig. S1A,B). Consistent with our RNA-seq data, 15/17 were expressed exclusively in the vegetal pole of stage IV oocytes (Fig. 3, Fig. S1A). Not surprisingly, the two low expressing mRNAs, *parn* and *wwtr1* (only 5- to 6.7-fold enriched compared with the animal pole) were not detected in stage IV oocytes but were detected later at blastula stage (Fig. S1B). Recently, another group analyzed transcript localization in the 8-cell *X. tropicalis* embryo by RNA-seq (De Domenico et al., 2015). Comparison of the vegetal/animal blastomeres revealed that 27 of our filtered 198 transcripts remain vegetally enriched after both fertilization and cortical rotation have occurred (De Domenico et al., 2015).

Spatial expression patterns were examined during oogenesis, the pre-midblastula transition (MBT), gastrula, neurula, and tailbud stages to determine whether localized mRNAs contribute to the future germline, soma, or both lineages. During oogenesis, three mRNA localization patterns were detected: the early or METRO pathway (Fig. 3A), the late Vg1-like pathway [Fig. 3B, Fig. S1A (*atrx*, *hook2*, *tob2*), Fig. S1B], or both [Fig. 3C, Fig. S1A (*rnf38*)] (King et al., 1999). Germ plasm-specific *xpat* served as a marker for germline expression (Hudson and Woodland, 1998). Regardless of the pathway used, all mRNAs were subsequently found in the germ plasm of embryos as well as in the soma (Fig. 3, Fig. S1A,B). 12/17 RNAs represented novel germline components, while *xpat*, *sybu*, *otx1*, *tob2* and *efnb1* were confirmed as previously described (Cuykendall and Houston, 2010; De Domenico et al., 2015).

Except for *sox7* and *efnb1*, germ plasm expression persisted through neurula (Fig. 3, Fig. S1A,B). During gastrulation, the germline segregates from endoderm and PGCs form a distinct lineage. By neurula, PGCs are transcriptionally active for the first time (Venkatarama et al., 2010). *sox7* is expressed through gastrula stages but is lost by neurula suggesting an early role in PGC development. *efnb1* is not expressed at gastrula but is re-expressed by neurula, probably as part of a new gene expression program in PGCs. PGC migration towards the dorsal mesentery and organogenesis occur at the early tailbud stage. Only 41% (7/17) of the vegetal transcripts, including *xpat*, remained expressed in PGCs during tailbud stage (Fig. 3). These transcripts included transcription factors (*otx1*, *e2f1*), RISC factor (*mov10*), actin regulator (*spire1*), Ser/Thr kinase (*wnk2*), and an adaptor protein that binds kinesin (*sybu*). They are likely to be zygotic transcripts and might play roles in migration and/or in preserving PGC totipotency.

In addition to PGC expression, 71% (12/17) are also expressed in the eye anlage and the future posterior region in neurula, including the somitogenic mesoderm (Fig. 3, Fig. S1A,B). In tailbud stages the most notable expression pattern was in neural regions including the eye, cranial ganglia, neural tube, nasal placodes, brain, otic vesicle and the intersegmental region between the somites. Vegetally enriched transcripts that have previously been shown to be involved in neural pathways included *efnb1*, *sybu*, *wnk2* and *otx1* (Colozza and De Robertis, 2014; Bovolenta et al., 2006; Rinehart et al., 2011; Zhang et al., 2015). Taken together, these data suggest that genes enriched at the vegetal pole of *Xenopus* oocytes contribute to both germline and neural specification during development.

Novel mRNAs enriched at the animal pole

We identified 15 mRNAs, six not previously reported, that were at least ten-fold enriched at the animal pole (Table 4). These mRNAs represent the following functional categories: signaling [*dand5*, *ifrd2*, *slc18a2*, *spata13*, *acaca*, *tmem192*, *ssr1*, *prr11*], gene expression (*pou2f1*), cell division (*rmdn3*) and metabolism (*adpgk*, *prrg4*, *prkag1*). Two mRNAs previously shown to be enriched at the animal pole were chosen for validation by WISH: *slc18a2* and *dand5* (Fig. S1C). *dand5* is a TGFβ and Wnt antagonist (Eimon and Harland, 2001; Bates et al., 2013). Slc18a2 transports monoamines into secretory vesicles for eventual exocytosis (Nikishin et al., 2012). As expected, both animal pole transcripts were expressed exclusively in the animal pole of pre-MBT embryos (Fig. S1C). Animal pole transcripts are expressed primarily in the

Development (2017) 144, 292-304 doi:10.1242/dev.139220

Table 1. Gene function of the 40 most highly enriched mRNAs identified at the oocyte vegetal pole

Gene ID	Gene name	JGI	Unigene	Fold change (V/A)	Protein function or functional feature
*grip2**	Glutamate receptor interacting protein 2	Xelaev16034487m.g	XI.64243	656.02	Scaffold protein with PDZ domain
nanos1	Nanos homolog 1	Xelaev16026544m.g	XI.1145	391.89	RNA binding
rras2‡	Related RAS viral (r-ras) oncogene homolog 2	Xelaev16056461m.g	XI.47312	360.11	Ras-related GTPase
*kif13b**	Kinesin family member 13B	Xelaev16027627m.g	XI.55605	267.80	Kinesin-like protein
pat	PGC-associated transcript protein	Xelaev16012806m.g	XI.38	226.17	Germ plasm formation, positioning and maintenance
dazl	Deleted in azoospermia-like	Xelaev16047101m.g	XI.311	217.44	RNA binding
*sybu**	Syntabulin (syntaxin-interacting)	Xelaev16066812m.g	XI.8441	175.30	Nucleolar GTPase/ATPase p130
rtn3-a	Reticulon 3	Xelaev16042166m.g	XI.57382	162.84	Reticulon
trank1‡	Tetratricopeptide repeat and ankyrin repeat containing 1	Xelaev16058483m.g	NA	154.58	Translocase of outer mitochondrial membrane complex
germes	RNA binding protein with multiple splicing	Xelaev16035521m.g	XI.17472	141.43	Nucleic acid binding
wnt11b	Wingless-type MMTV integration site family, member 11B	Xelaev16048139m.g	XI.24008	139.89	Secreted growth factor/developmental regulator
*trim36**	Tripartite motif containing 36	Xelaev16048562m.g	XI.6926	127.93	Zinc finger, B-box
farp2	FERM, RhoGEF and pleckstrin domain protein 2	Xelaev16049032m.g	XI.53782	111.93	Rho guanine nucleotide exchange factor that activates RAC1
cnppd1	Cyclin Pas1/PHO80 domain containing 1	Xelaev16027858m.g	XI.7190	93.24	Cyclin
slain1‡	SLAIN motif family, member 1	Xelaev16028782m.g	NA	89.96	Microtubule plus-end tracking protein
srgap1‡	SLIT-ROBO Rho GTPase activating protein 1	Xelaev16070447m.g	NA	78.71	CDC42- and RhoA-interacting protein CIP4
ddx25	DEAD (Asp-Glu-Ala-Asp) box helicase 25	Xelaev16075440m.g	XI.670	75.17	ATP-dependent RNA helicase
pcsk6	Proprotein convertase subtilisin/kexin type 6	Xelaev16063723m.g	XI.48635	71.99	protease, proprotein convertase
gk	Glycerol kinase	Xelaev16073000m.g	XI.17142	71.20	Ribulose kinase and related carbohydrate kinases
acsl1	Acyl-CoA synthetase long-chain family member 1 protein	Xelaev16052290m.g	XI.15591	52.76	Fatty acid CoA ligase
slc12a9	Solute carrier family 12, member 9	Xelaev16065485m.g	XI.59798	52.58	Amino acid transporter
wnk2	WNK lysine deficient protein kinase 2	Xelaev16065595m.g	XI.50299	51.69	Serine/threonine protein kinase
*vegt**	Vegt protein	Xelaev16018006m.g	XI.1775	51.34	T-box transcription factor, meso-endodermal determinant
cdr2l‡	Cerebellar degeneration-related protein 2-like	Xelaev16042542m.g	XI.9895	50.48	Myosin class II heavy chain
tesk2‡	Testis-specific kinase 2	Xelaev16001025m.g	XI.71668	46.74	Endosomal membrane proteins, serine/threonine protein kinase
cdk5r2	Cyclin-dependent kinase 5, regulatory subunit 2 (p39)	Xelaev16007720m.g	XI.31022	46.60	CDK5 kinase activator
otx1	Orthodenticle homeobox 1	Xelaev16005310m.g	XI.781	46.51	Homeodomain transcription factor
rhobtb2‡	Rho-related BTB domain containing 2	Xelaev16065046m.g	XI.12379	45.82	Ras-related small GTPase
tmem65	Transmembrane protein 65	Xelaev16028617m.g	XI.55768	45.36	Uncharacterized conserved protein
mov10	Mov10 RISC complex RNA helicase	Xelaev16074505m.g	XI.52320	45.22	RNA helicase
trib1	Tribbles pseudokinase 1	Xelaev16002738m.g	XI.75411	44.66	Serine/threonine protein kinase
ppp1r2	Protein phosphatase 1, regulatory (inhibitor) subunit 2	Xelaev16041705m.g	XI.52180	44.20	Protein phosphatase 1, regulatory (inhibitor) subunit PPP1R2
sox7	SRY (sex determining region Y)-box 7	Xelaev16000442m.g	XI.1241	42.25	HMG-box transcription factor
ldlrap1-b	Low density lipoprotein receptor adaptor protein 1	Xelaev16018841m.g	XI.8355	40.88	Adaptor protein of polarized cells, contains PTB domain
paqr6‡	Progestin and adipoQ receptor family member VI	Xelaev16007491m.g	XI.57243	39.20	Predicted membrane proteins, contain hemolysin III domain
sf3a1‡	Splicing factor 3a, subunit 1, 120 kDa	Xelaev16054742m.g	XI.34155	37.67	Required for pre-mRNA splicing
dnd1	DND microRNA-mediated repression inhibitor 1	Xelaev16016713m.g	XI.29785	37.50	Nuclear ribonucleoprotein R (RRM superfamily)

Continued

Development (2017) 144, 292-304 doi:10.1242/dev.139220

Table 1. Continued

Gene ID	Gene name	JGI	Unigene	Fold change (V/A)	Protein function or functional feature
fam168b	Family with sequence similarity 168, member B	Xelaev16019587m.g	XI.3240	36.47	RNA polymerase II, large subunit
hook2	Hook microtubule-tethering protein 2	Xelaev16057970m.g	XI.9790	34.05	Uncharacterized coiled-coil protein
bicc1	BicC family RNA binding protein 1	Xelaev16022482m.g	XI.641	25.70	Putative RNA-binding protein; acts as a negative regulator of Wnt signaling; and might be involved in regulating gene expression during embryonic development

GeneCards was used to indicate gene functions. ‡Nine novel mRNAs enriched at the vegetal pole and associated with germ plasm are identified. *The highest expressing isoform was used when multiple isoforms were identified. V/A, vegetal/animal.

neural ectoderm at later developmental stages, as previously described (Fig. S1C) (Grant et al., 2014). MetaCore (GeneGo) direct interaction pathway analysis did not reveal integrated networks among the animal pole-enriched RNAs.

Overexpression of *e2f1*, *otx1*, *parn*, *rras2* and *wwrt1* significantly reduces PGC number

As a first step towards functional analysis, six vegetally enriched transcripts expressed in PGCs were selected for overexpression studies. One-cell embryos were injected in the vegetal region with *in vitro* synthesized mRNA of selected transcripts or GFP. Tailbud embryos were collected and the number of PGCs per embryo was calculated and compared with GFP-injected controls. Overexpression of the transcription factors *e2f1* and *otx1*, the transcriptional co-activator *wwtr1*, a Ras-like small GTPase *rras2*, and the poly(A)-specific ribonuclease *parn* significantly reduced PGC number, whereas *spire1* had no effect (Fig. 4). Importantly, aside from the effects on PGC number, embryos appeared normal. Taken together, these results suggest a specific role in PGC development for *otx1*, *e2f1*, *wwtr1*, *rras2* and *parn*.

Embryos depleted of *otx1* and *wwtr1* are deficient in PGCs

To further test the function of *otx1* and *wwtr1*, we created loss-of-function morphants by injection of antisense morpholinos (MOs) into one-cell embryos. Both *otx1*-MO and *wwtr1*-MO blocked the translation of their respective proteins in a dose-dependent fashion

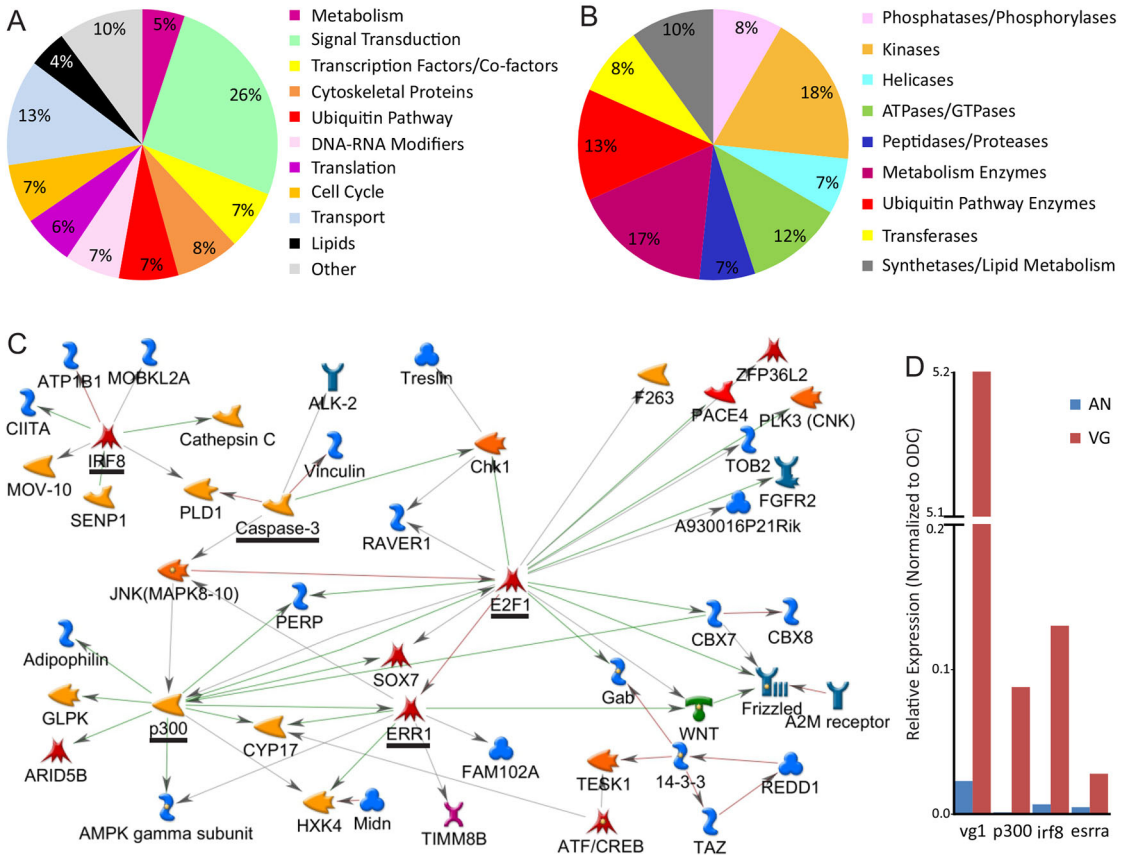

Fig. 2. Biological process and network analysis of vegetally enriched transcripts. (A) The 198 annotated transcripts upregulated in the vegetal pole categorized according to their biological function (GeneCards). (B) Sixty vegetally enriched transcripts with enzymatic activity grouped based on enzymatic function. (C) GeneGo pathway analysis using the direct interaction algorithm of vegetally enriched transcripts. (D) RT-qPCR verification of pathway hubs *p300*, *irf8* and *err1*/*esrra* at animal (An) versus vegetal (Vg) poles. *vg1*, control for vegetal pole.

Development (2017) 144, 292-304 doi:10.1242/dev.139220

Table 2. GO processes of vegetally enriched transcripts

GeneGo process	*P*-value	FDR	No. of genes
Developmental process	3.42E-05	3.28E-03	**78**
Cell development	2.58E-04	9.42E-03	32
Neurotrophin signaling pathway	2.38E-03	3.71E-02	9
Neuroblast proliferation	3.23E−03	4.56E−02	3
Regulation of signaling	9.13E−06	1.61E−03	**50**
Intracellular signal transduction	3.02E−08	9.56E−05	40
Transmembrane receptor protein tyrosine kinase signaling pathway	1.09E−04	6.11E−03	17
Signal transduction by p53 class mediator	1.83E−03	3.14E−02	6
Regulation of localization	2.80E−06	9.85E−04	**42**
Regulation of protein localization	4.15E−06	1.31E−03	19
Regulation of intracellular transport	7.49E−06	1.48E−03	16
Regulation of phosphate metabolic process	1.37E−05	2.18E−03	**38**
Regulation of phosphorylation	6.16E−05	4.55E−03	28
Regulation of protein serine/threonine kinase activity	2.07E−04	8.71E−03	14
Regulation of cellular protein metabolic process	5.22E−07	4.08E−04	**38**
Lipid transport	7.32E−05	5.15E−03	10
Positive regulation of deacetylase activity	8.68E−05	5.39E−03	3
Regulation of cell cycle	7.08E−06	1.48E−03	**24**
Negative regulation of cell cycle	1.61E−06	6.37E−04	14
Negative regulation of nuclear division	2.88E−05	3.06E−03	6
Cell cycle checkpoint	2.08E−03	3.37E−02	8
Gamete generation	6.10E−05	4.55E−03	**19**
Sexual reproduction	1.04E−04	6.00E−03	21
Female gamete generation	2.73E−04	9.75E−03	7
Spermatogenesis	1.10E−03	2.35E−02	14
Multicellular organismal reproductive process	1.70E−04	7.97E−03	22

All 198 vegetally localized genes were subject to GeneGo analysis. GO processes are categorized based on the greatest number of enriched transcripts identified. Seven primary categories are in bold.

(Fig. S2B). Injected embryos were collected at tailbud stages and the number of PGCs per embryo was calculated and compared with the control. Inhibition of *otx1* and *wwtr1* significantly increased PGC number (Fig. 5).

Misexpression of *sox7* reduces PGC number

The transcription factor *sox7* has been shown to play various roles in embryonic development, including proliferation, differentiation, hematopoiesis, cardiogenesis and vasculogenesis (Stovall et al., 2014). However, its role in PGCs is unknown. *sox7* expression is significantly upregulated in the vegetal compared with the animal pole in stage VI oocytes (Fig. 1C, Table 1) and its expression persisted in PGCs during gastrulation (Fig. 3B). To assess the role of *sox7* in PGC development, fertilized embryos were injected vegetally with either a *sox7*-targeted MO (Fig. S3) or the dominant-negative transcript *sox7*dCEnR (Fig. 6). Overexpression and rescue experiments were performed using *X. tropicalis sox7* (Xt*sox7*) mRNA (Zhang et al., 2005a). Initial injections of *sox7*dCEnR or Xt*sox7* mRNAs were performed at various concentrations to determine an effective dose that would not cause the phenotypic alterations observed by Zhang et al. (2005a) (data not shown). No notable changes in morphology were observed in embryos injected with *sox7*-MO, *sox7*dCEnR (200 pg) and/or Xt*sox7* (200 pg) mRNA (Fig. 6A-D, Fig. S3C,D). The number of PGCs per embryo was calculated and compared with uninjected controls. Both dominant-negative and MO-mediated inhibition and the overexpression of Xt*sox7* significantly reduced the number of PGCs in tailbud embryos (Fig. 6, Fig. S3). Expression of both *sox7*dCEnR and Xt*sox7* mRNAs together significantly rescued the effect that *sox7*dCEnR had on PGC number, presumably because Xt*sox7* restored function (Fig. 6D,E). These data suggest that the level of *sox7* expression must be tightly regulated for proper PGC development.

efnb1 plays an essential role in PGC specification and migration

Eph receptor tyrosine kinases and their ligands, ephrins, have been shown to be involved in the formation of tissue boundaries, including separation of the germ layers, by regulating migration, adhesion and repulsion during embryonic development (Rohani et al., 2014). However, their specific role(s) in germline development is unknown. Interestingly, *efnb1* expression is upregulated in vegetal versus animal poles of stage VI oocytes. *efnb1* is also expressed in PGCs at neurula (Fig. 3A), suggesting a role in this lineage. To assess the role of *efnb1* in PGC development, *efnb1* was overexpressed by injecting flag-tagged *efnb1* (*efnb1*-FL) mRNA into the vegetal region of fertilized embryos. No notable changes in morphology or PGC location were observed (Fig. 7A). However, overexpression of *efnb1* significantly reduced the number of PGCs compared with control (Fig. 7A). This effect is rescued by co-injection with *efnb1*-MO (Fig. 7A). These data suggest that PGC number is not maintained within an environment of excess Efnb1 protein.

We next assessed the effect of *efnb1* inhibition using an *efnb1*-MO as described (Moore et al., 2004). No notable changes in morphology or PGC number were observed in embryos injected with *efnb1*-MO compared with scrambled-MO or uninjected controls (Fig. 7B). *efnb1* inhibition significantly increased the total number of embryos containing mislocalized PGCs according to the boundaries designated for normal PGC localization, between somites 5 and 11, as described by Tarbashevich et al. (2011). Mislocalization was along the anterior/posterior (A/P) axis, primarily beyond the normal posterior boundary (Fig. 7B). PGC mislocalization was rescued by co-expression with an *efnb1* mRNA construct containing conservative mutations in the MO-binding region (*efnb1*-FL-rescue), rendering the MO ineffective (data not shown), confirming the specificity of the effect of *efnb1*-MO (Fig. 7B). These data suggest that *efnb1* is essential for the proper migration of PGCs.

p300 is required for normal PGC development

The transcriptional co-activator and histone acetyltransferase *p300* has been shown to be expressed during oogenesis (Kwok et al., 2006) and to regulate the metabolic state of mammalian germ cells (Boussouar et al., 2014). *p300* is enriched at the vegetal pole of stage VI oocytes (Fig. 2D) and represents one of the network hubs at the vegetal pole (Fig. 2C). We therefore examined whether p300 plays a role in *Xenopus* PGC development. Fertilized embryos were incubated with DMSO (control) or the p300 small molecule inhibitor C646 until they reached tailbud stages. Treated embryos were then collected, and the number of PGCs per embryo was calculated and compared with controls. Inhibition of p300 significantly reduced PGC number in a dose-dependent manner, suggesting a role in PGC development (Fig. 8).

DISCUSSION

Here we report the first interrogation of RNAs within the vegetal and animal poles by RNA-seq. WISH revealed all 17 selected mRNAs

Development (2017) 144, 292-304 doi:10.1242/dev.139220

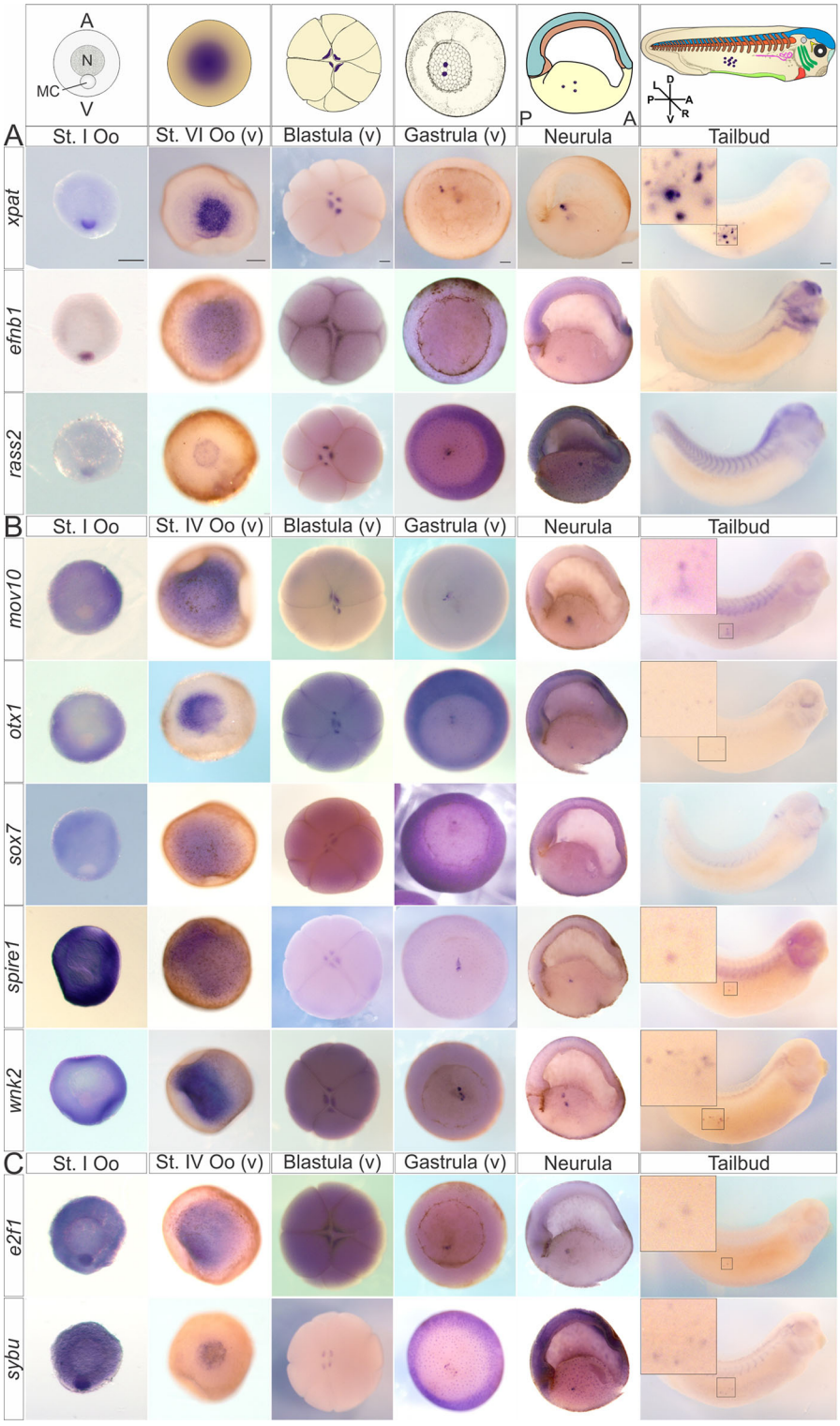

Fig. 3. WISH of vegetal pole transcripts. The expression of a subset of vegetally enriched mRNAs was analyzed during oogenesis and embryo development by WISH. Expression patterns are grouped according to mRNA pathway: early (A), late (B) or both (C). *xpat* expression marks germ plasm. Probes, developmental stages and developmental structures are indicated in the illustrations at the top: germ plasm/PGCs (purple), pronephros (pink), ventral blood islands (lime green), eye (black), lens (white), otic vesicle (gray), cranial ganglia (yellow), brachial arches (green), nasal placodes (teal), intersegmental region (brown), notochord (orange), brain and neural tube (blue). Transcripts detected in PGCs at the tailbud stage are shown at higher magnification in insets. WISH analysis was performed on ≥40 total embryos from at least two adult female frogs. Dorsal (D), ventral (V), left (L), right (R), posterior (P), anterior (A), vegetal view (v); MC, mitochondrial cloud; N, nucleus; Oo, oocyte. Scale bars: 100 µm for stage I oocyte; 200 µm in remaining panels.

to be localized, providing strong support for the accuracy of the data sets. Our findings underscore the dramatic transcript asymmetry along the A/V axis and the importance of the vegetal pole in initiating somatic and germline lineages in the early embryo. Importantly, as the annotation of the *Xenopus* genome improves, our data set can be continually mined to identify spliced variants and currently unknown transcripts.

Six important observations have emerged from our studies. (1) We identified 90 novel mRNAs that were over 4-fold enriched at the vegetal pole and six that were over 10-fold enriched at the animal pole. (2) GeneGo analysis revealed a network encompassing over 20% of the annotated vegetally enriched mRNAs, indicating great connectivity of gene function and localization. Transcription factors/co-factors *e2f1*, *irf8*, *err1* and *p300* defined four regulatory hubs for future analysis (Fig. 2C). (3) Unlike mRNAs, localization of maternal miRNAs does not appear to be a strategy employed to regulate gene expression along the A/V axis. (4) Enzymes represented 30% of the 198 enriched annotated mRNAs,

Development (2017) 144, 292-304 doi:10.1242/dev.139220

Table 3. miRNAs and their vegetally localized mRNA targets

Gene	FPKM	Fold change (V/A)	Refseq	Gene name	miRNAs						
					15c	19b	20a	363	427	429	92a
srgap1	28	57.3	NM_020762	SLIT-ROBO Rho GTPase activating protein 1			×		×	×	
rnf38	351	28.0	NM_022781	Ring finger protein 38		×	×	×	×		×
sulf1	198	26.6	NM_001128204	Sulfatase 1		×	×		×		
pgam1	420	20.6	NM_002629	Phosphoglycerate mutase 1 (brain)				×			×
fndc3a	102	18.2	NM_001079673	Fibronectin type III domain containing 3A		×	×		×		
tesk2	52	15.6	NM_007170	Testis-specific kinase 2		×					
chek1	219	14.7	NM_001114121.2	Checkpoint kinase 1	×						
tob2	346	14.1	NM_016272	Transducer of ERBB2, 2				×			×
wasl	130	13.2	NM_003941	Wiskott-Aldrich syndrome-like				×			×
papola	250	8.9	NM_032632	Poly(A) polymerase alpha			×		×		
cnot4	120	5.4	NM_001190847	CCR4-NOT transcription complex, subunit 4						×	
ube2d1	356	5.1	NM_001204880	Ubiquitin-conjugating enzyme E2D 1						×	
rbpms2	218	3.9	NM_194272	RNA-binding protein with multiple splicing 2 (Hermes)				×			×

Vegetally localized mRNA targets are shown with their corresponding non-localized miRNA(s).

underscoring the vegetal pole as a major platform for cell signaling. (5) Well over 10% of the vegetal mRNAs encode components with known functions in neurogenic pathways, including the transmembrane ligand Efnb1 and scaffold protein Grip2. We show that *efnb1* is required for proper PGC migration. (6) *sox7*, *otx1*, *e2f1*, *wwtr1*, *rras2* and *parn* are also germ plasm components and required for normal PGC development.

Biological process and network analysis

MetaCore (GeneGo) analysis placed over 20% (47/198) of the known vegetal pole mRNAs within a direct interaction network that identified four major hubs centered around transcription factors *e2f1*, *irf8*, *err1* and the histone acetyltransferase *p300* (Fig. 2C). Although the network is built based on data from different systems, it reveals novel regulatory pathways and candidates that can be tested for their functions in embryogenesis. For example, it remains unknown what restricts microtubule array formation to the vegetal pole during cortical rotation. Our RNA-seq analysis identified JNK (*mapk8*) and *slain1* as enriched at the vegetal pole and both have been implicated in microtubule dynamics (GeneCards).

e2f1 constitutes the largest hub by far, connecting 19 other localized mRNAs, including two additional hubs, *err1* and *p300*, and genes involved in cell cycle regulation, DNA replication, pluripotency/differentiation, and metabolism (Fig. 2C). Overexpression of *e2f1* resulted in a loss of PGCs (Fig. 4). This effect might be due to misregulation of the cell cycle and/or in initiating somatic differentiation. Consistent with this hypothesis, Zaragoza et al. (2010) showed that the balance of E2f transcription factors, E2f dimerization partners and C/EBPα is critical for proper cell cycle progression. E2f1 has also been shown to mediate proliferation through Wnt signaling by direct interaction with the *Fzd1* promoter (Yu et al., 2013). Furthermore, E2f1 activates Pcsk6, which allows for mesoderm induction by activation of Veg1

Table 4. Gene function of the 15 most highly enriched mRNAs identified at the oocyte animal pole

Gene ID	Gene name	v7.1 JGI	Unigene	Ratio (A/V)	Protein function or functional feature
ifrd2	Interferon-related developmental regulator 2	Xelaev16023388m.g	XI. 75341	22.84	Interferon-related protein PC4-like
slc18a2	Solute carrier family 18 (vesicular monoamine transporter), member 2	Xelaev16012438m.g	XI.13611	17.73	Transmembrane solute exchange of neurotransmitters
*spata13**	Spermatogenesis associated 13	Xelaev16016628m.g	XI.58825	16.42	Invasion-inducing protein TIAM1/CDC24 and related RhoGEF GTPases
adpgk	ADP-dependent glucokinase	Xelaev16007044m.g	NA	15.43	Catalyzes glucose to glucose-6-phosphate via ADP
*c7orf43**	Chromosome 7 open reading frame 43	Xelaev16021226m.g	XI.56042	15.20	Uncharacterized conserved protein
prrg4	Proline-rich Gla (G-carboxyglutamic acid) 4 (transmembrane)	Xelaev16039880m.g	XI.4519	14.75	Cytosolic, calcium ion binding protein
acaca	Acetyl-CoA carboxylase alpha	Xelaev16038229m.g	NA	12.36	Acetyl-CoA carboxylase
pou2f1	POU class 2 homeobox 1	Xelaev16026767m.g	XI.1265	12.28	Oct1 transcription factor
*prkag1**	Protein kinase, AMP-activated, gamma 2 non-catalytic subunit of AMPK	Xelaev16029267m.g	XI. 34576	12.00	AMPK helps modulate cellular energy metabolism
xkrx	XK, Kell blood group complex subunit-related, X-linked	Xelaev16047837m.g	NA	11.72	Uncharacterized conserved protein
tmem192	Transmembrane protein 192	Xelaev16056645m.g	XI.19528	11.71	Uncharacterized conserved protein
*ssr1**	Signal sequence receptor, alpha	Xelaev16076712m.g	XI.60983	11.48	Glycosylated endoplasmic reticulum receptor that mediates protein translocation
*prr11**	Proline-rich 11	Xelaev16035856m.g	NA	11.10	Rac1 GTPase effector FRL; cell cycle
dand5	DAN domain family member 5, BMP antagonist	Xelaev16057997m.g	XI.51884	11.04	Maternal BMP/TGFβ/Wnt inhibitor
*rmdn3**	Regulator of microtubule dynamics 3	Xelaev16028196m.g	XI.9583	10.15	Nuclear localization sequence binding; cellular calcium regulation

GeneCards was used to indicate gene functions. *Six novel mRNAs enriched at the animal pole are indicated.

Development (2017) 144, 292-304 doi:10.1242/dev.139220

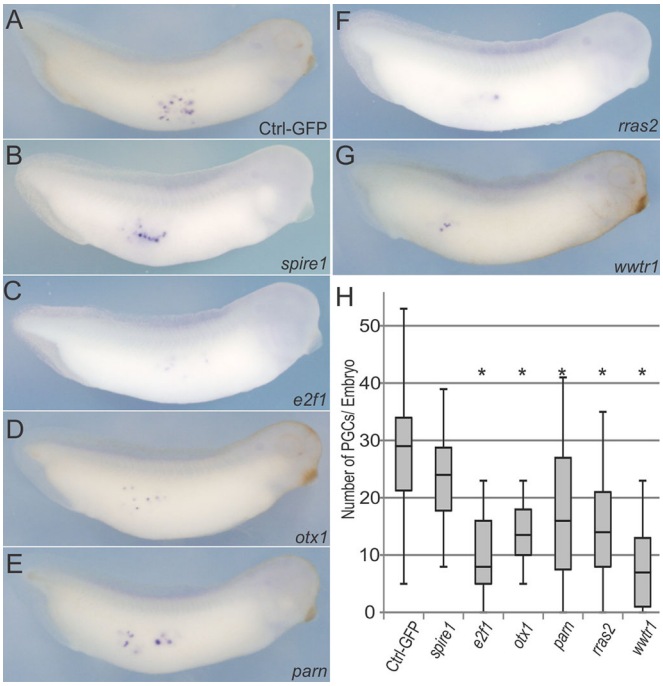

Fig. 4. Overexpression of five out of six selected vegetally enriched mRNAs reduces PGC number. One-cell embryos were injected in the vegetal region with GFP (control) or the indicated transcripts (0.5 ng). Tailbud embryos (stage 32-35) were analyzed for *xpat* expression by WISH, and representative images are shown (A-G). The number of PGCs per embryo was quantified (H). GFP, *n*=16; *spire1*, *n*=39; *e2f1*, *n*=64; *otx1*, *n*=27; *parn*, *n*=35; *rras2*, *n*=48; *wwtr1*, *n*=35. **P*<0.05, compared with GFP control. Analysis based on at least two independent experiments and shown as a box and whisker plot.

(Heasman, 2006). E2f1 also activates Cbx7, a member of the Polycomb repressor PRC1-like complex that plays a pivotal role in the transition from pluripotency to differentiation by regulating Cbx8 and Fzd1 (Klauke et al., 2013; Creppe et al., 2014; Mani et al., 2008; O'Loghlen et al., 2015). Thus, E2f1 in excess may tip the balance towards somatic differentiation in PGCs, causing their loss. Because of the dominant position that *e2f1* holds, we attempted to knockdown its activity by antisense MO injection into fertilized eggs (data not shown). Unfortunately, we could not detect a phenotype, most likely because of a pre-existing maternal supply of E2f1 protein (Peshkin et al., 2015). Future oocyte host transfer studies will investigate the function of E2f1 by depleting the maternal supply.

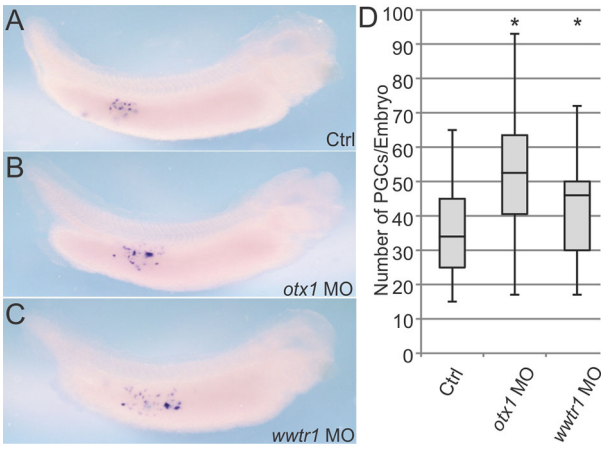

Fig. 5. MO-mediated knockdown of a subset of vegetally enriched mRNAs increases PGC number. One-cell embryos were injected in the vegetal region with MOs (15 ng) targeting *otx1* or *wwtr1*. Tailbud embryos (stage 32-35) were analyzed for *xpat*. Representative images are shown (A-C). The number of PGCs per embryo was quantified (D). Uninjected control (ctrl), *n*=32; *otx1*-MO, *n*=31; *wwtr1*-MO, *n*=25. **P*<0.05 compared with control. Analysis based on at least two independent experiments.

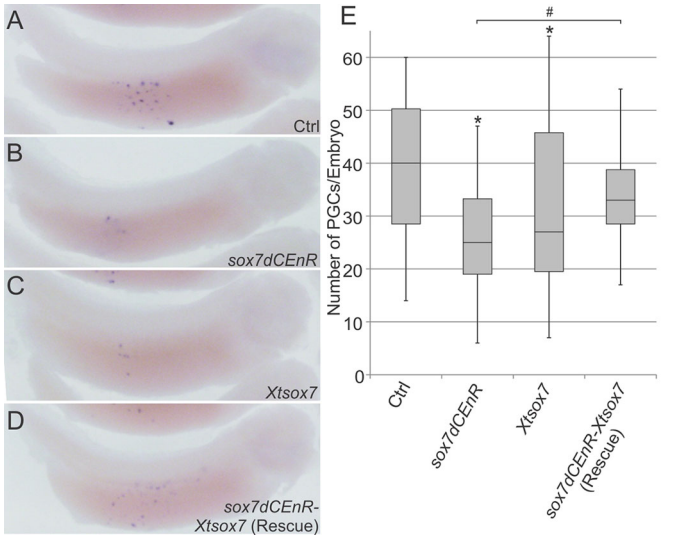

Fig. 6. Altering *sox7* expression reduces PGC number. One-cell embryos were injected in the vegetal region with *sox7*dCEnR RNA (200 pg) and *Xtsox7* RNA (200 pg), alone and in combination. Tailbud embryos were analyzed for *xpat* expression by WISH. Representative images are shown (A-D). The number of PGCs per embryo was quantified (E). Uninjected control (ctrl), *n*=28; *sox7*dCEnR, *n*=24; *Xtsox7*, *n*=27; *sox7*dCEnR and *Xtsox7*, *n*=28. **P*<0.05 compared with uninjected control. #*P*<0.05 compared with *sox7*dCEnR. Analysis based on at least two independent experiments.

Interestingly, overexpression of *rras2*, *parn*, *otx1* and *wwtr1* also resulted in PGC loss, while MO-mediated inhibition of *otx1* and *wwtr1* caused an increase in PGC number. Previous studies involving these transcripts all mention their possible roles in cell cycle regulation, proliferation and/or survival, consistent with our observations. The deadenylase Parn mediates progression through G0/G1 by regulating p53 and p21 expression (Zhang and Yan, 2015). Rras2 has been implicated in cell proliferation by regulating the PI3K (Murphy et al., 2002) and ERK pathways (Larive et al., 2012). Similarly, Otx1 regulates proliferation through the ERK/MAPK pathway, and is necessary for progression through S phase (Li et al., 2016). Wwtr1(Taz) is involved in cell cycle progression, proliferation and survival by regulating cyclin A and Ctgf expression and Casp3 activity (Wang et al., 2014). Taken together, our working hypothesis is that overexpression results in cell cycle checkpoint abnormalities and cell death, whereas loss of function releases the restrained PGC cell cycle clock resulting in more PGCs. Further investigation of these known downstream targets of *rras2*, *parn*, *otx1* and *wwtr1* in PGCs is necessary to deduce the exact mechanisms by which these genes regulate PGC number.

Sox7 is required for PGC development

Here we show a novel role for *sox7* in PGC development. *sox7* is expressed in germ plasm, and its expression persists in PGCs after segregation from the endoderm lineage (Fig. 3B). Both knockdown and overexpression of *sox7* in the fertilized embryo caused a significant decrease in the number of PGCs at the tailbud stage (Fig. 6, Fig. S3). These results strongly suggest that the level and

Development (2017) 144, 292-304 doi:10.1242/dev.139220

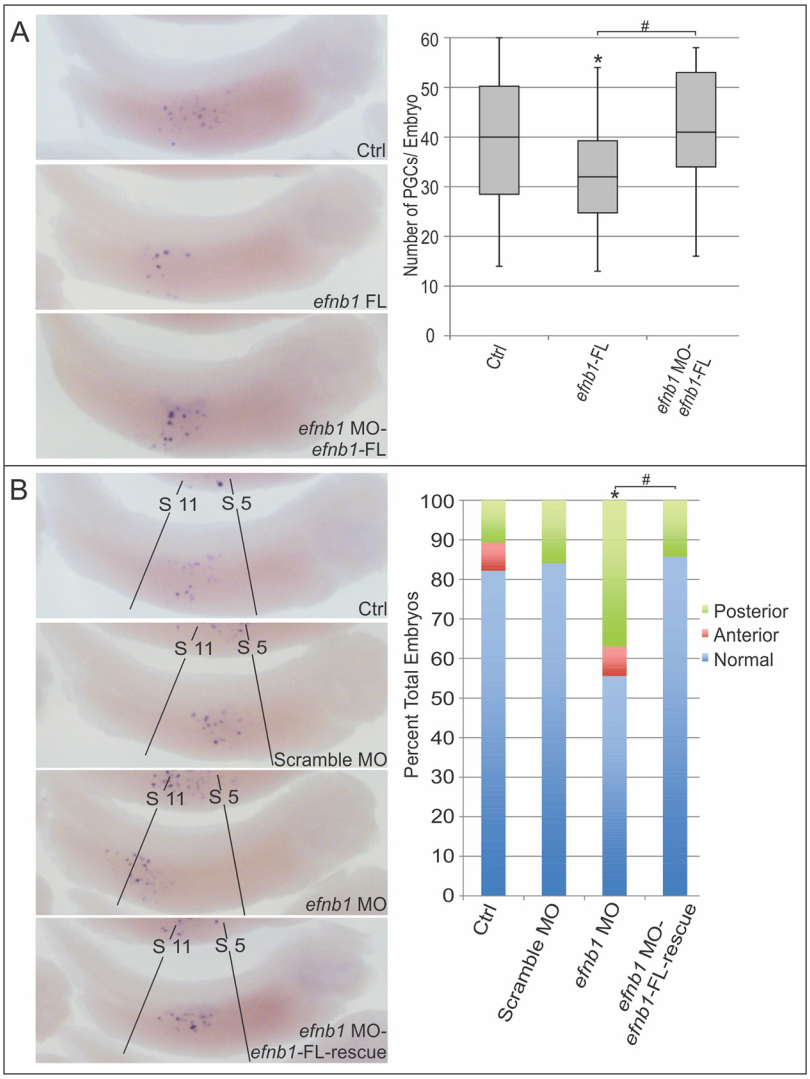

Fig. 7. *efnb1* is required for normal PGC number and migration. One-cell embryos were injected in the vegetal region with (A) *efnb1*-FL alone (200 pg), or *efnb1*-FL (200 pg) and *efnb1*-MO (16 ng); (B) scrambled-MO (16 ng), *efnb1*-MO alone (16 ng), or *efnb1*-MO (16 ng) and *efnb1*-FL-rescue (200 pg). Tailbud embryos were analyzed for *xpat* expression by WISH. Representative images are shown on the left. (A) The number of PGCs per embryo was quantified. (B) Percentage of embryos with mislocalized PGCs was calculated. Black lines indicate the boundaries of normal PGC location, between somites 5 and 11. Uninjected control (Ctrl), *n*=28; *efnb1*-FL, *n*=28; *efnb1*-FL and *efnb1*-MO, *n*=28; scrambled-MO, *n*=25; *efnb1*-MO, *n*=27; *efnb1*-MO and *efnb1*-FL-rescue, *n*=28. **P*<0.05 compared with uninjected control. $^{\#}P$<0.05 compared with *efnb1*-FL (A) or *efnb1*-MO (B). Analysis based on at least two independent experiments.

temporal regulation of *sox7* activity in PGCs are crucial to their normal development, most likely by activating the proper gene networks in the germline.

sox7 directly activates genes necessary for endoderm differentiation, including the nodal-related protein-encoding genes, and induces the expression of *endodermin* and *mixer* (Zhang et al., 2005a). Thus, overexpression of *sox7* in PGCs may cause ectopic expression of these endodermal differentiation genes, resulting in apoptosis of PGCs and ultimately a reduction in their number. Additionally, the Wnt/β-catenin signaling pathway has been reported to promote stem cell self-renewal and cell survival (reviewed by Mohammed et al., 2016), two characteristics necessary to preserve the germline. Xl*sox7* harbors the β-catenin-binding motif DRNEFDQYL (Guo et al., 2008), suggesting that *sox7* inhibition might cause reduced PGC number by influencing gene expression downstream of Wnt signaling.

Irie et al. (2015) have recently identified the F-type Sox family member *SOX17* as the primary regulator of human PGC-like fate. Interestingly, similar to the *Xenopus* F-type family member *sox7* (Zhang et al., 2005b), human *SOX17* has historically been reported as crucial for endoderm specification. F-type Sox genes regulate the expression of germline-specific genes, such as those of the Nanos family or *DND1*, and pluripotent genes such as *OCT4* (*POU5F1*) and *NANOG* (Irie et al., 2015). These recent findings, along with our observations, have led us to hypothesize that Sox7 is a key transcription factor necessary to specify PGCs in *Xenopus*. Future studies are necessary to establish when *sox7* is translated and whether it partners with another transcription factor(s) in PGCs to activate germline-specific gene expression programs.

Efnb1 is required for normal PGC development and migration

Our functional studies reveal a novel role for *efnb1* in PGC maintenance and migration (Fig. 7). *efnb1* is expressed in BBs, suggesting that it is an early component of germ plasm. After MBT, PGC-specific *efnb1* expression is lost but it is re-expressed at neurula (Fig. 2A). These observations suggested that *efnb1* might function in both the endoderm and germline lineages. MO-mediated knockdown of *efnb1* did not affect PGC number, but caused PGCs to migrate outside their normal boundaries, primarily into posterior endoderm. By contrast, *efnb1* overexpression decreased the number of PGCs but did not affect migration and development. These results suggest that *efnb1* is also involved in signaling pathways necessary for normal PGC development. Both the mismigration and loss of PGCs could be rescued, indicating specificity of the observed phenotypes (Fig. 7).

Ephrin ligands and Eph receptors are known to contribute to the maintenance of vertebrate tissue boundaries (Rohani et al., 2014), to regulate axon migration (Klein and Kania, 2014) and establish A/P

Development (2017) 144, 292-304 doi:10.1242/dev.139220

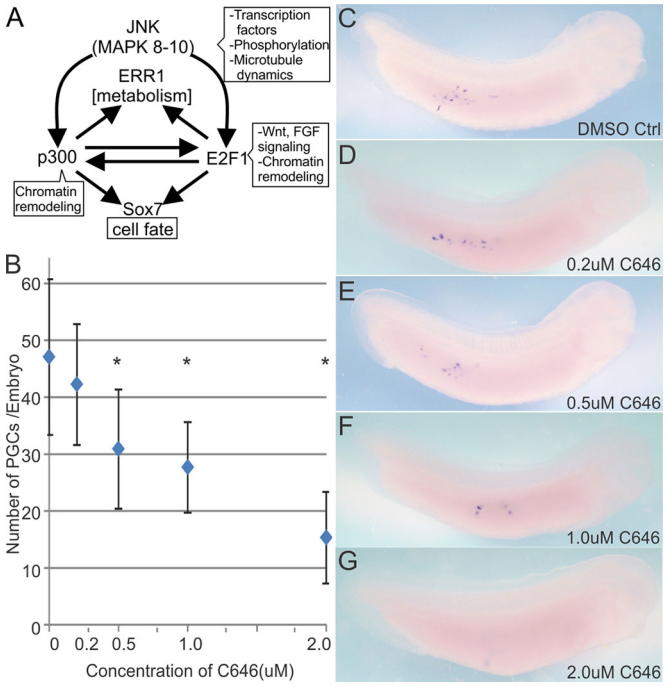

Fig. 8. Inhibition of p300 reduces PGC number. (A) Schematic of GeneGo-predicted p300 regulatory interconnections from our data set. (B-G) One-cell embryos were incubated with ≤0.04% DMSO (ctrl, *n*=31) or increasing concentrations of the p300 inhibitor C646: 0.2 µM, *n*=38; 0.5 µM, *n*=34; 1 µM, *n*=34; 2 µM, *n*=34. Tailbud embryos were analyzed for *xpat* expression by WISH. The average number of PGCs per embryo was quantified (B). **P*<0.05 compared with uninjected control.

gradients required for proper cell migration (Bush and Soriano, 2010). Therefore, *efnb1* knockdown-mediated PGC mismigration may be due to disruption of the A/P axis. Consistent with this interpretation, PGCs were not ectopically found outside of the endoderm nor was PGC number affected. Interestingly, Enfb1 associates directly with Dishevelled and is capable of recruiting it and Grip2 to the plasma membrane (Brückner et al., 1999; Moore et al., 2004; Lee et al., 2006). Grip2 encodes a scaffolding protein known to interact with receptors including Frizzled1 (Korkut et al., 2009; Ataman et al., 2006). Knockdown of Grip2 in the embryo disturbs normal PGC migration (Kirilenko et al., 2008; Tarbashevich et al., 2007), similar to our results with *efnb1* (Fig. 7). Taken together, these observations suggest a close physical association between Grip2, Efnb1 and the Wnt signaling components that facilitates correct PGC migration.

Unlike somatic cells, PGCs are known to divide symmetrically only two or three times before exiting the endoderm (Whitington and Dixon, 1975). Disruption of the cell cycle or inappropriate gene expression in PGCs would trigger their cell death (Lai et al., 2012). *efnb1* interacts with at least two signaling pathways: FGF (Moore et al., 2004; Lee et al., 2006) and Wnt (Lien and Fuchs, 2014; Lee et al., 2006). Disruption of these pathways could affect cell proliferation or differentiation, which might explain the effect on PGC number when *efnb1* is overexpressed.

Inhibition of p300 results in a loss of PGCs

Here we show for the first time that p300 is necessary for proper PGC development (Fig. 8). Pharmaceutical inhibition of p300 caused a significant reduction in PGC number, suggesting a role in proliferation, apoptosis and/or cell cycle regulation. Similar to what has been shown in retinal cells, p300 may promote PGC proliferation and protect PGCs from apoptosis by modulating the activity of Stat1 and Stat3 (Kawase et al., 2016). Alternatively, p300 might be necessary to allow PGCs to pass through G1 of the cell cycle in order to proliferate, consistent with its role in leukemia cells (Gao et al., 2013). Further investigation is necessary to determine the precise molecular mechanisms by which p300 regulates PGC number.

Our RNA-seq analysis has revealed interconnected pathways highlighting the vegetal pole as a major signaling center. Interestingly, *grip2*, which encodes an important scaffolding protein in the nervous system, was the most abundant vegetal mRNA identified in our analysis. Grip2 is likely to represent a key scaffolding protein for the assembly, through its PDZ domains, of multi-protein signaling complexes. The challenge now is to functionally test the pathways revealed by our comprehensive list of localized mRNAs and to define the maternal contributions to both germline and somatic cell fates.

MATERIALS AND METHODS

Isolation of animal and vegetal pole samples

X. laevis adult frogs were purchased from *Xenopus* Express. Ovarian tissue was surgically removed from anesthetized females, then oocytes were enzymatically released from ovarian tissue (Sive et al., 2000) and stage VI oocytes were selected. Oocyte-matched vegetal and animal poles (10-20% of total oocyte for each pole) were cut with a razor blade (Cuykendall and Houston, 2010) (for further details, see the supplementary Materials and Methods). Prior to collection, the germinal vesicle (GV) was manually removed from animal pole samples to ensure that GV retention of transcripts with different final locations would not contribute to false positives in our RNA-seq analysis. RNA was extracted (see the supplementary Materials and Methods) and 25 vegetal and animal oocyte-paired poles were collected per frog. Equal concentrations of RNA from the respective poles of oocytes from three different frogs were combined to make one vegetal and one animal pole paired sample. Oocytes from a total of nine frogs were used to form three vegetal and three animal pole samples that were submitted for RNA-seq analysis. All animal protocols were approved by the Institutional Animal Care and Use Committee of the University of Miami.

RNA preparation for Illumina sequencing

For each sample, 1 µg of total RNA was processed for RNA quality with the Agilent Bioanalyzer 2100. Samples were processed for both RNA-seq and small (sm)RNA-seq. RNA-seq samples were depleted of mitochondrial and ribosomal RNAs with the ScriptSeq Complete Gold Kit (Illumina) and subjected to ten cycles of PCR prior to RNA sequencing on the Illumina HiSeq 2000 using the reagents provided in the Illumina TruSeq PE Cluster Kit v3 and the TruSeq SBS Kit-HS (200 cycle) kits. Reads aligning to a ribosome-specific reference or mitochondrial sequences represented <5% and 1.28% of the total, respectively. For smRNA-seq, samples underwent 11 cycles of PCR and were then prepped using the Illumina TruSeq Small RNA Sample Preparation Guide (15004197 Rev. D). Cluster generation was performed on the Illumina cBot according to the manufacturer's recommendations. An average of 10.1 million (animal) and 19.4 million (vegetal) pass-filter paired-end 100 base reads were generated per sample (range: 9.0-11.2 million, animal; 17.8-22.7 million, vegetal).

Data processing and quantification

Quality control metrics were determined using FastQC software (Babraham Bioinformatics). The total library size was 2.8 Gb, with 95% of the total base pairs above internal FastQC thresholds. Variance between samples was minimal with sample reads ranging from 24.6-26.8 and 21.2-21.5 million read pairs for animal and vegetal pole samples, respectively. Raw reads were aligned using TopHat v2 RNA-seq analysis software. To assess the quality of the reference genome, reads were aligned to both *X. laevis* (v6.0 and v7.1) and *X. tropicalis* (v7.1) references. Alignment to the *X. laevis* genome

Development (2017) 144, 292-304 doi:10.1242/dev.139220

ranged from 66.42%-71.45% and the *X. tropicalis* genome ranged from 25.21%-39.27%. Therefore, reads from the *X. laevis* genome alignment were quantified.

Transcripts were quantified using Cufflinks v2.1. At least 21,555 transcripts were detected in each sample, and of those 13,930 had a fragments per kilobase of transcript per million mapped reads (FPKM) value ≥5. Differential expression analysis was performed using Cuffdiff to compare vegetal versus animal pole transcripts with an FPKM value ≥5 (Mortazavi et al., 2008). 5717 total transcripts (FDR<0.05, fold change ≥1.95) were differentially expressed. More stringent criteria were then employed to determine transcripts enriched in the vegetal or animal pole. The bulk of large yolk platelets are within the vegetal hemisphere, reducing the yolk-free cytoplasm content there (Danilchik and Gerhart, 1987; Callen et al., 1980; Rebagliati et al., 1985); therefore, we set a criterion of 4-fold enrichment for vegetal versus animal pole transcripts. An inherent bias exists towards the animal pole based on the 4-fold difference in RNA concentrations along this axis; thus, transcripts enriched 10-fold in the animal pole compared with the vegetal pole were considered localized to the animal pole.

Small RNA analysis and novel non-coding RNA identification

For small RNA analysis, reads were aligned to a *X. laevis*-specific miRNA reference and then counted. *X. tropicalis* homologs of *X. laevis* miRNAs were queried for miRNA targets. To identify novel long non-coding (lnc) RNAs, Trinity software (Trinity version 2012-06-08) was used for *de novo* assembly of the transcripts in each sample. For details, see the supplementary Materials and Methods.

Gene name identification and GO analysis

Joint Genome Institute (JGI) annotation was used for gene name identification (Xenbase.org). To verify gene identity, separate BLAST alignments (NCBI) for the scaffolds were also performed. Genes homologous to those of humans were submitted to GeneCards for summary information (genecards.org). GeneGo software (MetaCore MetaCore Bioinformatics software from Thomson Reuters, https://portal.genego.com/) was used for pathway analysis. For further details, see the supplementary Materials and Methods.

Whole-mount *in situ* hybridization (WISH)

X. laevis embryos were obtained as described by Sive et al. (2000) and WISH was performed as described in the supplementary Materials and Methods.

Functional analyses

mRNAs of select genes were transcribed from respective cDNA clones and 0.5 ng injected into fertilized embryos. Tailbud stage embryos were then analyzed for PGC number and location using WISH for the germ plasm marker *xpat*. *X. laevis otx1*, *wwtr1*, *efnb1* and *sox7* expression were inhibited by MO and/or by overexpressing a dominant-negative construct (*sox7*). *X. laevis* p300 expression was pharmaceutically inhibited in fertilized embryos by incubation with the small molecule inhibitor C646. Tailbud-stage embryos were then analyzed for proper PGC number and location using WISH. *P*-values were determined using a two-tailed unpaired Student's *t*-test. $P<0.05$ was considered significant. For further details, see the supplementary Materials and Methods.

Acknowledgements

We thank Drs Jing Yang for the *X. tropicalis* plasmids for *hook2*, *tob2* and *sox7* and Mike Klymkowsky for the *X. laevis sox7* dominant-negative construct. Drs Michael Gilchrist through the Bioinformatics Workshop (Cold Spring Harbor, NY) and Lingyu Wang provided excellent technical support for our RNA-seq analysis.

Competing interests

The authors declare no competing or financial interests.

Author contributions

D.A.O.: Performed experiments, data analysis, figure preparation, helped write manuscript. A.M.B.: Planned experiments, performed functional experiments, data analysis, helped write manuscript. T.H.A.: Functional experiments, figure preparation, data analysis. K.M.N.: Data preparation and analysis. D.V.B.: Bioinformatics and analysis of RNA-seq data. M.L.K.: Conceived and planned experiments, wrote manuscript.

Funding

This work was supported by the National Institutes of Health (HD072340, GM102397 to M.L.K.). Deposited in PMC for release after 12 months.

Data availability

All RNA-seq data are available at Gene Expression Omnibus with accession number GSE80971.

Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/doi/10.1242/dev.139220.supplemental

References

Aguero, T., Kassmer, S., Alberio, R., Johnson, A. and King, M. L. (2016). Mechanism of vertebrate germ cell determination. In *Vertebrate Development: Maternal to Zygotic Control* (ed. F. Pelegri, M. Danilchik and A. Sutherland), pp. 383-440. Berlin: Springer.

Ataman, B., Ashley, J., Gorczyca, D., Gorczyca, M., Mathew, D., Wichmann, C., Sigrist, S. J. and Budnik, V. (2006). Nuclear trafficking of Drosophila Frizzled-2 during synapse development requires the PDZ protein dGRIP. *Proc. Natl. Acad. Sci. USA* **103**, 7841-7846.

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

Bates, T. J. D., Vonica, A., Heasman, J., Brivanlou, A. H. and Bell, E. (2013). Coco regulates dorsoventral specification of germ layers via inhibition of TGFbeta signalling. *Development* **140**, 4177-4181.

Boussouar, F., Goudarzi, A., Buchou, T., Shiota, H., Barral, S., Debernardi, A., Guardiola, P., Brindle, P., Martinez, G., Arnoult, C. et al. (2014). A specific CBP/p300-dependent gene expression programme drives the metabolic remodelling in late stages of spermatogenesis. *Andrology* **2**, 351-359.

Bovolenta, P., Rodriguez, J. and Esteve, P. (2006). Frizzled/RYK mediated signalling in axon guidance. *Development* **133**, 4399-4408.

Brückner, K., Pablo Labrador, J., Scheiffele, P., Herb, A., Seeburg, P. H. and Klein, R. (1999). EphrinB ligands recruit GRIP family PDZ adaptor proteins into raft membrane microdomains. *Neuron* **22**, 511-524.

Bush, J. O. and Soriano, P. (2010). Ephrin-B1 forward signaling regulates craniofacial morphogenesis by controlling cell proliferation across Eph-ephrin boundaries. *Genes Dev.* **24**, 2068-2080.

Callen, J. C., Dennebouy, N. and Mounolou, J. C. (1980). Development of the mitochondrial mass and accumulation of mtDNA in previtellogenic stages of Xenopus laevis oocytes. *J. Cell Sci.* **41**, 307-320.

Claussen, M., Lingner, T., Pommerenke, C., Opitz, L., Salinas, G. and Pieler, T. (2015). Global analysis of asymmetric RNA enrichment in oocytes reveals low conservation between closely related Xenopus species. *Mol. Biol. Cell* **26**, 3777-3787.

Colozza, G. and De Robertis, E. M. (2014). Maternal syntabulin is required for dorsal axis formation and is a germ plasm component in Xenopus. *Differentiation* **88**, 17-26.

Costa, C., Santos, M., Martínez-Fernández, M., Dueñas, M., Lorz, C., García-Escudero, R. and Paramio, J. M. (2013). E2F1 loss induces spontaneous tumour development in Rb-deficient epidermis. *Oncogene* **32**, 2937-2951.

Creppe, C., Palau, A., Malinverni, R., Valero, V. and Buschbeck, M. (2014). A Cbx8-containing polycomb complex facilitates the transition to gene activation during ES cell differentiation. *PLoS Genet.* **10**, e1004851.

Cuykendall, T. N. and Houston, D. W. (2010). Identification of germ plasm-associated transcripts by microarray analysis of Xenopus vegetal cortex RNA. *Dev. Dyn.* **239**, 1838-1848.

Danilchik, M. V. and Gerhart, J. C. (1987). Differentiation of the animal-vegetal axis in Xenopus laevis oocytes: I. Polarized intracellular translocation of platelets establishes the yolk gradient. *Dev. Biol.* **122**, 101-112.

De Domenico, E., Owens, N. D. L., Grant, I. M., Gomes-Faria, R. and Gilchrist, M. J. (2015). Molecular asymmetry in the 8-cell stage Xenopus tropicalis embryo described by single blastomere transcript sequencing. *Dev. Biol.* **408**, 252-268.

Eimon, P. M. and Harland, R. M. (2001). Xenopus Dan, a member of the Dan gene family of BMP antagonists, is expressed in derivatives of the cranial and trunk neural crest. *Mech. Dev.* **107**, 187-189.

Gao, X.-N., Lin, J., Ning, Q.-Y., Gao, L., Yao, Y.-S., Zhou, J.-H., Li, Y.-H., Wang, L.-L. and Yu, L. (2013). A histone acetyltransferase p300 inhibitor C646 induces cell cycle arrest and apoptosis selectively in AML1-ETO-positive AML cells. *PLoS ONE* **8**, e55481.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, , D. A., , Amit, I., Adiconis, X., Fan, L., Raychowdhury, R. et al. (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat. Biotechnol.* **29**, 644-652.

Grant, P. A., Yan, B., Johnson, M. A., Johnson, D. L. E. and Moody, S. A. (2014). Novel animal pole-enriched maternal mRNAs are preferentially expressed in neural ectoderm. *Dev. Dyn.* **243**, 478-496.
Guo, L., Zhong, D., Lau, S., Liu, X., Dong, X.-Y., Sun, X., Yang, V. W., Vertino, P. M., Moreno, C. S., Varma, V. et al. (2008). Sox7 Is an independent checkpoint for beta-catenin function in prostate and colon epithelial cells. *Mol. Cancer Res.* **6**, 1421-1430.
Heasman, J. (2006). Maternal determinants of embryonic cell fate. *Semin. Cell Dev. Biol.* **17**, 93-98.
Houston, D. W. (2013). Regulation of cell polarity and RNA localization in vertebrate oocytes. *Int. Rev. Cell Mol. Biol.* **306**, 127-185.
Hudson, C. and Woodland, H. R. (1998). Xpat, a gene expressed specifically in germ plasm and primordial germ cells of Xenopus laevis. *Mech. Dev.* **73**, 159-168.
Irie, N., Weinberger, L., Tang, W. W. C., Kobayashi, T., Viukov, S., Manor, Y. S., Dietmann, S., Hanna, J. H. and Surani, M. A. (2015). SOX17 is a critical specifier of human primordial germ cell fate. *Cell* **160**, 253-268.
Johansen, A. K. Z., Dean, A., Morecroft, I., Hood, K., Nilsen, M., Loughlin, L., Anagnostopoulou, A., Touyz, R. M., White, K. and MacLean, M. R. (2016). The serotonin transporter promotes a pathological estrogen metabolic pathway in pulmonary hypertension via cytochrome P450 1B1. *Pulm. Circ.* **6**, 82-92.
Karpinka, J. B., Fortriede, J. D., Burns, K. A., James-Zorn, C., Ponferrada, V. G., Lee, J., Karimi, K., Zorn, A. M. and Vize, P. D. (2015). Xenbase, the Xenopus model organism database; new virtualized system, data types and genomes. *Nucleic Acids Res.* **43**, D756-D763.
Kawase, R., Nishimura, Y., Ashikawa, Y., Sasagawa, S., Murakami, S., Yuge, M., Okabe, S., Kawaguchi, K., Yamamoto, H., Moriyuki, K. et al. (2016). EP300 protects from light-induced retinopathy in zebrafish. *Front. Pharmacol.* **7**, 126.
King, M. L. (2014). Germ cell specification in Xenopus. In *Xenopus Development* (ed. M. Kloc and J. Z. Kubiak), pp. 75-100. Hoboken, New Jersey: John Wiley & Sons.
King, M. L., Zhou, Y. and Bubunenko, M. (1999). Polarizing genetic information in the egg: RNA localization in the frog oocyte. *BioEssays* **21**, 546-557.
Kirilenko, P., Weierud, F. K., Zorn, A. M. and Woodland, H. R. (2008). The efficiency of Xenopus primordial germ cell migration depends on the germplasm mRNA encoding the PDZ domain protein Grip2. *Differentiation* **76**, 392-403.
Klauke, K., Radulović, V., Broekhuis, M., Weersing, E., Zwart, E., Olthof, S., Ritsema, M., Bruggeman, S., Wu, X., Helin, K. et al. (2013). Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation. *Nat. Cell Biol.* **15**, 353-362.
Klein, R. and Kania, A. (2014). Ephrin signalling in the developing nervous system. *Curr. Opin. Neurobiol.* **27**, 16-24.
Korkut, C., Ataman, B., Ramachandran, P., Ashley, J., Barria, R., Gherbesi, N. and Budnik, V. (2009). Trans-synaptic transmission of vesicular Wnt signals through Evi/Wntless. *Cell* **139**, 393-404.
Kwok, R. P. S., Liu, X.-T. and Smith, G. D. (2006). Distribution of co-activators CBP and p300 during mouse oocyte and embryo development. *Mol. Reprod. Dev.* **73**, 885-894.
Lai, F. and King, M. L. (2013). Repressive translational control in germ cells. *Mol. Reprod. Dev.* **80**, 665-676.
Lai, F., Singh, A. and King, M. L. (2012). Xenopus Nanos1 is required to prevent endoderm gene expression and apoptosis in primordial germ cells. *Development* **139**, 1476-1486.
Larive, R. M., Abad, A., Cardaba, C. M., Hernandez, T., Canamero, M., de Alava, E., Santos, E., Alarcon, B. and Bustelo, X. R. (2012). The Ras-like protein R-Ras2/TC21 is important for proper mammary gland development. *Mol. Biol. Cell* **23**, 2373-2387.
Lee, H.-S., Bong, Y.-S., Moore, K. B., Soria, K., Moody, S. A. and Daar, I. O. (2006). Dishevelled mediates ephrinB1 signalling in the eye field through the planar cell polarity pathway. *Nat. Cell Biol.* **8**, 55-63.
Li, H., Miao, Q., Xu, C.-W., Huang, J.-H., Zhou, Y.-F. and Wu, M.-J. (2016). OTX1 contributes to hepatocellular carcinoma progression by regulation of ERK/MAPK pathway. *J. Korean Med. Sci.* **31**, 1215-1223.
Lien, W.-H. and Fuchs, E. (2014). Wnt some lose some: transcriptional governance of stem cells by Wnt/beta-catenin signaling. *Genes Dev.* **28**, 1517-1532.
Mani, K. M., Lefebvre, C., Wang, K., Lim, W. K., Basso, K., Dalla-Favera, R. and Califano, A. (2008). A systems biology approach to prediction of oncogenes and molecular perturbation targets in B-cell lymphomas. *Mol. Syst. Biol.* **4**, 169.
Minderman, H., Maguire, O., O'Loughlin, K. L., Muhitch, J., Wallace, P. K. and Abrams, S. I. (2016). Total cellular protein presence of the transcription factor IRF8 does not necessarily correlate with its nuclear presence. *Methods*, pii: S1046-2023(16)30266-3.
Mohammed, M. K., Shao, C., Wang, J., Wei, Q., Wang, X., Collier, Z., Tang, S., Liu, H., Zhang, F., Huang, J. et al. (2016). Wnt/beta-catenin signaling plays an ever-expanding role in stem cell self-renewal, tumorigenesis and cancer chemoresistance. *Genes Dis.* **3**, 11-40.
Moore, K. B., Mood, K., Daar, I. O. and Moody, S. A. (2004). Morphogenetic movements underlying eye field formation require interactions between the FGF and ephrinB1 signaling pathways. *Dev. Cell* **6**, 55-67.
Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621-628.
Mowry, K. L. (1996). Complex formation between stage-specific oocyte factors and a Xenopus mRNA localization element. *Proc. Natl. Acad. Sci. USA* **93**, 14608-14613.
Murphy, G. A., Graham, S. M., Morita, S., Reks, S. E., Rogers-Graham, K., Vojtek, A., Kelley, G. G. and Der, C. J. (2002). Involvement of phosphatidylinositol 3-kinase, but not RalGDS, in TC21/R-Ras2-mediated transformation. *J. Biol. Chem.* **277**, 9966-9975.
Nikishin, D. A., Kremnyov, S. V., Konduktorova, V. V. and Shmukler, Y. B. (2012). Expression of serotonergic system components during early Xenopus embryogenesis. *Int. J. Dev. Biol.* **56**, 385-391.
O'Loghlen, A., Martin, N., Krusche, B., Pemberton, H., Alonso, M. M., Chandler, H., Brookes, S., Parrinello, S., Peters, G. and Gil, J. (2015). The nuclear receptor NR2E1/TLX controls senescence. *Oncogene* **34**, 4069-4077.
Peshkin, L., Wühr, M., Pearl, E., Haas, W., Freeman, R. M., Jr, Gerhart, J. C., Klein, A. M., Horb, M., Gygi, S. P. and Kirschner, M. W. (2015). On the relationship of protein and mRNA dynamics in vertebrate embryonic development. *Dev. Cell* **35**, 383-394.
Rebagliati, M. R., Weeks, D. L., Harvey, R. P. and Melton, D. A. (1985). Identification and cloning of localized maternal RNAs from Xenopus eggs. *Cell* **42**, 769-777.
Rinehart, J., Vazquez, N., Kahle, K. T., Hodson, C. A., Ring, A. M., Gulcicek, E. E., Louvi, A., Bobadilla, N. A., Gamba, G. and Lifton, R. P. (2011). WNK2 kinase is a novel regulator of essential neuronal cation-chloride cotransporters. *J. Biol. Chem.* **286**, 30171-30180.
Rohani, N., Parmeggiani, A., Winklbauer, R. and Fagotto, F. (2014). Variable combinations of specific ephrin ligand/Eph receptor pairs control embryonic tissue separation. *PLoS Biol.* **12**, e1001955.
Sive, H. L., Grainger, R. M. and Harland, R. M. (2000). *Early Development of Xenopus Laevis: A Laboratory Manual*. Cold Spring Harbor, NY. CSHL Press.
Stovall, D. B., Cao, P. and Sui, G. (2014). SOX7: from a developmental regulator to an emerging tumor suppressor. *Histol. Histopathol.* **29**, 439-445.
Tarbashevich, K., Koebernick, K. and Pieler, T. (2007). XGRIP2.1 is encoded by a vegetally localizing, maternal mRNA and functions in germ cell development and anteroposterior PGC positioning in Xenopus laevis. *Dev. Biol.* **311**, 554-565.
Tarbashevich, K., Dzementsei, A. and Pieler, T. (2011). A novel function for KIF13B in germ cell migration. *Dev. Biol.* **349**, 169-178.
Venkatarama, T., Lai, F., Luo, X., Zhou, Y., Newman, K. and King, M. L. (2010). Repression of zygotic gene expression in the Xenopus germline. *Development* **137**, 651-660.
Wang, L., Chen, Z., Wang, Y., Chang, D., Su, L., Guo, Y. and Liu, C. (2014). TR1 promotes cell proliferation and inhibits apoptosis through cyclin A and CTGF regulation in non-small cell lung cancer. *Tumour Biol.* **35**, 463-468.
Whitington, P. M. and Dixon, K. E. (1975). Quantitative studies of germ plasm and germ cells during early embryogenesis of Xenopus laevis. *J. Embryol. Exp. Morphol.* **33**, 57-74.
Yamaguchi, T., Kataoka, K., Watanabe, K. and Orii, H. (2014). Restriction of the Xenopus DEADSouth mRNA to the primordial germ cells is ensured by multiple mechanisms. *Mech. Dev.* **131**, 15-23.
Yu, S., Yerges-Armstrong, L. M., Chu, Y., Zmuda, J. M. and Zhang, Y. (2013). E2F1 effects on osteoblast differentiation and mineralization are mediated through up-regulation of frizzled-1. *Bone* **56**, 234-241.
Zaragoza, K., Begay, V., Schuetz, A., Heinemann, U. and Leutz, A. (2010). Repression of transcriptional activity of C/EBPalpha by E2F-dimerization partner complexes. *Mol. Cell. Biol.* **30**, 2293-2304.
Zhang, L.-N. and Yan, Y.-B. (2015). Depletion of poly(A)-specific ribonuclease (PARN) inhibits proliferation of human gastric cancer cells by blocking cell cycle progression. *Biochim. Biophys. Acta* **1853**, 522-534.
Zhang, C., Basta, T., Fawcett, S. R. and Klymkowsky, M. W. (2005a). SOX7 is an immediate-early target of VegT and regulates Nodal-related gene expression in Xenopus. *Dev. Biol.* **278**, 526-541.
Zhang, C., Basta, T. and Klymkowsky, M. W. (2005b). SOX7 and SOX18 are essential for cardiogenesis in Xenopus. *Dev. Dyn.* **234**, 878-891.
Zhang, Y.-F., Liu, L.-X., Cao, H.-T., Ou, L., Qu, J., Wang, Y. and Chen, J.-G. (2015). Otx1 promotes basal dendritic growth and regulates intrinsic electrophysiological and synaptic properties of layer V pyramidal neurons in mouse motor cortex. *Neuroscience* **285**, 139-154.

Supplemental MATERIALS AND METHODS

Isolation of animal and vegetal pole samples (continued)

After surgical removal, ovaries were teased into small clumps and rinsed in Modified Barth's Solution. Oocytes were enzymatically released from ovarian tissue and from follicles with 0.2% collagenase (Type 1; Worthington Biochemical Corp.) in calcium free OR2 (Sive et al., 2000). Stage VI oocytes were selected and incubated overnight in OR2 with 1 mM $CaCl_2$ at 18°C. The following day, oocytes were transferred into P10EM solution (Elinson et al., 1993) and oocyte-matched vegetal and animal poles were cut (Cuykendall and Houston, 2010). Samples were immediately collected into Eppendorf tubes, frozen on dry ice, then stored at -80°C until RNA was extracted.

RNA isolation

Total RNA was isolated using Trizol reagent (Invitrogen) as per the manufacturer's protocol with the following modifications: samples were homogenized before adding Trizol reagent; RNA precipitation was done overnight in isopropanol and the RNA pellets were rinsed three times with ice-cold ethanol then air-dried and reconstituted in nuclease free water. RNA concentration was determined using Nanodrop 2000c (Thermo Scientific). For each sample, total RNA was used to synthesize cDNA via the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). All procedures for RNA isolation were done using filtered pipette tips. *X. laevis nanos* RNA is exclusively found within the germ plasm and associates with the vegetal cortex (Elinson et al., 1993). To assess both the purity and cortical quality of our samples, we determined the amount of *nanos* RNA in the vegetal and animal pole samples by semi-quantitative RT-PCR (Lai et al., 2011). Vegetal pole samples with the highest levels of *nanos* RNA and oocyte-matched animal pole samples with no detectable *nanos* RNA were selected for RNA-seq analysis.

Small RNA analysis and novel non-coding RNA identification (continued)

For small RNA analysis, Cutadapt was used to remove adapter sequences from read ends. Reads trimmed to < 20bp were discarded. Next, the reads were aligned to a *X. laevis* specific microRNA (miRNA) reference using TopHat v2. 23% alignment was achieved. Samtools was used to filter then count the aligned reads using the following command:

```
$ samtools view aligned_file.bam miR21 > miR21.sam | wc –l
```

Since *X. laevis* miRNAs are not well annotated, we identified miRNA homologues in *X. tropicalis*. We subsequently queried TargetScan for miRNA targets. Target mRNAs were identified based on highly conserved regions in homologues human genes. Thus *X. tropicalis* miRNAs that match with human mRNA target homologs of vegetally localized RNAs from our RNA-seq analysis were reported.

To identify novel long non-coding (lnc) RNAs, Trinity software was used for de novo assembly of the transcripts in each sample. These assemblies were analyzed by Coding-Potential Assessment Tools (CPAT) and Coding Potential Calculator (CPC), which determined that there were no viable transcripts. Further analysis of the potential function of the lncRNAs was done using advanced modeling, but nothing significant could be detected based on the current platforms. Further analyses of lncRNAs will have to await better annotation of the Xenopus genome.

Gene name identification and GO analysis (continued)

The total transcript set of all vegetal pole up-regulated genes (≥4 fold enrichment; total = 198) was submitted to GeneGO for enrichment and pathway analysis. Un-annotated sequences represented approximately half of the total RNAs enriched at either the animal (12/27) or vegetal

poles (213/411) (Fig. 1D) and these are listed in supplemental Table 3. Metacore Pathway Analysis (GeneGO v6.18, www.genego.com) was used as previously described (Egan et al., 2014). Pathway analysis was determined using the direct interaction and analyze network building algorithms.

Whole-mount *in situ* hybridization (WISH) (continued)

X. laevis embryos were obtained as described in Sive et al., 2000. WISH was performed as described in Agüero et al., 2012 with the exception of day 1, which was performed as described in Colozza and De Robertis, 2014. Plasmids containing full-length clones were purchased from GE Dharmacon (*spire1*, *atrx*, *wnk2*, *rras2*, *mov10*, *trank1*, *sybu*, *otx1*, *e2f1*, and *dand5*) and Transomic Technologies (*efnb1*, *slc18a2*), and provided by Dr. Jing Yang (*hook2*, *tob2*, and *sox7*). Inserts were verified by restriction digestion and sequencing. Inserts were PCR amplified using the following primer sets: T7/Sp6 (*efnb1*, *rras2*, *mov10*, *tob2*, *sybu*, *sox7*, *rnf38*, *spire1*, *hook2*, *e2f1*, *trank1*, *otx1*, and *slc18a2*) or T3/Sp6 (*wnk2*, *atrx*, and *dand5*). *Xpat* clone was synthesized as described in Lai et al., 2012. Antisense probes containing Digoxigenin-11-UTP were synthesized using the T7, SP6 or T3 RNA polymerase (NEB). Primer sequences: T7: 5'-taatacgactcactatag-3'; T3: 5'-aattaaccctcactaaag-3'; Sp6: 5'-atttaggtgacactatag-3'.

Gain-of-function analysis of vegetally enriched genes of interest

cDNA clones encoding full-length genes were obtained from GE-Dharmacon (*parn*, # MXL1736-202774978; *rras2*, # MXL1736-202773625; *wwtr1*, # MXL1736-202784975; *e2f1*, # MXL1736-202809814; *otx1*, # MXT1765-202789330; *spire1*, # MXL1736-202787536), transOMIC (*efnb1*, ID: TCL1007) or as a gift (pCS107-Xt*sox7*, from Dr. Jing Yang, University of Illinois at Urbana). Synthetic capped RNAs for microinjection were obtained by *in vitro* transcription using the mMESSAGE mMACHINE SP6, T7 or T3 Kit (Ambion). DNA templates were linearized or amplified by PCR, transcribed and 0.5 ng of each mRNA was injected in the vegetal pole of 1-cell

embryos as follows: pCS107-Xt*sox7* (SP6/NotI); pCMV-SPORT6-Xl*parn* (SP6/NotI), pCMV-SPORT6-Xl*rras2* (SP6/NotI); pExpress1-Xl*e2f1* (SP6/PCR: T7/SP6); pExpress1-Xl*otx1* (SP6/PCR: T7/SP6); pCR4-TOPO-Xl*efnb1* (T3/PCR: T3/T7); pCMV-SPORT6-Xl*spire1* (SP6/NotI). All clone identities were confirmed by sequencing. All results shown are representative of at least two independent experiments. All embryos were fixed for WISH at stage 32-35 (tailbud). Embryos were staged according to Nieuwkoop and Faber (1956). WISH was performed using the germ plasm marker *xpat* to see PGC differences of injected RNAs. Both sides of each embryo were used for counting PGC number.

Loss-of-Function analysis of novel germline RNAs

Morpholinos targeting the following RNAs were purchased from gene tools: *efnb1*-MO (5'-CCACCTGTGCGGATGGAAGGGCTCC), *sox7*-MO (5'-GTCATTATTCCAACTGACTTGCTGA), *otx1*-MO (5'-TAGGACATCATGCTCAAGGCTGGAT), and *wwtr1*-MO (5'-TGGTACAACAGCTACTTCCCAAGGC). To test MO efficiency full length *X. laevis efnb1* was PCR amplified from stage VI oocyte cDNA with the following forward and reverse primers: 5'-GGATAATACAAAGCTGGTTTCTGTG and 5'- CTTTGCTCCTGTGATTGGATTG, respectively, and subsequently a flag tag (FL) was added using the following primers: forward: 5'-CAGCATGAATTCGGATTTAGCAGCTGAGGGCAAG; reverse: 5'-ATATTCTCGAGTCACTTGTCGTCATCGTCTTTGTAGTCCTTGTAGTAAATGTTTGCAGG. The flag-tagged *efnb1* insert was then cloned into PCS2+ using EcoRI/XhoI to make *efnb1*-FL. Additionally, *otx1* and *wwtr1* (GE-Dharmacon) were cloned into flag(FL)-pCS2+ using BamHI/ClaI and ClaI/ClaI to generate *otx1*-FL and *wwtr1*-FL, respectively. Flag-tagged full length *sox7* (*sox7*-FL) was synthesized and purchased from Genewiz. *efnb1*-FL-rescue was generated from *efnb1*-FL, by introducing conservative mutations in the region that binds the morpholino (5'ATGGAgGGtCTtCGGCGTCTTCTC), rendering the morpholino ineffective.

Mutations were introduced using the Q5 Site-Directed Mutagenesis Kit (New England Biosciences, E0554) according to the manufacturer's protocol.

All morpholinos were tested for knockdown efficiency using the Wheat Germ Extract kit (Promega, REF: L4380) according to the manufacturer's protocol. Translation of *efnb1*-FL, *otx1*-FL, *wwtr1*-FL, and *sox7*-FL were detected by Western blot analysis for flag (Figs. S2 and S3). Primary antibody used: Monoclonal mouse ANTI-FLAG (Sigma, REF: F1804); secondary antibody used: Anti-Mouse IgG HRP Conjugate (Promega, REF: W402B).

Embryos were injected vegetally with *otx1*-MO (15ng), *wwtr1*-MO (15ng), *efnb1*-MO (16ng), *efnb1*-FL (200pg), *efnb1*-FL (200pg) and *efnb1*-MO (16ng) together, *efnb1*-FL-rescue (200pg) and *efnb1*-MO (16ng) together, or scramble-MO (16ng) at the one-cell stage. Embryos were also injected with *sox7*-MO (16ng), *sox7*dCEnR (*X. laevis*) (200pg), Xt*sox7* (*X. tropicalis*) (200pg), or *sox7*dCEnR (200pg) and Xt*sox7* (200pg) mRNAs at the one-cell stage. *X. laevis sox7* dominant negative construct (*sox7*dCEnR) and *X. tropicalis sox7* WT (Xt*sox7*) were generous gifts from Dr. Mike Klymkowsky (University of Colorado, Boulder) and Dr. Jing Yang (University of Illinois at Urbana), respectively. In sox7dCEnR the transacting domain was deleted and the engrailed transcriptional repression domain was inserted downstream of the HMG box (Zhang et al., 2005). Synthetic capped mRNAs for microinjection were obtained by *in vitro* transcription using the mMESSAGE mMACHINE SP6 or T3 Kit (Ambion). DNA templates were linearized then transcribed as follows: *efnb1*-FL (ApaI/SP6); *efnb1*-FL-rescue (ApaI/SP6); *sox7*dCEnR (XhoI/T3); Xt*sox7* (NotI/SP6).

To inhibit p300 activity, embryos were incubated with the small molecule inhibitor of p300, C646, or DMSO as a control. To prevent reported light induced retinopathy by C646 (Kawase et al., 2016), embryos were protected from light.

Injected and treated embryos were collected and fixed at stage 32-35 (tailbud), staging according to Nieuwkoop and Faber (1956), and PGCs were identified using WISH against *xpat* as previously described (Lai, et al., 2012). PGC number per embryo was calculated by counting PGCs on both sides of each embryo. An embryo was considered to have mis-localized PGCs if PGCs were located anterior to somite 5 or posterior to somite 11 as described in Tarbashevich et al. (2011). All results shown are representative of at least two independent experiments. The p(t)-values were determined using a two-tailed unpaired Student's t-test. P-values <0.05 were considered significant.

Ethical Statements

The animal protocols used were evaluated and approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Miami. All activities are in compliance with federal, state and institutional regulations. The University was granted full accreditation by the Association for Assessment and Accreditation of Laboratory Animal Care, International (AAALAC) in February 2005 and received its current re-accreditation in October 22, 2013. In addition, University of Miami is licensed by the U.S. Department of Agriculture (USDA) and has filed a Letter of Assurance with the Office of Laboratory Animal Welfare (OLAW), U.S. Department of Health and Human Services (DHHS).

Supplemental FIGURES

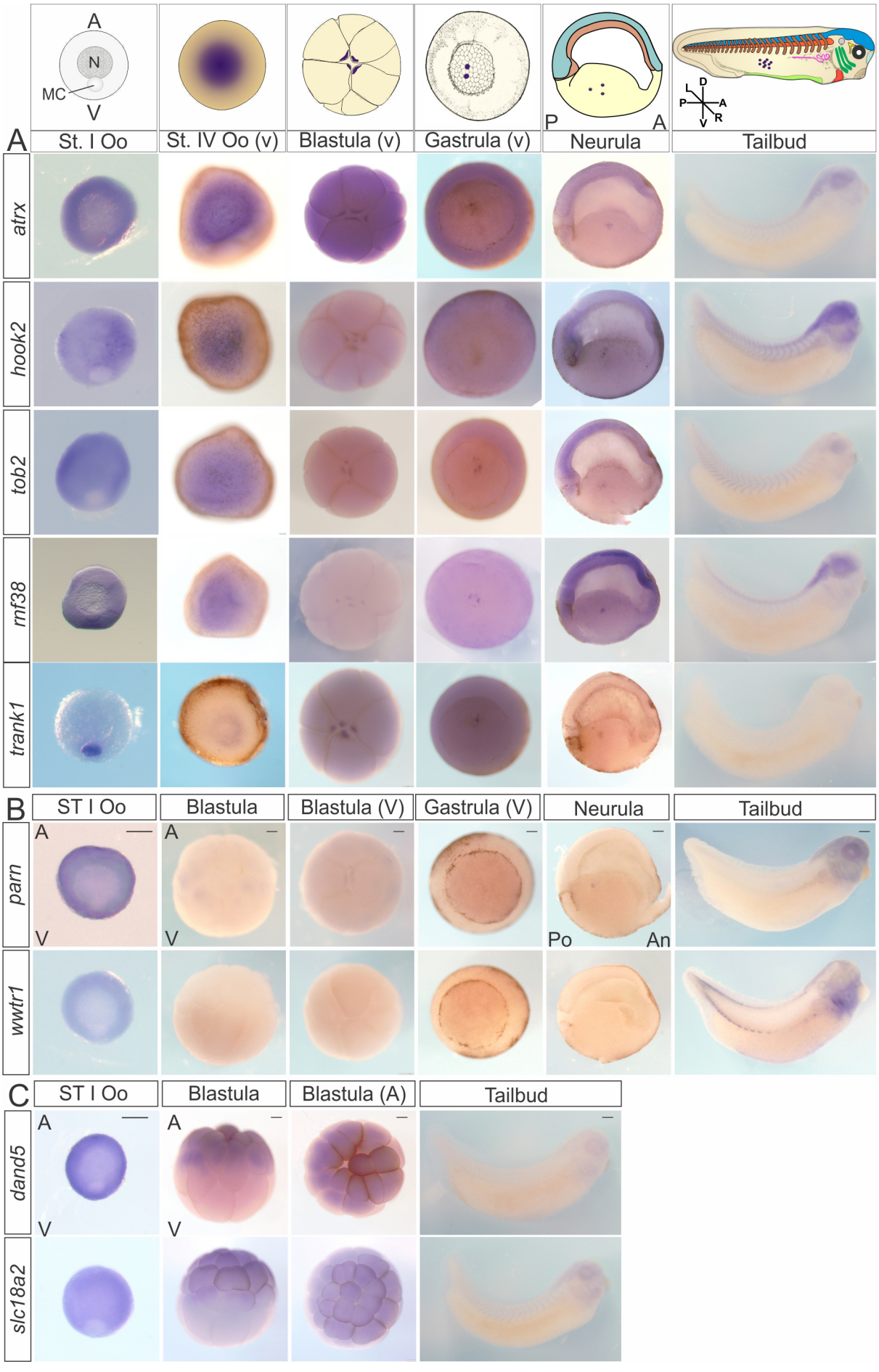

Figure S1. Whole mount in situ hybridization (WISH) of select vegetal and animal pole transcripts. Expression of a subset of RNAs enriched in the vegetal (A-B) and animal (C) poles were analyzed during oogenesis and development by WISH. Probes, developmental stages, and developmental structures are indicated: germplasm/PGCs (purple), pronephros (pink), ventral blood islands (lime green), eye (black), lens (white), otic vesicle (gray), cranial ganglia (yellow), brachial arches (green), nasal placodes (teal), inter-segmental region (brown), notochord (orange), brain and neural tube (blue). Transcripts detected in primordial germ cells at the tailbud stage are magnified and embedded in their respective images. Black bars represent 100um in st. I oocyte panels, and 200um in all other panels. Please refer to Fig. 3 for use of *xpat* probe as a positive control for germ plasm localization.

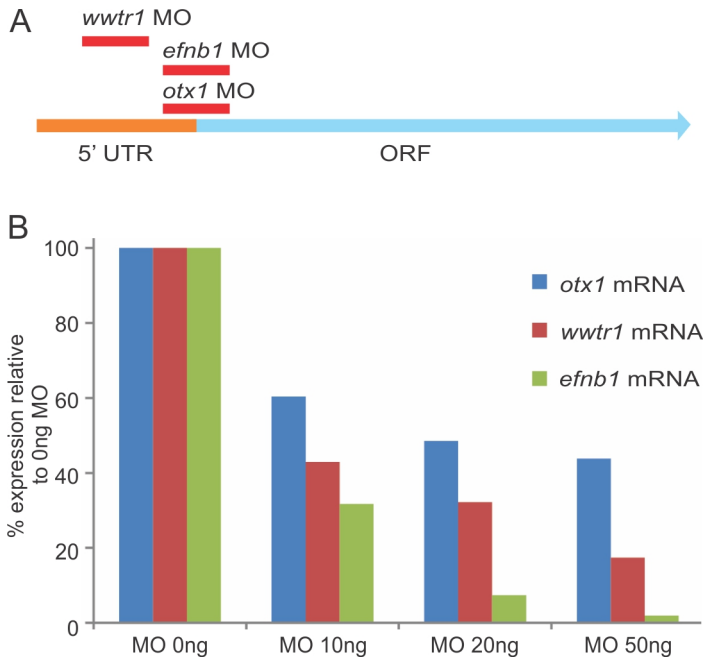

Figure S2. Morpholinos (MOs) inhibit expression of their target RNAs. A) Schematic of the *efnb1*, *otx1*, and *wwtr1* morpholino targeted regions are indicated in red. B) Wheat germ extracts were incubated with either *efnb1*-FL (1ug), *otx1*-FL (250ng), or *wwtr1*-FL (250ng) transcripts in the presence of increasing concentrations of their respective MOs and subject to anti-flag western blot analysis. Quantification of respective flag expression is shown.

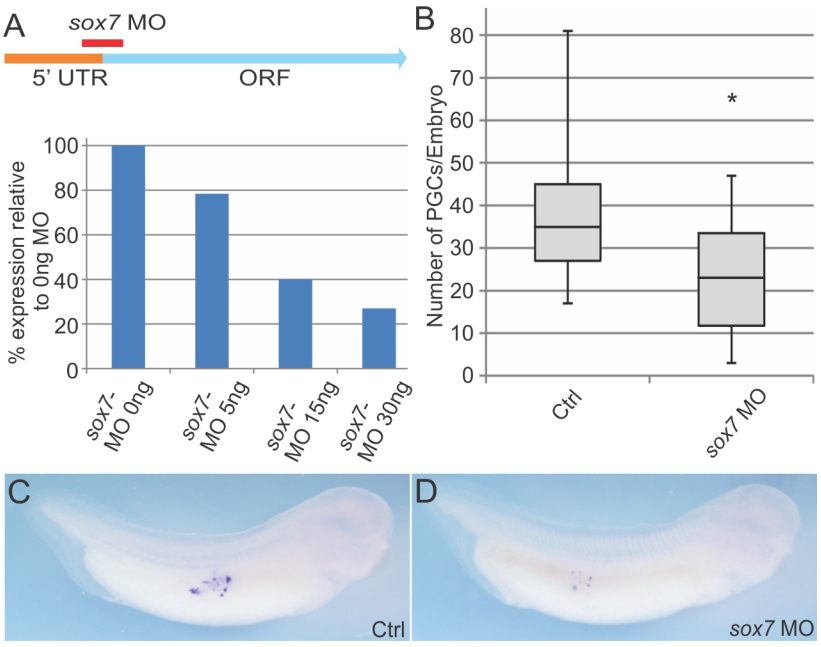

Figure S3. Morpholino-mediated *sox7* inhibition reduces PGC number in tailbud embryos. A) Schematic of the *sox7* morpholino target region (top). Wheat germ extract was incubated with *sox7*-FL (500ng) in the presence of increasing concentrations of *sox7*-MO and subject to anti-flag western blot analysis. Quantification of respective flag expression is shown (bottom). B-D) One-cell embryos were injected in the vegetal region with *sox7*-MO (16ng). Tailbud embryos were analyzed for *xpat* expression by WISH. The number of PGCs per embryo was quantified (B). Representative images are shown (C-D). Uninjected control (ctrl) n=36, *sox7*-MO n=43. * statistically significant compared to ctrl ($p<0.05$). Analysis based on at least two independent experiments.

Supplemental Table 1: Top 198 vegetally enriched transcripts identified by RNA-seq.

Click here to Download Table S1

Supplemental Table 2: All annotated transcripts identified using v6.0 and v7.1 *Xenopus laevis* scaffold sets.

Click here to Download Table S2

Supplemental Table 3: All un-annotated transcripts identified using v6.0 and v7.1 *Xenopus laevis* scaffold sets.

Click here to Download Table S3

References

Aguero, T. H., Fernandez, J. P., Lopez, G. A., Tribulo, C. and Aybar, M. J. (2012). Indian hedgehog signaling is required for proper formation, maintenance and migration of Xenopus neural crest. *Dev Biol* **364**, 99-113.

Colozza, G. and De Robertis, E. M. (2014). Maternal syntabulin is required for dorsal axis formation and is a germ plasm component in Xenopus. *Differentiation* **88**, 17-26.

Cuykendall, T. N. and Houston, D. W. (2010). Identification of germ plasm-associated transcripts by microarray analysis of Xenopus vegetal cortex RNA. *Dev Dyn* **239**, 1838-1848.

Egan, J. B., Barrett, M. T., Champion, M. D., Middha, S., Lenkiewicz, E., Evers, L., Francis, P., Schmidt, J., Shi, C. X., Van Wier, S., et al. (2014). Whole genome analyses of a well-differentiated liposarcoma reveals novel SYT1 and DDR2 rearrangements. *PloS one* **9**, e87113.

Elinson, R. P., King, M. L. and Forristall, C. (1993). Isolated vegetal cortex from Xenopus oocytes selectively retains localized mRNAs. *Dev Biol* **160**, 554-562.

Kawase, R., Nishimura, Y., Ashikawa, Y., Sasagawa, S., Murakami, S., Yuge, M., Okabe, S., Kawaguchi, K., Yamamoto, H., Moriyuki, K., et al. (2016). EP300 Protects from Light-Induced Retinopathy in Zebrafish. *Front Pharmacol* **7**, 126.

Lai, F., Singh, A. and King, M. L. (2012). Xenopus Nanos1 is required to prevent endoderm gene expression and apoptosis in primordial germ cells. *Development* **139**, 1476-1486.

Lai, F., Zhou, Y., Luo, X., Fox, J. and King, M. L. (2011). Nanos1 functions as a translational repressor in the Xenopus germline. *Mech Dev* **128**, 153-163.

Nieuwkoop, P. D. and Faber, J. (1956). *Normal Table of Xenopus laevis (Daudin): A systematical and chronological survey of the development from the fertilized egg till the end of metamorphosis.*: North-Holland Publ. Co.

Sive, H. L., Grainger, R. M. and Harland, R. M. (2000). *Early Development of Xenopus Laevis: A Laboratory Manual*: CSHL Press.

Tarbashevich, K., Dzementsei, A. and Pieler, T. (2011). A novel function for KIF13B in germ cell migration. *Dev Biol* **349**, 169-178.

Zhang, C., Basta, T., Fawcett, S. R. and Klymkowsky, M. W. (2005). SOX7 is an immediate-early target of VegT and regulates Nodal-related gene expression in Xenopus. *Dev Biol* **278**, 526-541.