Bacterial artificial chromosomes as recombinant reporter constructs to investigate gene expression and regulation in echinoderms

Katherine M. Buckley, Ping Dong, R. Andrew Cameron, and Jonathan P. Rast

Corresponding authors: Jonathan Rast, Emory University School of Medicine, 1462 Clifton Road NE, DSB, Room 403E, Atlanta, GA 30322, USA. Tel.: +1 404 727 5648; Fax: +1 404 727 8795; E-mail: jprast@emory.edu, Katherine M. Buckley, The George Washington University, 80022nd St. NW, Suite 6000, Washington, DC 20052, USA. Tel.: +1 202 994 9325; Fax: +1 202 994 6044; E-mail: kshank@gwu.edu

Abstract

Genome sequences contain all the necessary information—both coding and regulatory sequences—to construct an organism. The developmental process translates this genomic information into a three-dimensional form. One interpretation of this translation process can be described using gene regulatory network (GRN) models, which are maps of interactions among regulatory gene products in time and space. As high throughput investigations reveal increasing complexity within these GRNs, it becomes apparent that efficient methods are required to test the necessity and sufficiency of regulatory interactions. One of the most complete GRNs for early development has been described in the purple sea urchin, *Strongylocentrotus purpuratus*. This work has been facilitated by two resources: a well-annotated genome sequence and transgenes generated in bacterial artificial chromosome (BAC) constructs. BAC libraries played a central role in assembling the *S. purpuratus* genome sequence and continue to serve as platforms for generating reporter constructs for use in expression and regulatory analyses. Optically transparent echinoderm larvae are highly amenable to transgenic approaches and are therefore particularly well suited for experiments that rely on BAC-based reporter transgenes. Here, we discuss the experimental utility of BAC constructs in the context of understanding developmental processes in echinoderm embryos and larvae.

Key words: gene regulatory network; transgene; homologous recombination; large insert cloning

Introduction

As model organisms, sea urchins and other echinoderm species have yielded many important discoveries in the fields of developmental biology (e.g. pronuclear fusion during fertilization [1, 2]), cell biology (e.g. the prediction [3] and discovery [4] of cyclins), immunology (e.g. phagocytosis of foreign objects by mesenchymal cells [5]) and, more recently, regulatory biology [6] (reviewed in [7]). Echinoderms entered the genomic era with the

sequencing of the genome from the purple sea urchin (*Strongylocentrotus purpuratus*; [8]). This milestone genome was the first sequenced from a marine invertebrate and the largest non-mammalian assembled genome (814 Mb). To assemble this sequence, higher coverage whole-genome shotgun sequencing (WGS) was combined with lower coverage sequencing of a series of tiled bacterial artificial chromosomes (BACs) [8]. For this technique, known as clone-array pooled shotgun sequencing

Katherine M. Buckley is a Research Associate at the Department of Biological Sciences in the George Washington University. She uses recombinant bacterial artificial chromosome (BAC) constructs to investigate gut-associated immune responses in sea urchin larvae.

Ping Dong is a member of the research staff at the California Institute of Technology and has extensive expertise in BAC library construction, BAC clone mapping and BAC clone recombination to construct reporter transgenes.

R. Andrew Cameron is the Director of the Beckman Institute Center for Computational Regulatory Genomics at the California Institute for Technology, which is responsible for generating and maintaining echinoderm BAC libraries.

Jonathan P. Rast is a Senior Scientist at the Emory University School of Medicine. He generated the BAC libraries for the sea urchin genome project and has considerable experience in BAC recombination.

(CAPPS) [9], low-coverage sequence was generated from about 8000 ordered and oriented BACs to provide a structural framework onto which the WGS sequences could be confidently assembled [10]. As the purple sea urchin exhibits substantial regions of repetitive sequences and high heterozygosity and an outbred animal was used for the genome sequence, the CAPPS strategy was particularly useful for assembling this genome. Moreover, the BAC libraries that were originally generated as genomic resources continue to serve as invaluable tools for functional genomics.

BACs were developed as one of several strategies designed to increase the limited size of the inserts that could be cloned into genomic vectors. The first of these systems was yeast artificial chromosomes (YACs) [11]. Yeasts have very efficient double-stranded DNA break and repair pathways, which facilitate rapid YAC construction; YAC vectors are capable of carrying genomic fragments of up to 1Mb. However, these linear artificial chromosomes are fragile and difficult to separate from endogenous yeast DNA. Thus, new bacteria-based systems were developed using more stable vectors based on sequences from the P1 bacteriophage (P1-derived artificial chromosomes) or the functional fertility plasmid (F-plasmid) from the bacterium *Escherichia coli* (BACs) [12, 13]. BAC vectors can contain inserts of up to 350 kb, exhibit a reduced recombination rate compared to YACs and are significantly easier to manipulate and isolate. The introduction of phage-based homologous recombination systems (discussed in detail below) [14] further increases the efficiency of generating and modifying specific BAC sequences for functional genomic analyses.

BAC clones are particularly useful for experimental assays in echinoderm embryos, because sea urchins are highly amenable to transgenic manipulation. In this system, exogenous DNA fragments microinjected into zygotes are incorporated into the genomic DNA and replicated with the endogenous genome during mitosis. Before incorporation, the exogenous DNA fragments concatenate such that many copies of the transgene are inserted into a random site within the genome [15]. Although microinjection is typically performed within minutes of fertilization, individual blastomeres can be injected at up to the four-cell stage as an experimental chimeric platform [16, 17]. This method can be used to efficiently analyze many constructs, as thousands of zygotes can be injected in one sitting with nearly 100% incorporation. Because incorporation occurs during mitosis, transgenic sea urchin embryos are chimeric such that BACs are present in a proportion of cells (e.g. half the cells if incorporation occurs during first cleavage but only a quarter of the cells if the BAC is incorporated during second cleavage). Experimental evidence indicates that exogenous DNA is most commonly integrated during the 4th or 5th cleavages but may occur up to 7th cleavage (1/128 of cells) [18, 19].

The generation of transgenic embryos has served as the foundation for understanding the regulatory circuitry that controls embryonic development [20, 21]. Injecting constructs that couple predicted regulatory regions with sequences encoding reporters (e.g. fluorescent proteins) is the central method to assess the capacity of exogenous DNA segments to control transcription [15, 18]. Originally, this method relied on relatively small regions of regulatory DNA (3–5 kb) that were cloned into plasmid vectors. In embryos transgenic for artificial DNA constructs, often all or part of the spatial and temporal expression patterns of the endogenous sequences are recapitulated [22, 23]. However, these experiments are limited by the maximum insert size of standard plasmids. In the purple sea urchin genome, the average intergenic distance is 30 kb [8], and, as in other systems, regulatory regions can span large regions of genomic DNA [24].

Table 1. Arrayed BAC library resources available for echinoderms

Species	Library name	Estimated genome coverage (X) ^a	Characterized BACs	Recombinant BACs
<i>Strongylocentrotus purpuratus</i> ^b	Small ^c	4	311	196
	Large ^d	13		
	Mbol ^a	19		
	FR2 ^d	10		
<i>Alloccentrotus franciscanus</i>		n.d.	19	4
<i>Lytechinus variegatus</i>	Small ^c	5	82	9
	Large ^d	19		
<i>Paracentrotus lividus</i>		19	3	1
<i>Arbacia punctulata</i>		14	12	9
<i>Eucidaris tribuloides</i>		14	1	0
<i>Patiria (Asterina) miniata</i>	Small ^c	6	10	8
	Large ^d	33		
<i>Ptychodera flava</i>		n.d.	0	0

^agDNA inserts were prepared by Mbol restriction enzyme digestion and size-selected (150 kb) using pulsed-field gel electrophoresis (PFGE).

^bLibraries from each species were generated from sperm DNA isolated from a single male.

^cGenomic DNA (gDNA) inserts were prepared by EcoRI restriction/methylation and size-selected (50 kb) using PFGE.

^dgDNA inserts were prepared by EcoRI restriction/methylation and size-selected (150 kb) using PFGE. FR2 is a technical repeat of the Large *S. purpuratus* library.

^eGenome coverage is estimated from the genome size, the number of plates and the intended insert size from the library preparation. n.d., not determined; genome size estimates are unavailable for *Ptychodera flava* and *Alloccentrotus fragilis*.

In contrast to plasmids, the *S. purpuratus* BAC libraries harbor much larger insert sizes (on average, 140 kb; Table 1) [25]. As methods were developed to engineer specific BAC-based reporter constructs using sequences that encode fluorescent proteins or other tags, this became the construct of choice for characterizing gene regulatory logic and for efficient, precise localization of gene expression. Experiments to assess the fate of transgenic BAC DNA in sea urchin embryos indicate that most DNAs injected as large insert clones remain intact following genomic integration (J. Rast, unpublished data).

Here, we present a summary of echinoderm BAC resources and review the experimental strategies that have been used to investigate regulatory biology in echinoderm embryos and larvae. We do not present detailed protocols but instead refer to those that have been published elsewhere. The efficiency of transgenesis in many echinoderm embryos, in combination with these BAC library resources and recombination-based strategies for precise alteration of large insert clones, provides a wealth of approaches for reporter construction, genetic manipulation and regulatory analysis.

Echinoderm BAC libraries

As a resource for the Sea Urchin Genome Project [10], the purple sea urchin BAC libraries were constructed from the sperm of a single animal (the same animal from which the genomic sequence was generated) using several cloning strategies and target insert sizes into the pBACe3.6 vector [26]. Since then, additional BAC libraries have been generated from the same *S. purpuratus* individual to extend the original coverage and also from genomic DNA isolated from seven other echinoderm species (Table 1). The clones from these libraries provide an invaluable experimental platform that can be used in the following ways: (1) as a stable source of genomic DNA to confirm repetitive and/or complex regions of the assembled genome sequence

and (2) as starting material to generate specific constructs for *in vivo* localizing transcript expression, analyzing gene regulation or rescuing protein function (Table 2). These BAC libraries from *S. purpuratus*, as well as similar libraries from *Lytechinus variegatus*, have been used to identify regulatory regions crucial to the generation of the gene regulatory network model describing endomesodermal development in early embryogenesis of these species [21].

Details of the library coverage, construction and screening methodology can be found in Cameron et al. [25]. Several features of the pBACe3.6 vector used for these libraries are relevant for constructing reporter constructs using homologous recombination. The insert is flanked by *NotI* sites that can be used to release the insert from the vector for the majority of BAC clones. Data from digesting randomly selected BAC indicate that 10–15% of *S. purpuratus* BAC inserts (from the 150 kb library) contain endogenous *NotI* sites. In these cases, reporters can be linearized for microinjection using another rare-cutting restriction enzyme (two closely spaced sites *AscI* are present in pBACe3.6 near the cloning site). Alternatively, the pBACe3.6 vector contains a *PI-SceI* homing endonuclease recognition site that can be used to linearize the construct without affecting the insert. This low-copy vector confers antibiotic resistance to chloramphenicol.

Identifying BACs of interest using bioinformatics

In addition to traditional hybridization-based techniques for screening BAC libraries [25, 47, 57], BAC inserts that encompass genes of interest can be identified computationally. In the course of the ongoing *S. purpuratus* genome sequencing project, 7734 sequenced BAC clones (eBACs) that were used in the original CAPPS protocol have been mapped to the assembly (version 3.1). These BAC inserts overlap with 221 726 exons in 13 566 gene models (of 29 130 total). Additional information on the distribution and location of the BAC inserts was obtained by sequencing single ends (BAC-ends) of 182 536 randomly selected BACs. These sequences are also mapped to the genome assembly (v3.1). The positions of these sequences are shown on the genome browser hosted at Echinobase (www.echinobase.org) in relationship to other genomic features (e.g. genes or transcripts of interest). This ‘clone by computer’ method has greatly reduced the time to identify and use a genomic region for gene discovery or regulatory function studies.

BACs serve as a stable, invariant sequence for confirming the genome assembly

Echinoderms are outbred animals that exhibit high levels of heterozygosity (an estimated 4–5% in *S. purpuratus* [58]). Inbred lines are not currently available for any echinoderm species, which may reflect high numbers of deleterious recessive alleles [59]. Following their use as a framework for the original *S. purpuratus* genome assembly [8], BAC clones continue to provide additional resolution for complex regions that could not be properly assembled (e.g. the *SpTransformer* genes [51]) or as second sources of linkage information for regions with tandemly repeated genes (e.g. the unusually organized sea urchin *hox* gene cluster or the unexpectedly linked *Rag1* and *Rag2* genes [47, 50]; Table 2). These clones are particularly valuable as they originate from the same genomic DNA used for the genome assembly and thus have identical sequence. In a minority of cases, insert structure can also be confirmed using other sources of information, such as comparison with genomic DNA blots generated

using pulsed-field gene electrophoresis. These comparisons are sometimes complicated by the fact that a small percentage of genomic regions appear to be unstable in large BAC clones (about 3% of genes could not be recovered in large BAC inserts in attempts to isolate clones containing 96 transcription factor homologs). There is not a necessarily clear relationship, however, between genomic complexity and stability in BAC inserts (e.g. some clones with tandem arrays of highly similar genes appear to be stable). Furthermore, genes of interest for which clones are difficult to identify in large insert BAC libraries are often recoverable in BACs that harbor smaller inserts; small BAC libraries (average insert size is 50 kb) are available for several echinoderm species (Table 1).

Recombineering: methods and pitfalls

Large insert clones provide more realistic genomic backgrounds for experimental analysis of regulatory control. However, precisely manipulating sequences in this context presents challenges. For example, restriction digests, ligation and transformation are significantly more complex for BAC clones relative to smaller plasmids. These difficulties are overcome through the use of homologous recombination (reviewed in [60, 61]). Recombineering (recombination-mediated genetic engineering), which has been refined over the past two decades, yields precisely modified BAC inserts by inducing the expression of phage-derived recombination machinery in bacteria that harbor both the BAC clone of interest and a ‘recombination cassette’ that is used to guide the modifications. Many strategies can be used to generate BAC-based reporter constructs; a general overview is shown in Figure 1.

The region of interest (this is typically an exon but may also be intronic or regulatory sequence) is targeted using short (50–200 nucleotide) ‘arms’ that are identical to the sequence flanking the portion of BAC that will be eliminated. Although short homology arms (30–50 nt) are sufficient to effect recombination [62], longer arms increase the efficiency of this process [63]. A recombination cassette is constructed in which the recombination arms flank the sequence to be added (e.g. coding sequence for a fluorescent protein) as well as a sequence used for selection (e.g. antibiotic resistance). The inserted sequence often includes a complete coding sequence, including a stop codon and a 3′ untranslated region (e.g. SV40 3′ UTR) for messenger RNA (mRNA) stability. The BAC of interest and the recombination cassette are introduced into one of several bacterial strains modified for this technique using electroporation. Here, we will describe EL250 and SW105, although several others are available (reviewed in [61]).

These two strains are derivatives of DY380, a modified DH10B strain that contains a defective prophage λ with the recombination genes *bet*, *exo* and *gam* under the transcriptional control of the temperature-sensitive *cI857* repressor [64]. In this system, expression of the prophage recombination system is transiently induced by raising the temperature from 32 °C to 42 °C for 15 min, which de-represses the *bet*, *exo* and *gam* genes. This tightly controlled system reduces off-target effects and viability problems that result from high *gam* expression. DY380 also contains a tetracycline resistance cassette. EL250 is a modified DY380 strain that includes an arabinose-inducible flippase (*flpe*) gene. This FLP recombinase mediates specific recombination between two flippase-recognition target (FRT) sites [65]. This offers one strategy to remove the selection marker from recombinant BAC clones (Figure 1). Here, the selection gene in the recombination cassette is flanked by FRT sites. During FLP-FRT

Table 2. Recombinant BACs have been used in a variety of experimental contextsRecombinant BACs used to localize expression^a

Gene	FP location	Experimental strategy	Ref.
<i>SpAlx1</i>	GFP at TSS	Fluorescence, qPCR	[27]
<i>SpBlimp1/krox</i> , <i>LvBlimp1/krox</i>	GFP at TSS	WMISH and fluorescence	[28]
<i>SpCyclinD</i>	GFP at TSS	Fluorescence	[29]
<i>SpDelta</i> , <i>SpHesC</i>	GFP at TSS	Fluorescence	[30]
<i>SpE-protein</i>	GFP at TSS; mCherry at start of transcription	Fluorescence, qPCR, WMISH, used to mark expression in perturbed cells	[31]
<i>SpFoxa</i> , <i>SpNodal</i>	GFP/RFP at TSS	Fluorescence co-expression; used as a marker in tandem with deletion mutants	[32]
<i>SpFoxY</i>	GFP at TSS	Fluorescence	[33]
<i>SpGatac</i> , <i>SpScl</i>	Rescue	Rescue protein function; used to mark expression in perturbed cells	[34]
<i>SpGcm</i>	GFP fusion (exon 3)	Reporter construct for regulatory analysis	[35]
<i>SpGatae</i> , <i>LvGatae</i>	GFP at TSS	Fluorescence	[36]
<i>SpH2A.Z</i>	GFP/RFP translational fusion	Fluorescence	[37]
<i>SpHesC</i>	GFP at TSS	Fluorescence; used to mark expression in perturbed cells	[38]
<i>SpNodal</i>	GFP at TSS	Fluorescence	[39]
<i>SpOnecut</i>	GFP at TSS	WMISH and fluorescence	[40]
<i>PmOtxβ1/2</i>	GFP at TSS	Fluorescence	[41]
<i>SpTbrain</i>	GFP at TSS	Fluorescence	[42]
<i>SpTbx2/3</i> , <i>SpDlx</i>	GFP at TSS	Fluorescence	[43]
Regulatory analyses			
Gene	Strategy		Ref.
<i>SpAlx1</i>	GalK-positive/negative selection for seamless removal of predicted transcription factor binding sites		[27]
<i>SpGatae</i> , <i>LvGatae</i>	Flippase-mediated CRM deletions		[36]
<i>SpH2A.Z</i>	Deletion-based regulatory analysis		[37]
<i>SpHesC</i>	Regulatory analysis in the BAC (mutation of TFBS)		[38]
<i>LvNodal</i> , <i>PlNodal</i>	BAC clones provided constant regulatory region for CRM amplification and mutation		[44]
<i>SpOnecut</i>	Flippase-mediated CRM deletions		[40]
<i>SpSix1/2</i>	Reporter construct for regulatory analysis		[35]
<i>SpTbrain</i>	Flippase-mediated CRM deletions; GalK positive/negative selection for seamless removal of predicted transcription factor binding sites		[42]
	Validation of GRN circuitry by misexpression of <i>gcm</i>		[45]
BACs used to confirm gene organization			
Gene	Species		Ref.
<i>Blimp1/krox</i>	<i>S. purpuratus</i> , <i>L. variegatus</i>		[28, 46]
Hox cluster	<i>S. purpuratus</i>		[47, 48]
Parahox cluster	<i>P. miniata</i>		[49]
<i>Rag1/2</i>	<i>S. purpuratus</i>		[50]
<i>SpTransformer</i>	<i>S. purpuratus</i>		[51, 52]
Evolutionary analyses			
Gene	Strategy		Ref.
<i>Alx1</i>	<i>L. variegatus</i> BAC used to confirm genome and in phylogenetic footprinting		[27]
<i>Blimp1</i> , <i>bra</i> , <i>otx</i> , <i>wnt8</i>	Cross-species BAC injections (<i>S. purpuratus</i> and <i>L. variegatus</i>) to assess evolutionary flexibility of CRMs		[53]
<i>H2A.Z</i>	Injected <i>S. purpuratus</i> BAC into <i>L. variegatus</i>		[37]
Other techniques			
	BACs used as a fluorescence source for FACs and transcriptomic analysis		[54, 55]
	High-throughput identification of CRMs using barcoding		[56]

^aSpecies names are indicated by initials preceding gene name; *Lv*, *L. variegatus*; *Pm*, *P. miniata*; *Pl*, *Paracentrotus lividus*; *Sp*, *S. purpuratus*.

Figure 1: Homologous recombination can be used to generate BAC-based fluorescent reporter constructs. Two strategies for modifying BAC sequences through homologous recombination are shown; specific protocols are detailed in Sharan et al. [61]. A BAC is identified that contains the gene or region of interest (indicated by pink rectangles) using either bioinformatics or hybridization-based library screens. BAC stocks are typically maintained in DH10B *Escherichia coli* strains. A recombination cassette is designed in which homology arms (narrow, light blue) flank the exon or sequence to be replaced. The BAC and recombination cassette are introduced via electroporation into a modified bacterial strain (either EL250 or SW105; shown as green ovals) that contains the defective prophage λ in which the recombination machinery is repressed. Raising the temperature to 42°C de-represses these genes to initiate recombination. To identify recombinant clones, the recombination cassette may include an antibiotic resistance gene (e.g. kanamycin resistance; kan^R) flanked by FRT sites (orange ovals). The antibiotic resistance is removed upon arabinose-dependent activation of the *flp* recombinase. Alternatively, a two-step process in which *galK* expression is first used to positively select recombinant clones grown on minimal media and then used as a means of negative selection to seamlessly construct BAC recombinants. (A colour version of this figure is available online at: <https://academic.oup.com/bfg>.)

recombination, the intervening sequence is removed, and only a single FRT site remains.

An alternative recombinering strategy employs SW105 cells (EL250Δ*galK*; Figure 1) [66]. In this two-step process, two recombination cassettes are constructed using identical homology arms. In the first cassette, the arms flank a sequence encoding galactokinase (*galK*). This enzyme catalyzes two reactions involved in galactose metabolism that enable positive and negative selection steps [67]. This *galK* recombination cassette

is introduced with the target BAC into SW105 cells, and the temperature-controlled recombination machinery is activated. Bacteria are grown on minimal media in which galactose is the only carbon source [66]. Colonies containing recombinant BACs (*galK*⁺) are able to initiate galactose metabolism by phosphorylating galactose to galactose-1-phosphate, which is then further metabolized into glucose. The second step utilizes an additional recombination cassette in which the homology arms flank the sequence encoding a reporter protein or otherwise modified sequence. This cassette is introduced and recombined as above, but negative selection is used to identify recombinant (*galK*⁻) BACs. Galactokinase also catalyzes the phosphorylation of the galactose analog 2-deoxy-galactose into the toxic compound 2-deoxy-galactose-1-phosphate, which cannot be further metabolized and leads to cell death [67]. Since the homology arms are identical during each recombination, this strategy generates 'seamless' BAC recombinants with no remaining selection cassettes or FRT sites.

Functional genomics using BACs

BACs as reporter constructs

One of the most powerful uses of recombinant BACs is to trace gene expression in living embryos or larvae. Other techniques for localizing transcripts (*in situ* hybridization) or proteins (immunohistochemistry) involve a fixation step that destroys fine morphological details. Consequently, imaging embryos *in vivo* provides far superior resolution for localizing gene expression. This is particularly true in echinoderms, which have optically transparent embryos and larvae that can be imaged at single-cell resolution. For this purpose, recombinant BACs can be generated as described above by inserting the coding sequence of a fluorescent protein into one of several locations within the gene of interest: (1) the start of transcription [31, 42], (2) start of translation [27, 32, 34, 37, 40, 41] or (3) into an exon to generate a fusion protein [68] or to localize a specific isoform [28, 31] (Table 2). When linearized and microinjected into fertilized eggs, recombinant BACs integrate into the genomic DNA and express the sequence encoding the fluorescent protein according to the regulatory information contained within the BAC, thereby faithfully recapitulating the expression pattern of the endogenous gene.

The Center for Computational Regulatory Genomics (CCRG) has generated a total of 108 recombinant BACs from 6 echinoderm species (www.echinobase.org; Table 1). These BACs reproduce the expression patterns of 66 different genes (mostly transcription factors) using 5 different fluorescent proteins [green fluorescent protein (GFP), red fluorescent protein, mCherry, Cerulean or the photoswitchable Dendra]. For some BACs, the Cerulean protein is fused to the H2B histone, which localizes the fluorescent signal to the nucleus [69]. Stocks of these clones are maintained by the CCRG and are freely available on request.

An elegant implementation of this strategy is to insert multiple fluorescent proteins into a single BAC construct to assess alternative splicing or differential isoform expression (Figure 2). This technique has been used for the analysis of an alternatively spliced bHLH transcription factor expressed in the purple sea urchin known as *E-protein* (an ortholog of the vertebrate *HEB*, *E2A* and *E2-2* proteins) [31]. This gene is expressed as two splice variants: a longer form known as *E-Can* and an alternative, shorter form (*E-Alt*). The *E-Can* transcript is the product of 22 exons within the *E-protein* locus. Transcription of *E-Alt* initiates at a short, unique (*Alt*) exon located between exons 8 and 9 that is spliced to the remaining 13 exons (10–22) of *E-Can*. This large

Figure 2: Recombinant BACs can be generated using multiple fluorescent proteins to trace expression of alternatively spliced genes. An example of this technique is shown based on the strategy in Schrankel et al. [31]. Exons that are specifically expressed in distinct isoforms serve as targets for recombination using fluorescent proteins. Through this strategy, 'double-fluorescent protein' BAC constructs can be generated such that GFP indicates expression of isoform 1, whereas RFP specifically reflects expression of isoform 2. (A colour version of this figure is available online at: <https://academic.oup.com/bfg>.)

region of shared sequence interferes with *in situ* hybridization strategies to localize the two transcripts. Using homologous recombination, a 'double-fluorescent protein' BAC construct was generated such that the GFP coding sequence replaced the 1st exon of *E-Can*, and the mCherry coding sequence replaced the Alt exon. Image analysis of embryos injected with this single-BAC construct revealed that, whereas *E-Can* (GFP) is expressed ubiquitously, *E-Alt* (mCherry) expression is restricted to mesodermal immune cells. Furthermore, although the two isoforms are initially co-expressed during mid-blastula stage, this double-fluorescent BAC enabled the visualization of two mutually exclusive expression domains [31].

BACs as rescue constructs

Recombinant BACs also provide an experimental strategy to rescue gene expression in functional perturbation experiments (Figure 3). One common technique to interfere with protein function in echinoderm embryos is the use of Morpholino antisense oligonucleotides (MASOs) [70]. These oligonucleotides anneal to the transcript of interest either near the start of translation (which prevents protein synthesis) or near a splice site (which inhibits correct splicing to cause a premature stop codon). To demonstrate causality between an observed phenotype and perturbation of a specific gene, a rescue experiment can be performed in which the gene of interest is experimentally reintroduced in the context of the perturbation. Often, however, this is done in an unregulated manner by injecting mRNA or using a promoter that drives ubiquitous expression at much higher levels than the endogenous gene, which may complicate interpretation of results, particularly for genes expressed at low

levels in few cells. Here, recombinant BACs provide an elegant solution. BACs can be engineered in which the MASO target site is mutated to prevent annealing (Figure 3A). This technique was used to analyze the function of the transcription factors *gata1/2/3* and *scl* in mesodermal differentiation in the sea urchin embryo [34]. To rescue the phenotypic consequences of *gata1/2/3* perturbation, embryos were microinjected with the following: (1) a MASO that interfered with *gata1/2/3* translation; (2) a recombinant BAC that expressed GFP under the regulatory control of *gata1/2/3* to trace cells that endogenously express this transcript (which was resistant to the effects of the MASO) and (3) a recombinant BAC in which the *gata1/2/3* MASO binding site was mutated (the 'rescue' BAC) or an arbitrary control BAC. Analysis of transgenic embryos indicate that when fluorescent signal was present in developing mesoderm (indicating that the BAC integrated correctly to rescue expression; Figure 3B), in the presence of the rescue BAC, development proceeded normally. In contrast, in embryos that either received the control BAC or in which fluorescence was not evident in the mesoderm, development was impaired [34]. Similar results were obtained in parallel experiments that targeted *scl*. Although this strategy may be insufficient to rescue phenotypes for widely expressed genes, it has been successful for genes with limited expression domains.

Using BACs to establish and validate gene regulatory network circuitry

Recombinant BACs are invaluable tools for elucidating the transcription factor inputs that activate, maintain and repress expression of a specific gene, a process known as gene regulatory network (GRN) analysis. The genomic program of development can be described by a series of interacting networks of regulatory genes that lead to the elaboration of structure and more cell types over time [71]. BAC engineering provides an efficient method to identify and manipulate *cis*-regulatory modules (CRMs). Specific transcription factor binding sites can be assayed by mutating these sites and analyzing the subsequent effects on spatiotemporal expression of the reporter gene relative to wild type [42]. Alternatively, deletion of complete CRM sequences can be achieved using homologous recombination to replace the region of interest with an antibiotic resistance gene. Flanking the resistance gene with *frt* sites enables its elimination through the inducible expression of FLP recombinase [42]. Using a recombinant BAC is particularly advantageous for performing this type of experiment in the highly outbred echinoderm species. Individual purple sea urchin genomes exhibit an estimated 4–5% heterozygosity [72], which is significant when analyzing a large (up to 30 kb) regulatory region. In contrast to the variable environment generated by modifying the germ line DNA directly, BACs provide a constant sequence for experimental manipulation.

The regulatory causality that is experimentally established with the aid of recombinant BACs is used to form the network models that describe development. Circuitry within these networks can be validated *in vivo* using BAC-based misexpression experiments [45]. For example, GRN analysis predicted that the aboral mesoderm (pigment cell precursors) program was initiated by the transcriptional activation of the transcription factor *gcm* as a consequence of Delta/Notch (D/N) signaling. The GRN indicated that D/N signaling was the only time-dependent *gcm* input and that this factor was sufficient to induce the pigment cell fate. To test this *in vivo*, a recombinant BAC was generated in which the coding sequence of *gcm* replaced the first exon of *tbrain* (*tbr*). *Tbr* is exclusively expressed in the primary

Figure 3: Recombinant BACs can be used to rescue protein function in the context of endogenous perturbation. (A) Constructing reporter and rescue BACs. The *GeneX* structure is shown within the original BAC (coding sequence, red; noncoding sequence, white). The most common technique for functional perturbation in echinoderm embryos is the use of MASO reagents [70], which often anneal to the transcript near the start of translation (indicated by blue line). To generate a MASO-resistant BAC, the annealing site can be modified by introducing a complementary oligonucleotide containing point mutations into the EL250 or SW105 cells and activating the recombination machinery as shown in Figure 1. The recombination frequency is much higher than introducing a larger DNA fragment and recombinants can be isolated using a pooled polymerase chain reaction strategy. Based on the location of the fluorescent protein, these reporter BACs may or may not be resistant to the MASO. If not, they can be modified as described above (dashed green arrows). (B) An experimental strategy for perturbation and functional rescue of *GeneX*. The endogenous expression pattern of hypothetical *GeneX* is shown in red [in mesenchyme blastula sea urchin embryos, this territory corresponds to aboral non-skeletal mesoderm (NSM)]. Microinjection of the *GeneX::gfp* BAC recapitulates the *GeneX* expression pattern but is mosaic so is only expressed in a portion of the endogenous pattern. In this context, striped cells express both endogenous *GeneX* and the GFP reporter. In contrast, MASO perturbation affects embryos globally (indicated by light pink cells). Protein function can be rescued in a mosaic subset by co-injecting the *GeneX::MASO^R* BAC (blue), which is resistant to MASO perturbation, as well as a *GeneX::gfp* BAC to trace integration. Here, the endogenous *GeneX* translation is blocked by the MASO, but GFP and *GeneX* are expressed from the BAC clones in a portion of the endogenous territory (striped cells). (A colour version of this figure is available online at: <https://academic.oup.com/bfeg>)

mesenchyme cells (PMCs; skeletogenic lineage). Endogenous *tbr* expression begins ~8 hpf, which is several hours prior to *gcm* (which is expressed beginning ~12 hpf). The temporal delay between *tbr* and *gcm* expression provides the opportunity to synthetically 'rewire' the network by artificially inducing ectopic *gcm* expression and monitoring expression levels of downstream genes. The recombinant BAC (designated *tbr::GCM*) was used in combination with a *tbr* GFP reporter BAC (*tbr::GFP*) to trace transgenic cells. The results of this experiment not only confirmed the GRN prediction (*gcm* expression was sufficient to induce pigment cell differentiation genes) but also revealed a previously unknown repressive function (activation of *gcm* led to a downregulation of the skeletogenic program in GFP⁺ cells [45]). This experimental strategy highlights the utility of recombinant BAC constructs to not only express reporter genes but also express other functional sequences.

High-throughput GRN analysis using BACs

The purple sea urchin uses the majority of the transcription factors encoded in its genome during the first 72 h of development [73]. The recombinant BACs discussed here offer efficient high-throughput opportunities for parallel cis-regulatory analyses. A recent case study on a diverse suite of previously unstudied

sea urchin effector genes expressed in skeletogenic cells illustrates this approach [56]. Positively active cis-regulatory modules in each BAC containing a gene of interest were located in single Nanostring quantification experiments through the use of 'barcode' tag expression vectors of which >100 can be analyzed simultaneously. When the DNA sequences that drive expression are narrowed down in this way, computational analysis based on the known skeletogenic regulatory state can lead to the effective identification of functional target site clusters. Subsequent deletion of these subregions from the parent BACs revealed the necessity for a particular module, whereas simultaneous tests of the same regions in short construct revealed sufficiency. The predicted functional inputs are then confirmed by site mutations. The described steps are all conducted in multiplex formats, which permit analysis of a cohort of genes at once. These experiments revealed the simple conclusion that each effector gene utilizes a small subset of inputs from the previously well-defined skeletogenic GRN.

BACs in comparative genomics

Finally, cross-species transfer of recombinant BACs provides a novel mechanism to understand the evolution of regulatory

networks [53]. The CCRC maintains BAC libraries from eight species that span a range of evolutionary distances (Table 1). Several studies indicate that CRM architecture can change dramatically even among closely related species [74]. This is illustrated by an analysis using recombinant BACs from two sea urchin species, *S. purpuratus* and *L. variegatus*, which shared a last common ancestor about 50 million years ago [75]. Recombinant BACs were generated for four transcription factors (two from each species) and micro-injected into fertilized eggs from the opposite species. Among embryos with fluorescent signal, 89–100% exhibited the correct spatial expression domains [53]. This level of conservation may be somewhat higher than other genes, as the genes tested have regulatory functions that are evolutionarily stabilized within the network architecture.

Conclusions: BAC clones in an era of genome editing

Large-insert BAC libraries were integral in assembling the first complex animal genome sequences. Even in an era of rapid electronic access to genome sequences and *in vivo* genomic modification strategies (e.g. the CRISPR/Cas9 system) [76, 77], these libraries remain valuable experimental resources, particularly for outbred, polymorphic organisms. BAC clones provide a constant source of DNA that can be used to assemble repetitive genomic regions or as a defined haplotype in regulatory analysis. Homologous recombination in the context of a BAC is more tightly controlled than genome editing, which eliminates potentially confounding off-target effects. This is particularly important when working with nonmodel organisms or those that lack a sequenced genome, in which the specificity of guide RNAs cannot be predicted computationally. Finally, introducing an exogenous BAC construct to analyze gene expression using a reporter provides a sufficiently large genomic context necessary for investigating gene regulatory control but does not affect the regulatory state of the endogenous genes. In fact as CRISPR/Cas9 systems become commonplace, engineered transgenic BACs offer a complementary tool to provide highly regulated reporters or wild-type controls as well as modified gene products for testing in the context of altered genetic backgrounds. Thus, BAC clones continue to provide stable, defined resources that enable novel experimental pursuits in a wealth of genetically polymorphic organisms to shed light on fundamental questions in developmental and evolutionary biology.

Key Points

- Bacterial artificial chromosome (BAC) libraries were integral in the sequencing and assembly of the purple sea urchin genome. These resources continue to aid in genomic analysis of echinoderm species even as sequencing strategies evolve.
- BACs provide reproducible and stable sources of DNA for experiments in polymorphic animals.
- Homologous recombination can be used to construct functional reporter transgenes with little foreknowledge of gene regulatory control mechanisms.
- Recombinant BACs serve as platforms to precisely modify regulatory DNA in the context of large genomic regions and to express modified coding regions under strict genetic control.
- Data from these experiments can be used to efficiently generate gene regulatory network models.

Funding

The Center for Computational Regulatory Genomics is supported by funding from the National Institute of Child Health and Human Development (P41HD071837 to R.A.C.).

References

1. Hertwig O. Beiträge zur Kenntniss der Bildung, Befruchtung und Theilung des thierischen Eies. *Morphol Jahrb* 1876;1: 347–434.
2. Clift D, Schuh M. Restarting life: fertilization and the transition from meiosis to mitosis. *Nat Rev Mol Cell Biol* 2013;14: 549–62.
3. Hultin T. The effect of puromycin on protein metabolism and cell division in fertilized sea urchin eggs. *Experientia* 1961;17: 410–1.
4. Evans T, Rosenthal ET, Youngblom J, et al. Cyclin: a protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division. *Cell* 1983;33:389–96.
5. Metchnikoff E. Lectures on the comparative pathology of inflammation: delivered at the Pasteur Institute in 1891. London: Kegan Paul, 1893.
6. Oliveri P, Carrick DM, Davidson EH. A regulatory gene network that directs micromere specification in the sea urchin embryo. *Dev Biol* 2002;246:209–28.
7. Ernst SG. Offerings from an urchin. *Dev Biol* 2011;358:285–94.
8. Sodergren E, Weinstock GM, Davidson EH, et al. The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science* 2006; 314:941–52.
9. Cai WW, Chen R, Gibbs R. a, et al. A clone-array pooled shotgun strategy for sequencing large genomes. *Genome Res* 2001; 11:1619–23.
10. Cameron RA, Mahairas G, Rast JP, et al. A sea urchin genome project: sequence scan, virtual map, and additional resources. *Proc Natl Acad Sci USA* 2000;97:9514–8.
11. Murray AW, Szostak JW. Construction and behavior of circularly permuted and telocentric chromosomes in *Saccharomyces cerevisiae*. *Mol Cell Biol* 1986;6:3166–72.
12. Shizuya H, Birren B, Kim UJ, et al. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci USA* 1992;89:8794–7.
13. Monaco AP, Larin Z. YACs, BACs, PACs and MACs: Artificial chromosomes as research tools. *Trends Biotechnol* 1994;12: 280–6.
14. Thomason LC, Sawitzke JA, Li X, et al. Recombineering: genetic engineering in bacteria using homologous recombination. *Curr Protoc Mol Biol* 2014;106:1.16.1–39.
15. McMahon AP, Novak TJ, Britten RJ, et al. Inducible expression of a cloned heat shock fusion gene in sea urchin embryos. *Proc Natl Acad Sci USA* 1984;81:7490–4.
16. Peterson RE, McClay DR. Primary mesenchyme cell patterning during the early stages following ingress. *Dev Biol* 2003; 254:68–78.
17. Coffman JA, Coluccio A, Planchart A, et al. Oral–aboral axis specification in the sea urchin embryo. *Dev Biol* 2009;330: 123–30.
18. Hough-Evans BR, Britten RJ, Davidson EH. Mosaic incorporation and regulated expression of an exogenous gene in the sea urchin embryo. *Dev Biol* 1988;129:198–208.
19. Guay CL, McQuade ST, Nam J. Single embryo-resolution quantitative analysis of reporters permits multiplex spatial cis-regulatory analysis. *Dev Biol* 2017;422:92–104.

20. Materna SC, Oliveri P. A protocol for unraveling gene regulatory networks. *Nat Protoc* 2008;**3**:1876–87.
21. Davidson EH, Rast JP, Oliveri P, et al. A genomic regulatory network for development. *Science* 2002;**295**:1669–78.
22. Flytzanis CN, Britten RJ, Davidson EH. Ontogenic activation of a fusion gene introduced into sea urchin eggs. *Proc Natl Acad Sci USA* 1987;**84**:151–5.
23. McMahon AP, Flytzanis CN, Hough-Evans BR, et al. Introduction of cloned DNA into sea urchin egg cytoplasm: replication and persistence during embryogenesis. *Dev Biol* 1985;**108**:420–30.
24. Spitz F. Gene regulation at a distance: from remote enhancers to 3D regulatory ensembles. *Semin Cell Dev Biol* 2016;**57**:57–67.
25. Cameron RA, Rast JP, Brown CT. Genomic resources for the study of sea urchin development. *Methods Cell Biol* 2004;**74**: 733–57.
26. Frengen E, Weichenhan D, Zhao B, et al. A modular, positive selection bacterial artificial chromosome vector with multiple cloning sites. *Genomics* 1999;**58**:250–3.
27. Damle S, Davidson EH. Precise cis-regulatory control of spatial and temporal expression of the *alx-1* gene in the skeletogenic lineage of *S. purpuratus*. *Dev Biol* 2011;**357**:505–17.
28. Livi CB, Davidson EH. Expression and function of *blimp1/krox*, an alternatively transcribed regulatory gene of the sea urchin endomesoderm network. *Dev Biol* 2006;**293**:513–25.
29. McCarty CM, Coffman JA. Developmental cis-regulatory analysis of the cyclin D gene in the sea urchin *Strongylocentrotus purpuratus*. *Biochem Biophys Res Commun* 2013;**440**:413–8.
30. Smith J, Davidson EH. Gene regulatory network subcircuit controlling a dynamic spatial pattern of signaling in the sea urchin embryo. *Proc Natl Acad Sci USA* 2008;**105**:20089–94.
31. Schrankel CS, Solek CM, Buckley KM, et al. A conserved alternative form of the purple sea urchin HEB/E2-2/E2A transcription factor mediates a switch in E-protein regulatory state in differentiating immune cells. *Dev Biol* 2016;**416**:149–61.
32. Ben-Tabou de-Leon S, Davidson EH. Information processing at the *foxa* node of the sea urchin endomesoderm specification network. *Proc Natl Acad Sci USA* 2010;**107**:10103–8.
33. Materna SC, Swartz SZ, Smith J. Notch and Nodal control forkhead factor expression in the specification of multipotent progenitors in sea urchin. *Development* 2013;**140**:1796–806.
34. Solek CM, Oliveri P, Loza-Coll M, et al. An ancient role for Gata-1/2/3 and Scl transcription factor homologs in the development of immunocytes. *Dev Biol* 2013;**382**:280–92.
35. Ransick A, Davidson EH. Cis-regulatory logic driving glial cells missing: self-sustaining circuitry in later embryogenesis. *Dev Biol* 2012;**364**:259–67.
36. Lee PY, Nam J, Davidson EH. Exclusive developmental functions of *gatae* cis-regulatory modules in the *Strongylocentrotus purpuratus* embryo. *Dev Biol* 2007;**307**:434–45.
37. Hajdu M, Calle J, Puno A, et al. Transcriptional and post-transcriptional regulation of histone variant H2A.Z during sea urchin development. *Dev Growth Differ* 2016;**58**:727–40.
38. Smith J, Davidson EH. Regulatory recovery in the sea urchin embryo and the stabilizing role of fail-safe gene network wiring. *Proc Natl Acad Sci USA* 2009;**106**:18291–6.
39. Nam J, Su Y-H, Lee PY, et al. Cis-regulatory control of the nodal gene, initiator of the sea urchin oral ectoderm gene network. *Dev Biol* 2007;**306**:860–9.
40. Barsi JC, Davidson EH. cis-Regulatory control of the initial neurogenic pattern of *onecut* gene expression in the sea urchin embryo. *Dev Biol* 2016;**409**:310–8.
41. Hinman VF, Nguyen A, Davidson EH. Caught in the evolutionary act: precise cis-regulatory basis of difference in the organization of gene networks of sea stars and sea urchins. *Dev Biol* 2007;**312**:584–95.
42. Wahl ME, Hahn J, Gora K, et al. The cis-regulatory system of the *tbain* gene: alternative use of multiple modules to promote skeletogenic expression in the sea urchin embryo. *Dev Biol* 2009;**335**:428–41.
43. Ben-Tabou de-Leon S, Su Y-H, Lin K-T, et al. Gene regulatory control in the sea urchin aboral ectoderm: spatial initiation, signaling inputs, and cell fate lockdown. *Dev Biol* 2013;**374**: 245–54.
44. Range RC, Lapraz F, Quirin M, et al. Cis-regulatory analysis of nodal and maternal control of dorsal-ventral axis formation by *Univin*, a TGF-beta related to *Vg1*. *Development* 2007;**134**: 3649–64.
45. Damle SS, Davidson EH. Synthetic in vivo validation of gene network circuitry. *Proc Natl Acad Sci USA* 2012;**109**:1548–53.
46. Smith J, Kraemer E, Liu H, et al. A spatially dynamic cohort of regulatory genes in the endomesodermal gene network of the sea urchin embryo. *Dev Biol* 2008;**313**:863–75.
47. Martinez P, Rast JP, Arenas-Mena C, et al. Organization of an echinoderm Hox gene cluster. *Proc Natl Acad Sci USA* 1999;**96**: 1469–74.
48. Cameron RA, Rowen L, Nesbitt R, et al. Unusual gene order and organization of the sea urchin hox cluster. *J Exp Zool B Mol Dev Evol* 2006;**306**:45–58.
49. Annunziata R, Martinez P, Arnone M. Intact cluster and chordate-like expression of *ParaHox* genes in a sea star. *BMC Biol* 2013;**11**:68.
50. Fugmann SD, Messier C, Novack LA, et al. An ancient evolutionary origin of the *Rag1/2* gene locus. *Proc Natl Acad Sci USA* 2006;**103**:3728–33.
51. Miller CA, Buckley KM, Easley RL, et al. An *Sp185/333* gene cluster from the purple sea urchin and putative microsatellite-mediated gene diversification. *BMC Genomics* 2010;**11**:575.
52. Oren M, Barela Hudgell MA, D'Allura B, et al. Short tandem repeats, segmental duplications, gene deletion, and genomic instability in a rapidly diversified immune gene family. *BMC Genomics* 2016;**17**:900.
53. Cameron RA, Davidson EH. Flexibility of transcription factor target site position in conserved cis-regulatory modules. *Dev Biol* 2009;**336**:122–35.
54. Barsi JC, Tu Q, Davidson EH. General approach for in vivo recovery of cell type-specific effector gene sets. *Genome Res* 2014;**24**:860–8.
55. Barsi JC, Tu Q, Calestani C, et al. Genome-wide assessment of differential effector gene use in embryogenesis. *Development* 2015;**142**:3892–901.
56. Nam J, Dong P, Tarpine R, et al. Functional cis-regulatory genomics for systems biology. *Proc Natl Acad Sci USA* 2010;**107**: 3930–5.
57. Multerer KA, Smith LC. Two cDNAs from the purple sea urchin, *Strongylocentrotus purpuratus*, encoding mosaic proteins with domains found in factor H, factor I, and complement components C6 and C7. *Pharmacologia* 2004;**89**:9–106.
58. Grula JW, Hall TJ, Hunt JA, et al. Sea urchin DNA sequence variation and reduced interspecies differences of the less variable DNA sequences. *Evolution (NY)* 1982;**36**:655–76.
59. Anderson D, Hedgecock D. Inbreeding depression and growth heterosis in larvae of the purple sea urchin *Strongylocentrotus purpuratus* (Stimpson). *J Exp Mar Bio Ecol* 2010;**384**:68–75.
60. Court DL, Sawitzke JA, Thomason LC. Genetic engineering using homologous recombination. *Annu Rev Genet* 2002;**36**: 361–88.

61. Sharan SK, Thomason LC, Kuznetsov SG, et al. Recombineering: a homologous recombination-based method of genetic engineering. *Nat Protoc* 2009;4:206–23.
62. Zhang Y, Buchholz F, Muirers JP, et al. A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat Genet* 1998;20:123–8.
63. Yu D, Ellis HM, Lee EC, et al. An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci USA* 2000;97:5978–83.
64. Lee EC, Yu D, Martinez de Velasco J, et al. A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* 2001;73:56–65.
65. Schlake T, Bode J. Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci. *Biochemistry* 1994;33:12746–51.
66. Warming S, Costantino N, Court DL, et al. Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res* 2005;33:e36.
67. Alper MD, Ames BN. Positive selection of mutants with deletions of the gal-chl region of the *Salmonella* chromosome as a screening procedure for mutagens that cause deletions. *J Bacteriol* 1975;121:259–66.
68. Buckley KM, Ho ECH, Hibino T, et al. IL17 factors are early regulators in the gut epithelium during inflammatory response to *Vibrio* in the sea urchin larva. *Elife* 2017;6:e23481.
69. Kanda T, Sullivan KF, Wahl GM. Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. *Curr Biol* 1998;8:377–85.
70. Materna SC. Using morpholinos to probe gene networks in sea urchin. *Methods Mol Biol* 2017;1565:87–104.
71. Davidson EH. Emerging properties of animal gene regulatory networks. *Nature* 2010;468:911–20.
72. Britten RJ, Cetta A, Davidson EH. The single-copy DNA sequence polymorphism of the sea urchin *Strongylocentrotus purpuratus*. *Cell* 1978;15:1175–86.
73. Howard-Ashby M, Materna SC, Brown CT, et al. Gene families encoding transcription factors expressed in early development of *Strongylocentrotus purpuratus*. *Dev Biol* 2006;300:90–107.
74. Hinman VF, Nguyen AT, Cameron RA, et al. Developmental gene regulatory network architecture across 500 million years of echinoderm evolution. *Proc Natl Acad Sci USA* 2003;100:13356–61.
75. Smith AB, Pisani D, Mackenzie-dodds J. a, et al. Testing the molecular clock: molecular and paleontological estimates of divergence times in the Echinoidea (Echinodermata). *Mol Biol Evol* 2006;23:1832–51.
76. Lin C-Y, Su Y-H. Genome editing in sea urchin embryos by using a CRISPR/Cas9 system. *Dev Biol* 2015;1–9.
77. Oulhen N, Wessel GM. Albinism as a visual, in vivo guide for CRISPR/Cas9 functionality in the sea urchin embryo. *Mol Reprod. Dev* 2016;83:1046–7.