A single-cell RNA-seq analysis of *Brachyury*-expressing cell clusters suggests a morphogenesis-associated signal center of oral ectoderm in sea urchin embryos

Noriyuki Satoh a,*, Kanako Hisata a, Stephany Foster b, Shumpei Morita b, Koki Nishitsuji a, Nathalie Ouhlen b, Hitoshi Tominaga a, Gary M. Wessel b,**

a Marine Genomics Unit, Okinawa Institute of Science and Technology Graduate University, Onna, Okinawa, 904-0495, Japan
b Department of Molecular and Cellular Biology, Division of BioMedicine, Brown University, Providence, RI, 02912, USA

ARTICLE INFO

Keywords:
- Sea urchin embryos
- Brachyury-expressing cell clusters
- Endoderm
- Archenteron invagination
- Oral ectoderm
- Stomodeum invagination
- The ventral organizer

ABSTRACT

*Brachyury* is a T-box family transcription factor and plays pivotal roles in morphogenesis. In sea urchin embryos, *Brachyury* is expressed in the invaginating endoderm, and in the oral ectoderm of the invaginating mouth opening. The oral ectoderm is hypothesized to serve as a signaling center for oral (ventral)-aboral (dorsal) axis formation and to function as a ventral organizer. Our previous results of a single-cell RNA-seq (scRNA-seq) atlas of early *Strongylocentrotus purpuratus* embryos categorized the constituent cells into 22 clusters, in which the endoderm consists of three clusters and the oral ectoderm four clusters (Foster et al., 2020). Here we examined which clusters of cells expressed *Brachyury* in relation to the morphogenesis and the identity of the ventral organizer. Our results showed that cells of all three endoderm clusters expressed *Brachyury* in blastulae. Based on expression profiles of genes involved in the gene regulatory networks (GRNs) of sea urchin embryos, the three clusters are distinguishable, two likely derived from the Veg2 tier and one from the Veg1 tier. On the other hand, of the four oral-ectoderm clusters, cells of two clusters expressed *Brachyury* at the gastrula stage and genes that are responsible for the ventral organizer at the late blastula stage, but the other two clusters did not. At a single-cell level, most cells of the two oral-ectoderm clusters expressed organizer-related genes, nearly a half of which coincidently expressed *Brachyury*. This suggests that the ventral organizer contains *Brachyury*-positive cells which invaginate to form the stomodeum. This scRNA-seq study therefore highlights significant roles of *Brachyury*-expressing cells in body-plan formation of early sea urchin embryos, though cellular and molecular mechanisms for how *Brachyury* functions in these processes remain to be elucidated in future studies.

1. Introduction

Sea urchin embryos provide a model experimental system to explore gene regulatory networks (GRNs) responsible for specification and differentiation of early embryonic cells (Davidson, 2006; McClay, 2011). Detailed descriptions of embryonic cell lineages, experimental manipulations of embryos, decoded genomes, a list of well-characterized genes encoding transcription factors and signaling pathway molecules, observation of spatio-temporal expression of genes, and functional manipulation of genes, have created sophisticated approaches to explore details of interaction of GRN component genes in each embryonic territory (EchinoWiki; https://wiki.echinobase.org).

A recently developed technology of single cell RNA sequencing (scRNA-seq) provides a powerful tool to categorize genes that are expressed in constituent cells of embryos, tissues, or organs on a cell-by-cell basis (e.g., Stuart and Satija, 2019; Cao et al., 2019). In a previous study, we reported a scRNA-seq resource for early sea urchin development (Foster et al., 2020), in which early embryos of *Strongylocentrotus purpuratus* were interrogated at eight developmental stages, 8-cell, 64-cell, morula, early blastula, hatched blastula, mesenchyme blastula, early gastrula and late gastrula. This resource includes mRNA information in embryonic cells delineated into 22 clusters: two clusters for the neural lineage, one for germline, four for oral ectoderm, two for aboral ectoderm, two for the skeleton originating from the primary...
mesenchyme cells (PMCs), one for secondary mesenchyme cells (SMCs), one for pigment cells, two for ciliated cells, and three for endoderm (Fig. 1A and B) (Foster et al., 2020). Each cluster possesses marker gene expression: the ectoderm clusters were identified by expression of Sox2B, Foxq2, NK2.1 and Arakt1; the endoderm clusters by FoxA and Endo16; the mesodermal PMCs by Alx1, SM50 and SM37; SMCs by Six1/2 and Eya; and the germ cell cluster by Nanos2, Vasa and Seaw2 (Foster et al., 2020). In addition, scRNA-seq resources for 48-hpf gastrula and 72-hpf larva are also available (Perillo et al., 2020).

Taking advantage of this scRNA-seq resource, the present study aimed to identify coordinate gene activity for transcription factors and signaling molecules that may be involved in early morphogenesis of sea urchin embryos. We sought to test coordinate expression of the transcription factor member of the T-box family, Brachury, with key signaling molecules to posit functional organizer testing in the future. Brachury (Bruc) is expressed in the invaginating endoderm and in the oral ectoderm (Gross and McClay, 2001; Croce et al., 2001; Peter and Davidson, 2010) and suppression of Bruc function resulted in a failure of gastrulation and gut formation (Gross and McClay, 2001; Rast et al., 2002).

Here we test the following: (1) The scRNAseq analysis identified three clusters of endoderm cells in late gastrulae (Foster et al., 2020). Do cells of all three clusters express SpBra (Brachury of S. purpuratus)? The endoderm of sea urchin embryos consists of cells from two different origins, Veg1 and Veg2 (e.g., Ransick and Davidson, 1998; Howard-Asbury et al., 2006). What are these three clusters in relation to the Veg1 and Veg2 lineages? The larval digestive system is tripartite, composed of a foregut or muscular esophagus, a midgut or large spherical stomach, and a hindgut or short tubular intestine, and various sets of developmentally relevant genes are dynamically expressed in each region (reviewed by Annuini et al., 2014; Annuini et al., 2019). How are the three endoderm clusters related to the tripartite structures of the larval digestive system? (2) The oral ectoderm of sea urchin embryos serves as a signaling center or ventral organizer for oral (ventral) - aboral (dorsal) axis formation (Dubuc et al., 2004; Lapraz et al., 2009, 2015). The oral (ventral) ectoderm expresses nodal, BMP2/4, chordin and ADM1 whereas ADMP2 is expressed in the aboral (dorsal) ectoderm (Dubuc et al., 2004, 2008; Lapraz et al., 2009, 2015; Wei et al., 2012). The embryonic oral ectoderm consists of two parts with different properties, one remains at the outer surface to form the oral epidermis and the other invaginates to form the mouth opening (stomodeum) (Su et al., 2009; Li et al., 2013; Materna et al., 2013). The oral ectoderm was shown to consist of four cell clusters by the scRNA-seq analysis (Foster et al., 2020). Which of these clusters represent the oral epidermis and which ones represent the stomodeum? Is the ventral organizer a property of the surface ectoderm cells or of cells invaginating to form the stomodeum, or both? Results of our present analyses provide insights into molecular mechanisms involved in the body plan formation of early sea urchin embryos.

2. Results

2.1. SpBra is expressed in cells solely of the endoderm and oral ectoderm

Of 22 (0–21) clusters (Fig. 1A), Foster et al. (2020) have assigned clusters 6, 8 and 14 to be endoderm and those 1, 4, 12 and 13 to be oral ectoderm (Fig. 1B). We found that five out of the seven clusters contained cells with significant SpBra expression (Fig. 1C). These clusters are 1, 6, 8, 12 and 14, the latter four of which showed high level of SpBra expression. Clusters 6, 8, and 14 are endoderm (Fig. 1B), indicating that all three endoderm clusters express SpBra in blastula stages (Figs. 1C and 2). On the other hand, two of the four oral-ectoderm clusters, 12 and 1, expressed SpBra in later blastula stage and gastrula stage (Figs. 3 and 4), whereas the other two, 4 and 13, did not express SpBra (Figs. 3 and 4).

We examined whether other clusters of cells also express SpBra, especially PMCs, SMCs and aboral ectoderm. PMCs give rise to the larval skeleton and are represented by clusters 16 and 19 (Fig. 1B), based on the fact that cells of the clusters express several skeletogenic maker genes, including Alx1, SM50 and SM37 (Foster et al., 2020), and msp130 and SM30 (Perillo et al., 2020). The present analysis confirmed the expression of these markers in clusters 16 and 19 (msp130 in Fig. 1C; Data in Brief Tables 1–3) but the expression of SpBra was at undetectable level in the two PMC clusters (Figs. 1C and 4). Pigment cells originate from SMCs and are represented by cluster 11 based on the expression of gcm, six1/2 and pks1 in the cells (Foster et al., 2020; Perillo et al., 2020). The present independent analysis also found that cells of cluster 11 expressed pks1 (Fig. 1C), six1 and gcm (Data in Brief Tables 1 and 2). Six1 was reported to be expressed in mesodermal domain of the tip of the archenteron (Andrikou et al., 2013) and gcm in cells of pigment lineage (Calestani and Rogers, 2010), confirming that cluster 11 is of SMC lineage. The expression of SpBra was at undetectable level in this cluster (Figs. 1, 2 and 4). Whole mount in situ hybridization (WMISH) analyses in previous research showed SpBra expression in SMCs (Harada et al., 1995; Peterson et al., 1999). Therefore, results of the present scRNA-seq analysis were inconsistent with these earlier observations, but recent WMISH analyses showed that SpBra is not expressed in SMCs (Peter and Davidson, 2010; S. Yaguchi, personal communication). Clusters 0 and 5 are assigned to be aboral ectoderm (Fig. 1B), since cells of these clusters express SPEC1, Hox1, NK2.2, spec2c, lrcA, and EGF2 (Data in Brief Tables 1-5) (Perillo et al., 2020). The expression of SpBra was at undetectable level in clusters 0 and 5 (Figs. 1C and 3). Therefore, we concluded that SpBra is expressed in cells of only endoderm and oral ectoderm.

2.2. Property of the three endoderm clusters

Studies of GRN in sea urchin embryos show that (a) the Veg2 tier includes Veg2-Endo and Veg2-Meso (SMC), the latter of which occupies the vegetal pole (Fig. 5A), (b) the Veg1 tier forms Veg1-Endo and Veg1-Ecto, the latter of which is adjacent to the animal ectoderm (Fig. 5A), (c) genes involved in Veg2-Endo specification include hox11/13b, eve, bilm1/krox, foxA, and gates-e (Fig. 5B; wnt genes that act upstream of these genes are not discussed here) (e.g., Sethi et al., 2012; Cui et al., 2014; Li et al., 2014; Erkenbrach, 2016; Echinobase, https://www.echinobase.org/common/jsp/showWiki.jsp?Davidson_Lab_Gene_Regulatory_Networks; Ettenshohn, 2020), (d) genes involved in Veg1-Endo specification include hox11/13b, eve, and hnf1 (Sethi et al., 2012; Cui et al., 2014). Since our scRNA-seq studies (Foster et al., 2020; this study) demonstrated that the endoderm is comprised of three clusters (6, 8 and 14), it is likely that two of them belong either to Veg2-Endo or Veg1-Endo. Therefore, taking these facts in mind, we examined which transcription factor genes are expressed in each of clusters 6, 8 and 14 (Figs. 1C and 2).

We found two discrete profiles of gene expression in the three clusters, namely, clusters 14 and 6 exhibited a similar profile, and cluster 8 a different profile (Fig. 2). Expression profiles of SpBra, hox11/13b and eve were similar among the three clusters except at early blastula stage, although cluster 8 retained expression of the three genes in later stages (Figs. 1C and 2; Figs. S1 and S2). Although Cui et al. (2014) showed that hnf1 is expressed in Veg1-Endo but not Veg2-Endo, this study showed that hnf1 is expressed in the three clusters (Fig. 2). On the other hand, bilm1/krox, foxA, and gates-e were highly expressed in cells of clusters 14 and 6 but not in cluster 8 (Fig. 2; Fig. S1). Since bilm1/krox, foxA, and gates-e are not included in Veg1-Endo GRN (see Fig. 5B) and since this pattern was seen only in cluster 8, we concluded that cluster 8 corresponds to Veg2-Endo. Accordingly, clusters 14 and 6 correspond to Veg2-Endo (Fig. 5B). The transcription factors of the Ets family are posited as activators of SpBra (Rast et al., 2002; Oliveri et al., 2006; Li et al., 2013). Since this scRNA-seq analysis showed that Ets1 and Ets4 expression in blastulae and gastrulae are quite low, it is not clear whether Ets could be an activator.

In addition to the transcription factor genes mentioned above, Endo16, which encodes a large, secreted protein of the embryonic and larval midgut (Nocente-McGrath et al., 1989; Soltysik-Espanola et al., 1994), has been used as an endoderm and archenteron marker.
Fig. 1. Identification of cell clusters with expression of Brachyury (Bra) in early embryos of Strongylocentrotus purpuratus. (A) t-SNE plot of integrated dataset with clustering, showing 22 cell-clusters numbered from 0 to 21. t-SNE plot of integrated dataset with clustering was performed at a resolution of 0.5 (from Foster et al., 2020). (B) Drawing of late gastrula and mapping cell clusters identified by scRNA-seq to the embryo. Colors match cell clusters seen in A (from Foster et al., 2020). Clusters 6, 8 and 14 are endoderm and clusters 1, 4, 12 and 13 oral ectoderm. (C) Transcriptome trajectories for 22 clusters (from 0 to 21) comprising blastula and gastrula stages of Strongylocentrotus purpuratus embryos. Major genes that are preferentially expressed in endoderm lineages include Bra, hox11/13b, eve, blimp1/krox, foxA, gata-e, hnf1, Endo16, SLBP, Cyl, msp130, and pks1. Dot size represents the percentage of cells that express genes for transcription factors, signaling molecules, and structural proteins. Dot color shows the averaged level of expression. EB, early blastula; HB, hatched blastula; MB, mesenchyme blastula; EG, early gastrula; and LG, late gastrula. Clusters enclosed by green belong to the endoderm with SpBra expression, numbers 6, 8 and 14, and those enclosed by brown belong to the oral ectoderm with SpBra expression, number 1 and 12.
with SLBP from clusters 6 and 8. We tentatively assigned cluster 14 to Veg2-Endo1 SLBP type actin. Aforementioned, cluster 11 is SMC gut subcluster and also locomotive activity, and thereby they express high level of cytoplasmic 11 (Fig. 2; Fig. S1). Since cells of clusters 14 and 6 are involved in clusters 6, 14 and 16, and considerable level expression in cluster 8 and (EchinoWiki) (Ransick et al., 1993; Yuh et al., 2001). We detected high expression. EB, early blastula; HB, hatched blastula; MB, mesenchyme blastula; EG, early gastrula; and LG, late gastrula.

Fig. 2. Enlargement of clusters 14, 6 and 8 to show transcriptome trajectories of Bra, hox11/13b, eve, blimp1/krox, foxA, gata-e, hnf1, Endo16, SLBP, and Cyl. Dot size represents the percentage of cells that express genes for transcription factors, signaling molecules, and structural proteins. Dot color shows the averaged level of expression. EB, early blastula; HB, hatched blastula; MB, mesenchyme blastula; EG, early gastrula; and LG, late gastrula.

2.3. Co-expression of SpBra with hox11/13b, eve, foxA, gata-e, and Endo 16 at single cell level

SpBra is specifically expressed in cells of endoderm clusters 8, 6 and 14, coincident with several genes encoding transcription factors (Fig. 2). We examined the grade of co-expression of these genes at the single-cell level by three methods. First, we superimposed t-SNE plots of the two genes, that is, combinations of SpBra and hox11/13b (Supplementary Fig. S3A), SpBra and eve (Fig. S3B), SpBra and blimp1/krox (Fig. S3C), SpBra and foxA (Fig. S3D; Fig. 6A), SpBra and gata-e (Fig. S3E) or SpBra and Endo16 (Fig. S3F; Fig. 6B). In Fig. 6, SpBra-positive cells of clusters 8, 6 and 14 are seen at left-upper part of feature plots, marked with red (left panels). Cells that express foxA or Endo16 are marked by green (middle panels), and therefore, if single cells express both genes simultaneously, they appeared yellow (right panels). These superimposed t-SNE plots showed that single cells co-expressed SpBra and hox11/13b (Fig. S3A), SpBra and eve (Fig. S3B), SpBra and blimp1/krox (Fig. S3C), SpBra and foxA (Fig. 6A) and SpBra and Endo16 (Fig. 6B). Most yellow cells appeared in clusters 14, 6 and 8 (left-upper part of the panels) except for co-expression of SpBra and foxA in clusters 12 and 1 (Fig. 6A; ectoderm clusters, left-lower part of the panels).

Second, the average expression of genes was measured as count data of mRNAs that were normalized to library size and log transformed. Heatmap and clustered matrix were constructed to determine co-expression of SpBra with other genes at single cell level (Supplementary Fig. S4). Results of co-expression of SpBra with foxA or Endo16 are shown in Fig. 7A. If single cells express these genes coincidently, color-coded bars (yellow, brown and red) appear on the same vertical mass. Several regions with such feature were evident in the three clusters 14, 6, and 8 (Fig. 7A, boxes). On the other hand, the heatmap also showed the presence of cells that expressed high level of foxA but not SpBra (Fig. 7A).

Third, we counted the numbers of cells that highly expressed SpBra, foxA or Endo16 with a criterion of 2+$\sigma$ or higher standardized expression level, as to each cluster of 14, 6 and 8 at four embryonic stages (Supplementary Table S2). For example, cluster 14 at hatched blastula stage comprised of 326 cells, in which high levels of expression of SpBra, foxA and Endo16 were detected in 182 (56%), 288 (88%) and 94 (29%) cells, respectively. Of them, 164 (50%) cells simultaneously expressed SpBra and foxA, 32 (10%) cells SpBra and Endo16, and 31 cells (10%) SpBra,
Fig. 3. Transcriptome trajectories for 22 clusters comprising blastula and gastrula stages of Strongylocentrotus purpuratus embryos. Major genes that are preferentially expressed in oral ectoderm lineages include Bra, SoxB1, chordin, bmp2/4, nodal, Not, Lim1, ADMP1, ADMP2, bmp1 gsc, emx, hnf6, fox5, otx, hes, dri, FGFR, DVR1, nk2.1, foxq2, univin, and six3/6. Dot size represents the percentage of cells that express genes for transcription factors, signaling molecules, and structural proteins. Dot color shows the averaged level of expression. EB, early blastula; HB, hatched blastula; MB, mesenchyme blastula; EG, early gastrula; and LG, late gastrula. Clusters enclosed are ectoderm-lineage, 1 and 12 with SpBra expression (brown) and 4 and 13 with no SpBra expression (yellow).
2.4. SpBra expression in the oral ectoderm clusters is associated with the organizer activity. The scRNA-seq analysis has assigned clusters 1, 4, 12 and 13 to central, and Veg1 oral ectoderm (Fig. 9A). Consistent with this interpretation, the oral ectoderm of the sea urchin embryo is derived from the apical ectoderm organizer (Fig. 6B), stomodeal, and/or composed of four regions: near apical ectoderm, stomodeal, and Veg1 oral ectoderm (Fig. 9A). Therefore, we concluded that the ventral organizer activity is a property of cells of clusters 12 and 1.

Antagonistic interactions between the ventral organizer ADMP1 and the dorsal signaling center ADMP2 has been proposed to play a key role in ventral-dorsal axis formation in sea urchin embryos (Lapraz et al., 2015). The expression of ADMP1 was undetectable in blastulae and early gastrulae, and a few cells with ADMP1 expression appeared at late gastrula stage in clusters 12 and 1 (Figs. 3 and 4; Supplementary Fig. S6). On the other hand, ADMP2 was detected in a few cells of clusters 0, 3, and 5 at early and late gastrula stages (Fig. 3; Fig. S6). The clusters 0 and 5 have been assigned to aboral ectoderm (Fig. 1B) (Foster et al., 2020). Therefore, although the number of cells with ADMP1 expression was comparatively small, a combinatorial expression between oral ectoderm bmp2/4, chordin and ADMP1 and aboral ectoderm ADMP2 was supported by this analysis.

2.5. SpBra is co-expressed with genes important for an ectodermal fate

As in the case of SpBra-positive endoderm clusters, we used three methods to examine co-expression at a single-cell level of SpBra and other genes in the ectoderm clusters. First, we constructed superimposed t-SNE plots in combinations of SpBra and BMP2/4 (Supplementary Fig. S7A; Fig. 6D), SpBra and chordin (Fig. S7B; Fig. 6E), SpBra and nodal (Fig. S7C; Fig. S7D), and SpBra and not (Fig. S7E). Cells of clusters 12 and 1 are seen as dots located at left lower side of the panels (Fig. S7-E). Superimposed t-SNE plots showed clear co-expression of SpBra and chordin (Fig. 6C), and SpBra and not (Fig. 6E) at a single cell level (dots appeared as yellow). Co-expression was found in combinations of SpBra and bmp2/4 (Fig. 6D), and SpBra and Lim1 (Fig. 5F). Because of the scarcity of cells with nodal expression (Fig. S7-C, middle), the presence of cells with simultaneous expression of SpBra and nodal was not evident (Fig. S7-C).

Second, a heatmap of gene expression in the clustered matrix also showed co-expression of SpBra and other genes at a single cell level (Supplementary Fig. S8; Fig. 7B). As in the case of the endoderm clusters, cells of several regions in clusters 12 and 1 showed coincidental expression of SpBra, chordin, BMP2/4 and not (Fig. 7B, boxes). Third, a
Fig. 5. (A) Diagrams to show specification in the sea urchin embryo. Color-coded tracings from photomicrographs of the embryo of Strongylocentrotus purpuratus are shown. veg1 and veg2 are rings of eight cells each, arising from their parental cells at the horizontal 6th cleavage. From veg1 derives ectoderm (mainly) plus hindgut endoderm; and from veg2 non-skeletogenic (secondary) mesenchyme (mesodermal cell types) plus gut endoderm. Skeletogenic mesenchyme lineage, red; endoderm, blue; secondary mesenchyme, violet; oral ectoderm, yellow; apical oral ectoderm, hatched yellow; aboral ectoderm, green; unspecified cells, white. 15 h, blastula stage; 20 and 24 h, mesenchyme blastula; 30, 33, 38 h, gastrula stages. The original drawing of embryos by A. Ransick and E. Davidson is cited (reproduced from Davidson, 2006, copyright Elsevier Inc.). SpBra expression was shown by red dots. (B) Gene regulatory network (GRN) in endoderm specification of sea urchin embryos. The original drawing of embryos by Cui et al. (2014) is cited. Endoderm comprises of Veg2-derived (Veg2-Endo) and Veg1-derived (Veg1-Endo). The present scRNA-seq analysis suggests two classes of Veg2-derived cells, Veg2-Endo1 (cluster 14) and Veg2-Endo2 (cluster 6), which is distinguishable by exclusive expression SLBP in the former. Genes enclosed are findings of this study.
Fig. 6. t-SNE plots showing co-expression of Brachyury with (A) foxA, (B) Endo16, (C) chordin, (D) bmp2/4, and (E) not. Cells with SpBra expression are shown by red dots, the other genes by green dots, and co-expression of the two genes by yellow dots. Clusters 14, 6 and 8 are of endoderm and clusters 12 and 1 of oral ectoderm.
direct count of cells with high levels of SpBra expression also supported the results (Fig. 10; Supplementary Table S3). In clusters 12 and 1, between 60 and 80% of constituent cells highly expressed chordin and Not (Fig. 10A). Most of SpBra-positive cells expressed simultaneously chordin and/or not (Fig. 10A), although the number of cells with simultaneous SpBra and BMP2/4 expression was small (less than 5%) (Fig. 10A). Cells of cluster 1 also exhibited a similar expression profile (Fig. 10B).

These results indicate that most of the constituent cells of clusters 12 and 1 are involved in the ventral organizer, a portion of which expressed SpBra and invaginated to form the stomodeum. In other words, the ventral organizer property overlaps morphogenesis of the oral invagination at a single-cell level, i.e. the two are not independent.

2.6. GRN of oral ectoderm clusters with and without SpBra expression

GRN studies have demonstrated pathways of transcription factors and signaling molecules in the specification of the oral ectoderm (e.g., Li et al., 2014) (Fig. 9B). In addition to BMP2/4, chordin, nodal, not, and ADMP1 described above, we examined expression profiles of the developmentally relevant genes for the oral ectoderm, including bmp1, goosecoid (gsc), emx, hnf6 (onecut), foxG, otx, hes, dri, FGFR, DVR1, nk2.1, foxq2, univin and six3/6 in oral ectoderm clusters (Figs. 3 and 4; Figs. S9 and S10). Of those, otx, hes and dri showed high levels of expression in cells of oral ectoderm clusters 1, 4, 12, and 13 at gastrula stage, though the expression of these genes was found in other clusters as well (Fig. 3). foxG expression was detected in all the four clusters at gastrula stage, and therefore does not provide a cue to distinguish clusters 12 and 1 from clusters 4 and 13 (Figs. 3 and S9B). nk2.1 and foxq2 were highly expressed in cells of clusters 9 and 17 in blastulae and gastrulae (Figs. 3; Fig. S10), and both clusters are assigned to apical and neural lineages (Fig. 1B).

gsc was highly expressed in clusters 12 and 1 but not in clusters 4 and 13 (Figs. 4 and S9B), suggesting an involvement of gsc in specification of the two former clusters (stomodeum-related ectoderm), although gsc was reportedly expressed in the apical ectoderm and the central ectoderm (Li et al., 2014). On the other hand, expression of emx and hnf6 (onecut) was higher in cells of clusters 4 and 13 than in clusters 12 and 1 (Fig. 4). Li et al. (2014) showed the involvement of emx and hnf6 (onecut) in animal ectoderm specification (not for Veg1-Ecto), and our data suggests that in clusters 4 and 13, maybe one is animal and the other vegetal, and are not distinguished by the expression profiles of emx and hnf6 (onecut) (Fig. 9B). Lim1 is proposed to be a marker of central oral ectoderm (including stomodeum) and Veg1-derived oral ectoderm (Li et al., 2014). This study showed high levels of Lim expression in cluster 12 and 1 and moderate levels in cluster 13 but undetectable levels in cluster 4 (Fig. 4). Altogether, if the lineage scheme is forced into an scRNA-seq scheme, it is likely that the stomodeal ectoderm corresponds to cluster 12, near apical to cluster 1, central to cluster 13 (Fig. 9).

Of further note is higher expression of FGFR in cells of cluster 12 in gastrulae (Fig. 4), although moderate levels of FGFR expression was found in other clusters as well (Fig. 3). This FGFR corresponds to FGFR1.
Fig. 8. Cells highly expressing Bra, FoxA or Endo16, and those co-expressing Bra and FoxA, Bra and Endo, or Bra, FoxA and Endo16 are shown as percentages of the cells among constituent cell numbers of clusters (A) 14, (B) 6 and (C) 8, respectively.
reported by Röttinger et al. (2008). On the other hand, FGF is reportedly expressed rather specifically, first in cells of cluster 13 at hatched blastula stage, and 4, 13, 16 and 19 at mesenchyme blastula stage, and 1, 4, 13, 16 and 19 at early gastrula stage (McCoon et al., 1996; Röttinger et al., 2008) (Fig. S9). Clusters 13 and 4 are SpBra-negative oral ectoderm (Fig. 1B) (Foster et al., 2020), suggesting a possible scenario in which FGF signaling is transmitted from SpBra-negative cells of clusters 4 and 13 and received by SpBra-expressing cells of clusters 12 and 1 clusters, although such details should be examined experimentally in future studies.

2.7. Genes proposed as possible downstream targets of SpBra

Rast et al. (2002) have succeeded in the identification and characterization of possible target genes of SpBra, by subtraction of mRNAs of normally SpBra expressed embryos minus SpBra morpholino-knockdown embryos. Eleven genes are listed in Supplemental Table S4, representatives of which include kakapo and gelsolin for blastopore formation, APOBEC (a cytidine deaminase) and OrCt for gut formation, and P1103, PKS (polyketide synthase), DopT and CAPK for pigment cell formation (Rast et al., 2002). If SpBra is involved in positive regulation of these genes, they might be expressed in SpBra-positive cells. Therefore, we examined here the specificity and expression level of these genes by the sc-RNAseq method. High levels, but rather broad gene expression was found for Hes, CAPK, kakapo, gelsolin, and ephx2 (Table S4; Supplementary Fig. S11). The expression level was very low and without specificity in hyalin-like, OrCt, EH1, and Nk1 (Table S4; Fig. S11). On the other hand, specific and high level of expression was...
detected in APOBEC in endoderm clusters 6 and 14 (Table S4; Fig. S11). Therefore, APOBEC provides an experimental system to explore molecular mechanisms of how SpBra activates this gene in endoderm.

3. Discussion

The present scRNA-seq deep-analysis unambiguously showed that SpBra is expressed in endoderm cells (clusters 14, 6, and 8) to form the archenteron, and in ectoderm cells (clusters 12 and 1) to form the stomodeum. Cells of cluster 12 and 1 exclusively express genes responsible for the ventral organizing activity, nearly a half of which coincidently express SpBra. This dataset thereby is well situated to identify and to prioritize exactly which cells express SpBra, and what genes SpBra regulates.

3.1. Three endoderm lineages

The endoderm of a sea urchin embryo is divided into Veg2-Endo derived from Veg2 tier and Veg1-Endo from Veg1 tier of blastulae (Fig. 5A) (Ransick and Davidson, 1998). We propose here that clusters 14 and 6 correspond to Veg2-Endo and that cluster 8 corresponds to Veg1-Endo. Veg1-Endo and Veg2-Endo differ from each other by gene expression profiles, in which Veg2-Endo exclusively exhibits expression of blimp1/krox, foxA, and gata-e (Fig. 5B) (Cui et al., 2014). hnf1 is reportedly expressed only in Veg1-Endo (Fig. 5B) (Cui et al., 2014). However, the present results showed that this gene is also expressed in Veg2-Endo as well (Fig. 2), being cautious to use hnf1 as a marker of Veg1-Endo. Veg2-Endo comprises of two clusters 14 and 6, and the two clusters are distinguished by specific expression of SLBP in cluster 14 but not in cluster 6 (Figs. 1C and 2). Based on these results, we proposed here cluster 14 as Veg2-Endo1 and cluster 6 as Veg2-Endo2 (Fig. 5A). However, this scenario should be confirmed by future studies.

It should be mentioned here that the three clusters of endoderm cells commence the expression of SpBra, hox11/13b, eve, blimp1/krox, and foxA at blastula stages, before these progenitor cells form a gut at the gastrula stage. In other words, cell fate specification of the endoderm cells occurs earlier than actual morphogenesis of gut formation, and cells precisely follow their fate to form the endoderm-derived larval structures (Davidson, 2006). This scRNA-seq study provides additional data to support this notion on the early and lineage-dependent specification mode of embryogenesis in sea urchins.

3.2. Relationship between three endoderm clusters and three gut regions

The larval digestive system is tripartite, composed of a foregut or muscular esophagus, a midgut or large spherical stomach, and a hindgut or short tubular intestine (reviewed by Annunziata et al., 2014; Annunziata et al., 2019). Therefore, a question raised here is a compositional relationship between the three embryonic cell clusters and the tripartite structure of the larval gut. Various developmentally relevant genes are dynamically expressed in each of the three regions of the larval digestive system (see Fig. 3 of Annunziata et al., 2014). These genes include Bra, foxA, Blimp1, gata-e, Endo16, Limp1, Xios, and Cdx (Olivier et al., 2008; Cui et al., 2014; Annunziata et al., 2014). First, there is a difference in...
temporal SpBra expression profile among the three clusters. Although SpBra expression commences almost simultaneously in the three clusters at hatched blastula, the gene expression was downregulated in cluster 14 and 6 by early gastrula stage, while SpBra expression was retained in cluster 8 by late gastrula stage. This transient expression profile of SpBra coincides with the result shown by Gross and McClay (2001), in which Bra protein is expressed in a circumferential region of presumptive endoderm cells neighboring the vegetal plate and the expression becomes undetectable when those same endoderm cells move inside the gastrula to form the archenteron. The last component of the tripaptite gut to invaginate is the hindgut and SpBra expression remained there (Annunziata et al., 2014). The hindgut does not express gata-e, which coincides to our result that cluster 8 exhibits less expression of gata-e (Fig. 2). Lower levels of expression of foxA and Endo16 in cluster 8 also suggests that Veg1-Endo-derived cluster 8 gives rise to the hindgut.

Midgut morphogenesis is concluded by identifying higher levels of expression of bmp1/krox, foxA, gata-e and Endo16 in clusters 14 and 6 coincident with higher levels of expression of these genes in the midgut by in situ hybridization (Annunziata et al., 2014, Fig. 5B). We posit that cluster 14 and 6 give rise to anterior and posterior parts of midgut, respectively, based on an assumption in which cluster 14 is close to the Veg2-meso lineage exhibits high levels of gata-e in blastulae and Endo16 in blastulae and early gastrulae (Fig. 1C), it is likely that a portion of cluster 11 and cluster “SMC (gut subcluster)” as well (Fig. 1B) (Foster et al., 2020) is involved in the formation of foregut. These are forerunners in archenteron invagination. However, cells of these mesoderm clusters do not express SpBra. It is likely that sea urchin gastrulation is accomplished not only by physical changes of first-invaginating cells but also by physical changes caused by SpBra in second-invaginating cells (Veg2-Endo and Veg1-Endo), that might push Veg2-meso inside the embryo.

3.3. Ectoderm cluster with SpBra expression and oral organizer activity

According to Li et al. (2015), the oral ectoderm of the sea urchin embryo is composed of four regions, stomodeal, near-apical, central, and veg1 oral ectoderm (Fig. 9A). This scRNA-seq study showed that the oral ectoderm is represented by four clusters, cluster 12 with cells of high level of SpBra expression, cluster 1 with cells of moderate level of SpBra expression, and cluster 4 and 13 with cells lacking SpBra expression (Figs. 3 and 4). Genes involved in the ventral organizer activity are expressed in cells of cluster 12 and 1 (Fig. 4). At least three results should be in mind to interpret the relationship of the ventral organizer and morphogenesis of oral invagination to form stomodeum. First, developmentally relevant genes that are involved in the ventral organizer activity such as chordin, BMP2/4 and not commence their expression at hatched blastula stage earlier than SpBra expression at gastrula stage (Figs. 3 and 4) (e.g., Lapraz et al., 2015). Second, in contrast, nearly 80% of the constituent cells in clusters 12 express chordin and not (Figs. 4 and 10), and only 25% and 40% of the cells express SpBra at early and late gastrula, respectively (Fig. 10). Third, most SpBra-positive cells are simultaneously expressing chordin and not (Fig. 10), indicating co-expression of SpBra, chordin and not in the same individual cells. The portion of cells with the three gene expression did not change when embryogenesis proceeds from early to late gastrula stage. Therefore, it is likely that (1) the ventral organizer activity is confined to cells of clusters 12 and 1, (2) the organizer activity might commence by late blastula stage, and (3) a portion of the organizer cells begin to express SpBra at gastrula stage and invaginate inside the embryo to form stomodeum. That is, the stomodeum formation is accomplished by cells with the ventral organizer activity.

In this study, we assigned cluster 1 to be the central oral ectoderm (Fig. 9B). However, it is likely that a portion of cluster 1 forms near the surface stomodeum and the other portion forms the surface ectoderm, and these questions will be answered by future studies. Our scheme of relationship between lineage-based and scRNA-seq-based specification of embryonic cells also should be examined in future studies. Nevertheless, we would like to emphasize that the ventral organizer activity resonates well with morphogenesis during stomodeum formation.

3.4. SpBra; upstream cascades and downstream targets

SpBra is expressed in two different regions of early sea urchin embryos, namely in endoderm and oral ectoderm. Because the GRN for endoderm (or endomesoderm) specification differs from GRN for oral ectoderm specification, the genetic cascade leading to SpBra likely differs between the two lineages. The genes that possibly act upstream of SpBra expression in endoderm have been studied extensively (Fig. 5B) (Rast et al., 2002; Davidson et al., 2002; Oliveri et al., 2006; Li et al., 2013; Cui et al., 2014; EchinoWiki). The present scRNA-seq analysis focused on hox11/13b, eve, bmp1/krox, foxA, gata-e, and hnf1 (Figs. 1C and 2). First, the timing of commencement of gata-e and hnf1 expression was later than that of SpBra (Fig. 2), suggesting that gata-e and hnf are not upstream components of SpBra expression (Fig. 5B). Second, hox11/13b and eve are expressed commonly in all three endoderm clusters, suggesting that these two transcription factors share upstream components of SpBra (Figs. 2 and 5B). On the other hand, the expression level of bmp1/krox and foxA in cells of cluster 6 (Veg1-Endo) was quite low compared with that in cells of clusters 14 and 6 (Veg2-Endo). This suggests that the function of bmp1/krox and foxA as upstream regulators of SpBra is limited to the cells of Veg2-Endo (Cui et al., 2014). However, so far, no studies have reported sequence-specific cis-regulatory modules upstream of sea urchin Brachyury, which is essential for future discussion of SpBra expression control in the endodermal cells.

In endoderm cells, Ben-Tabou et al. (2010) reported the presence of a SpBra-binding site on the foxA enhancer, which supports the upregulation of foxA. In the absence of Bra, foxA continued to be expressed at a low level, indicating that other transcription factors operated prior to SpBra, and as it turns out, after SpBra is extinguished from those cells. Superimposed t-SNE plot analysis indicated co-expression of SpBra and foxA in cells of clusters 12 and 1 (Fig. 6D), suggesting a possibility that foxA acts upstream of SpBra in cells of oral stomodeal region. On the other hand, foxA expression was evident in cells of clusters 12 and 1 at late blastula and gastrula stages (Fig. 4), suggesting that it could be a possible upstream regulator of SpBra (Fig. 7B). In addition, gsc is highly likely to act upstream of SpBra (Fig. 7), since gsc is specifically expressed in cells of clusters 12 and 1 (Figs. 3 and 4). chordin, bmp2/4, nodal, not and Lim are highly expressed in cells of clusters 12 and 1 but not in cells of cluster 4 and 13. How do the genes responsible for the ventral organizer interact with SpBra? Whether the expression of chordin, bmp2/4, nodal, not and Lim is essential for SpBra transcription or independent of SpBra transcription, is one of key questions to be answered to understand the molecular mechanisms involved in the specification of early sea urchin embryonic cells.

4. Conclusions

A deep-analysis of scRNA-seq datasets identified and characterized clusters of cells expressing SpBra in sea urchin embryos. SpBra expression was transient in blastulae, within cells of three clusters of endoderm and in gastrulae in cells of two of the four clusters of oral ectoderm. Presumptive endodermal cells derived from Veg2 (clusters 14 and 6) simultaneously expressed SpBra, hox11/13b, eve, bmp1/krox, foxA, gata-e, hnf1, and Endo16, whereas those derived from Veg1 (cluster 6) simultaneously expressed SpBra, hox11/13b, eve, and hnf1, but the expression of bmp1/krox, foxA, gata-e and Endo16 was present in low levels. Cells of cluster 14 and 6 are likely involved in mid-/or/foregut formation and cluster 8 in hindgut formation. Two (12 and 1) of the four clusters of oral ectodermal cells express SpBra together with chordin, bmp2/4, nodal, not, Lim1, and gsc. At a single-cell level, half of the cells with the ventral-organizer gene expression co-express SpBra, indicating that the two clusters serve as the ventral organizer and invaginate inside the
the embryo to form the stomodeum. In other words, the oral organizer activity was endowed in cells that can achieve morphogenetic movements to form the stomodeum. Two other clusters of oral ectoderm without SpBra expression give rise to the oral surface epithelium. The present scRNA-seq analysis therefore highlights a significant role of Brachyury in early morphogenesis of sea urchin embryos at single cell exposure and opens a plethora of new candidates to test for SpBra expression and function.

5. Materials and methods

All scRNA-seq data used in this study have been deposited in NCBI database under accession no. GSE149221. Clustering of cells have already been reported by Foster et al. (2020), but the dataset was reanalyzed here. In this study, CellRanger gene expression matrices were analyzed using the R (v4.1.2; http://www.R-project.org/) package Seurat v 4.0.5 (Satija et al., 2015; Hao et al., 2021). First the number of analyzed using the R (v4.1.2; http://www.R-project.org/) package Seurat v 4.0.5 (Satija et al., 2015; Hao et al., 2021). First the number of constituent cells in each cluster at five different stages was counted (Table S1).

The t-SNE (t-distributed stochastic neighbor embedding is a machine learning algorithm for visualizations) projection and clustering analysis for visualization of the integrated data such as Violin plots and Feature matrix (Fig. 7; Supplementary Figs. S4 and S8) were created using R with the package Bioconductor (v3.14) and pheatmap (v1.0.12). Numbers were converted to bar graph (Figs. 8 and 10). The heatmap and clustered matrix (Fig. 7; Supplementary Figs. S4 and S8) were created using R with the package Bioconductor (v3.14) and pheatmap (v1.0.12). Clustering for Fig. 7 was done with Ward's method (Ward, 1963), and Figs. S4 and S8 were sorted by log-normalized count of SpBra.

Author contributions

NS and GW conceived the experiment. KN, KH, HT, SM, NO and SF performed data analyses. NS and GW wrote the manuscript which was approved by all authors.

Declaration of competing interest

The authors declare no competing or financial interests.

Acknowledgements

OIST internal fund to Marine Genomics Unit (NS) supports this analysis. Members of Brown University are grateful for support from the National Institutes of Health (1R35GM140897 and 1P20GM119943).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ydbio.2022.01.005.

References


