UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
On Monsters, Patterns and Appendages

Permalink
https://escholarship.org/uc/item/89s7x7tc

Author
Alberstat, Erin Jarvis

Publication Date
2018

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

On Monsters, Patterns and Appendages:
Targeted CRISPR/Cas9 genome editing in the amphipod crustacean *Parhyale hawaiensis* to explore Hox gene interactions in the establishment of appendage diversity

By

Erin Jarvis Alberstat

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Integrative Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Nipam H. Patel, Chair
Professor Michael W. Nachman
Professor John Ngai

Summer 2018

On Monsters, Patterns and Appendages

© 2018

by

Erin Jarvis Alberstat

Abstract

Targeted CRISPR/Cas9 genome editing in the amphipod crustacean *Parhyale hawaiensis* to explore Hox gene interactions in the establishment of appendage diversity

Erin Jarvis Alberstat

Doctor of Philosophy in Integrative Biology

University of California, Berkeley

Professor Nipam H. Patel, Chair

The diversity of animals is a topic of childlike fascination, poetic wonder, and scientific discovery; but how did such great diversity of form and function come to be? To uncover the possible mechanisms by which diverse new body plans are generated over evolutionary time, we turn to understanding the patterns and processes behind how individual bodies are constructed — and deviated — during embryonic development. Critically important to the specification of body plan diversity is the patterning of the anterior-posterior axis and establishment of regional identities by the Hox genes.

Hox genes have a revolutionizing history, complete with striking transformation of form upon their mutation. The rapid discovery of orthologous sets of Hox genes in a diversity of species led to the exciting realization that not only are individual Hox genes conserved across Bilateria, so too are their patterns of expression, mutation, arrangement, and properties in aggregate—changes to which have been implicated as important generators of evolutionary diversification. This apparent conservation enabled many useful generalizations to be gleaned from a limited number of representative species, and paradigms such as posterior prevalence became widespread. Increasingly, however, new technologies enable us to expand the scope of our study to a wider array of species and cases breaking the "rules" of these paradigm are accumulating. This warrants the comprehensive study of Hox genes in aggregate in a greater number of representative species to resolve a fuller picture of the Hox complex and its interactions during the specification of regionalized bodies and the diversification of specialized forms.

By resolving Hox expression patterns in transformed individuals following Hox gene knockout in the appendage-diverse amphipod crustacean *Parhyale hawaiensis*, we have revealed a number of cross-regulatory interactions among Hox genes that influence the patterning of Hox domains. Furthermore, by amassing a fuller picture of the Hox expression underlying the appendage transformations, we make a more informed model of the functional roles and interactions of the different Hox genes in the designation of limb identities. We have shown that posterior prevalence is not sufficient to predict the appendage transformations in *Parhyale* and that a modular 'Hox code' expressed in individual segments functions to establish segmental identity. We believe this study demonstrates that interactions among Hox genes provide a mechanism for the diversification of appendage morphologies and arrangements over evolutionary time.

Dedication

Bryan, my husband, for all of his love, support and pride; and Caspian, who's been with me the most.

You are my family, no matter what, and each have a piece of my heart.

I'll cherish the adventures we share always.

Table of contents

Chapter 0. Segmentation of Dissertation Bauplan (Preface)

Chapter 1. Introduction: Hox Genes and the DiversiĀicaŧon of Regionalized Animal Forms

Chapter 2. Characterizing Hox gene expression in the emerging model *Parhyale hawaiensis*

- Comprehensive analysis of Hox gene expression in the amphipod crustacean *Parhyale hawaiensis* (Dev Bio 409 (2016) 297–309)

Chapter 3. Linking the putative *Parhyale* Hox gene cluster: chromosomal co-localization by double fluorescent in-situ hybridization

- The genome of the crustacean Parhyale hawaiensis, a model for animal development, regeneration, immunity and lignocellulose digestion (eLife 2016;5:e20062)

Chapter 4. Exploring the roles of individual Hox genes in the establishment of diverse appendage types (CRISPR/Cas9 knockout of individual Hox genes)

- CRISPR/Cas9 Mutagenesis Reveals Versatile Roles of Hox Genes in Crustacean Limb Specification and Evolution (Current Biology 26, 14–26)

Chapter 5. Exploring the interactions among Hox genes in the establishment of diverse appendage types (CRISPR/Cas9 knockout of multiple Hox genes simultaneously)

- Hox-Hox Manuscript (unpublished)

Chapter 6. Concluding remarks: Next steps and Where Does it all Fit in?

References

Segmentation of Dissertation Bauplan (Preface)

Here I present my dissertation work on the Genetic interactions among Hox genes in the specification of segmental and appendage diversity in the amphipod crustacean *Parhyale hawaiensis*. The following dissertation can be "tagmatized" into three main sections: Chapter one introduces the significance and study of embryological patterning, setting an exciting stage of advancements in the study evolutionary developmental biology and mapping where my small contributions integrate into the larger context for understanding embryonic development and evolutionary diversification of biological form. Chapters 2 - 5 follows the emergence of *Parhyale* as a model for the study of appendage diversification in arthropods, interweaving four publications from which I contributed to characterizing *Parhyale* Hox genes and their patterns of expression (Chapters 2 and 3), function (Chapter 4), and interactions (Chapter 5) in the specification of appendage identity. Finally, I give a summary of where my dissertation fits into the greater picture (Chapter 6). Appendexis include additional narration, media attachments, and protocols/materials that I hope that some will find useful.

Chapter 1: Sets the story of the evolution of animal forms, embryonic patterning, and contemporary history of "pre-Hox" thoughts on hereditary, evolution, and development.

Chapter 2: Introduces to you the amphipod crustacean model *Parhyale hawaiensis* and gives a comprehensive sketch of *Parhyale* Hox genes and their patterns of expression.

Chapter 3: Linking the *Parhyale* Hox Cluster and a collaborative genome & repository of resources

Chapter 4: CRISPR/Cas9 in *Parhyale* for the first time and the most comprehensive Hox gene study on a single non-model at the date of its compilation. I am proud to have had a substantial role in this. The beginnings of my dissertation project incorporated here.

Chapter 5: My major contributions on interactions among Hox genes in establishing appendage identities.

Chapter 6: Concluding remarks

Acknowledgements

No one does science alone, and I have been lucky to have such a great number of people in my life to guide me and inspire, to support me gladly, and to accept me just the way I am.

Bryan Alberstat, my husband, for your proud excitement and support, and **Caspian**, for being with me throughout it all. Someone once said that while each dog in one's life will be special in their own unique way, there will always be *that special one* to stand out from them all; I'm both happy and sad to know that you are that one.

Mom & Dad, my biggest advocates. You supported me always and let me be me, even when that "me" was a little bit weird and the paths that I chose eccentric. You provided both a boost and a safety net, and it is this that any "bravery" (ranging from intentional naivety to rash decision making) is founded. **Dad**, for pushing me to work hard and do well because you really and truly believed in me and my potential; making you proud was a major driver for achievement when I was too young to realize I wanted such things on my own. You gave me a sense that I could do anything grounded in the firm reality that I must work hard for my aspirations to come true. Through both reason and example, you instilled a strong sense of radical self reliance and work ethic, and although the budding tips of our ideologies are not always aligned, the values you imbue will always be a core part of who I am (and foundation for any deviance to a radically "Bezerklian" identity). Your strong sense of integrity remains a constant guide. **Mom**, who protected me fiercely and has been there for me always. You gave me the parts of me I like most and are most important to who I am — a penchant for novelty (and an inability to settle in one place too long), a sometimes inexplicable resiliency (founded in the refusal to be brought down by circumstance), and unbounded excitement (unsullied by worldly realism). **My siblings**, to whom I share a special closeness and understanding that can only be achieved through shared genetics and life history, and yet this is a very different thing for each of you. **Heather**, my sister; for your rock solid support no matter what, undeserved admiration since we were little, and constant forgiveness of my shortcomings. I share a special closeness with you, in mindset and experience. **Chris**, my brother; whose brilliance is revealed in glimpses, and whose opus I can't wait to one day read. We share a certain philosophical something and I cherish the conversations we have, albeit few and far between. **Aunt Mary** for also being so proud. I looked up to you from the start, and aspire to be more like you in many ways. You loved me then as you do now. **Cate** and **Lindsay French**, my most favoritest cousins. I remember throwing you on my shoulders and now you're both more adults than I am. I resonate in the "weirdness" of Cate and am in constant awe of Lindsay's abilities. And **Kevin**; you're pretty cool, too — thanks for putting up with us Jarvis girls.

Becky! There for me always. The voice of reason who loves me not in spite of but because of precisely who I am. You assure me that my quirks are my best features and you make me more me. Thanks for your continued advice even though I never listen. You are that one most specialist friend who can never be replaced.

Nipam; I'm truly, truly lucky to have serendipitously fallen into your lab and cannot imagine a better advisor. You've created and allowed so many things for me (museum exhibits and science communication, Russia, Japan, Africa, and a very generous number of conferences). You let me follow my distractions while trying to keep my progress in check. You know my failures, my inabilities to follow through, my stumbling in discussions and presentations, and my stark holes in knowledge and competence, and yet for some reason you keep putting your trust in me. Your confidence in me fuels my confidence to reach to that potential. You introduced me to Woods Hole & Evo-Devo and now I'm hooked. Thank you. **My Lab Mates**; I couldn't have asked for a more likeable and genuinely supportive lab family. **Arnaud Martin**; You taught me welcoming collaboration and infectious enthusiasm, and have become for me both a mentor and a friend. **Heather Bruce**; my lab sista; you've always totally got my back and are like the most wanting to help out your fellow comrade-ist-person ever. **Ryan Null**; for being super awesome and ever helpful. **Rachel Thayer**, your humble competence and covert powerfulness are an inspiration. **Angela Kaczmarczyk**; for always being so kindly helpful. **Dennis Sun, Aaron Pomerantz, Kyle DeMarr,** and **Jenny McCarthy**; You guys are all awesome and I'm glad I got to get to know you! My core group of amazing undergraduates **Jennifer Wang**, **Liam O'Connell**, **Panda Kreidler** and **Kevin Chung**; Thanks for putting up with me as a mentor — I couldn't have done all this work without you! Each of you is special in your own *very* unique way. **Jacques Bothma**; your assistance in opening my blender was instrumental in earlier years. **Matt Ronshaugen**; for making everything just … better. You're one of the good ones.

My committee **Michael Nachman** and **John Ngai** — Thanks for everything and for making all my last minute thesis deadlines happen! **Tom Carlson** and **David Presti**; it was such a pleasure and an honor to teach for you, and I learned so many interesting things. **Kevin Padian**, for showing me Darwin; You're an enigma that I'm both terrified and fascinated by. My former mentors; I have been so lucky that each of you has been caring, instructive, and fantastic to work with. **Joe Polacco**, my first research advisor, for being such a smart, funny, and crass character; you're definitely one of a kind and may we always keep in touch. **Tippi MacKenzie**, you're an inspiration, a role model, and a friend; thanks for telling me I needed to go to graduate school. **Rebecca Abergel**, for always being supportive and letting me stalk professors down the hill. **Bob Spencer**; for being my voice when I was too timid to stand up for the rats myself. **Chris Todd**; for teaching me molecular biology and Canadianisms.

And of course to **Craig**; for being a grounding beacon of happiness as the end and all of its stresses drew near ;)

And to all the *Parhyale*, whether you wanted to be my experimental subjects or not.

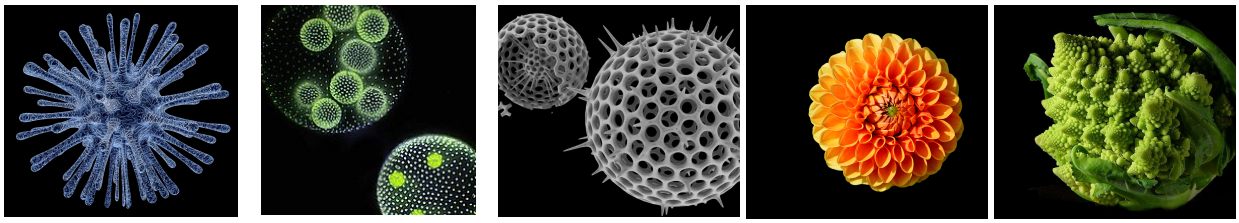
And finally, I thank the person(s) instrumental but who I've failed to list. Not because you were not instrumental but because … of course I'm going to accidentally leave out someone obvious and amazing and feel really terrible for doing so.

Chapter 1: Hox Genes And The Diversification Of Regionalized Form

From so simple a beginning, endless forms most beautiful, and most wonderful, have been, and are *being, evolved.*

-Charles *Darwin "On the Origin of a Species"*

Our planet is filled with a kaleidoscope of lifeforms, ever changing shapes and patterns reflecting beautiful symmetries and form. Plants branch out in a multitude of complexities and sizes, giant redwoods, spine-covered cactuses, brightly colored carnivorous plants, crawling vines, floating lily pads, and fields of little clovers, housing leaves and petals of a medley of shapes and arrangements from simple to complex to whorles spiraling into astoundingly fractal-like patterns. Increasingly advanced microscopy reveals the striking geometries and individual details of tiny life forms belonging to an exquisite microworld of viruses, volvox, radiolarians, and so many other micro-lifeforms with intricate structural designs otherwise invisible to our eye.



Virus (Public Domain) | Volvox by Frank Fox (CC BY-SA 3.0) | Radiolarian by Äickleandfeckled (CC2.0) | Dahlia (Public Domain) | Broccoli fractals by Jon Sulliva (Public Domain)

Fig 1.1. Diversity of Structures & Patterns. Microorganisms (A-C) and Land Plants (D-E)

But of all of life's great diversity, it is often the diversities of animal form that we are most drawn to and most adept at distinguishing. It is when observing this diversity that we have a particularly heightened awareness — and appreciation — and capacity for discrimination. From this cornucopia arise the characters of our myths, our poetry, our art, and our story; and it is this kingdom to which we ourselves belong despite our colloquial separation founded in a worldview that is inescapably humanocentric. Nautilus and jellyfish, giraffes and manta rays, elephant shrews and walking sticks, platypus and komodo dragons, earthworms and hyenas — whether lauded or forgotten, imagined or never found, the animal clade is comprised of a medley of life forms so peculiar that one cannot help but resonate with Sean Carroll's distillation of wonder in his book "Endless forms most beautiful,"

"If these creatures did not exist, they would be almost too incredible to believe[1]."

[1] "Endless Forms Most Beautiful" — Sean B. Carroll

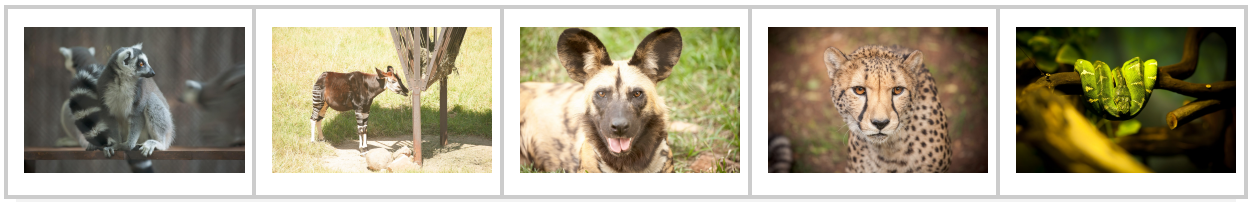


Fig 1.2. Diversity of Animal Forms
Authors photos while performing NSF-sponsored conservation genomics research in South Africa

How did such great diversity of form and function come to be, each represented by a distinct "kind" of animal whose very members are recognized through their shared uniformity of type? There are a myriad of distinct groupings of animals, each representing a particular form (or set of forms) that is strictly adhered to by its related members, rather than a cacophony of individuals blending into one another. Indeed, the formation of the shapes and structures of all these different organisms is one of the most fundamental questions we have about the natural world, and is a topic that must be addressed from the perspectives of multiple fields.

"It's survival of the fittest!", you may be tempted to answer, "Specialized life forms evolve from the natural selection of body plan variations and novel structures that arise from the accumulation of random mutations in our DNA!". But while studying the patterns of genetic variation among different organisms is useful for informing relatedness and building phylogenetic trees (and can certainly guide our inquiry), it does not tell us *how* the changes in genotype (revealed by DNA sequencing) led to the changes in phenotype (physical attributes) that we see, or even which genotypic mutations are responsible for which phenotypic variations. In order to attribute function to a gene (or verify its role predicted by corollary studies), the function of the gene must be experimentally tested by removing its influence (e.g., by breaking the gene or its transcript) or "misexpressing" it when or where it should not be. One can then extrapolate gene function by observing what happens as a result of its loss or gain.

Even as more and more genes are identified, however, and their functions discovered, the relationship between genotype and phenotype is not 1:1, making our comprehension of an organism from its DNA incredibly difficult to decipher. Who you are is determined not simply by the genes you have, but in how they are expressed. Rather than simply being present or absent to "do a thing" or not, genes are regulated and the ultimate impact that a gene has on an organism is dependent on when, where, and to what degree a gene is activated. New structures and body plan configurations can therefore be established not only by changing the genes or altering their function, but by changing their regulation and shifting the patterns of their expression.

And when it comes to building an organism, no genes, patterns or processes are as important, nor changes to as impactful, as those involved in the individual development of each new organism. Thus, in order to understand how different body plans are diversiĀied over evolutionary time, we must first understand how different bodies are built during embryonic development.

Part I. From So Simple a Beginning; Endless Forms Most Beautiful

A brief primer in the embryonic development and regionalization of a complex multicellular organism

Each and every one of us[2] starts life as a single cell, a diploid fusion of two gametes (egg and sperm), which must divide into a specific number more of cells that must grow, differentiate, and arrange themselves just so to form the particular size, shape, and organization of tissues that cell was "fated" to become according to its unique set of genetic instructions handed down and recombined with each successive generation. The initial zygote (aptly named from *zugōtos* 'yolked') is undifferentiated and totipotent, and has the capacity to give rise to all cell types required to build the particular organism it is to become. A series of rapid cellular divisions in the absence of significant growth (cleavage) gets the (non-proverbial) embryonic ball of cells rolling. Slight asymmetries in the embryo (which at some point forms some version of a more hollow ball of cells or "blastula" (*blastos* 'sprout')) polarize its development along three directional axes, initiating the determination of positional identities that establish which "ends" of the embryo are to become the back and front (dorsal-ventral) and head and tail (anterior-posterior). The burgeoning embryo is progressively regionalized by patterns of cells whose fates and specific identities become more and more determined with each series of "decisions". This beginning of shape is **morphogenesis**. Distinct lineages of cells differentiate into increasingly restricted cell types and three germ layers form during gastrulation, which give rise to the specific types of somatic tissues and organs: the ectoderm (e.g., epidermis and nervous system), mesoderm (e.g., muscles and connective tissue), and endoderm (e.g., soft organs). The bulk of allometric growth and specific anatomical details are added in the final stages of embryonic development, which ends at varying levels of completion in respect to the final adult form. The newly developed life form can range from an immature adult of matching form all set to go and grow, a premature version with varying degrees of development still remaining, or an altogether different form that must undergo one or several transformational rounds of development (metamorphosis) before it reaches the mature form.

Axial Patterning & The Shape of Things to Come

Embryonic development progresses from general to complex so that only in later stages are the characteristic attributes of type prominently distinguishable. But while the specific details of structures and distinctive features of an organism appear to be shaped very late in development, their particular arrangements and structural fates were in fact determined

[2] The non-inclusive, sexually reproducing triploblastic bilateral-normative version of "us"

early on by mechanisms that coordinate the positional organization and fates of initially equivalent cells that will take on complex forms and functions (Ball 2011, page 261-290).

Morphogenesis proceeds by a variety of molecular, chemical, and mechanical mechanisms that organize the spatial distributions and specific cell types during embryonic development. The initial embryonic axes are "patterned" by chemical and molecular mechanisms that are genetically controlled (and therefore subject to encoded variation and diversification) that can include the selection of cell types by cycles of systematic stochasticity, positional field sensing and graded cellular responses to extracellularly diffused gradients of signalling molecules ("morphogens"), and the coordinated control of cell fates by a core set of regulatory genes that are themselves patterned in time and space.

Axial patterning and the progressively complex organization and specification of cell fates proceeds by a series of hierarchical decisions heavily informed by relative positional "awareness" of cells and fields of cells and by the progressive subdivision of positional cues. The progressive subdivision of positional coordinates and fate specification can be achieved by cascades of inductive processes initiated by pre-localized determinants and asymmetries (e.g., maternally-deposited factors and sperm entry)(Rivera-Pomar and Jäckle 1996; St Johnston and Nüsslein-Volhard 1992; Nüsslein-Volhard and Wieschaus 1980) (Spemann and Mangold 1924; (Translated by Spemann and Mangold 2001)) and/or by intrinsic systems of self-organization and emergent patterning (Reviewed by Gerhart 1999; Meinhardt 2001; De Robertis 2006; Deglincerti et al. 2016).

Classic models for patterning include: Turing's reaction-diffusion model (Turing 1952) and Lewis Wolpert's "French flag" model (whereby cell state is selected by field sensing and threshold responses to discrete ranges of morphogen concentrations (Wolpert 1969)). Classic studies in *Drosophila melanogaster* show that polarity of the AP axis is initiated by gradients of the maternal effect genes Bicoid (the first identified morphogen), Hunchback, Nanos, and Caudal, which set off a cascade of progressively-regulated/refined patterns of zygotically expressed **segmentation genes** (Gap genes → Pair-rule genes → segment polarity genes) that determine the number, size, and polarity of segments (Nüsslein-Volhard and Wieschaus 1980). The Gap and Pair-rule genes establish the patterns of a set of **homeotic genes** — the Hox Genes — that specify segment identity (Lewis 1978). Hox genes are perhaps the most iconic of a core set of "master" regulatory genes that specify cell state through their regulation of a cascade of downstream gene targets and regulatory pathways, and are the focus of this dissertation.

So important are the core regulatory genes in setting up a proper embryo, that they are highly conserved among disparate organisms (Reviewed by St Johnston and Nüsslein-Volhard 1992; McGinnis and Krumlauf 1992). Deviations to the core mechanisms responsible for axial patterning and regional specification result more often than not in early termination of an embryo or in the birth of a grotesque "monster" unfit (in terms of natural selection) for passing along any deviant genes or aberrant regulation encoded by

genetic mutation. It is these deviations, however, the rare few that offer an advantage that sticks, that are the fodder for the macroevolutionary diversification of form and function.

And it is these developmental anomalies, both those that occur naturally and those that are experimentally contrived, that inform our understanding of how an entire organism is formed from a single cell and a packet of genetic instructions. In other words, we can understand how bodies are built by understanding the mechanisms behind the different ways in which they break. And by understanding how individual forms are built and how they are "broken", we can better understand how mutational changes in an organism's DNA can lead to macroevolutionary change of an actual organism.

Selection of regional identities

The Hox genes are a very special set of clustered homeotic genes that contain a region of DNA called the "Homeobox", which encodes homeodomain transcription factors that regulate gene transcription by binding specific regions of DNA (most often a promoter region) to block or activate the transcription of their target gene. Homeodomains are encoded by genes involved in morphogenic patterning and direct the development of specific anatomical structures by acting on specific target sequences — typically switching on a cascade of downstream genes, during embryonic development (Corsetti et al. 1992; Abzhanov and Kaufman 1999; Graba et al. 1997). Homeodomains are encoded by many different genes throughout the genome, though they were first identified in, and most rigorously studied in, homeotic and segmentation proteins in *Drosophila*.

Hox genes regionalize animal forms by acting as modular "switches" that select region-specific identities through the transcriptional activation and suppression of downstream gene targets within the boundaries of their expression in partially overlapping domains along the AP axis (Reviewed in Hughes and Kaufman 2002). If an organism were a house, and embryonic development its construction, Hox genes would be the foremen in charge of directing construction of the specific rooms after the initial framing of the house (i.e., the establishment of the embryonic axis). It may be the final details added by the interior designer that really makes the kitchen and bathrooms "pop", but you can't very well add a venetian soaking tub to a room erected as a bedroom and left without water pipes and a proper drain.

SpeciĀically, this thesis addresses the genetic and functional interactions among the different Hox genes, and how these shape regional identities (specifically, what type of appendage goes where in a crustacean model), and how shifts in their expression boundaries may have led to the diversification of animal forms. But before stepping through the hours to weeks of embryonic patterning that shape a complex multicellular organism from a single cell, let's take a quick run through of roughly 3+ billion years of evolving forms that sculpted a complex multicellular life form from a very different sort of single cell.

A brief primer in the evolution of complex multicellular lifeforms

The rise of complex multicellularity

Life began as simple unicellular microorganisms that lacked complex cellular components (the Prokaryotes). The growing populations of Bacteria and Archaea did, however, roll out the carpet that would support this advancement. The unlikely fusion of one of these into another (Margulis 1970) led to the subfunctionalization of cellular compartments bound within membranes (e.g. a membrane bound nucleus to house the chromosomes and a membrane bound mitochondria whose role could be devoted to the production of energy), launching an entire new domain of organisms (the Eukaryotes) able to support — both energetically and structurally — complex multicellular configurations begetting specialized tissues and coordinated roles.

Primitive multicellular entities emerged from the colonization of single-celled organisms and from the fusion of cells dividing incompletely. Among the assorted collections of shape-shifting amoebas and amorphous slimes (simple multicellular aggregations found to exhibit collective cellular "intelligence" and complexity in their own right (Dickinson et al. 2011)), a handful of complex, 3-dimensional multicellular life forms took shape, ready to divide and conquer and fabricating a conceptual line between a multicellular individual and a community of cells.

Seaweeds and land plants arose from different groups of algaes (red, green and brown) and like their unicellular relatives, harness the sun to produce their own energy. Mushrooms are the most iconic (and visible) multicellular structures of fungi, a major clade of principle ecological decomposers that also includes single-celled yeasts and multinucleated molds that propagate outward as networks of filamentous threads. While phylogenetically distinct from slime molds, algaes and other major taxons, the life histories, life cycles and life styles of fungi are often deeply (and quite literally in the very forms they take) intertwined with other groups of life whose energy they directly absorb. And then came the metazoans — the animals. Least by number and last to arrive, animals evolved 600 million years ago from a colonial flagellate that came from a unicellular flagellate — free swimming organisms wielding lasso-like flagellum used for locomoting, sensing, and capturing bacteria that are a shared last common ancestor with an extant group of choanoflagellates actively studied (King 2004; Brunet and King 2017) to learn more about our transition into the most complex multicellular arrangements to be packaged in discrete, autonomous, mobile units.

The mobile animals

All animals are motile (meaning that they can move about on their own), if only for a brief stage during their life cycles; sponges and barnacles, for example, swim about in their immature, larval form, while certain jellyfish get their starts on life as sessile polyps. Locomotion is central to the life histories of most animals, providing a key strategy for

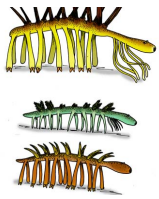
acquiring food, avoiding becoming someone else's food, and reproduction. Indeed, even our "higher" processes of cognitive thought under the domain of the prefrontal cortex are hypothesized to have been co-opted from machinery and mechanisms thought to have first evolved for movement (Loxdale 1991).

All but the sponges (Porifera), drawn at the lowest branch of the animal tree, are comprised of specialized tissues, distinct layers of specialized cells tasked to perform specific functions such as muscular tissue to drive movement and networks of nervous tissue that coordinate their control. Sponges, however, are boring, the only animal without even a primitive nerve network. For everyone else, different tissue types are derived from two (ectoderm and endoderm) or three (adding in mesoderm) germ layers formed during embryonic development. The predominant branches of diploblasts (those with only two) are the ctenophores (comb jellies) and cnidarians (sea anemones, corals, jellyfish, and hydras) have nerve nets but no muscles (no mesoderm!). These are the radiata, those with indistinct cleavage patterns during embryonic development and radial symmetry (Reviewed in Martindale and Henry 1998). Move up past the ghostly ctenophores and radiating cnidarians and you'll arrive at Bilateria, the major clade of (primarily) triploblastic animals that develop all three germ layers during embryonic development.

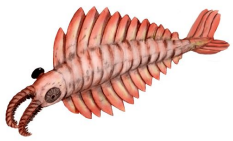
The key feature of Bilateria is their bilateral symmetry of form (at the very least their embryonic form). Unlike the stationary sponges (the mature ones, anyway) and "floaty" natures of radially-formed jellyfish, bilaterians are built to move *forward,* and are founded on a set of body axes polarized by a distinct head and tail (anterior-posterior), back and belly (dorsal-ventral), and as a consequence, two lateral sides reflecting outward from the midline (medio-lateral in both directions). All but the relatively slowly-moving echinoderm adults (sea stars and urchins) are bilaterally symmetrical (or nearly so), with a left side that mirrors the right divided along the sagittal plane; this symmetry is not easily broken (Palmer 2004). These body axes so principle to the bilaterian animal form are patterned early during embryonic development as described above. The most iconic of these are those attributed to specifying the unique regional identities of these forms — the Hox genes, which are expressed in all bilaterian animals (Minelli 2008).

Cambrian Explosion and the diversification of bilaterian form

For a few tens of thousands of years after the evolution of the animal form, there were mostly soft bodied floaty things hanging about in the ocean and small metazoan forms revealed by microfossils "Acritarchs". And then came an explosion of Bilaterian form, shells and skeletons and most abundant of all, segmented impressions of arthropod bodies and their ancestors. The Cambrian (542 mya - 488 mya) (Maloof et al. 2010) was a time known for being "uniquely experimental" (Gould 1990); a sandbox of tinkering that led to so many "weird wonders" that were gone just as soon as they arrived and from which our modern critters are derived, including the Annelida, Arthropoda, Branchiopoda, Chaetognatha, Chordata, Echinodermata, Entoprocta, Hemichordata, Mollusca, Nematoda, Onychophora, and Priapulida.



Hallucigenia
by Apokryltaros
(CC BY-SA 4.0)



Anomalocaris
by Nobu Tamura (CC BY 3.0)

Opabinia regalis
by Nobu Tamura (CC BY 3.0)

Acanthostega
by Nobu Tamura
(CC BY-SA 3.0)

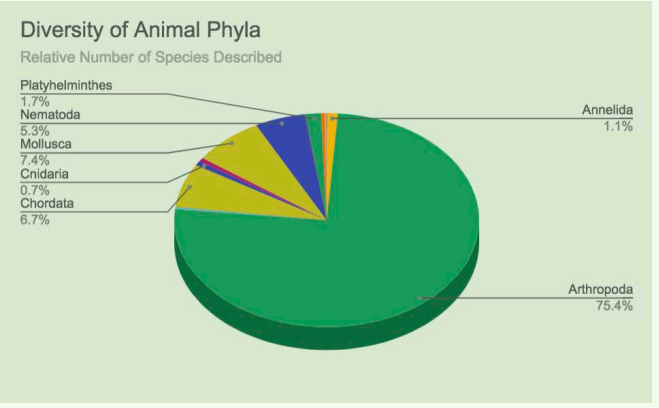
Fig 1.3. "Weird Wonders" of the Cambrian Explosion
(A) *Hallucigenia* — Ancestor to velvet worms (Onychophora) or more closely related to arthropods? Either way, they're panarthropod! (B) *Anomalocaris* are "abnormal shrimp", a giant (1m long!) proto-arthropod thought to be among the first predators to evolve. *Anomalocaris* undulated through the water by a fan-shaped tail and flexible body lobes acting as a simple fin and not requiring much of a complex brain to manage. In addition, two large appendages (~18cm in length) barbed with spikes extended from in front of its mouth (Whittington and Briggs 1985; Usami 2006; Gould 1990). (C) *Opabinia* cannot be placed in by modern phylogenetic groups. Its presence informed Eldredge and Gould's theory of "Punctuated Equilibrium" and led Gould to believe that the Cambrian explosion was indeed a dramatic, "experimental" event possibly driven by unusual evolutionary mechanisms (Gould 1990; Whittington 1975; Eldredge and Gould 1972). (D) The recognizable tetrapod-like limbs of *Acanthostega* did not evolve until much later (Devonian).

Bilateral symmetry of form

Bilaterian animals can be broadly lumped into three major "superphylum" named for the feature of embryonic development that the original taxons were based (though just how distinguishing these features really are have come into question and particular branches are still subject to phylogenetic debate and rearrangement). These are as follows, broken down further by examples of representative phyla: Spiralia, named for the spiral pattern of cleavage made with the initial divisions of their embryos; these include molluscs (octopus, cuttlefish, snails, slugs, mussels, and the like), annelids (segmented worms), and platyhelminthes (flatworms). Ecdysozoa are named for their rigid exoskeleton that they must molt in order to grow larger, and includes nematodes (roundworms), onychophora (velvet worms), tardigrades (water bears), and arthropods (e.g., spiders, millipedes, insects and crustaceans). Finally, there are the deuterostomes, which includes echinoderms (sea stars and urchins) and chordates — the phyla that we vertebrates are under. Deuterostomia means "second mouth", an important distinction set by which orifice forms first from the initial embryonic opening during development; for us, this is our anus, which comes before the development of our mouth (in contrast to protostomes, the now-contested major grouping of Spiralia and Ecdysozoa, which is characterized by the formation of the mouth first then anus).

Of all the animal phyla (a higher-level grouping of animals that imbues no special categorical value beyond its historical convenience), the most diverse (in terms of number

of species) by a very wide margin are the arthropods, representing a whopping ¾ of animal species described. One major feature of the arthropod body plan (that may likely have contributed to its evolutionary success) is that it is segmented along the AP axis, dividing the body into repeating units (known as "metameres") that can be distinctly diverse from one another, but all based on the same modular foundation. Unsurprisingly, this segmentation was patterned during embryonic development. Perhaps surprisingly, segmentation in animals is not represented by a monophyletic branch; rather, there is one major lineage of segmented animals represented in each of the three major divisions described above: the annelids (segmented worms) of Spiralia, arthropods (Ecdysozoa), and chordates (Deuterostomia), all of which were segmented as an embryo during early axial patterning.



Movement of Diversity

When we think of animal diversity we often think of lions, cheetahs, zebras and giraffes, a safari of running spots and stripes; osprey and ostriches, those with wings that take flight (or fumble with them awkwardly); and crocodiles and dart frogs, those equipped with claws or webbed feet to match their respective locomotion on land or in water — revealing a worldview that is four-legged and vertebrate-centric. But with four limbs of limited variety always set on the same bed-post structure, the tetrapod animals we are most familiar with are somewhat lacking in body plan diversity, and their representative percentage of animal diversity is well within single digits.

Compare this to arthropods, whose representative body plans can include from three (insects) to many hundreds (millipedes) of pairs of legs, all set on a modularized, segmented body. Arthropods offer the greatest diversity of body plans in terms of both number and types of appendages specialized for crawling, jumping, swimming, flying, feeding, fighting, camouflage and display. The very success of arthropods is in fact founded on their locomotion and presumably rapid diversification of appendages, allowing their expansion into a great number of niches. The focus on appendages is even in their name, arthropod = jointed "foot" or leg.

The rapid diversification of arthropod body plans has been frequently suggested to be due to the modularity of the regulatory elements of key developmental genes that pattern these repeated segments and their associated appendages. During evolution this modular organization of regulatory elements may have allowed for rapid diversification of the function of these developmental gene regulatory networks in *specific* segments or regions. This mosaic pleiotropy is encoded into the gene regulatory pathways and may even

contribute to subfunctionalization of these gene networks, allowing them to acquire roles in novel processes.The patterns of pigmentation in the wings of butterflies and bodies of spots and stripes may be a novel co-option of the fundamental patterning genes into such novel roles (Gompel et al. 2005). When it comes to the specification of a particular regional identity during embryonic development, there is no evidence of a gene's apparent role more striking than that generated by the mutation or functional manipulation of Hox genes in arthropods. Hox genes have been shown to specify segment and appendage identity in arthropods by acting as modular "switches" of segment-specific development once the AP axis has been set and segmented. This is also the case for vertebrates, where segment-specific wiring of Hox genes has been shown to specify the identities of many metameric structures such as ribs and vertebrae (Reviewed in Abzhanov and Kaufman 2000; Lewis 1978; Morata and Kerridge 1981; Lawrence and Morata 1994).

The extraordinary diversity of appendage form and function among the arthropods in general and even within a particular species exposes the spectacular capacity for small changes in the Hox function to sculpt morphology. So important are Hox patterns of expression and modulation of Hox function to the specification of appendage identities in arthropods, that shifts in patterns of Hox expression have been hypothesized (and functionally demonstrated) to be a mechanism for the macroevolutionary diversification of appendage types and arrangements (Averof and Patel 1997; Reviewed in Hughes and Kaufman 2002).

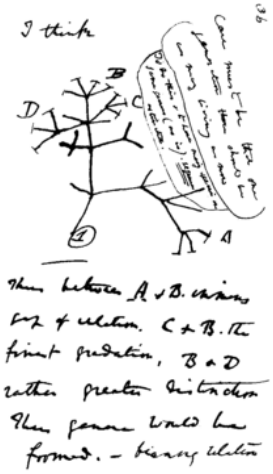
Part II of this introduction briefly explores the contemporary history of inheritance and development thought leading to the discovery of Hox genes and the "Evo-Devo" revolution, detailing the conserved patterns, functions, and phenomena of Hox genes and the impact of these discoveries.

Intro Part II. Most Wonderful Are Being Evolved; Hox genes & the diversification of regional form

Pre-History of Evo-Devo Thought

The Modification and Inheritance of Traits

On the Origin of Species … by a Mix of Particles Subject to Environmental Change



One cannot write a dissertation on the evolution and development of diversity without mention and a nod of high regard to Charles Darwin. Our current knowledge of the processes of development and evolution of natural forms is profoundly more sophisticated than what Darwin could have ever hoped to understand, and yet the primary tenants of his theories have stood strong, and are indeed still the founding tenants for which contemporary mechanistic studies are based: Descent with Modification molded by the forces of Natural Selection (Darwin 1859).

Embryology was an important informer of evolutionary and heredity studies, but there was no knowledge of genes and genetics at this time. Darwin knew that there that there was *something* physical passed down from each parent that mixed in *some way* to define characteristic attributes *somehow*. But even Darwin didn't quite have it, and the best he could come up with was pangenesis ("whole origin"), his idea that minute particles are shed from the cells of an organism's body and that these "gemmules" migrate through the blood (where they are subject to change based on environmental factors) to the gonads, where they are passed along to the offspring (Reviewed in Liu 2008).

The Conceptualization of "Genes" as Discrete Units of Heritable Traits

Although the physical makeup of a gene would not be discovered until the mid 1900s, the concept of a gene was put forth by Gregor Mendel after his meticulous crossing of pea plants led him to the realization that traits are inherited as discrete units rather than by previous notions of blended inheritance (Mendel 1865). "Mendelian genetics" resolved the independent assortment of dominant and recessive traits that combine to form a genotype that is homozygous or heterozygous — concepts critically informative to the understanding of the mechanisms of heredity and evolution that would go unnoticed for decades.

The term 'gene' was first coined by Danish botanist Wilhelm Johannsen in 1905, who would go on further to distinguish the concept of genotype (heritable instructions) and phenotype

(its physical manifestation) (Reviewed in Wanscher 1975; Johannsen 2014 (Translation from 1911)).

The Modern Synthesis (1918 - 1942)

Evolutionary scientists continued to reconcile Darwin's theory of evolution (descent with modification by natural selection) and Mendel's concepts of hereditary (by discrete units that combined in a particular way) under a mathematical framework coined "The Modern Synthesis" by Julian Huxley in his similarly titled book (Huxley 1942), which heavily informed, and was informed by "Neo-Darwinian" thought during this period; primarily, that evolution proceeds gradually by the accumulation of very small genetic mutations.

Evolutionary Changes by Monsters, Leaps, and Bounds

Bateson's Natural Collection of Homeotic Mutants (1894)

In 1894, English biologist William Bateson first described a particular mutation of a curious type in which one distinct body part is transformed toward the identity of another (such as a leg instead of an antenna, or a hindwing for a leg) in his book "Materials for the study of variation" (Bateson 1894). He coined this type of transformation "homeotic" meaning, "in the likeness of another".

Goldschmidt's Theoretically Hopeful Monsters (1940)

Bateson was a proponent that biological variation of characters exists both continuously ("meristic") and discontinuously ("substantive"), depending on the trait in question. In contrast to the neo-darwinian views on natural selection that would soon go on to form the modern synthesis, he believed that it was the latter and not the former that was predominantly acted upon by natural selection, i.e., that evolution proceeds in large leaps (saltation, though he did not use this word) rather than by an accumulation of small mutations (gradualism) (Reviewed in Gillham 2001). This aligned Bateson with the rather unorthodox (regarding his views of inheritance) Richard Goldschmidt, who believed that "small gradual changes could not bridge the hypothetical divide between microevolution and macroevolution" and that the evolution of a distinct new species (rather than merely a particular variety of the same) necessitated an unmistakable change of primary pattern rather than by anatomical bits. Goldschmidt believed that large genetic changes necessitated profound mutations such as could be founded in rare special cases of a "Hopeful Monsters" (Goldschmidt 1940; Goldschmidt 1933), similar to those found by Bateson. His views on macroevolutionary theory were dismissed on the grounds that mutants enter a population as inviable heterozygotes unable to establish themselves however well adapted (Dobzhansky and Dobzhansky 1982).

The Mechanistic Basis for Inheritance

Morgan's Fly Room (1915) - Genes mapped on chromosomes are the physical basis of hereditary

Initially, like Goldschmidt and Bateson, Thomas Hunt Morgan argued that a new species could not evolve by selection acting on slight variation of characters — there was no mechanism. But not only did he began to see otherwise when his lab produced many small, stable heritable mutations in *Drosophila*, they then found that mechanism. Only *heritable* characters are in reach of evolution, and Morgan and team solved this through genetics that linked particular character attributes (that were not in it of themselves heritable) to a physical region on a chromosome (a gene) that was (Morgan 1916).

Not only did Morgan and his team come up with the mechanical basis of heredity (Morgan 1915), they also grandfathered in *Drosophila melanogaster*, the first and most widely used model organism for some time and the workhorse of genetic discovery.

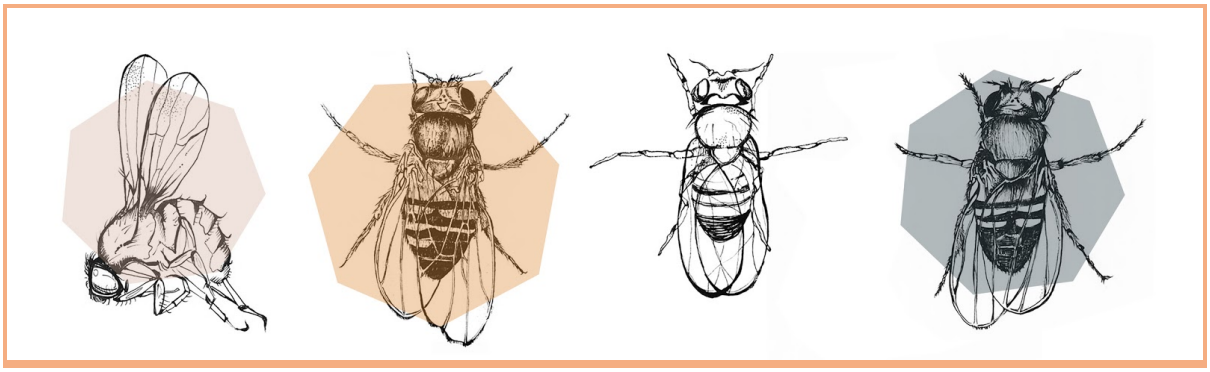


Image used with permission from "The Fly Room" (http://theĀlyroom.com)

Watson and Crick (1953) - Discovery of the double helix

Although genes as the physical unit of hereditary had been physically mapped onto chromosomes, the structural properties and chemical composition was still unknown. No mechanistic insight could be made based on the nature of its physical properties. Watson and Crick did not discover DNA nor were they first to suggest it as the chemical constituent of genes. What they did was discover the code of life, the simple magic of how all the pieces fit together in two complementary strands of nucleotides wrapped in a double helix *that could be copied*, providing a mechanism for genetic replication; one strand serving as the template for the other in their unassuming but monumental paper (Watson and Crick 1953) that changed the course of genetics and our study of life from having no code, to copy, to read-write.

Regulating Development with a Phenomenal Gene Cluster

Discovery of the Homeotic Gene Clusters in *Drosophila*

Bithorax Complex (BX-C) - Ed Lewis (1978)

In his seminal paper in (1978), Ed Lewis genetically described the Bithorax complex (BX-C), a single cluster of genes imparting segment-specific identities in the thorax and abdomen of *Drosophila*. It was believed that 2-winged flies had evolved from 4-winged insects, and that 3-pairs of legs in insects had evolved from those with many more, so Ed Lewis presumed that, during the course of evolution, flies must have evolved 'leg-suppressing' and 'haltere-promoting' (full wing suppressing) genes to achieve its dipteran body plan. If such 'suppressing' genes existed, it followed that their functional mutation and corresponding release of the suppressor should 'reverse' the fly phenotype to the more basel form. And that's exactly what he found; most iconically, the mutation of *Ubx* (and presumed loss of of wing repression of the T3 haltere) results in a four-winged fly (Fig. 1.5).

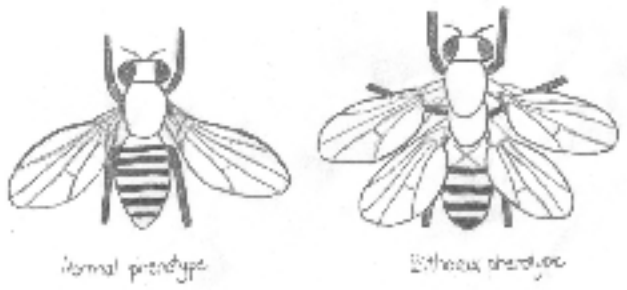


Fig. 1.5. Normal and mutant bithorax Drosophila
Loss of *Ubx* repression of T3 wing development leads to a haltere to wing transformation.
Image credit Rachgo20 (CC BY-SA 4.0)

These genes that had the apparent function of controlling the diversification of fly appendages during development met the criteria of those he predicted to be an evolutionary mechanism of macroevolutionary change surprisingly all came from the same cluster of "pseudogenes". The proximity of location in the genome and similarity of function in the fly led Lewis to further predict that genes in this complex arose from multiple duplications of a single gene and that their mutation permitted by their redundancy conveyed a new (but similar) function (Lewis 1951). "It is as if during ontogeny the BX-C genes recapitulate their own phylogeny"[3]

Antennapedia Complex (ANT-C)

Meanwhile, back in another lab, Kaufman and team were working to describe the genes, cytogenetic organization and developmental fates of another homeotic gene complex

[3] Ed Lewis Nature. A gene complex controlling segmentation in Drosophila. 1978 Dec 7;276(5688):565-70

similar to and more proximal the bithorax locus of chromosome 3R in *Drosophila melanogaster* (Duncan and Kaufman 1975; Kaufman et al. 1980; Scott et al. 1983; Kaufman 1978). Like the bithorax complex, the Antennapedia Complex (ANT-C) (so named for the *Antennapedia (Antp)* mutation it was first defined by) was identified by the genomic clustering of homeotic mutations. While the Bithorax complex (the more distal locus) controlled segmental identities of the more posterior thorax and abdomen, genes of the Antennapedia Complex controlled the fates of the more anterior segments of the head and prothorax.

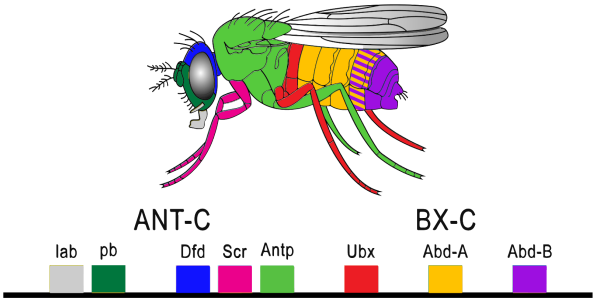


Fig. 1.6. ANT-C and BX-C genes and segment-selected gene expression (Modified from Public Domain)

Hox Gene Regulation

Discovery of a Conserved Homeodomain and the Launch of a New Field

Lewis had suggested that BX-C genes controlled segmental identities by regulating other genes that actually build the structures (and genetic studies in *Drosophila* and other insects would indeed later demonstrate that the products of Hox genes control the transcription of numerous downstream genes within their domains of expression (Abzhanov and Kaufman 1999; Graba et al. 1997)). But by what mechanism?

Two different labs simultaneously discovered a short (180 bp) repetitive nucleotide sequence localized in genes required for correct segmental development (McGinnis et al. 1984; Scott and Weiner 1984). Named for its association with homeotic mutations, this "homeobox" encodes a 60 AA DNA binding protein domain and is highly conserved among many of the genes of the homeotic complexes. The conservation of the homeobox in otherwise highly converged sequence suggested a common evolutionary origin for at least some of the genes spread across the bithorax and Antennapedia complexes, supporting Lewis's proposal of the duplication and subfunctionalization of pseudogenes as a mechanism for segmental diversification and the possibility that the principle of "duplicate and diverge" is a fundamental mechanism for the specialization of segmental form and function, and further suggesting that the homeobox may be conserved among different species.

So researchers looked for a homeobox in something that wasn't a fly or even an insect, and found one in a clawed frog (*Xenopus*) (Carrasco et al. 1984). This was perhaps the first development regulating gene identified in a vertebrate, launching a wild search for the

discovery of Hox genes and other putative regulatory genes in a all sorts of other animals, and the comparison of inter-specific notes to understand the function and evolution of these genes, but also using the patterns of these genes to understand more about the development and phylogenetic history of that organism--and thus the field of Evolutionary Developmental Biology or "Evo-Devo" began, once again reuniting the fields of embryology and evolutionary studies, but this time with a third player: genetics, and this time with an understanding (and methods of visualizing) the physical properties and interactions of genes and gene regulation.

So conserved are is the function of the homeodomain, that the ectopic expression of vertebrate Hox genes in a *Drosophila* embryo yields phenotypic outcomes similar to those produced by the ectopic expression of the respective ortholog of its own (*Drosophila*) Hox genes overexpression (Malicki et al. 1990; Malicki et al. 1993; McGinnis et al. 1990; Zhao et al. 1993). Even more strikingly, Lutz et al. (1996) demonstrated that vertebrate hox genes expressed under the control of *Drosophila* regulatory sequences can substitute the function of its *Drosophila* ortholog. For this, they fused the protein-coding region of gHoxb-1 (the chick ortholog to *lab*, which diverges most of its sequence beyond the Homeodomain and one other small region) behind known *lab* regulatory sequences in order to express gHoxb-1 in its correct spatiotemporal pattern in *lab* null mutant *Drosophila* and showed that, for most cases tested, the functional replacement of gHoxb-1 rescued the null mutant *Drosophila* (Lutz et al. 1996).

Mechanisms of Macroevolution

We can uncover mechanisms of development from its deviations, both by studying the cause of developmental anomalies of individuals and comparing the differences between developmental pathways and outcomes among different groups. The newish science of evolutionary developmental biology, known fondly as "Evo-devo" explores just that. It is the fusion of the study of genetics, development, and evolution.

Following neo-darwinian belief that evolution proceeds by the accumulation of very small mutations (and that those substantially large result in fatal dead ends), it followed that those genes responsible for building the different body types would be ever-modified by small changes and very different indeed. This turned out to be not the case, however, and the striking conservation and morphological impact of core regulatory genes, particularly those informed by Hox mutations, led to a paradigm shift that it is changes in gene regulation — rather than mutations affecting the function of those genes directly responsible for building the different structures — that are a major mechanism of macroevolutionary change (Ronshaugen et al. 2002; Galant and Carroll 2002; Reviewed in Hughes and Kaufman 2002; Akam 2000; Akam et al. 1994).

As will be discussed in much greater detail in the chapters to follow, shifting boundaries of Hox gene expression are one such example of regulatory changes thought to be a major driver of macroevolutionary diversification (Averof and Patel 1997). This makes sense as large effects could be contained within a redundant module rather than by breaking genes

(however gradually) that are directly involved in the building of body structures in general, or are critical to multiple processes due to pleiotropic roles.

Mutations affecting the expression domains of different Hox genes are likely the culprit behind Bateson's naturally occuring homeotic variants, as is so strikingly demonstrated by the experimentally induced transformations of Lewis, Krumlauf and so many others. Suddenly, Goldschmidt's Monsters seem just a little bit more hopeful. The difference is in the genes: changes in region-specific regulation provides a mechanism for inducing large changes in the structures of select segments without impacting all parts of the body. In fact Stephen Jay Gould used the discovery of regulatory genes to rebrand Goldschmidt's theories of evolution by jumps in an extended version of the Modern Synthesis that he proposed. He put forth that the mechanisms that enabled small changes to the constraint systems during embryogenesis that lead to such grand morphological transformation could also possibly be a mechanism of macroevolutionary change (Gould 1977). Gould advocated his theory of Punctuated Equilibrium, an alternative to evolution by gradualism that was not quite to the extreme of Goldschmidt (Eldredge and Gould 1972).

But what regulates the patterns of Hox gene expression? Lewis had further suggested that the Bithorax complex was in turn regulated by the re-repression of at least four BX-C genes by *cis*-regulatory elements, and that polycomb codes for the repressors of this complex (Lewis 1978; Wedeen et al. 1986). It was further suggested that the spatially restricted expression of the Hox genes is maintained by hierchierchacal cross-regulation among the genes of the homeotic clusters themselves (Harding et al. 1985). Thus, not only do Hox genes regulate the transcription of downstream genes, they regulate the transcription — and patterns of expression — of other Hox genes. The particulars of this cross-regulation is a primary component to my dissertation research, and will be addressed in much greater detail in Chapter 5.

Generalization of Hox Phenomena — Impact & Models

A Phenomenal Complex: Posterior Prevalence and Collinearity

Lewis noticed a particular pattern of transformation in his mutants, that loss-of-function (LOF) experiments led to the transformation of posterior segment identity toward the anterior and, conversely, gain-of-function (GOF) led to anterior transformation of posterior segments. He noticed one other particularity about the directionality of the mutations that he didn't yet understand but expected to be important: the mutations he mapped along the chromosome were collinear with their effect along the AP axis (Lewis 1978). Indeed, the collinearity between the arrangement of Hox genes on the chromosome and the pattern of their expression along the embryonic axis in time and space is more often than not conserved, as is these same patterns of phenotypic transformation direction upon their mutation (Akam 1989). Indeed, Hox mutations in a mouse showed similar same pattern of homeotic transformation and collinearity (Reviewed in Krumlauf 1994).

It cannot be overemphasized the impact that this had. That a single gene was found to be responsible for such a striking transformation was remarkable. That there was an entire complex of these homeotic genes whose position on the chromosome was recapitulated by the patterns of expression on the embryo was phenomenal. That these genes were found in an entire suite of disparate animals and that this "collinearity" was tightly conserved was astonishing, leading to a myriad of evolutionary questions and implications.

Arthropod and Vertebrate Models

Although Hox genes are studied in a number of representatives from an assortment of lineages, the two primary groups studied are the arthropods and vertebrates, two of the three major segmented lineages. A segmented body comprised of repeating modules that are visually distinct provides a series of very readable outputs post Hox gene manipulations, a happy fortuity and retro-justification for models chosen for a different reason: Arthropods because of the history of *Drosophila* as *the* genetic workhorse in Thomas Hunt Morgan's "Fly Room"[4] and vertebrates models by proxy. It is likely, at least in part, that the serendipitous feature of segmentation in arthropods facilitated the discovery of the Hox complex, and it is this major commonality between arthropods and us that makes the study of Hox genes in arthropods particularly useful for making generalizations and understanding more about our own development.

Fig 1.7. A vertebrate and a fly;
Two representative examples from the major segmented lineages of Ecdysozoa (arthropods) and Deuterostomes (chordates). *Image used with permission from the movie "The Fly Room"*

The many demonstration of conservation of the specific genes, patterns and ordered function was crucial for expanding generalized model of Hox function founded in detailed studies in *Drosophila melanogaster* into other animals where such detailed studies would not be feasible. While the availability of tools and procedures in *Drosophila* greatly facilitated the initial study of the Hox complex and other regulatory genes responsible for building a body, the gravitation toward *Drosophila* as a primary model for Hox gene research was grandfathered in due to its pre-existing model status rather than made as a deliberate choice based on any great attributes it possessed for the study of Hox genes. In respect to other possible arthropod choices, it is non-ideal for the generalization of Hox

[4] Now illustrated by the lovely, slightly creepy scenes of "The Fly Room" that gives you a shallow glimpse of Morgan's Fly Room and what went on there (lots and lots of fly sorting) in the form of a character sketch of a child (Betsy) and her strained relationship with her father Calvin Bridges.

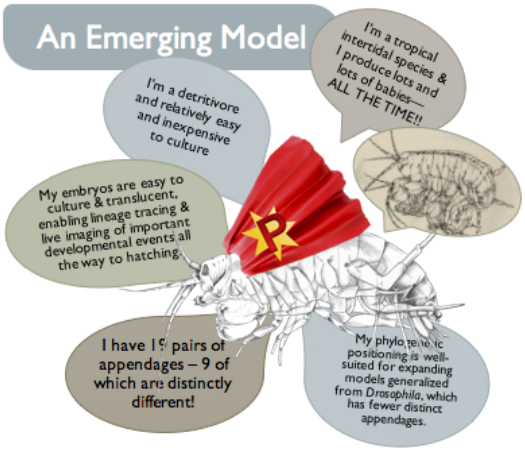
models overall: *Drosophila* and its fellow dipterans (an orders of insects with the typical fly body plan of 3 pairs of legs and a pair each of antenna, wings, and halteres) are relatively derived and lacking in appendage diversity. Indeed, new technologies are now allowing us to expand these studies to much greater depths in a greater diversity of species, and contemporary research is showing us more and more that these canonical "rules" made by initial studies in flies may not be quite so easily generalized. Our understanding of the function and regulatory interactions of the Hox complex in the specification of segmental and appendage diversity, as well as our understanding of regulatory mechanisms in general, would thus be greatly improved by comprehensive studies in an arthropod model speciĀically chosen for its diversity of appendages and phylogenetic standing.

My primary research studies the interactions among Hox genes effecting the specification of appendage identities in the amphipod crustacean *Parhyale hawaiensis*, an up-and-coming model whose body plan boasts a much greater diversity of appendage types than *Drosophila* or any other insect, and with a phylogenetic standing close enough to flies to make easy comparisons, but far enough away to provide a greater range of generalizations.

Using the newly developed gene editing technology CRISPR/Cas9 (a system discovered, in large part, at UC Berkeley), I systematically "break" the three most-posterior Hox genes individually and in tandem (*Ultrabithorax, abdominal-A*, and *Abdominal-B*) in newly fertilized, single-celled *Parhyale* embryos, resulting in striking seawater monsters that try to swim with walking legs protruding where there should be swimming appendages or the modification of legs to jaw-like feeding appendages all the way down their bodies. Using confocal microscopy and immunofluorescent staining, I then analyze the structural changes of the transformed appendages and the misexpressed patterns of Hox genes that underlie their altered arrangements in order to better understand the interactions among Hox genes in establishing — and diversifying — the types and arrangement of appendages.

Chapter 2: Characterizing Hox gene expression in the emerging model *Parhyale hawaiensis*

Non-model organisms have the advantage of expanding our breadth of knowledge and are often chosen because of wanting to study a feature that is only, or most strikingly, found in a particular organism. Choosing to work with a non-model is exciting, but requires laborious study to develop and optimize protocols that work well in that organism, finding and making tools to manipulate them, understanding their life-cycles and standardizing stages, or even just in how to raise them successfully before one can even begin the research the new model was chosen for in the first place.



Parhyale met all the basic requirements desired by the Patel lab to be chosen as the model they would develop for functional Hox studies: a crustacean with good phylogenetic positioning for expanding and connecting generalizations, many, many easily distinguished appendage types for surveying morphological transformation from one type to another, and the constant production of easily accessible embryos all year round. It was there serendipitous attachment to the filters of tanks not meant for them, however, that led to *Parhyale* becoming the "chosen ones" (and very easy keepers) brought back by Bill Browne from the Chicago Shedd Aquarium.

Parhyale entered the Patel lab with the intention that it would be used as an experimental model to functionally test the hypothesis that shifts in Hox domains could provide a mechanism for macroevolutionary change, specifically, to functionally test the correlation between phylogenetically observed correlations between shifts in the anterior boundary of *Ubx/abd-A* and the transition of thoracic legs to modified jaw-like "maxillipeds" fused to the head (Averof and Patel 1997). But first, the groundwork of establishing a new model.

Parhyale husbandry, life cycle, embryonic development and staging (Browne et al. 2005) as well as *Parhyale*-specific tools and techniques (Rehm et al. 2009; Rehm et al. 2009; Rehm et al. 2009) were established over the next several years, laying the groundwork for more in depth studies in *Parhyale*. LOF (Liubicich et al. 2009) and GOF (Pavlopoulos et al. 2009) indeed recapitulated the correlation of phylogenetic shifts of the anterior *Ubx* with the loss or gain of maxillipeds from (Averof and Patel 1997). It was now time to identify and sequence the rest of the *Parhyale* Hox, and to perform a comprehensive examination of the expression patterns of Hox genes in time and space during the different stages of *Parhyale* embryonic development. This was phase I of the goal of performing a comprehensive functional analysis of each of the *Parhyale* Hox genes in order to better understand Hox gene function and evolution in crustaceans overall.

By the time I entered the Patel lab, all *Parhyale* Hox genes had been identified, largely sequenced, and incompletely mapped. *In-situ* probes for each *Parhyale* Hox gene had been designed and HOX-specific antibodies partially gathered. My contributions to this publication included further optimization of the *Parhyale*-specific antibodies and parameters, the generation of embryonic expression data during later stages of appendage development (including fine tuning the posterior expression boundaries), the generation of Hox expression schematics and input into the final manuscript.

The identification of all 9 Parhyale Hox genes and the analysis of their expression patterns relative other developmental milestones are described in detail in the following co-authored publication "Comprehensive analysis of Hox gene expression in the amphipod crustacean *Parhyale hawaiensis*" (Developmental Biology 409 (2016) 297–309) copied and re-formatted in its entirety. In the following section I highlight and expand upon the results most relevant to the remainder of my dissertation.

The "Hox code" of limb identities in Parhyale

The following study characterizes in great detail the temporal and spatial expression of *Parhyale* Hox genes throughout their embryonic development. In particular, this study highlights the developmental milestones that correspond to the Hox patterns it establishes. Of potential significance are the particular patterns of Hox expression that correspond to the development of different limb types.

SigniĀicantly, this study finds that there are 11 unique combinations of Hox genes expressed in *Parhyale* appendages, and that each of the distinct appendage identities corresponds with the expression of a unique combination of Hox genes (Fig 2.i). Moreover, the minor phenotypic differences among appendages we consider the same "type" are correlated with differences in Hox expression *levels* though the overall combination of Hox genes expressed remained the same. For example, the expression of *Ubx* in T2 is lower than that of T3 (which is lower than its strong expression throughout the rest of the thorax T4 - T8). The graded anterior to posterior decrease in size of the A4 - A6 uropods (which strongly express *Abd-B*) is correlated with low levels of anteriorly-restricted *abd-A* expression in A4 that is lower still in A5 and not discernable in A6 (see Chapter 5; Fig 5.2. and 5.3 for my detailed analysis).

This comprehensive study provides the most complete sketch of Hox genes and expression patterns in a crustacean to date, providing an essential framework for systematically testing the function of each of the Hox genes individually (Chapter 4) and as part of an entire complex (Chapter 5) in order to complete a more comprehensive examination of the function and evolution of Hox genes in the generation of limb diversity.

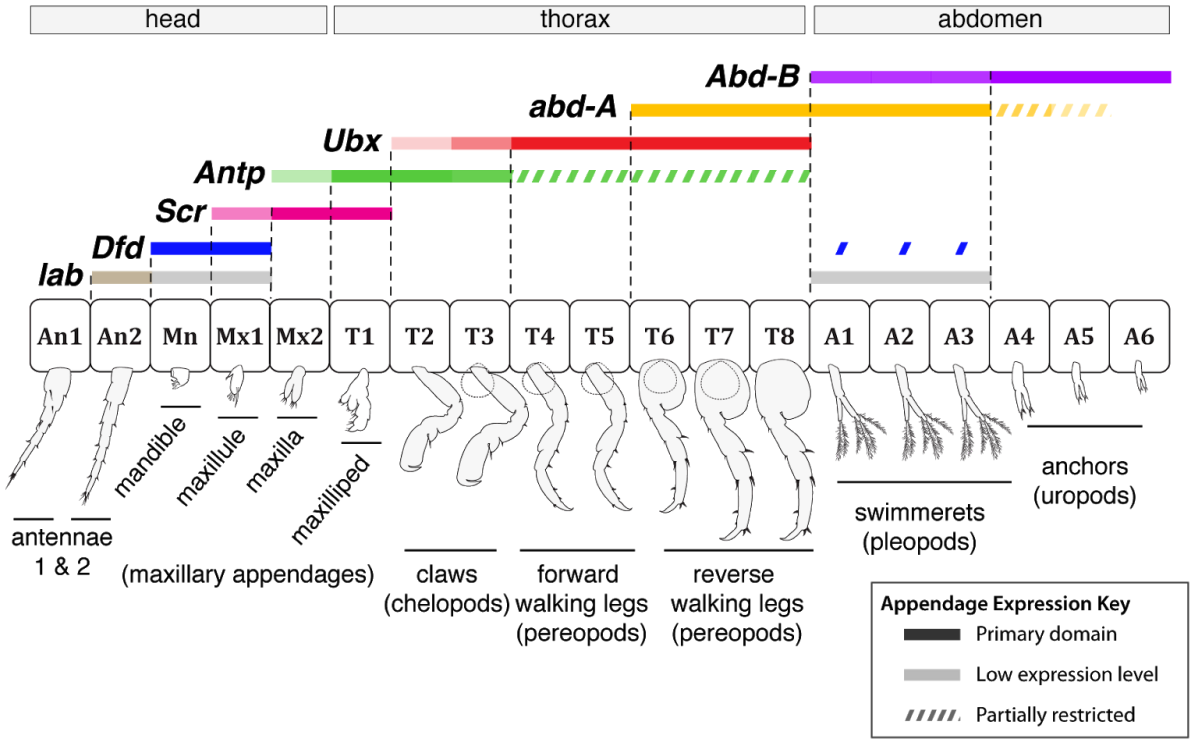


Figure 2.i. Hox Expression in *Parhyale* limbs (Mn - A6)
Parhyale have 19 pairs of appendages (one pair per segment) that can be grouped into distinct appendage sets across all three body tagma: The antennae and maxillary appendages of the head, maxillipeds (a modified thoracic leg that is fused to the head and used for feeding), claws (chelopods), forward walking legs and reverse jumping legs (pereopods) of the thorax, and swimming appendages (pleopods) and uropods of the abdomen. Each of the antennae (An1 and An2) and maxillary appendages (mandible, maxillule, and maxilla) have distinctly different appendage phenotypes and are each considered here to be a separate identity. While the T2 and T3 claws can be distinguished from one another by the presence of a gill on T3, they are considered to share the same "claw" identity for this dissertation. Likewise, the T6 walking leg is distinguishably smaller than that of T7 and T8 (which also lacks a gill) but they are considered to share the same type overall. In the abdomen, A4 - A6 uropods become progressive smaller in the posterior direction but share the same basic shape and identity overall. Thus, Parhyale have 11 clearly distinct appendage phenotypes. (Partial expression of *pb* and *Hox3* in An2 and Mn appendages not shown)

Comprehensive analysis of Hox gene expression in the amphipod crustacean *Parhyale hawaiensis*

Julia M. Serano[1], Arnaud Martin[1], Danielle M. Liubicich[2,3], Erin Jarvis[2], Heather S. Bruce[1], Konnor La[2], William E. Browne[4], Jane Grimwood[5], and Nipam H. Patel[1,2]

[1] Department of Molecular Cell Biology, University of California, Berkeley, CA 94720-3200 USA.
[2] Department of Integrative Biology, University of California, Berkeley, CA 94720-3140, USA.
[3] Los Medanos College, 2700 East Leland Rd., Pittsburg, CA 94565, USA.
[4] Department of Biology, University of Miami, 1301 Memorial Drive, Coral Gables, FL 33146, USA.
[5] HudsonAlpha Genome Sequencing Center, 601 Genome Way, Huntsville, AL 35806, USA

Published in **Developmental Biology** 409 (2016) 297–309

Highlights

- Comprehensive analysis of Hox gene expression in a crustacean amphipod.
- All canonical Hox genes except *fushi tarazu* isolated from *Parhyale* genome.
- Expression of the nine Hox genes shows both spatial and temporal colinearity.
- Hox gene anteroposterior expression boundaries correlate with limb identity.
- Co-expression domains suggest a combinatorial mode of limb specification.

Abstract

Hox genes play crucial roles in establishing regional identity along the anterior–posterior axis in bilaterian animals, and have been implicated in generating morphological diversity throughout evolution. Here we report the identification, expression, and initial genomic characterization of the complete set of Hox genes from the amphipod crustacean *Parhyale hawaiensis. Parhyale* is an emerging model system that is amenable to experimental manipulations and evolutionary comparisons among the arthropods. Our analyses indicate that the *Parhyale* genome contains a single copy of each canonical Hox gene with the exception of *fushi tarazu*, and preliminary mapping suggests that at least some of these genes are clustered together in the genome. With few exceptions, *Parhyale* Hox genes exhibit both temporal and spatial colinearity, and expression boundaries correlate with morphological differences between segments and their associated appendages. This work represents the most comprehensive analysis of Hox gene expression in a crustacean to date, and provides a foundation for functional studies aimed at elucidating the role of Hox genes in arthropod development and evolution.

2.1 Introduction

Few developmental gene families have been studied as extensively as the Hox genes, which encode homeodomain-containing transcription factors that determine regional identity along the anterior–posterior (AP) axis in bilaterian animals (Pourquié, 2009). Hox genes are typically organized within the genome in conserved clusters that display spatial colinearity—that is, their position along the chromosome correlates with the positions of their expression domains along the AP axis. Perturbation of Hox gene expression is associated with homeotic transformations, wherein body parts in one region of an animal are transformed to more closely resemble those of another region. Given the importance of Hox genes in establishing regional and segmental identity, it is not surprising that numerous lines of molecular, genetic, and developmental evidence have implicated Hox genes in generating morphological diversity during animal evolution.

Arthropods have proven to be especially well suited for studying the contribution of Hox genes to evolutionary changes in morphology (reviewed in Angelini and Kaufman, 2005; Hughes and Kaufman, 2002a). Arthropods have clearly segmented body plans, and each segment typically possesses a single, often unique, pair of appendages and other characteristics that provide useful markers for segmental identity. Furthermore, homologous segments between related arthropod species often display significant differences in appendage morphology. This allows one to ask questions about how differences in Hox gene expression, and the response of downstream target genes, generate segmental and appendage diversity, both within a single organism as well as between different species. Analyses of Hox gene expression in various arthropod lineages have led to a number of intriguing hypotheses linking observed shifts in Hox expression domains to specific morphological differences (Abzhanov and Kaufman, 1999a, 2000a; Averof and Patel, 1997; Hughes and Kaufman, 2002a; Hughes et al., 2004; Mahfooz et al., 2004; Peterson et al., 1999; Rogers et al., 1997). As a member of a major lineage, Malacostraca, within the Pancrustacean clade (Misof et al., 2014; Regier et al., 2010), which comprises all crustaceans and hexapods, *Parhyale* is well positioned as both an outgroup of insects and as a reference crustacean for evolutionary developmental studies of the diversity of arthropod *bauplans* (Rehm et al., 2009a).

The *Parhyale* body plan is fairly typical of the crustacean order Amphipoda. From anterior to posterior, the head segments include a pre-antennal segment (which lacks paired appendages), the first antennal (An1), second antennal (An2), mandibular (Mn), first maxillary (Mx1), and second maxillary (Mx2) segments. The thorax is comprised of eight segments, T1 through T8. The T1 appendages are maxillipeds, which, along with the mandible and the two maxillae, facilitate feeding. The gnathopods on T2 and T3 have a distinctive distal claw; the T3 claw grows especially large in adult males. Segments T4 through T8 contain walking legs. The walking legs on T4 and T5 are oriented anteriorly, while those on T6 through T8 are larger, have thicker bristles, and are oriented posteriorly (the name of the order, "Amphipoda", derives from having these two distinctly different walking legs. The abdomen consists of six segments, A1 through A6, each bearing a pair of

biramous appendages. The pleopods (swimmerets) of A1 through A3 are indistinguishable from one another, whereas the uropods (anchor legs) of A4 through A6 significantly decrease in size from anterior-to-posterior.

As a first step toward a comprehensive analysis of Hox gene function and evolution in crustaceans, we have cloned and examined the expression patterns for the entire *Parhyale* Hox gene suite. We have identified a single copy of each canonical Hox gene with the exception of *fushi tarazu (ftz).* Our preliminary BAC data indicate that many, if not all, *Parhyale* Hox genes are clustered in the genome. Expression analyses demonstrate that, with a couple of exceptions, they exhibit both temporal and spatial colinearity. Many of the Hox expression boundaries coincide with obvious morphological differences between appendage types and/or tagmata. The work presented here forms the foundation for future studies aimed at examining the role that Hox genes play in crustacean development and arthropod evolution.

2.2 Materials and methods

Cloning *Parhyale* Hox genes

Parhyale embryos of mixed ages were collected to make cDNA as previously described (Price and Patel, 2008). We used a variety of degenerate PCR primers targeting either the homeodomain or hexapaptide regions of Hox genes. The forward primers we initially used were 5'-YTIGARYTNGARAARGARTT-3', 5'-ACITAYACNCGNTAYCARAC-3', 5'-ACICGITAYCARACNYTNGA-3', 5'-CARACIYTIGARYTNGARAA-3', 5'-ATITAYCCNTGGATGMGN-3', 5'-ATHTAYCCNTGGATGGCN-3', 5'-ATHTAYCCNTGGATGAAR-3', 5'-CARATHTAYCCNTGGATG-3', and reverse primers were 5'-CKRTTYTGRAACCANATYTT-3' and 5'-CATWCKWCKRTTYTGRAACCA-3'. The method we found to be most useful for isolating the broadest possible range of Hox genes involved creating a series of less-degenerate primers targeting the highly conserved LELEKEF and KIWFQNR motifs in the Hox homeodomain. Four LELEKEF forward primers (5'-YTNGARYTNGARAAAGAATT-3', 5'-YTNGARYTNGARAAAGAGTT-3', 5'-YTNGARYTNGARAAGGAATT-3', 5'-YTNGARYTNGARAAGGAGTT-3') were used in every possible combination with four KIWFNQR reverse primers (5'-CKRTTYTGRAACCASATCTT-3', 5'-CKRTTYTGRAACCASATTTT-3', 5'-CKRTTYTGRAACCAWATCTT-3', 5'-CKRTTYTGRAACCAWATTTT-3'). Because the 3′-most nucleotides of each primer are not degenerate, PCR amplifications tended to be specific for certain Hox genes but not others—this circumvented the problem of repeatedly re-isolating the same few (perhaps most highly abundant) Hox genes. Each primer was used at a final concentration of 5 μM in 50 μl PCR reactions that included 2.5 units of Taq DNA Polymerase (New England BioLabs), 1× ThermoPol buffer (New England BioLabs), 1.5 mM MgCl2, 0.2 mM dNTPs, and embryonic cDNA. Forty cycles of amplification were carried out at a melting temperature of 94 °C, a “step-down” annealing temperature starting at 63 °C and decreasing 1 °C per cycle to 52 °C, and an extension temperature of 72 °C. PCR products were run on a 2% agarose gel, 120 nt fragments were isolated *via* low melting point

agarose and directly cloned into pBluescript II (Stratagene). Miniprep DNA was directly sequenced. In some case, colonies were pre-screened by PCR using oligos specific for previously isolated *Parhyale* Hox genes in order to identify those clones that would most likely represent novel Hox genes.

5′and 3′ flanking sequences for each identified Hox gene were isolated using the GeneRacer (Invitrogen) and/or GenomeWalker (Clontech Laboratories) kits. When possible, primers were designed to the 5′- and 3′most sequences of the untranslated regions (UTRs) for each Hox gene and used to PCR-amplify full-length cDNAs. Accession numbers for *Parhyale* Hox cDNA sequences are as follows: *labial*, JQ952576; *proboscipedia*, JQ952587; Hox3, JQ845948; *Deformed*, JQ952571; *Sex comb reduced*, JQ952579; *Antennapedia variant I*, JQ952581; *Antennapedia variant II*, JQ952582; *Ultrabithorax isoform I*, FJ628448; *Ultrabithorax isoform II*, FJ628449; *abdominal-A*isoform I, Q952572; *abdominal-A* isoform II, JQ952573; *Abdominal-B* isoform I, JQ952574; *Abdominal-B* isoform II, JQ952575.

BAC library screening

Parhyale BAC screening using radiolabeled Hox probes was carried out according to (Parchem et al., 2010). Inverse PCR was employed to generate sequence data for the 5′ and 3′ ends for many of the Hox-positive BAC clones. In some cases, inverse PCR-derived probes were used to re-screen the *Parhyale* BAC library. Miniprep DNA from 96 unique Hox-positive BACs was spotted onto "dot-blots," and each blot was hybridized to probes derived either from portions of Hox cDNAs and from the 5′ and 3′ ends of specific BAC clones. This dot-blot data was used to construct a rough map of the *Parhyale* Hox complex(es), and this map was refined and verified using BAC sequencing data. Accession numbers for the *Parhyale* Hox BAC sequences are as follows: PA31-H15, KR869963; PA24-C06, KR869964; PA93-L04, KR869965; PA264-B19, KR869966; PA179-K23, KR869967; PA40-O15, KR869968; PA81-D11, KR869969; PA272-M04, KR869970; PA92-D22, KR869971; PA221-A05, KR869972; PA284-I07, KR869973; PA76-H18, KR869974; PA120-H17, KR869975; PA268-E13, KR869976; PA222-D11, KR869977.

In situ hybridization and imaging

Embryo dissection, fixation, and *in situ* hybridization were carried out according to Rehm et al. (2009b,2009c) and Vargas-Vila et al. (2010). Images were analyzed on a Zeiss Axiophot microscope, captured in color with a Diagnostic Instruments Spot Camera, and figures assembled using Adobe Photoshop. False-color overlays of *in situ* hybridization images were carried out as follows: The DAPI signal was photographed and shifted to cyan by adjusting the hue to −35. A bright field image of the histochemical staining (from the BCIP/NBT/alkaline phosphatase reaction; Roche) was inverted, and the blue and green channels were eliminated so that the staining appears red. In double-staining experiments, the fluorescent signal from the Fast Red (Roche) product was photographed and shifted to yellow by adjusting the hue to +50. The "screen" command was applied to the red, yellow and DAPI images to allow all signals to be visualized simultaneously.

Immunohistochemistry

Antibody staining was carried out following previously published protocols (Liubicich et al., 2009; Patel, 1994; Rehm et al., 2009d). For detection of *Parhyale* Ubx we used a rat polyclonal antibody at a 1:4000-6000 dilution (Liubicich et al., 2009). Antp was detected using the previously described crossreactive mouse monoclonals 4C3 and 8C11 at dilutions of 1:30–1:50 (Condie et al., 1991; Hayward et al., 1995; Saenko et al., 2011); the two monoclonals displayed identical patterns in *Parhyale* embryos. Embryos were imaged using either an LSM 700 or LSM 780 (Zeiss) confocal microscope, and processed using Volocity software (Perkin-Elmer). For all *in situ* and immunohistochemical embryo preparations, all images are ventral views unless stated otherwise.

CRISPR/Cas9-mediated homologous recombination

We synthesized a sgRNA targeting the codons 7–14 of *Antp* following the procedure of (Bassett and Liu, 2014). To generate a donor plasmid for homologous recombination around the cleavage site, we cloned a genomic DNA fragment of 2.1 kb approximately centered on the *Antp* start codon (forward primer: 5′-CCCGAACTGTAAAAGGCAAA; reverse primer: 5′-TGCTGGGCAAAAGAAAAACT-3′) into the pGEM-T vector (Promega). A linear fragment was created by inverse PCR and Gibson Assembly used to integrate an *EGFP* sequence followed by the T2A ribosome skipping peptide signal (Diao and White, 2012) to the 5′ coding region of *Antp*, generating a *pHR_EGFP-T2A-Antp* donor for fluorescent tagging of *Antp*-expressing cells. Using modified oligonucleotides for fragment amplification, we also altered the PAM sequence in the donor plasmid to avoid secondary cleavage after successful homologous repair. We injected embryos at the one-cell stage with 40–60 picoliters of an injection mix containing the *Antp* sgRNA (200 ng/µL), the *AntpHR1_EGFP-T2A_AntpHR2* donor plasmid (200 ng/µL), a recombinant Cas9 protein (333 ng/µL, PNABio), and Phenol Red (0.05% final concentration). Embryos were screened for eGFP signal with a Lumar.V12 stereomicroscope (Zeiss) and imaged with a LSM780 confocal microscope (Zeiss).

2.3 Results

Isolation and genomic mapping of Parhyale Hox genes

We used a comprehensive degenerate PCR strategy to identify the entire set of *Parhyale* Hox genes. This approach resulted in the isolation of the homeobox sequences from all of the canonical Hox genes with the exception of *ftz* (see Supplemental Fig. 2.1). In subsequent cloning, we were able to isolate full-length cDNAs for *labial (lab), Deformed (Dfd), Ultrabithorax (Ubx), abdominal-A (abd-A)* and *Abdominal-B (Abd-B),* and the entire open reading frame for *Sex combs reduced (Scr)*. For *proboscipedia (pb)* and *Antennapedia (Antp),* we obtained cDNA sequence 5′ and 3′ to the homeobox, but not a full-length cDNA, and for *Hox3* we only isolated sequence 3′ to the homeobox. Our analyses suggest that *Parhyale* has a single copy of each of these Hox genes. Multiple alternatively spliced variants for *Ubx* (Liubicich et al., 2009), *abd-A*, and *Abd-B* were identified; for each gene, these alternatively

spliced versions are referred to as *isoform-1* and *isoform-2*. For *Antp*, two classes of cDNA differing from one another at three distinct polymorphic regions within the coding region were identified. Since these polymorphic variants (alleles) encode slightly different proteins, we refer to them here as *Antp variant I*and *Antp variant II*. We also found that a subset of transcripts transcribed from the *Ubx* promoter(s) are spliced to exons from the *Antp* transcription unit, indicating that transcriptional read-through occurs between the two genes. Similar *Ubx-Antp* read-through transcripts have been reported for five other crustaceans, two myriapods, and an onychophoran (an outgroup to arthropods), suggesting that this may be a relatively ancient feature of arthropod Hox complexes (Brena et al., 2005; Janssen and Budd, 2010; Shiga et al., 2006).

To determine whether *Parhyale* Hox genes are organized into clusters, we screened a *Parhyale* BAC library (Parchem et al., 2010) using probes specific for each Hox gene. Sequences from the ends of each positive BAC clone were also used as templates for probes to both re-screen the BAC library and to map individual BACs in relation to one another. Once sequenced, we were able to confirm our BAC map and precisely define the overlap between BACs (Fig. 2.1). As shown in Fig. 2.1, one set of BACs together span four linked Hox genes (5′-*Ubx, Antp, Scr, Dfd*-3′) and a second spans two linked Hox genes (5′-*pb, lab*-3′). In both cases, the Hox genes are organized in a linear, 5′-to-3′ orientation consistent with what has been reported for vertebrate and other Hox clusters. BAC clones corresponding to the three remaining Hox genes *(Abd-B, abd-A* and *Hox3)* were also isolated, but we were unable to determine whether they are located near or adjacent to each other and/or the aforementioned Hox mini-clusters. Given that *Parhyale* has a relatively large genome size (Parchem et al., 2010), additional genome data will be required to determine whether all nine Hox genes are organized within a single, large genomic cluster.

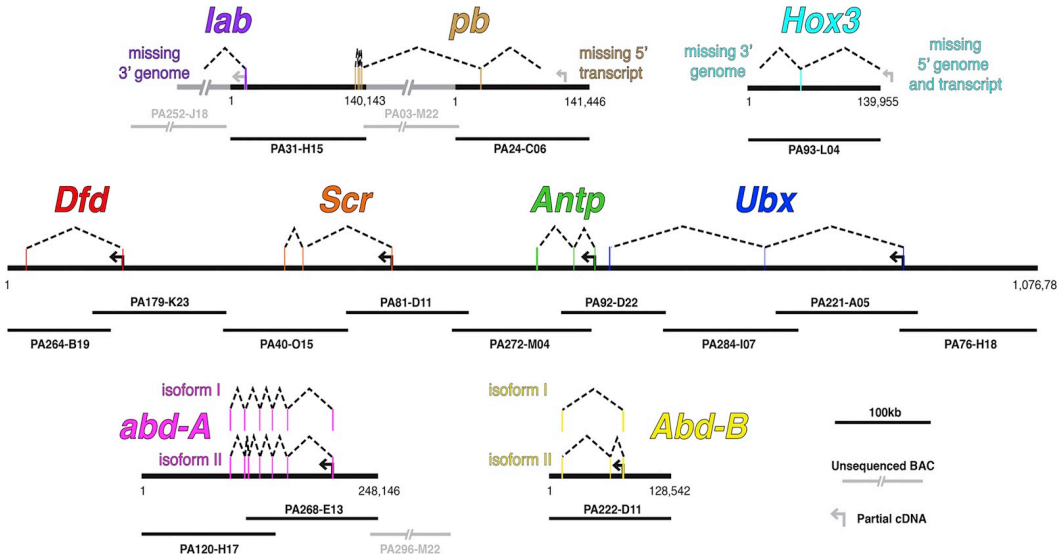


Fig. 2.1. Genomic organization of *Parhyale* Hox genes
The nine *Parhyale* Hox genes, along with the putative position of their respective promoters and intron/exon structure, are shown in relation to 18 BACs. Each BAC ranges from 100–150 kb in size. Based on our genomic mapping, two sets of genes, *Ubx-Antp-Scr-Dfd* and *pb-lab*, are contiguous with one another. Our data is inconclusive as to whether the remaining three Hox genes (*Abd-B, abd-A* and *Hox3*) are linked to one another or to the aforementioned Hox genes.

Hox gene expression during Parhyale embryogenesis

To gain insight into the potential function of *Parhyale* Hox genes in patterning and segmental identity, we examined their embryonic expression patterns *via in situ* hybridization. We started by examining the expression of all nine Hox genes at stage 22–24, a time when all segments are clearly visible and the different appendage types have taken on their unique morphologies (Fig. 2.2). Each Hox gene is expressed in a unique, but often overlapping, domain along the anterior–posterior (AP) axis. Hox genes generally display spatial colinearity, with the relative domains of expression along the embryonic A–P axis aligned to the relative genomic position of the individual genes within the complex. This spatial colinearity appears conserved in *Parhyale,*at least within the limits of our preliminary data on clustering: *lab, pb, Hox3* and *Dfd* expression is restricted to head segments, *Scr*and *Antp* span the head and thoracic segments, *Ubx* is expressed predominantly in the thorax, *abd-A* spans the thoracic and abdominal segments, and *Abd-B* is restricted to the abdomen (Fig. 2.2).

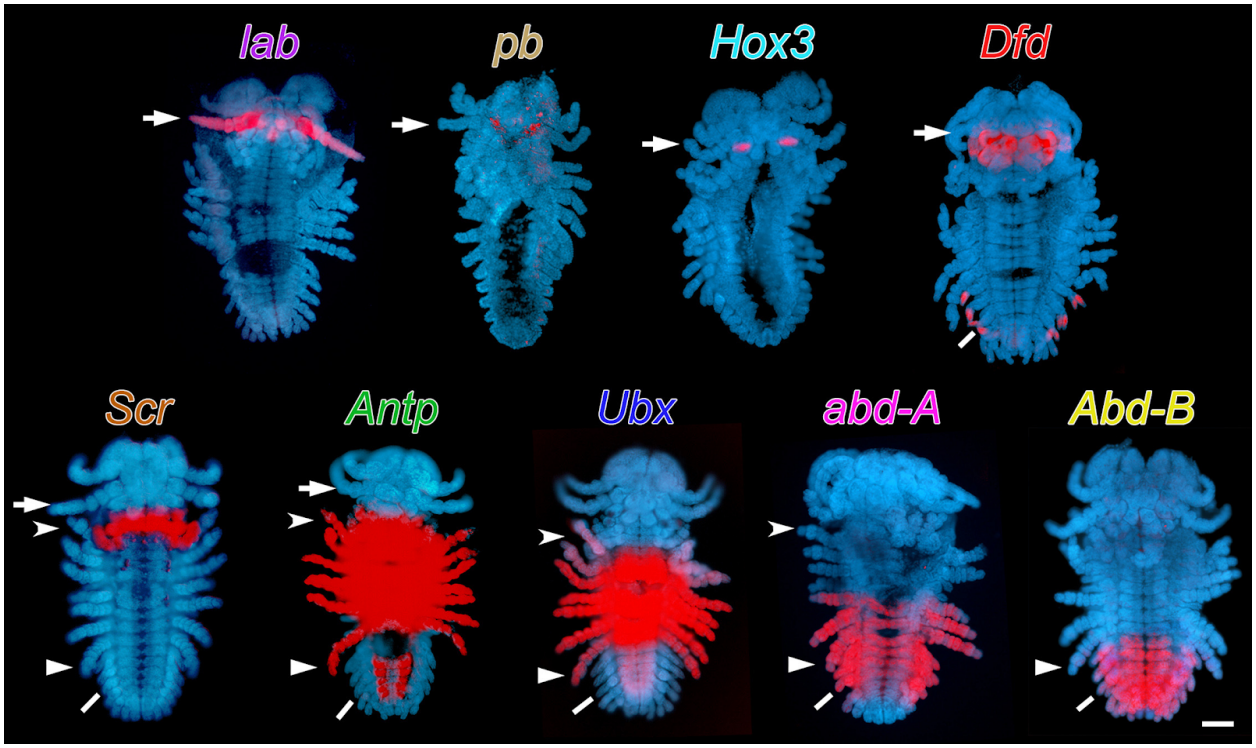


Fig. 2.2. *Parhyale* Hox gene expression
Anterior is at top. *In situ* hybridization patterns for each Hox gene in ventral views of stage 22–24 embryos. Nuclei are stained with DAPI (blue) and the corresponding Hox *in situ* pattern is in red. Specific appendages are indicated by the following symbols, arrow: antennae 2; arrowhead: T2 leg; triangle: T8 leg; bar: A3 pleopod. Scale bar 100 μm. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

For each Hox gene, we then analyzed its expression through all stages of development (see below). The stages of *Parhyale* embryogenesis discussed here have been previously described (Browne et al., 2005). The germband forms shortly after gastrulation, when most of the cells in the embryo migrate to the anterior end and condense to form the early germ disc. Over time, ectodermal cells organize themselves into a germband composed of an orderly grid beginning in the posterior part of the head. We refer to the first row that forms as "row 1," and the row that eventually forms just anterior to it as "row 0." Rows 0 and 1 together appear to comprise the future parasegment 1, which ultimately gives rise to the posterior of the Mn and the anterior of Mx1. Each transverse cell row that forms posterior to row 1 will give rise to a single parasegment, and is thus termed a parasegmental precursor row (PSPR). In succession, PSPR 2 (the cell row immediately posterior to row 1) through PSPR 16 (which gives rise to the posterior-most parasegment) undergo two rounds of mitotic division that proceed in waves from medial to lateral and from anterior to posterior. The first mitotic wave (FMW) results in two rows of progeny cells: a/b and c/d. These cells divide again during the second mitotic wave (SMW) to yield four rows of cells (a, b, c, d).

We also employed two additional markers to assist in assigning the precise boundaries of Hox gene expression relative to segmentation and appendage development: *engrailed* (*en*) and *Distal-less* (*Dll*). *Engrailed* expression can be detected in the "a" cell of each parasegment (Browne et al., 2005; Patel et al., 1989). As in other arthropods, *engrailed* expression marks the anterior boundary of each parasegment and the posterior boundary of each morphological segment. Cells anterior to parasegment 1 (*e.g.*, those that primarily give rise to the An1, An2, and the anterior majority of Mn) do not seem to organize into precise rows, nor do they exhibit regular division patterns as far as we have observed. Individual segments begin to express *Distal-less-early (Dll-e)* shortly after the onset of *engrailed* expression and, for most segments, *Dll-e* expression is initially restricted to a subset of the daughters of the "d" row, but then expands to include more anterior row progeny that contribute to the developing appendages (Browne et al., 2005; Liubicich et al., 2009). By stage 15, developing limb buds are evident in An1, An2 and Mn, and by stages 23–24 (just prior to cuticle secretion and, as such, the latest stages we can assay by *in situ* hybridization) a full complement of limbs are present and readily distinguished from one another (Fig. 2.2).

Head-specific Hox genes (*lab, pb, Hox3*, and *Dfd*)

The first Hox genes to be expressed during *Parhyale* embryogenesis are *lab* and *Dfd*. Beginning at stage 8, they are co-expressed in a single, imperfectly aligned row of cells (row 1) prior to grid formation (Fig. 2.3A-A'). After the onset of *lab* and *Dfd* expression, the ectodermal cells that will form PSPR 2 begin to organize just posterior to row 1 (see progression in Fig. 2.3A'–D'). Around this same time, cells that will become PSPR 0, just anterior to row 1, begin to weakly express *lab* (Fig. 2.3B and B'). As *lab* expression levels increase in these cells, they decrease within row 1 (Fig. 2.3C and C'). By the time that the FMW is initiated in PSPR 2, row 1 typically lacks any detectable *lab* expression—the sole

exception to this is the midline cell of row 1, which sometimes remains *lab*-positive during the initial progression of the FMW (Fig. 2.3D and D').

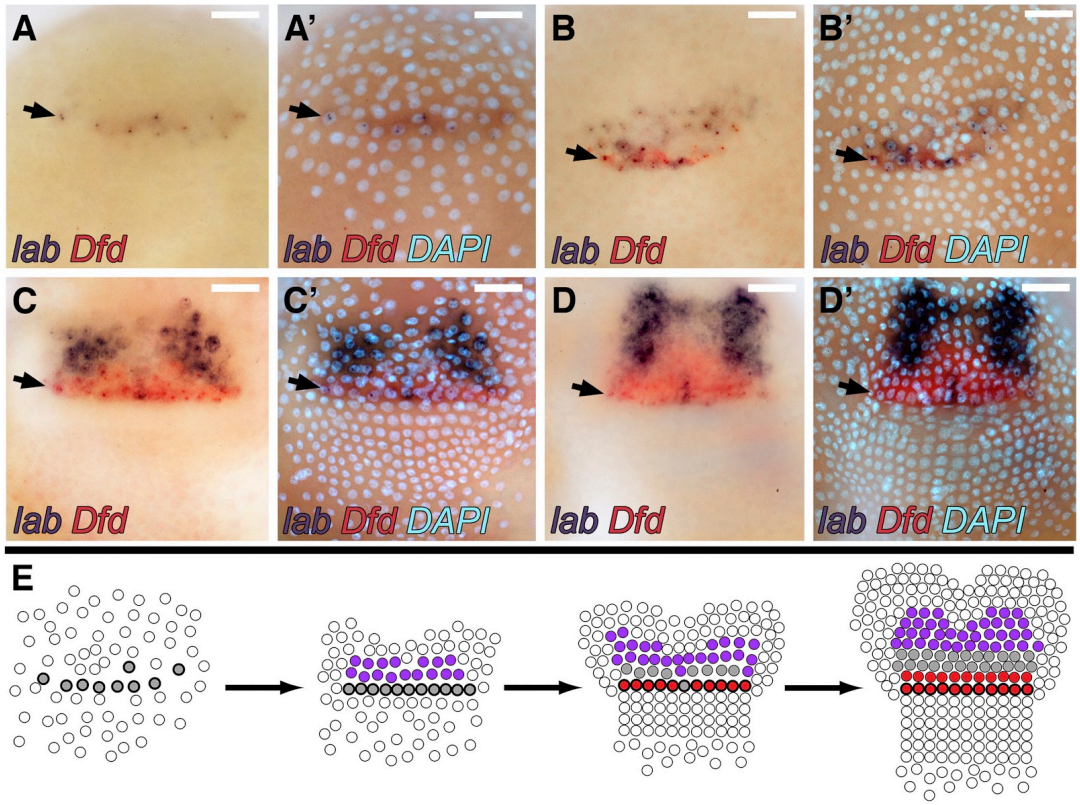


Fig. 2.3. *lab* and *Dfd* expression during early germband formation and elongation.
Anterior is at top. (A–D) Brightfield images of progressively older embryos ranging from stage 8 to 12 following double *in situ* hybridization for *lab* (purple) and *Dfd* (red). Arrows mark row 1 of the developing germband grid. (A'–D') The same four brightfield images shown in A–D, respectively, additionally showing DAPI staining in order to show *lab* and *Dfd* gene expression in relation to the organization of nuclei in the head. (E) Schematic illustration of early *lab* (purple) and *Dfd* (red) expression. Row 1 is shown in bold, and cells that express both *lab* and *Dfd* are colored gray. Scale bars 100 μm in all panels. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

When the *lab* and *Dfd* expression domains first become distinct from one another, *Dfd* is strongly expressed in row 1 and *lab* is strongly expressed in two anterior patches on either side of the midline (Fig. 2.3C and E). There is a one- to two-cell wide (in the A–P direction) region that is composed of cells that express both *lab* and *Dfd*. This co-expression region persists as the exclusively *lab* and *Dfd* expression domains on either side of it begin to expand (Fig. 2.3D and E). We observe very few cell divisions in this region during this period (based on fixed DAPI preparations and preliminary time-lapse movies), suggesting that the expansion might not be simply due to proliferation of the initial *lab* and *Dfd* expressing cells. This conjecture could be tested in the future with live imaging of GFP tagged versions of these genes.

With the onset of *en1* expression (when it becomes possible to clearly demarcate parasegmental boundaries), it is clear that *Dfd* expression occurs throughout parasegments 0 and 1 (Fig. 2.4A and B). Segmentally, this corresponds to the very posterior of the An2 segment (in the ventral body region, but not the appendages), the entire Mn segment, and most of M×1 (with the exception of the *en1*-expressing cells at the very posterior of this segment). By stage 19, as the head appendages are extending, *Dfd* expression in the Mn segment is strongest in the ventral-most region of the segment (where the paragnaths will form) and in the distal tip of the extending mandible, whereas *Dfd* expression throughout the anterior of M×1 is more uniform (Fig. 2.4C). At stage 21, this pattern remains relatively unchanged, but there is additional expression in the abdomen, specifically, in the distal regions of the A1-3 appendages (Fig. 2.4D). By stage 24, *Dfd* expression is localized to the two biramous branches of each developing pleopod (swimmeret) (Fig. 2.4E). For *lab*, during most of germband elongation, expression levels in An2 are stronger in the lateral regions of the segment, where limb buds will eventually form, than in the ventral region (Fig. 2.4F). Expression of *lab* can be observed throughout the An2 appendages as they grow, and is also detected at the posterior base of the labrum (Fig. 2.4G).

The third head-specific Hox gene to be expressed during *Parhyale* embryogenesis is *Hox3*. Its transcripts are first detected around stage 12 in two patches in the head lobes just anterior to the *Dfd* expression domain (Fig. 2.4H). These *Hox3* domains appear to be predominantly in the posterior half of the future An2 segment, although we also observe *Hox3*-positive cells throughout and immediately posterior to the An2 engrailed stripe (Fig. 2.4I). Around stage 17, there is a significant decrease in *Hox3* expression anterior to the An2 engrailed stripe (Fig. 2.4K), such that, by stage 20, *Hox3* is only found in posterior and medial (with regards to the proximal-distal axis) regions of the developing mandibles (Fig. 2.4J). By stage 24, *Hox3* expression becomes localized to a small patch of mesodermal cells in the mandible (Fig. 2.2).

Compared to the other three head Hox genes, *pb* comes on relatively late in development. Transcripts are first observed around stage 17 at the anterior base of the second antennae (not shown). This expression domain then extends posteriorly as embryogenesis progresses, such that, by stage 19, *pb* transcripts can be found at the ventral-most base of antenna 2 (Fig. 2.4L). This staining is still observed through stages 22–24, at which point additional *pb* staining can be detected at the ventral, anterior-most region of the paragnaths of the Mn segment (Figs. 2 and 4M).

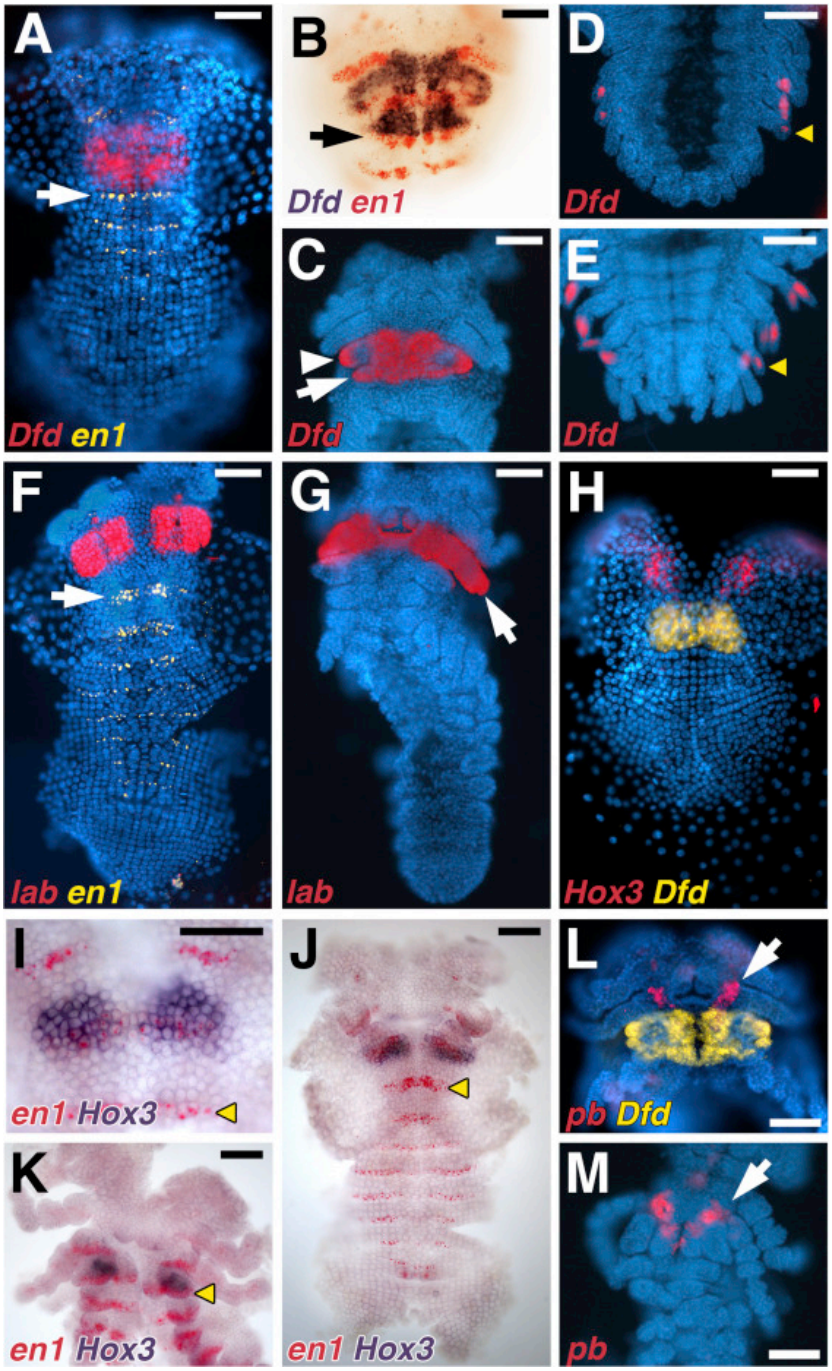


Fig. 2.4. *Dfd, lab, Hox3* and *pb in situ* hybridization of stage 14–24 embryos.
Anterior is at top. A, C–H, and L–M are DAPI stained with *in situ* patterns shown in red or yellow; B and I-K are brightfield images with *in situ* hybridization patterns in purple and red. In panels A, B, F and I–J, *en1* is used as a marker for the posterior boundary of each segment. (A) *Dfd* (red) and *en1* (yellow) expression in a stage 14 embryo. Arrow marks Mx1 *en1* stripe. (B) *Dfd* (purple) and *en1* (red) expression in a stage 17 embryo; arrow marks Mx1 en1 stripe. In this embryo, *Dfd* transcripts are observed throughout parasegments 0 and 1, which corresponds to the posterior of An2 through the anterior of M×1. (C) *Dfd* (red) expression in a stage 19 embryo. Arrowhead, Mn segment; arrow, M×1 segment. (d-E) *Dfd* (red) expression in the developing pleopods during stages 21 and 24, respectively. Arrowheads mark the developing A3 pleopod. (F) *lab* (red) and *en1*(yellow) expression in a stage 16 embryo; arrowhead marks Mn *en1*stripe. (G) *lab* (red) expression in An2 (arrow) of a stage 19 embryo. (H) *Hox3* (red) and *Dfd* (yellow) expression in a stage 12 embryo. (I) Close up of *Hox3* (purple) and *en1* (red) expression in a stage 13 embryo; arrowhead marks Mn *en1* stripe. *Hox3* expression is observed in the posterior half of the An2 segment and the anterior-most cells of the Mn segment. (J) *Hox3* expression (purple) is restricted to the developing mandibles in a stage 17 embryo. Expression of *en1* shown in red; arrowhead marks Mn *en1* stripe. (K) *Hox3* (purple) and *en1* (red) expression in a stage 20 embryo; arrowhead marks Mn en1 stripe. (L) *pb* (red) and *Dfd* (yellow) expression in a stage 19 embryo. *pb* is restricted to the base of the An2 appendages (arrow). (M) *pb* expression (red) at the base of the An2 appendages (arrow) and in the paragnaths of a stage 22 embryo. Scale bars 100 μm in all panels. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Hox genes spanning the head and thorax (Scr and Antp)

Two Hox genes span head and thoracic segments: *Scr* and *Antp. Scr* is first expressed around stage 12 in the PSPR 2 cells as they undergo their first round of division to produce the a/b and c/d daughter rows (Fig. 2.5A). By stage 15, *Scr* expression expands anteriorly into parasegment 1 and posteriorly into the a and b cells of parasegment 3 (Fig. 2.5B and C). Together, these cells will give rise to Mx1 through the anterior region of the first thoracic segment (T1). It should be noted that *Scr* expression levels tend to be weaker in Mx1 than in Mx2 throughout development (Figs. 5C, D and 2). During stage 19, as the limb buds begin elongating, *Scr* expression appears more intense in the most distal and proximal regions of the maxillipeds, with only weak or no expression in the medial region, whereas expression throughout the maxillary appendages appear more uniform (Fig. 2.5D). By stage 23, *Scr* expression is present in both maxillary segments and their appendages, as well as in the T1 appendages (maxilliped), but not the T1 ventral body wall (Fig. 2.2).

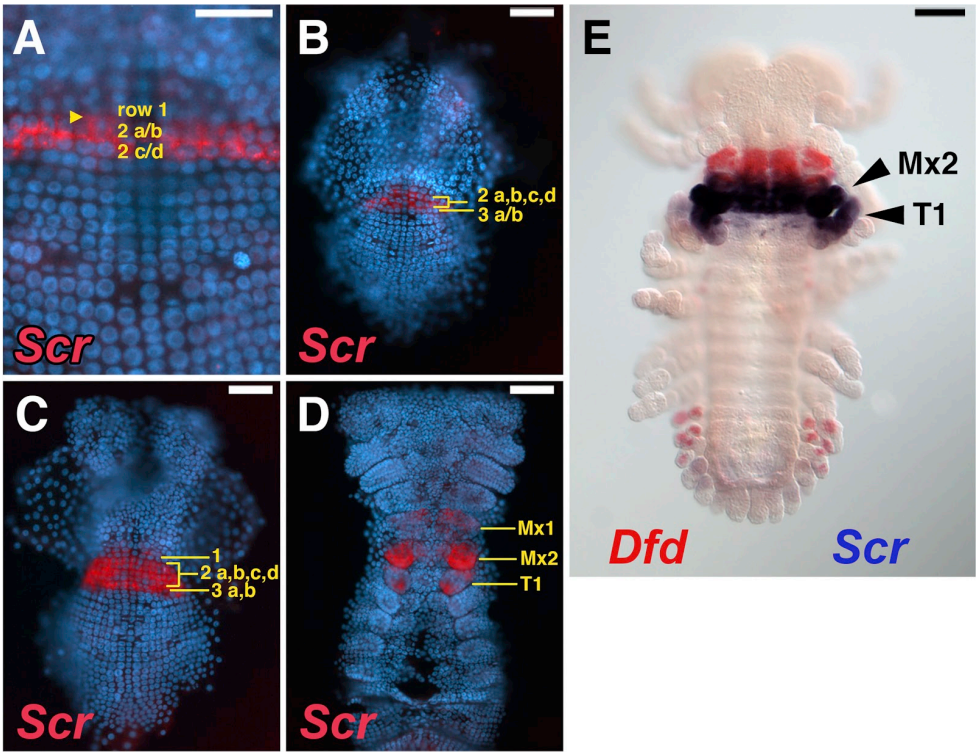


Fig. 2.5. *Scr* in situ hybridization to stage 12–19 embryos
Anterior is at top. Embryos in A–D are DAPI stained with *in situ* patterns in red. (A) Close up of *Scr* expression in rows 2 a/b and c/d in a stage 12 embryo. Arrowhead: posterior edge of row 1 (B) Stage 14 embryo with *Scr* transcripts in 2 a, b, c, d and 3 a/b. (C) The *Scr* domain expands anteriorly to include cells in parasegment 1 in stage 15 embryos. (D) Stage 19 embryo showing *Scr* expression in M×1, M×2 and in the T1 appendage. (E) Stage 24 embryo showing *Dfd* (red) and *Scr* expression (black). *Scr* expression is seen throughout the M×1 and M×2 segments and the more distal portions of the T1 appendages. *Dfd* is as far anterior as the Mn segment (expression in more posterior head segments is obscured by the *Scr* signal); expression of *Dfd* is also seen in the A1-A3 pleopods. Scale bars 100 μm in all panels. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Antp is initially expressed during stage 12 in PSPR 3 cells as they undergo their first division to give rise to a/b and c/d daughter cells (Fig. 2.6A). The expression boundaries of *Antp* then expand in both the anterior and posterior directions. Anteriorly, in the second parasegment, the c/d cells begin to express *Antp* at stage 13; these cells will eventually be a part of the Mx2 segment (Fig. 2.6B). At stage 15, the *Antp* expression domain extends from the 2c and d rows through at least parasegment 7 (Fig. 2.6C). By stage 17–18, *Antp* expression can be found in the very posterior, ventral region of Mx1, throughout all of Mx2 and the developing thoracic segments, and at lower levels in the most ventral (neurogenic) region of the abdominal ectoderm (Fig. 2.6D). *Antp* expression in Mx2 appears weaker in general than *Antp* expression in the thoracic segments. This expression pattern persists through at least stage 24, the latest time at which we can carry out *in situ* hybridization of whole mount embryos, and the late abdominal *Antp* expression is restricted to the nervous system (Figs. 2 and 6E).

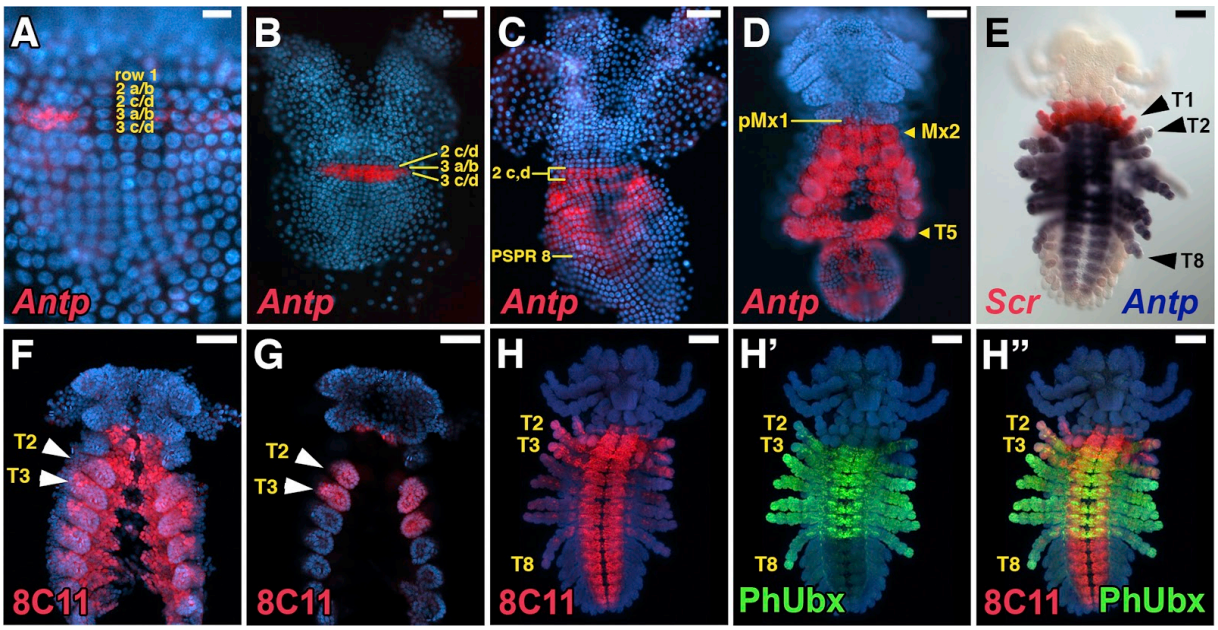


Fig. 2.6. *Antp in situ* hybridization and immunolocalization.
(A) Close up of a stage 12 embryo with *Antp* transcripts accumulating in rows 3 a/b and c/d. (B) *Antp* expression expands anteriorly into 2 c/d by stage 13. (C) Stage 15 embryo displaying *Antp* expression from parasegments 2 through 8. (D) Stage 19 embryo. The most anterior domain of expression is in the CNS in the posterior portion of the Mx1 segment. Expression is rather uniform throughout M×2 and the entire thorax (thoracic segments posterior to T5 are not visible in this view due to the flexure of the embryo at this stage). Within the abdomen, *Antp* transcripts are expressed in primarily in the neurogenic region. (E) Stage 24 embryo with *Scr* in red and *Antp* in purple. At this stage, *Antp*expression in the head is largely confined to the CNS, is and detectable in the T1 appendage where it overlap with *Scr* expression. (F–G) Immunodetection of Antp proteins with the 8C11 monoclonal antibody suggests post-transcriptional repression in the T4-T8 appendages. (F) Extended focus view of a stage 20 embryos shows expression throughout the thorax in the nervous system and body wall, but (G) optical sections through the developing limbs reveals that Antp protein is not detected in the T4–T8 appendages. (H–H") At stage 24, neuronal expression of Antp (red) is seen from the posterior of the head through to the end of the abdomen. In the appendages and body wall, however, widespread expression is seen in T2 and T3, but in the remaining thoracic appendages, expression is restricted to a small subset of neurons and mesodermal cells. Ubx protein (green) is detected from T2 through T8. Scale bar 20 μm in panel A, 100 μm in panels B–H". (For interpretation of the references to color in this Āigure legend, the reader is referred to the web version of this article.)

The description provided above for *Antp* transcript distribution in *Parhyale* is based on *in situ* hybridization data, but relating this to where *Antp* functions presents an unusual problem in crustaceans. In at least five previously studies crustacean species, a *Ubx-Antp* bi-cistronic transcript is produced, presumably the result of transcriptional read-through from the *Ubx* promoter (Shiga et al., 2006). In the case of *Daphnia* and *Artemia*, these bi-cistronic transcripts are incapable of generating functional Antp protein (Shiga et al., 2006). Thus, *in situ* probes for *Antp* actually detect the combined expression domains of *Antp* plus *Ubx*. In the case of *Daphnia* and *Artemia*, it was shown that Antp protein was not produced by the bi-cistronic transcripts, and indeed Antp protein was produced in a domain that did not overlap with Ubx.

Parhyale also produces *Ubx-Antp* bi-cistronic transcripts. *Ubx* transcript and protein in *Parhyale* are expressed at moderate levels in T2 and T3, and higher levels from T4-T8 (Liubicich et al., 2009; also summarized below). Therefore, it seemed possible that *Antp* transcripts produced from the *Antp* promoter—*i.e.*, those that make functional Antp protein—are not expressed in the more posterior thoracic segments and appendages, and that the staining observed in this region is solely due to our *Antp* probe hybridizing to *Ubx-Antp* hybrid transcripts. To test this possibility, we used two approaches to detect the distribution of Antp protein. First, we used a previously characterized pair of monoclonal antibodies (MAb 8C11 and MAb 4C3) that have been shown to detect Antp protein across various arthropods (Condie et al., 1991; Hayward et al., 1995; Saenko et al., 2011). Immunostaining reveals that Antp protein is initially broadly expressed in the neurogenic region from Mx2 on posterior, but within the developing appendages, expression is not detected posterior to T3 (Fig. 2.6F and G). At later stages, the discrepancy between the *in situ* results and antibody staining is even more marked (compare Fig. 2.6E and H). Antp protein is detected throughout the entirety of the T1–T3 limbs, but restricted to just a small subset of what appear to be neurons and muscle within the developing limbs in T4–T8 (Fig. 2.6H). Within the nervous system, expression is detected throughout the entire CNS starting at M×2 (Fig. 2.6H). Thus, there is overlap of Ubx and Antp protein expression in the CNS, but within the limbs, the two overlap in T2 and T3, but in the remaining thoracic segments, Ubx is expressed throughout the appendages, but Antp is restricted to a small subset of cells (Fig. 2.6H' and H").

To further confirm the restriction of Antp to the anterior part of the thorax, we used a CRISPR/Cas9 homology mediated knock-in strategy to create embryos in which GFP is inserted in frame with the *Parhyale Antp* coding sequence (see Section 2 and Fig. 2.7A). This approach yielded GFP- expressing G0 embryos that confirmed the results seen with the 8C11 and 4C3 monoclonal antibodies. Especially striking were late stage embryos (stage 24 and beyond that are stages in which we cannot carry out routine *in situ* or immunohistochemistry due to the cuticle) where we see GFP thoracic appendage expression restricted T1-T3 (Fig. 2.7B and C). Thus, we are able to conclude that, as in *Daphnia* and *Artemia*, Antp protein does not appear to be expressed from the bi-cistronic transcripts in *Parhyale*, and Antp and Ubx overlap less than expected from the *in situ* data.

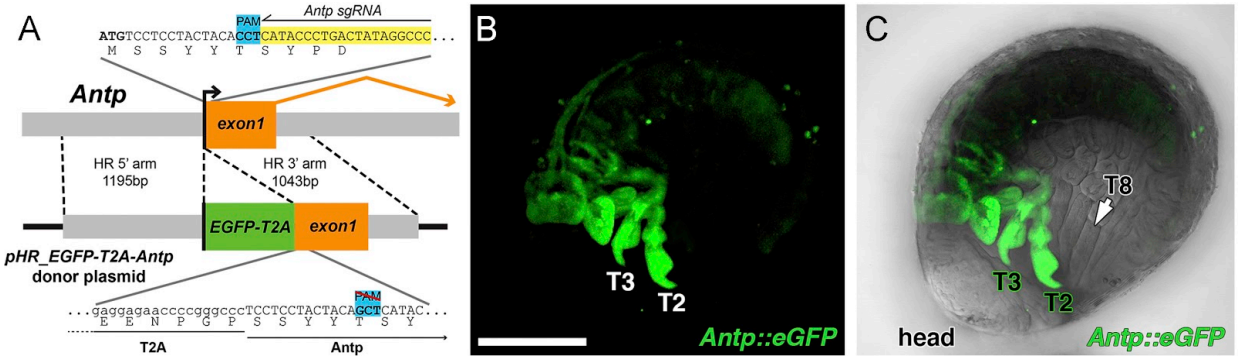


Fig. 2.7. GFP reporter of *Antp* via CRISPR/Cas9 genome editing.
(A) Diagram of the strategy for somatic fluorescent tagging of Antp. A CRISPR gRNA targets the 5′ region of the Antp coding sequence (yellow) and was co-injected with a Cas9 recombinant protein at the one-cell stage. A donor plasmid was provided as a repair template for homologous recombination, generating alleles that incorporate eGFP and the T2A ribosome-skipping site to the N-terminus of the Antp protein. Secondary cleavage of repaired alleles is avoided by a degenerate protospacer-adjacent motif (PAM) sequence (blue). (B, C) Live fluorescent imaging of an eGFP-positive G0 CRISPR-induced somatic transgenic animal at stage 27. Note the strong signal in T2-T3 limbs, while little signal is seen in the remaining more posterior thoracic limbs. Scale bar 100 μm. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Thoracic Hox gene (Ubx)

We previously published a detailed account of *Ubx* transcript and protein expression during *Parhyale* embryogenesis (Liubicich et al., 2009). By stage 23, when *Ubx* expression reaches its full fruition, it is expressed at lower levels in the gnathopods of T2 and T3, and more robustly in the walking legs of T4 through T8 (Figs. 2, 6H′, and 8D). There is transient expression of *Ubx* transcripts in the neurogeneic region of the abdomen at stages 17–18, which then resolves to a small subset of neuronal cells before ceasing expression in the abdomen.

Abdominal Hox genes (abd-A and Abd-B)

The Hox genes that are expressed in the *Parhyale* abdomen are *abd-A* and *Abd-B*. In *Parhyale*, we first detected *abd-A* transcripts at stage 17 in the a/b and c/d cells of parasegments 12 and 13—these cells will eventually give rise to the posterior of A1 through the anterior of A3 (Fig. 2.8A). At stage 18, we also observe weaker levels of *abd-A* in parasegments 8 through 11, which will go on to form the very posterior of T5 through the anterior of A1. As the limb buds start protruding during stage 19, *abd-A* is expressed strongly in A1-3, more weakly in T6-8, and it is just starting to come on in A4 (Fig. 8B). During stages 21-22, *abd-A* expression appears in the T6-8 appendages, but is absent from the distal most tips, whereas the strong A1-3 and very weak A4 expression appears more homogeneously throughout the limbs (Fig. 2.8C). By stage 24, we consistently observe *abd-A* expression in the ventral ectoderm from the posterior of T5 through the anterior of

A3, and in the appendages of T6 through A3, with very weak staining in the limbs of A4, and no staining in A5, A6 or the telson (Figs. 2 and 8D).

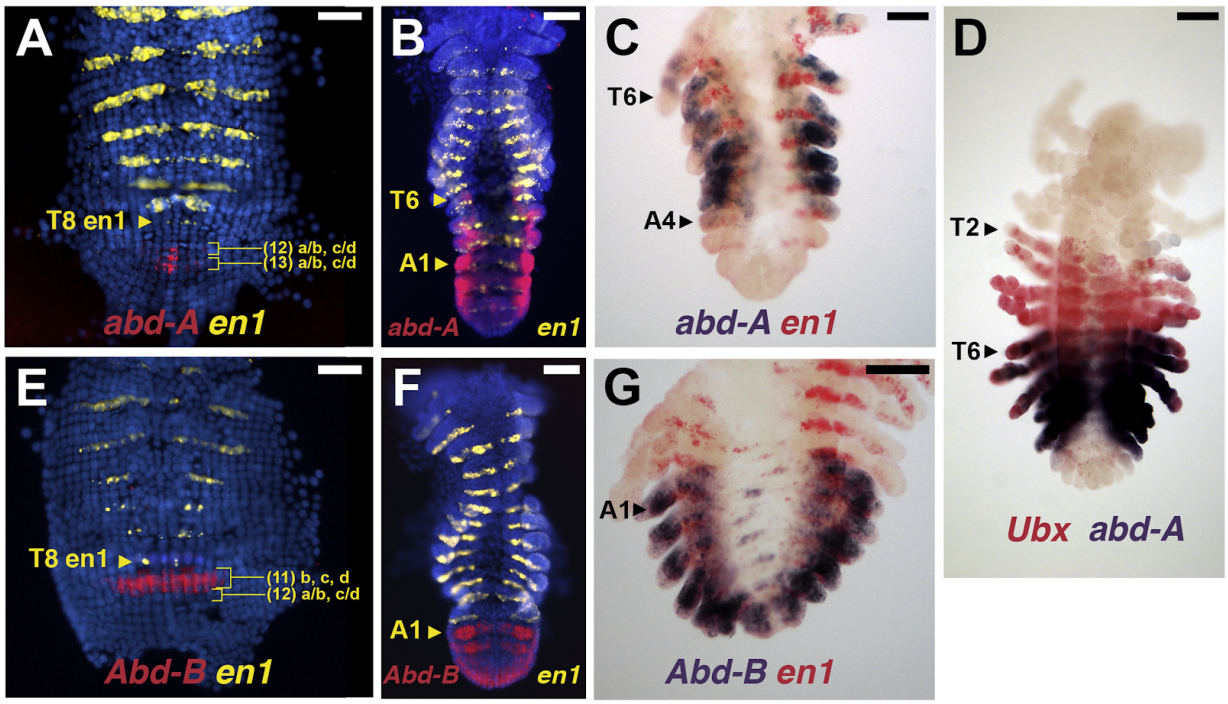


Fig. 2.8. *abd-A* and *Abd-B* in situ hybridization in stage 17–22 embryos.
Anterior is at top. A, B, E and F are DAPI stained with *in situ* patterns in red and yellow; C, D and G are brightfield images with *in situ* hybridization patterns in purple and red. (A) Close up of a stage 17 embryo with *abd-A* expression in developing parasegments 12 and 13; *en1* expression is shown in yellow. (B) *abd-A* (red) and *en1* (yellow) expression in a stage 19 embryo. (C) Stage 21 embryo with *abd-A* expression (purple) at moderate levels in the T6-T8 appendages, at higher levels in the A1–A3 appendages, and at low levels in the A4 appendage; *en1* staining shown in red. (D) Stage 24 embryo with *Ubx*in red and *abdA* in purple. The anterior boundary of *Ubx* is at T2, while *abdA* is expressed thought T6-A3, with low levels in A4. (E) Stage 17 embryo with *Abd-B* transcripts first detectable in rows 11b, c and d, and in 12 a/b and c/d; *en1* expression is shown in yellow. (F) *Abd-B* (red) and *en1* (yellow) expression in a stage 19 embryo. Expression is seen throughout the entire abdomen. (G) Stage 22 embryo with *Abd-B* expression (purple) throughout segments A1-6; *en1* staining shown in red. Scale bars 100 μm in all panels. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Abd-B is first expressed at stage 17–18 in the b, c, and d cells of parasegment 11, and in the a/b and c/d cells of parasegment 12 (Fig. 2.8E). By stage 19, *Abd-B* expression remains absent from the T8 en1 stripe, but is present in all of the other cells of parasegment 11, plus the developing parasegments 12, 13 and 14, which together give rise to the anterior of A1 through the anterior of A4 (Fig. 2.8F). As the germband continues to expand, *Abd-B* is expressed in all of the developing abdominal segments and appendages (Fig. 2.8G). Finally, at stage 24, *Abd-B* is rather homogeneously expressed throughout the ventral ectoderm from the very posterior of T8 through A6, and in the developing limbs of A1 through A6, but is excluded from the developing telson (Fig. 2.2).

2.4 Discussion

Previous studies have examined the expression of subsets of crustacean Hox genes, most notably in the branchiopods *Artemia* and *Daphnia*, the maxillopod *Sacculina*, and in the malacostracans *Porcellio*, *Procambarus* and *Asellus* (Abzhanov and Kaufman, 1999a, 1999b, 2000a, 2000b; Averof and Akam, 1995; Blin et al., 2003; Brena et al., 2005; Copf et al., 2003; Mouchel-Vielh et al., 2002; Papillon and Telford, 2007; Shiga et al., 2002; Vick and Blum, 2010). The work reported here represents the most thorough cloning and expression analysis of Hox genes in a crustacean to date. We have isolated sequences (in most cases, full-length cDNAs) for nine *Parhyale* Hox genes and have shown that at least six of them are linked to other Hox genes in the genome. Further experiments are necessary to determine whether all of these genes lie together in a single cluster, or whether the *Parhyale* Hox complex has been split, as is the case for other arthropod species (Negre and Ruiz, 2007; Yasukochi et al., 2004).

Parhyale has at least eleven distinct appendage types, and each type develops in the presence of a unique set of Hox genes (Fig. 2.9). For example, as the appendages are developing in the head, the first antenna does not express any Hox genes. The second antenna expresses *lab* and *pb*. For the mandibular (Mn) segment, *Dfd* is expressed throughout the segment, with additional expression of *pb* in the paragnaths and *Hox3* in the mandibles. The first maxillae express both *Dfd* and *Scr*; the second maxillae express *Scr* strongly and *Antp* weakly. Within the anterior thorax *Antp* is the only Hox gene expressed in the T1 body wall, although both *Scr* and *Antp* are strongly expressed in the T1 appendages (maxillipeds), with *Scr* appearing significantly later in embryogenesis; the gnathopods of T2 and T3 are unique in that they express *Antp* plus moderate levels of *Ubx*. High levels of *Ubx* are found in the walking legs of T4–T8, with the posteriorly oriented T6-T8 walking legs additionally expressing *abd-A*. Analysis of Antp protein expression by immunohistochemistry and GFP knock-in reveals that there is initial Antp expression throughout the thorax in the neurogenic region, but expression in the developing appendages is largely restricted to T1-T3, with T4-T8 expression in the limbs restricted to a small subset of cells, most likely neurons and muscle cells. Within the abdomen, the pleopods (swimmerets) of A1-3 express both *abd-A* and *Abd-B*; and *Abd-B* is the only Hox gene expressed at high levels in the uropods of A4-6. Antp protein is expressed in the abdomen, but is restricted to the nervous system. All these patterns are summarized schematically in Fig. 2.9. These patterns suggest that a complex "Hox code" may be responsible for establishing the morphological differences that exist between these appendage types (Fig. 2.9). We also note that while the initial expression patterns of the Hox genes follow parasegmental boundaries, and such parasegmental boundaries are maintained for expression in the central nervous system, the expression boundaries in the developing and mature limbs are instead segmental. We previously showed for *Ubx* that the initial parasegmental expression boundary retracted back to a segmental boundary in the ectoderm before limb growth began (Liubicich et al., 2009), and this behavior is also seen here for several of the other Hox genes.

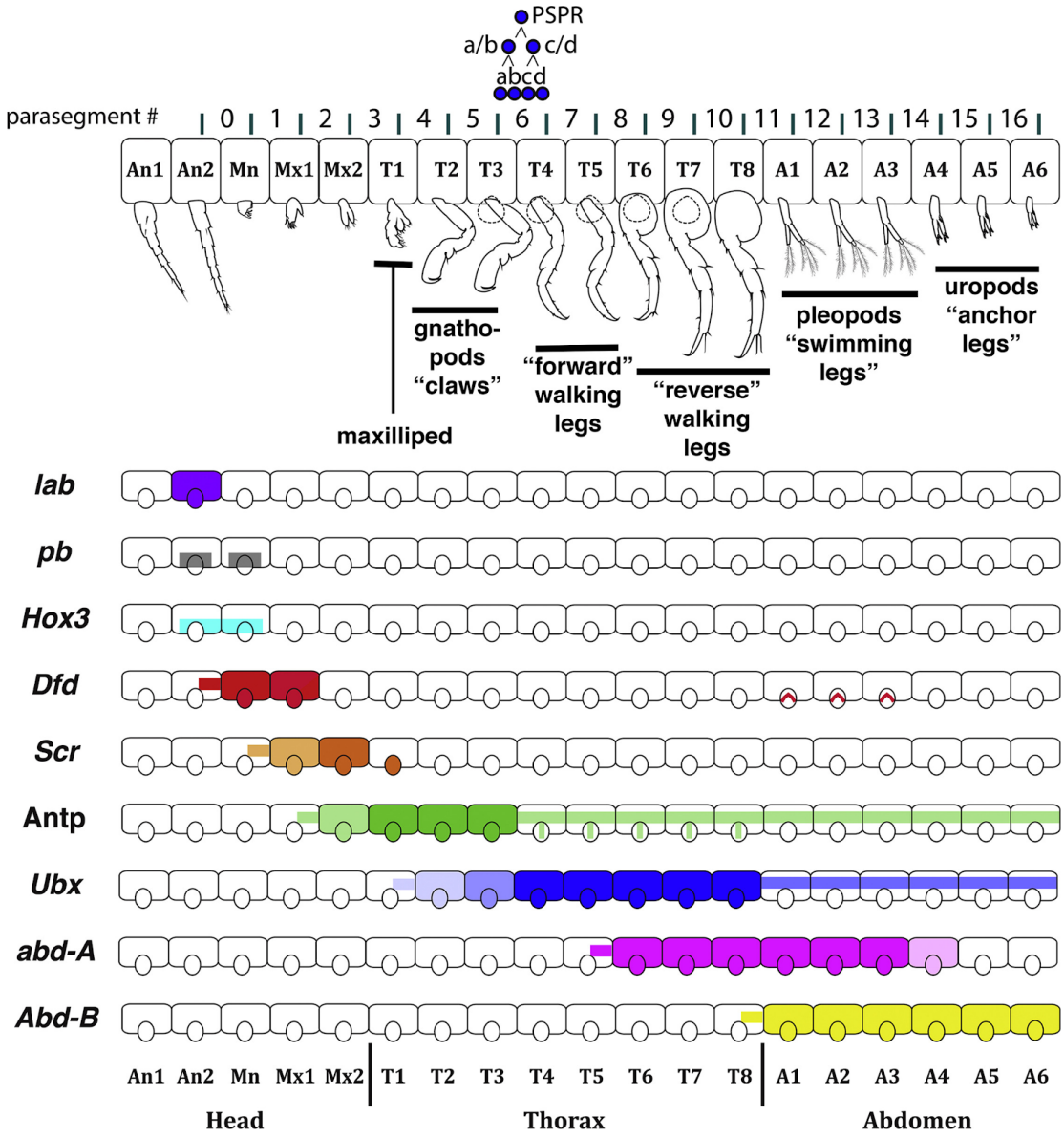


Fig. 2.9. Summary of Hox gene expression domains in Parhyale.
At the top is an illustration of the different limb morphologies seen along the *Parhyale* body axis, with an alignment to the numbered Parasegment Precursor Rows (PSPR) and progeny rows (a, b, c, d) that contribute to each segment. Colored bars through the middle of a segment indicate that expression is largely confined to the CNS for these segments (and in the case of *Ubx* the expression in the abdominal CNS is relatively transitory). In many cases, the anterior limit of expression is parasegmental for the CNS, but segmental for the ectoderm and appendages, although in a few cases the anterior CNS domain is transient. The Antp domain shown here represents expression of the protein, as *in situ* hybridization detects the *Antp-Ubx*bi-cistronic message, and bars through the T4-T8 limbs represent expression in a subset of neurons and mesodermal cells in these limbs. For *Ubx*, the CNS expression in the abdomen is transient as shown in Liubicich et al. (2009). For *pb* and *Hox3*, expression occurs within very limited domains that does not include the entire appendage. The chevrons in the A1-3 segments for *Dfd* represent the late patterns seen within these appendages. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

As in many other animals, *Parhyale* Hox genes also exhibit temporal colinearity—that is, those genes located at the proximal or 3′ end of the Hox cluster (*e.g.*, *lab*) are transcribed prior to more distal, 5′ end genes (*e.g.*, *Abd-B*). The only exceptions to this temporal colinearity in *Parhyale* are *pb*, which is located on the proximal end of the Hox cluster but is one of the last Hox genes to be expressed, and *Hox3*, which is first expressed after *Dfd*, but before (or perhaps coincident with) *Scr*.

The only Hox ortholog that we were unable to isolate from *Parhyale* is ftz. Typically, ftz protein sequence, function and expression profiles are highly divergent in arthropods, where they appear to play a role in neurogenesis, segmentation, and/or segment identity depending on the species (Damen, 2002; Heffer et al., 2010,2013). While we cannot rule out the existence of a *ftz*ortholog in *Parhyale*, our degenerate PCR primers targeted the LELEKEF and KIWFNQR motifs that are conserved in all pancrustaceans (insects and crustaceans) where *ftz* has been isolated (Heffer et al., 2013). Furthermore, genomic sequences obtained from contiguous *Parhyale* BAC clones show that there is no *ftz* between *Scr* and *Antp*, its canonical position in arthropods (Chipman et al., 2014). It is worth mentioning that while *ftz* genes have been isolated from brachiopod and maxillopod crustaceans, no *ftz* ortholog has been reported for any malacostracan crustacean—this includes both *Porcellio* and *Procambarus*, for which numerous Hox genes (including *Scr* and *Antp*) have already been isolated (Abzhanov and Kaufman, 1999a,1999b,2000a,2000b). Whether *ftz* was lost in the entire malacostracan lineage poses an interesting question that awaits further analysis.

The other Hox gene that has diverged significantly during arthropod evolution is *Hox3*. While *Hox3* appears to play a segment identity role in chelicerates and myriapods, in insects (where it is often called *zen*) it typically serves a function in extra-embryonic tissues (Abzhanov and Kaufman, 1999b; Falciani et al., 1996; Hughes and Kaufman, 2002b; Janssen and Damen, 2006; Panfilio et al., 2006; Telford and Thomas, 1998). Prior to this report, only one *Hox3* ortholog had been identified in a crustacean: *Daphnia Hox3*. Its expression pattern, initially expressed in An2 and Mn before ultimately becoming restricted to the mesoderm of the developing mandible (Papillon and Telford, 2007), is strikingly similar to what we observe for *Parhyale Hox3*. Interestingly, *Hox3* is also expressed in the mandibular mesoderm in the centipede *Lithobius* and in *Thermobia,* a member of the early diverging insect lineage Thysanura (Hughes and Kaufman, 2002a; Hughes et al., 2004). It has been suggested that this conserved pattern may have arisen early within Mandibulata, and perhaps may have even played a role in the origin of mandibles (Papillon and Telford, 2007). Our results demonstrate that *Hox3*expression in the mandibular mesoderm also occurs in at least one malacostracan, lending further credence to this hypothesis.

Prior to this work, the isopod *Porcellio* was the only crustacean for which *lab*, *pb* and *Dfd* expression patterns were known (Abzhanov and Kaufman, 1999b). Given that *Porcellio* are isopods (closely related to amphipods), it is not surprising that such expression patterns are similar to, albeit slightly different from, those we observe in *Parhyale*. In both species, *lab* is expressed exclusively in An2 and its appendages, while *pb* is expressed at the base of

the second antennae as well as in the paragnaths, a pair of lobes that act as the lower lips of the mandibular apparatus (Wolff and Scholtz, 2006). Both species also exhibit *Dfd* expression in Mn (including the paragnaths and mandibles) and the anterior portion of the Mx1 segment. However, they differ in that *Parhyale Dfd* is additionally expressed in the first maxillae and in the developing swimmerets of A1-A3, while *Porcellio Dfd* is not. To the best of our knowledge, *Parhyale* is the only arthropod for which *Dfd*expression has been observed in abdominal appendages. It will be interesting to examine whether *Dfd* plays a significant role in pleopod development.

Compared to the head-specific Hox genes, *Scr*, *Antp* and *Ubx*have all received considerably more attention in crustaceans. *Ubx*expression is typically restricted to the thorax, and its anterior boundary marks the transition between feeding appendages (*e.g*., maxillae or maxillipeds) and locomotory appendages (Averof and Patel, 1997). Our previously published expression and functional data on *Parhyale Ubx* confirms that this gene plays a role in specifying this transition (Liubicich et al., 2009; Pavlopoulos et al., 2009).

The expression patterns associated with *Scr* have been previously examined in two malacostracan crustaceans: *Porcellio* (which, like *Parhyale*, has a single pair of maxillipeds on T1) and the crayfish *Procambarus* (which have three pairs of maxillipeds located on segments T1-T3). In *Porcellio*, *Scr* transcripts are expressed in Mx1 and Mx2 segments and their appendages, and in the T1 maxilliped (Abzhanov and Kaufman, 1999a); these expression boundaries coincide with what we see in *Parhyale*. The *Procambarus Scr* pattern is similar to this except for the absence of *Scr* transcripts in Mx1 (Abzhanov and Kaufman, 2000a). In both *Porcellio* and *Procambarus*, Scr protein is detected early in second maxillae, but comes on much later in the T1 maxilliped, presumably as a result of post-transcriptional regulation (Abzhanov and Kaufman, 1999a). Whether a similar post-transcriptional regulation of Scr protein occurs in *Parhyale* awaits future experiments.

Our *in situ* hybridization experiments using probes for *Parhyale Antp* reveal a pattern similar to that reported for *Porcellio Antp* in that both are observed in Mx2 through most of the thorax, with additional staining in the abdominal neuroectoderm and CNS. A similar pattern of *Antp* mRNA expression (*i.e*., from Mx2 through most or all of the thorax) has been observed in the branchiopods *Artemia* and *Daphnia*, and during the early stages of *Procambarus*development (Abzhanov and Kaufman, 2000a; Averof and Akam, 1995; Shiga et al., 2002). However, it has been shown that all of these species produce *Ubx-Antp* read-through transcripts, and in at least two of these species (*Artemia* and *Daphnia*), Antp protein is not translated from these chimeric transcripts (Shiga et al., 2002, 2006); we find the same appears to be true in *Parhyale*. This suggests that the previously described posterior boundaries for *Antp* expression (*i.e*., at the border between the thorax and abdomen) may merely reflect transcription of the *Ubx-Antp*readthrough transcripts from the *Ubx* promoter(s) rather than bona fide *Antp* expression. Consistent with this, Antp protein expression has been found to be more anteriorly restricted in several of these species than the initial *in situ* hybridization analyses suggested. For example, Antp protein is only expressed in Mx1 in *Artemia*, and in Mx1 and T1 of *Daphnia* (Shiga et al., 2002,

2006). Furthermore, during later stages in *Procambarus*, *Antp* transcripts are restricted to Mx1 through T3, suggesting that no *Ubx-Antp*read-through transcripts are being produced at this late stage (Abzhanov and Kaufman, 2000a). Intriguingly, in all three of these cases, *Antp*-specific transcripts and/or Antp protein are only found in Mx2 and in those thoracic appendages that do not express Ubx protein, or, in the case of *Parhyale*, express Ubx at lower levels than the rest of the thorax. This raises the possibility that *Ubx* may negatively regulate *Antp* in these species, perhaps by acting as a transcriptional repressor of the *Antp* promoter as it does in *Drosophila* (Beachy et al., 1988). Alternatively, the production of *Ubx-Antp* readthrough transcripts might somehow physically interfere with the transcription of *Antp*-specific transcripts from the *Antp* promoter in these species.

We found that *Parhyale abd-A* is expressed in T6–8 and A1–A3, and is detected at very low levels in A4. This general pattern of being expressed in the posterior-most thoracic segments and the anterior-most abdominal segments has also been observed in the three other malacostracans for which *abd-A* expression has been examined: *Asellus, Porcellio* and *Procambarus. In situ* hybridizations for *Asellus abd-A* revealed staining in T6-T8, in A1-A2, but not in A3-A6 (Vick and Blum, 2010). *Porcellio abd-A* has three splicing variants: *abd-A1* and *abd-A2* are expressed in a graded manner (stronger in anterior, weaker toward the posterior) in A1-5, but not A6. *abd-A3* also shows a graded expression pattern in A1-A5, with additional expression in T6–T8 (Abzhanov and Kaufman, 2000b). *Procambarus abd-A* is similarly expressed in A1-A5, and shows additional expression in the appendages of T7 and T8 (Abzhanov and Kaufman, 2000a). The authors suggested that *abd-A* may play two roles in *Procambarus*, namely, specifying abdominal identity and differentiating posterior thoracic appendages from more anterior ones. Consistent with this idea, in *Parhyale*, the *abd-A* expressing walking legs (T6-T8) differ significantly in morphology from their more anterior counterparts (T4–T5)—they are larger, have thicker bristles, and are positioned in the opposite orientation (*i.e*., curved posteriorly rather than anteriorly).

There also appears to be a correlation between *abd-A* expression and morphological differences in the abdomen. For instance, in *Asellus, abd-A* is expressed in the two free abdominal segments (A1–A2), but not in the pleotelson, a structure specific to some groups of isopods and comprised of A3-6 fused with the telson (Vick and Blum, 2010). In both *Porcellio* and *Procambarus, abd-A*is expressed in A1–A5, which give rise to pleopods (swimmerets), but not in A6, which gives rise to a single pair of uropods. *Parhyale* and other amphipods have a different arrangement of abdominal appendages: they bear pleopods on A1-A3 and uropods on A4-A6. Intriguingly, *Parhyale abd-A* is expressed strongly in the pleopod-bearing segments (A1-A3), and either very weakly or not at all in the uropod-bearing segments (A4–A6). This raises the possibility that, in malacostracans, evolutionary changes in the posterior boundary of *abd-A* have facilitated morphological diversity in abdominal appendage type (*e.g*., number of pleopods versus uropods) in a manner similar to how changes in the anterior boundary of *Ubx* lead to diversity in thoracic appendage type (*e.g*., number of maxillipeds versus walking legs) (Averof and Patel, 1997; Liubicich et al., 2009; Pavlopoulos et al., 2009).

To date, *Abd-B* expression has only been examined in two malacostracans, *Porcellio* and *Parhyale*, and in both cases, it is expressed in all abdominal segments and their appendages (Brena et al., 2005; this work). This suggests that *Abd-B* may be involved in specifying abdominal identity in these species. A hallmark feature of malacostracan abdominal appendages (*i.e.*, pleopods and uropods) is their biramous nature. It will be interesting to determine whether *Abd-B* plays a role in directing biramous appendage development, and/or in specifying uropod identity (as *Abd-B* is the only Hox gene expressed in the uropods of *Porcellio* and *Parhyale*).

In summary, our analysis provides the most complete picture to date of Hox gene expression in a crustacean. Hox gene expression in *Parhyale* generally follows the properties of spatial and temporal colinearity, with a complex pattern of overlapping domains that suggests a "Hox code" that may specify the eleven distinct appendage types observed in *Parhyale*. Given that techniques for gene misexpression and knock down have already been established in *Parhyale*, these and other questions regarding Hox gene function, and the role of Hox genes in generating the diverse arrangements of crustacean body plans, can be addressed in future experiments in this emerging model organism.

Acknowledgments

We thank Ron Parchem for the *Parhyale en1* clone, Jay Rehm for thoughtful discussions regarding early germband cell divisions in *Parhyale*, Rachel Thayer for the *abd-A in situ* in Fig. 2, and Rob White for providing the FP6.87 monoclonal antibody. The BAC sequencing work conducted by the U.S. Department of Energy Joint Genome Institute was supported by the Office of Science of the U.S. Department of Energy under Contract no. DE-AC02-05CH11231. This work was supported by NSF Grant IOS-1257379 to NHP.

References

Abzhanov, A., and Kaufman, T.C. (1999a). Novel regulation of the homeotic gene *Scr* associated with a crustacean leg-to-maxilliped appendage transformation. Development *126*, 1121–1128. Scopus

Abzhanov, A., and Kaufman, T.C. (1999b). Homeotic genes and the arthropod head: expression patterns of the *labial*, *proboscipedia*, and *Deformed* genes in crustaceans and insects. Proc. Natl. Acad. Sci. *96*, 10224–10229. CrossRef Scopus

Abzhanov, A., and Kaufman, T.C. (2000a). Embryonic expression patterns of the Hox genes of the crayfish Procambarus clarkii (Crustacea, Decapoda). Evol. Dev. *2*, 271–283. CrossRef Scopus

Abzhanov, A., and Kaufman, T.C. (2000b). Crustacean (malacostracan) Hox genes and the evolution of the arthropod trunk. Development *127*, 2239–2249. Scopus

Angelini, D.R., and Kaufman, T.C. (2005). Insect appendages and comparative ontogenetics. Dev. Biol. *286*, 57–77. Article PDF Scopus

Averof, M., and Akam, M. (1995). Hox genes and the diversification of insect and crustacean body plans. Nature *376*, 420–423. CrossRef Scopus

Averof, M., and Patel, N.H. (1997). Crustacean appendage evolution associated with changes in Hox gene expression. Nature *388*, 682–686. CrossRef Scopus

Bassett, A., and Liu, J.-L. (2014). CRISPR/Cas9 mediated genome engineering in *Drosophila*. Methods *69*, 128–136. Article PDF Scopus

Beachy, P.A., Krasnow, M.A., Gavis, E.R., and Hogness, D.S. (1988). An Ultrabithorax protein binds sequences near its own and the *Antennapedia P1* promoters. Cell *55*, 1069–1081. Article PDF Scopus

Blin, M., Rabet, N., Deutsch, J.S., and Mouchel-Vielh, E. (2003). Possible implication of Hox genes *Abdominal-B* and *abdominal-A* in the specification of genital and abdominal segments in cirripedes. Dev. Genes Evol. *213*, 90–96. Scopus

Brena, C., Liu, P.Z., Minelli, A., and Kaufman, T.C. (2005). *Abd-B* expression in Porcellio scaber Latreille, 1804 (Isopoda: Crustacea): conserved pattern versus novel roles in development and evolution. Evol. Dev. *7*, 42–50. CrossRef Scopus

Browne, W.E., Price, A.L., Gerberding, M., and Patel, N.H. (2005). Stages of embryonic development in the amphipod crustacean, *Parhyale hawaiensis*. Genesis *42*, 124–149. CrossRef Scopus

Chipman, A.D., Ferrier, D.E.K., Brena, C., Qu, J., Hughes, D.S.T., Schröder, R., Torres-Oliva, M., Znassi, N., Jiang, H., Almeida, F.C., et al. (2014). The first myriapod genome sequence reveals conservative arthropod gene content and genome organisation in the centipede *Strigamia maritima*. PLoS Biol *12*, e1002005. CrossRef

Condie, J.M., Mustard, J.A., and Brower, D.L. (1991). Generation of anti-Antennapedia monoclonal antibodies and Antennapedia protein expression in imaginal discs. Dros Inf Serv *70*, 52–54. Scopus

Copf, T., Rabet, N., Celniker, S.E., and Averof, M. (2003). Posterior patterning genes and the identification of a unique body region in the brine shrimp *Artemia franciscana*. Development *130*, 5915–5927. CrossRef Scopus

Damen, W.G. (2002). *fushi tarazu*: a Hox gene changes its role. BioEssays *24*, 992–995. CrossRef Scopus

Diao, F., and White, B.H. (2012). A novel approach for directing transgene expression in *Drosophila*: T2A-Gal4 in-frame fusion. Genetics *190*, 1139–1144. CrossRef Scopus

Falciani, F., Hausdorf, B., Schröder, R., Akam, M., Tautz, D., Denell, R., and Brown, S. (1996). Class 3 Hox genes in insects and the origin of *zen*. Proc. Natl. Acad. Sci. *93*, 8479–8484. CrossRef Scopus

Hayward, D.C., Patel, N.H., Rehm, E.J., Goodman, C.S., and Ball, E.E. (1995). Sequence and Expression of Grasshopper Antennapedia: Comparison to *Drosophila*. Dev. Biol. *172*, 452–465. Article PDF Scopus

Heffer, A., Shultz, J.W., and Pick, L. (2010). Surprising flexibility in a conserved Hox transcription factor over 550 million years of evolution. Proc. Natl. Acad. Sci. *107*, 18040–18045. CrossRef Scopus

Heffer, A., Xiang, J., and Pick, L. (2013). Variation and constraint in Hox gene evolution. Proc. Natl. Acad. Sci. *110*, 2211–2216. CrossRef Scopus

Hughes, C.L., and Kaufman, T.C. (2002a). Hox genes and the evolution of the arthropod body plan. Evol. Dev. *4*, 459–499. CrossRef Scopus

Hughes, C.L., and Kaufman, T.C. (2002b). Exploring the myriapod body plan: expression patterns of the ten Hox genes in a centipede. Development *129*, 1225–1238. Scopus

Hughes, C.L., Liu, P.Z., and Kaufman, T.C. (2004). Expression patterns of the rogue Hox genes *Hox3/zen* and *fushi tarazu* in the apterygote insect *Thermobia domestica*. Evol. Dev. *6*, 393–401. CrossRef Scopus

Janssen, R., and Budd, G.E. (2010). Gene expression suggests conserved aspects of Hox gene regulation in arthropods and provides additional support for monophyletic Myriapoda. EvoDevo *1*, 4. CrossRef

Janssen, R., and Damen, W.G. (2006). The ten Hox genes of the millipede *Glomeris marginata*. Dev. Genes Evol. *216*, 451–465. CrossRef Scopus

Liubicich, D.M., Serano, J.M., Pavlopoulos, A., Kontarakis, Z., Protas, M.E., Kwan, E., Chatterjee, S., Tran, K.D., Averof, M., and Patel, N.H. (2009). Knockdown of Parhyale *Ultrabithorax* recapitulates evolutionary changes in crustacean appendage morphology. Proc. Natl. Acad. Sci. *106*, 13892–13896. CrossRef Scopus

Mahfooz, N.S., Li, H., and Popadić, A. (2004). Differential expression patterns of the hox gene are associated with differential growth of insect hind legs. Proc. Natl. Acad. Sci. U. S. A. *101*, 4877–4882. CrossRef Scopus

Misof, B., Liu, S., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., Frandsen, P.B., Ware, J., Flouri, T., and Beutel, R.G. (2014). Phylogenomics resolves the timing and pattern of insect evolution. Science *346*, 763–767. CrossRef Scopus

Mouchel-Vielh, E., Blin, M., Rigolot, C., and Deutsch, J.S. (2002). Expression of a homologue of the *fushi tarazu* (ftz) gene in a cirripede crustacean. Evol. Dev. *4*, 76–85. CrossRef Scopus

Negre, B., and Ruiz, A. (2007). HOM-C evolution in *Drosophila*: is there a need for Hox gene clustering? Trends Genet. *23*, 55–59. Article PDF Scopus

Panfilio, K.A., Liu, P.Z., Akam, M., and Kaufman, T.C. (2006). *Oncopeltus fasciatus zen* is essential for serosal tissue function in katatrepsis. Dev. Biol. *292*, 226–243. Article PDF Scopus

Papillon, D., and Telford, M.J. (2007). Evolution of *Hox3* and *ftz* in arthropods: insights from the crustacean *Daphnia pulex*. Dev. Genes Evol. *217*, 315–322. CrossRef Scopus

Parchem, R.J., Poulin, F., Stuart, A.B., Amemiya, C.T., and Patel, N.H. (2010). BAC library for the amphipod crustacean, *Parhyale hawaiensis*. Genomics *95*, 261–267. Article PDF Scopus

Patel, N.H. (1994). Imaging neuronal subsets and other cell types in whole-mount Drosophila embryos and larvae using antibody probes. Methods Cell Biol. *44*, 445–487. Article PDF View Scopus

Patel, N.H., Kornberg, T.B., and Goodman, C.S. (1989). Expression of engrailed during segmentation in grasshopper and crayfish. Development *107*, 201–212. Scopus

Pavlopoulos, A., Kontarakis, Z., Liubicich, D.M., Serano, J.M., Akam, M., Patel, N.H., and Averof, M. (2009). Probing the evolution of appendage specialization by Hox gene misexpression in an emerging model crustacean. Proc. Natl. Acad. Sci. *106*, 13897–13902. CrossRef Scopus

Peterson, M.D., Rogers, B.T., Popadić, A., and Kaufman, T.C. (1999). The embryonic expression pattern of *labial*, posterior homeotic complex genes and the *teashirt* homologue in an apterygote insect. Dev. Genes Evol. *209*, 77–90. CrossRef Scopus

Pourquié, O. (2009). Hox genes (Elsevier Academic Press).

Price, A.L., and Patel, N.H. (2008). Investigating divergent mechanisms of mesoderm development in arthropods: the expression of *Ph-twist* and *Ph-mef2* in *Parhyale hawaiensis*. J. Exp. Zoolog. B Mol. Dev. Evol. *310*, 24–40. CrossRef Scopus

Regier, J.C., Shultz, J.W., Zwick, A., Hussey, A., Ball, B., Wetzer, R., Martin, J.W., and Cunningham, C.W. (2010). Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. Nature *463*, 1079–1083. CrossRef Scopus

Rehm, E.J., Hannibal, R.L., Chaw, R.C., Vargas-Vila, M.A., and Patel, N.H. (2009a). The crustacean *Parhyale hawaiensis*: a new model for arthropod development. Cold Spring Harb Protoc *1*, 1–10.

Rehm, E.J., Hannibal, R.L., Chaw, R.C., Vargas-Vila, M.A., and Patel, N.H. (2009b). In situ hybridization of labeled RNA probes to fixed *Parhyale hawaiensis* embryos. Cold Spring Harb Protoc *2009*, pdb.prot5130.

Rehm, E.J., Hannibal, R.L., Chaw, R.C., Vargas-Vila, M.A., and Patel, N.H. (2009c). Fixation and dissection of *Parhyale hawaiensis* embryos. Cold Spring Harb Protoc *2009*, pdb.prot5127.

Rehm, E.J., Hannibal, R.L., Chaw, R.C., Vargas-Vila, M.A., and Patel, N.H. (2009d). Antibody staining of *Parhyale hawaiensis* embryos. Cold Spring Harb Protoc *2009*, pdb.prot5129.

Rogers, B.T., Peterson, M.D., and Kaufman, T.C. (1997). Evolution of the insect body plan as revealed by the Sex combs reduced expression pattern. Development *124*, 149–157. Scopus

Saenko, S.V., Marialva, M.S., and Beldade, P. (2011). Involvement of the conserved Hox gene *Antennapedia* in the development and evolution of a novel trait. EvoDevo *2*, 1–10. Scopus

Shiga, Y., Yasumoto, R., Yamagata, H., and Hayashi, S. (2002). Evolving role of Antennapedia protein in arthropod limb patterning. Development *129*, 3555–3561. Scopus

Shiga, Y., Sagawa, K., Takai, R., Sakaguchi, H., Yamagata, H., and Hayashi, S. (2006). Transcriptional readthrough of Hox genes *Ubx* and *Antp* and their divergent post-transcriptional control during crustacean evolution. Evol. Dev. *8*, 407–414. CrossRef Scopus

Telford, M.J., and Thomas, R.H. (1998). Of mites and zen: expression studies in a chelicerate arthropod confirm *zen* is a divergent Hox gene. Dev. Genes Evol. *208*, 591–594. CrossRef Scopus

Vargas-Vila, M.A., Hannibal, R.L., Parchem, R.J., Liu, P.Z., and Patel, N.H. (2010). A prominent requirement for *single-minded* and the ventral midline in patterning the dorsoventral axis of the crustacean *Parhyale hawaiensis*. Development *137*, 3469–3476. CrossRef Scopus

Vick, P., and Blum, M. (2010). The isopod *Asellus aquaticus*: a novel arthropod model organism to study evolution of segment identity and patterning. Palaeodiversity *3*, 89–97. Scopus

Wolff, C., and Scholtz, G. (2006). Cell lineage analysis of the mandibular segment of the amphipod *Orchestia cavimana* reveals that the crustacean paragnaths are sternal outgrowths and not limbs. Front. Zool. *3*, 19. CrossRef

Yasukochi, Y., Ashakumary, L.A., Wu, C., Yoshido, A., Nohata, J., Mita, K., and Sahara, K. (2004). Organization of the Hox gene cluster of the silkworm, *Bombyx mori*: a split of the Hox cluster in a non-*Drosophila* insect. Dev. Genes Evol. *214*, 606–614. CrossRef Scopus

Appendix A. Supplementary material

```
labial        HEXAPEPTIDE                          HOMEODOMAIN
Ph-lab        PTYKWMQVKRNVPKP.APK          AGGTGRTNFSTKQLTELEKEFHFNKYLTRARRIEIASALQLNETQVKIWFQNRRMKQKKRLKEGL
Dp-lab        PTYKWMQVKRNVPKP.APR          ANNTGRTNFTTKQLTELEKEFHFNKYLTRARRIEIAAALQLNETQVKIWFQNRRMKQKKRMKEGL
Sc-lab        PTYKWMQVKRNNPRP.VPK          LNNTGRTNFTTKQLTELEKEFHFNKYLTRARRIEIATALALNETQVKIWFQNRRMKQKKRIKEGL
Cs-lab        PTYKWMQVKRSLPKP.APK          SNGSGRTNFTTKQLTELEKEFHYNKYLTRARRIEIATALQLNETQVKIWFQNRRMKQKKRMKEGL
Tc-lab        PTYKWMQVKRNVPKP.TPA          CLNTGRTNFTNKQLTELEKEFHFNKYLTRARRIEIASALQLNETQVKIWFQNRRMKQKKRMKEGL
Dm-lab        PTYKWMQLKRNVPKP.APK          TNNSGRTNFTNKQLTELEKEFHFNRYLTRARRIEIANTLQLNETQVKIWFQNRRMKQKKRVKEGL

proboscipedia
Ph-pb         VPEYPWMKEKKPVRK        ENGLPRRLRTAYTNTQLLELEKEFHFNKYLCRPRRIEIAASLDLTERQVKVWFQNRRMKHKRQVL
Dp-pb         SSDFAWMKDKKVGRK        -EGLPRRLRTAYTNTQLLELEKEFHFNKYLCRPRRIEIAASLDLTERQVKVWFQNRRMKHKRQAL
Cs-pb         ?????WMKEKKTTRK        DNGMPRRLRTAYTNTQLLELEKEFHFNKYLCRPRRIEIAASLDLTERQVKVWFQNRRMKHKRQTS
Of-pb         ???YPWMKEKKTARK        ENGLPRRLRTAYTNTQLLELEKEFHFNKYLCRPRRIEIAASLDLTERQVKVWFQNRRMKHKRQTL
Tc-pb         VPEYPWMKEKKTTRK        ENGLPRRLRTAYTNTQLLELEKEFHFNKYLCRPRRIEIAASLDLTERQVKVWFQNRRMKHKRQTL
Dm-pb         VPEYPWMKEKKTSRK        ENGLPRRLRTAYTNTQLLELEKEFHFNKYLCRPRRIEIAASLDLTERQVKVWFQNRRMKHKRQTL

Hox3
Ph-Hox3                              TTKRQRTQYSTFQLIELEKEFHYNSYLCRPRRAELAKTLSLSDRQVKIWFQNRRMKEKKT
Dp-Hox3                              PAKRARTAYTSAQLVELEKEFHFNRYLCRPRRIEMATLLNLTERQIKIWFQNRRMKFKKE
Sm-Hox3                              TNKRSRTAYTQSQLLELEKEFHFNRYLCRPRRLELASLLNLTERQIKIWFQNRRMKIKKI
Cs-Hox3                              PAKRARTAYTSAQLVELEKEFHFNRYLCRPRRIEMANLLNLSERQIKIWFQNRRMKYKKE
Fc-Hox3                              PTKRARTAYTSAQLVELEKEFHYNRYLCRPRRIEMASLLSLTERQIKIWFQNRRMKYKKE
Tc-zen                               AGKRARTAYTSAQLVELEREFHHGKYLSRPRRIQIAENLNLSERQIKIWFQNRRMKHKKE
Dm-zen                               KSKRSRTAFSSLQLIELEREFHLNKYLARTRRIEISQRLALTERQVKIWFQNRRMKLKKS

Deformed
Ph-Dfd        VIYPWMKKIHIAGVANG      PGCEPKRQRTAYTRHQILELEKEFHFNRYLTRRRRIEIAHSLCLSERQIKIWFQNRRMKWKKDNKLPNTKNVRRK
Dp-Dfd        VIYPWMKKIHVAG.ANG      PGMEPKRQRTAYTRHQILELEKEFHFNRYLTRRRRIEIAHSLCLSERQIKIWFQNRRMKWKKDNKLPNTKNVRRK
Cs-Dfd        ??YPWMKKVHVGSVAAN      PGVEPKRQRTAYTRHQILELEKEFHFNRYLTRRRRIEIAHALCLSERQIKIWFQNRRMKWKKDNKLPNTKNVKKK
Bm-Dfd        VIYPWMKKIHVAGASNG      PGMEPKRQRTGYTRHQILELEKEFHYNRYLTRRRRIEIAHTLVLSERQIKIWFQNRRMKWKKDNKLPNTKNVRRK
Tc-Dfd        QIYPWMRKVHVAGASNG      PGMEPKRQRTAYTRHQILELEKEFHYNRYLTRRRRIEIAHTLVLSERQIKIWFQNRRMKWKKDNKLPNTKNVRRK
Dm-Dfd        IIYPWMKKIHVAGVANG      PGMEPKRQRTAYTRHQILELEKEFHYNRYLTRRRRIEIAHTLVLSERQIKIWFQNRRMKWKKDNKLPNTKNVRKK

Sex combs reduced
Ph-Scr        PQIYPWMKRVHLGTNST      VNSNGETKRQRTSYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEHKMASMN
Dp-Scr        PQIYPWMKRVHLGQNA-      VNANGETKRQRTSYTRYQTLELEKEFHFNRYLTRRRRIEIAHSLCLSERQIKIWFQNRRMKWKKEHKIATMN
Cs-Scr        PQIYPWMRKVHVGQNG-      VNSMGETKRQRTSYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLSERQIKIWFQNRRMKWKKEHKMASPI
Of-Scr        PQIYPWMKRVHLGQ-ST      VNANGETKRQRTSYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEHKMASMN
Tc-Scr        PQIYPWMKRVHLGQ-ST      VNANGETKRQRTSYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEHKMASMN
Dm-Scr        PQIYPWMKRVHLGT-ST      VNANGETKRQRTSYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEHKMASMN

Antennapedia
Ph-Antp       SPLYPWMRSQF            ERKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKENKTK
Dp-Antp       SPLYPWMRSQF            ERKRGRQTYTRFQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKENKAK
Af-Antp       SNLYPWMKSQF            ERKRGRQTYTRFQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKENKSK
Cs-Antp       SPLYPWMRSQF            ERKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKENKAK
Tc-Antp       SPLYPWMRSQF            ERKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKENKTK
Dm-Antp       SPLYPWMRSQF            ERKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKENKTK

abdominal-A
Ph-abdA       PRYPWMSITD             SPNGCPRRRGRQTYTRFQTLELEKEFHFNHYLTRRRRIEIAHALCLTERQVKIWFQNRRMKLKKELRAVKEINEQVRRERE
Dp-abdA       PRYPWMSITG             GPNGCPRRRGRQTYTRYQTLELEKEFHFNHYLTRRRRIEIAHALCLTERQIKIWFQNRRMKLKKELRAVKEINEQARRERE
Af-abdA       PRYPWMSIT-             GPNGCPRRRGRQTYTRYQTLELEKEFHFNHYLTRRRRIEIAHALCLTERQIKIWFQNRRMKLKKELRAVKEINEQARKDREE
Sm-abdA       PRYPWMSIT-             GPNGCPRRRGRQTYTRFQTLELEKEFHFNHYLTRRRRIEIAHALCLTERQIKIWFQNRRMKLKKEMRAVKEINEQARREAAE
Tc-abdA       PRYPWMSITD             GPNGCPRRRGRQTYTRFQTLELEKEFHFNHYLTRRRRIEIAHALCLTERQIKIWFQNRRMKLKKELRAVKEINEQARREREE
Dm-abdA       PRYPWMTLTD             GPNGCPRRRGRQTYTRFQTLELEKEFHFNHYLTRRRRIEIAHALCLTERQIKIWFQNRRMKLKKELRAVKEINEQARRDREE

Abdominal-B
Ph-AbdB                        SWTGNVTVRKKRKPYSKFQTLELEKEFLYNAYVSKQKRWELARNLNLTERQVKIWFQNRRMKNKKNSQRQAAQ
Dp-AbdB                        EWTGQVTVRKKRKPYSKFQTLELEKEFLFNAYVSKQKRWELARNLNLTERQVKIWFQNRRMKNKKNSQRQQAG
Sc-AbdB                        EWTGNVSVRKKRKPYSKYQTLELEKEFLFNAYVSKQKRWELARNLNLTERQVKIWFQNRRMKNKKAQQRSQPP
Sm-AbdB                        EWTGNVTVRKKRKPYSKFQTLELEKEFLFNAYVSKQKRWELARNLNLTERQVKIWFQNRRMKNKKNSQRNQTD
Cs-AbdB                        EWTGTVTVRKKRKPYSKFQTLELEKEFLFNAYVSKQKRWELARNLNLTERQVKIWFQNRRMKSKKTSQRNAEN
Tc-AbdB                        EWTGQVTVRKKRKPYSKFQTLELEKEFLFNAYVSKQKRWELARNLNLTERQVKIWFQNRRMKNKKNSQRQAAQ
Dm-AbdB                        EWTGQVSVRKKRKPYSKFQTLELEKEFLFNAYVSKQKRWELARNLQLTERQVKIWFQNRRMKNKKNSQRQANQ
```

Supplemental Figure 2.1:
Alignment of the homeodomain (underlined) and adjacent sequences of Parhyale hawaiensis (Ph) Hox proteins to their orthologs from the following arthropod species: Anopheles gambiae (Ag), Artemia franciscana (Af), Bombyx mori (Bm), Cupiennius salei (Cs), Daphnia pulex (Dp), Drosophila melanogaster (Dm), Oncopeltus fasciatus (Of), Porcellio scaber (Ps), Procambarus clarkii (Pc), Sacculina carcini (Sc), Strigamia maritima (Sm), Thermobia domestica (Td), and Tribolium castaneum (Tc). In some instances, dashes have been inserted in order to make the sequences align; periods signify that nonconserved amino acids have been omitted. Amino acids that differ from their counterparts in the Parhyale ortholog are shown in red. Colored highlighting represents amino acids that are characteristic for, or indicative of, that particular Hox ortholog. Grey highlighling shows amino acids that are well conserved across all Hox genes

Chapter 3: Elucidation of the *Ph*-Hox Complex

One of the most remarkable properties of Hox genes emerges from their genomic arrangement into cluster(s) in a conserved order that is mirrored by their temporal and spatial expression along the AP axis. The phenomena (and conservation) of collinearity has many exciting evolutionary and mechanistic implications. Do *Parhyale* Hox genes reflect this canonical clustering and orders? Serano et al. (2016) (Chapter 2) identified the entire set of Hox genes in *Parhyale* and found the overall order of their embryonic expression to be temporally and spatially conserved. The study was able to link these genes only in part, however, through the assembly of incomplete genomic sequence data.In order to visually assess the clustering of Hox genes along a single chromosome, I performed a series of double fluorescent *in-situ* hybridizations to co-localize their physical proximity. The visualization of nascent nuclear signal over mature cytoplasmic transcript was enhanced by staining for *lab*, *Dfd*, and *Scr* at the beginning of their temporal expression (early germband stages S11-S15) and by using exclusively intronic probes for *Ubx, abd-A, Abd-B*. By using multiple combinations of double *in-situ* probe sets, we were able to spatially co-locate the clustering of the *lab-pb* and *Dfd-Scr-Antp-Ubx* assemblies (*lab* + *Dfd* probes) and connect these to *abd-A* (*Ubx* + *abd-A* probes) and *Abd-B* (*abd-A* + *Abd-B* probes).

Characterizing the Parhyale *Hox Complex*

The ability to quickly find (and conceptually localize) your gene or region of interest (GOI/ROI) from a variety of well-managed repositories of annotated sequences and genome assemblies is something only those studying something more "model-y" like a fruit fly, mouse, or human can fully take advantage of, while those studying a new organism or a model that is less developed must isolate, clone and sequence desired transcripts and gene regions "by hand" or save up for a major effort of whole genome sequence (GWS) and analysis. Indeed, Serano et al. (2016) (Chapter 2) identified the entire set of Hox genes in *Parhyale* with the use of degenerate PCR primers to target conserved homeodomain regions followed by gene walking to extend the isolated sequences, finding that the *Parhyale* genome includes all canonical Hox genes with the exception of *fushi tarazu (ftz)*.

Further analyses indicated that the *Parhyale* genome contains only a single copy of each Hox gene and the following variants: *Ubx* (Liubicich et al. 2009), *abd-A*, and *Abd-B* have multiple alternatively spliced variants (*isoform-1* and *isoform-2* for each gene); *Antp* has two classes of cDNA that exhibit distinct regions of polymorphism in the CDS, encoding two slightly different proteins (*Antp variant I* and *Antp variant II*); *Ubx* and *Antp* have read-through transcripts that result from when a subset of transcripts transcribed from the *Ubx* promoter(s) are spliced to exons from the *Antp* transcription unit. Transcriptional read-through between *Ubx* and *Antp* has been described for five other crustaceans, two myriapods, as well as in the arthropod outgroup onychophora (Brena et al. 2005; Janssen and Budd 2010; Shiga et al. 2006).

A rough map of the putative *Parhyale* Hox complex(es) was generated from the hybridization of Hox probes to *Parhyale* BAC clones, which offered some evidence of chromosomal clustering of the Hox genes, at least in part. The fragmented sequence

assembly of the Hox-specific BACs resulted in only partial linkage of the Hox cluster; *lab* and *pb*, as well as *Dfd, Scr, Antp*, and *Ubx*; both in the same linear, 5-'to-3' orientation as shown in vertebrate and other Hox clusters (Krumlauf 1994) (See Fig 2.1 in previous chapter for a schematic of genomic organization). Genomic linkage, if any, among these two linked sets and the altogether disconnected *Hox3, abd-A,* and *Abd-B* remained inconclusive. And because of the relatively large genome size of *Parhyale* (Parchem et al. 2010), it would take a lot more genomic data to resolve.

Collaborative Generation of Parhyale *Data & co-localization of Hox clustering by double fluorescent in-situ hybridization*

The *Parhyale* community is small but tightly knit and growing, with the frequent sharing of protocols, resources, and transgenic lines as they are developed. As the community and depth of studies continued to grow, it became apparent that *Parhyale* was establishing itself as a very useful model indeed and that a more collaborative repository for genomic data and resources was necessary, and there was a push for a large-scale community effort to sequence, assemble, annotate and fund the *Parhyale* genome. Preliminary assembly of the *Parhyale* genome still did not result in any further linkage of the Hox complex, however, and in fact raised concerns that the *Parhyale* Hox complex(es) may have undergone one or more duplication event(s), the outcome of which would significantly impact both the feasibility and interpretation of functional Hox gene experimentation.

As part of the collaborative effort to sequence and characterize the *Parhyale* genome and assemble a substantial manuscript with tools and resources, I was tasked to visualize the chromosomal (co)localization of the Hox genes by a series of double *in-situ* hybridizations using paired fluorescent probes in order to physically link the putative clustering of *Parhyale* Hox genes and to confirm (or refute) the presence of a single copy of each Hox gene. These data were published as part of the assembly and extensive analysis of "The genome of the crustacean *Parhyale hawaiensis*" (Kao et al. 2016 , Fig. 10), a collaborative effort among multiple *Parhyale* groups providing a compilation of genome and transcriptome assemblies, extensive genomic analyses, and genetic resources for *Parhyale*.

Kao et al. (2016) is a substantial manuscript addressing a multitude of different areas of *Parhyale* biology, which I re-publish here only in part. Selected sections or excerpts are arranged in two parts:

> **Part 1.** Are the data, text, figures, and methods that I directly contributed, pulled from the respectively labeled sections of the manuscript.
>
> **Part 2.** Following the pulled excerpts that are "my" part of the manuscript, I include the title page, abstract, eLife digest, and introduction in full to provide the narrative and overall context in which to place my specific contributions, and as a courtesy to the reader by providing convenient access to an excellent *Parhyale* primer with reference to a number of *Parhyale* resources and data repositories. The Introduction section provides a well-written description of *Parhyale* as a model as well as many useful links to *Parhyale* tools and resources. I also include a select few subsections from the Results and

Discussion, Materials and Methods and Conclusion sections that are relevant to my dissertation research. Other than minor editorial remarks, I did not directly contribute to the generation and writing of these sections, but was part of the overall effort and discussions leading to its manifestation, and feel warmly a part of and among the *Parhyale* community and core network of *Parhyale* researchers actively contributing to the development, optimization, back-and-forth communication and discussion of *Parhyale* methods, tools and techniques.

Linking the putative *Parhyale* Hox gene complex by double fluorescent in-situ hybridization

The following sections are "my" direct contributions excerpted from the respectively labeled section **eLife** 2016;5:e20062 DOI: 10.7554/eLife.20062

3I.2: Results and discussion (excerpted from **eLife** 2016;5:e20062)

Major signaling pathways and transcription factors in Parhyale

. . .

Parhyale contains a complement of 9 canonical Hox genes that exhibit both spatial and temporal colinearity in their expression along the anterior-posterior body axis (Serano et al., 2015). Chromosome walking experiments had shown that the Hox genes *labial* (*lab*) and *proboscipedia* (*pb*) are linked and that *Deformed* (*Dfd*), *Sex combs reduced* (*Scr*), *Antennapedia* (*Antp*) and *Ultrabithorax* (*Ubx*) are also contiguous in a cluster (Serano et al., 2015). Previous experiments in *D. melanogaster* had shown that the proximity of nascent transcripts in RNA fluorescent in situ hybridizations (FISH) coincide with the position of the corresponding genes in the genomic DNA (Kosman et al., 2004; Ronshaugen and Levine, 2004). Thus, we obtained additional information on Hox gene linkage by examining nascent Hox transcripts in cells where Hox genes are co-expressed. We first validated this methodology in *Parhyale* embryos by confirming with FISH, the known linkage of *Dfd* with *Scr* in the first maxillary segment where they are co-expressed (Figure 10A–A"). As a negative control, we detected no linkage between *engrailed1* (*en1*) and *Ubx* or *abd-A* transcripts (Figure 10B - B",C - C"). We then demonstrated the tightly coupled transcripts of *lab* with *Dfd* (co-expressed in the second antennal segment, Figure 10D - D"), *Ubx* and *abd-A* (co-expressed in the posterior thoracic segments, Figure 10E - E"), and *abd-A* with *Abd-B* (co-expressed in the anterior abdominal segments, (Figure 10F - F"). Collectively, all evidence supports the linkage of all analysed Hox genes into a single cluster as shown in

(Figure 10G - G"). The relative orientation and distance between certain Hox genes still needs to be worked out. So far, we have not been able to confirm that *Hox3* is also part of the cluster due to the difficulty in visualizing nascent transcripts for *Hox3* together with *pb* or *Dfd*. Despite these caveats, *Parhyale* provides an excellent arthropod model system to understand these still enigmatic phenomena of Hox gene clustering and spatio-temporal colinearity, and compare the underlying mechanisms to other well-studied vertebrate and invertebrate models (Kmita and Duboule, 2003).

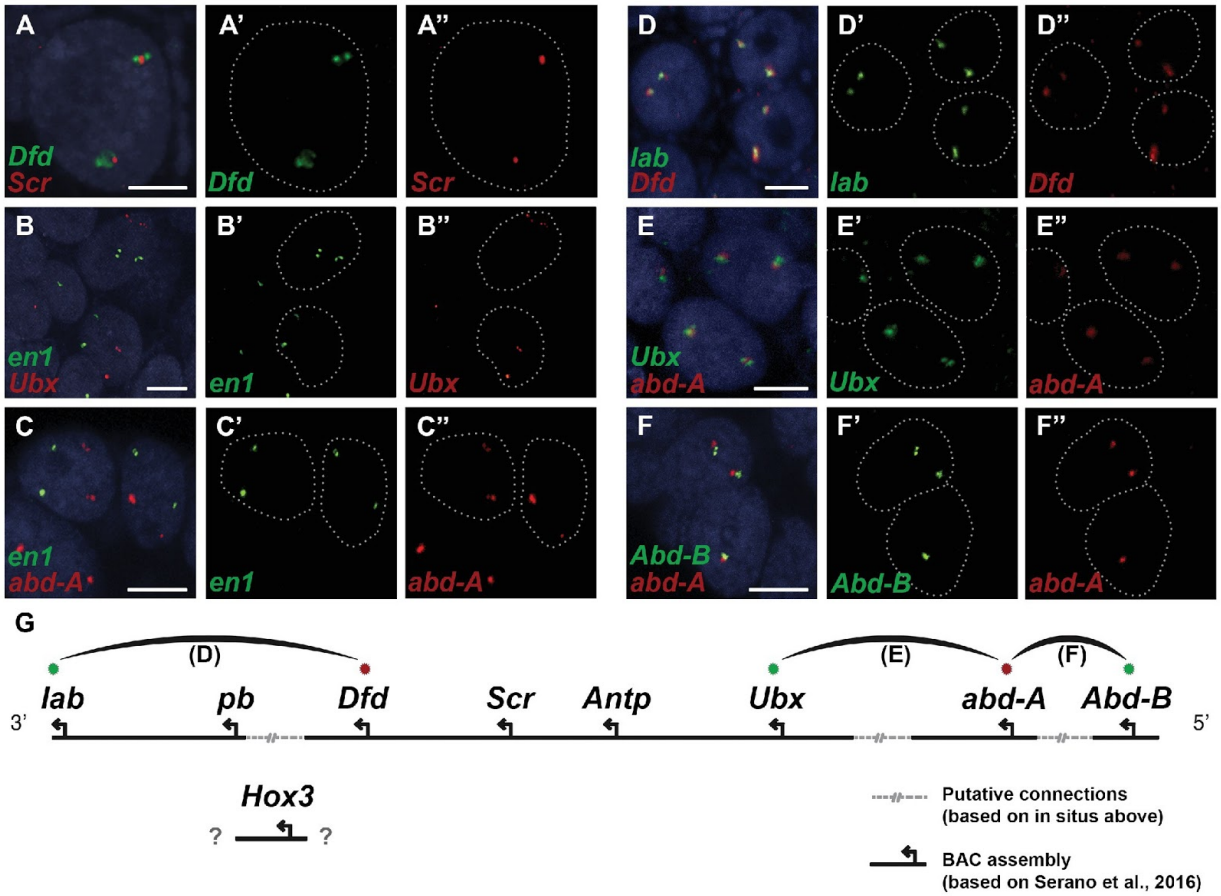


Figure 3.10. Evidence for an intact Hox cluster in *Parhyale*

(**A–F"**) Double fluorescent in situ hybridizations (FISH) for nascent transcripts of genes. (**A–A"**) Deformed (Dfd) and Sex combs reduced (Scr), (**B-B"**) engrailed 1 (en1) and Ultrabithorax (Ubx), (**C–C"**) en1 and abdominal-A (abd-A), (**D–D"**) labial (lab) and Dfd, (**E–E"**) Ubx and abd-A, and (**F–F"**) Abdominal-B (Abd-B) and abd-A. Cell nuclei are stained with DAPI (blue) in panels A–F and outlined with white dotted lines in panels A'–F' and A''. Co-localization of nascent transcript dots in A, D, E and F suggest the proximity of the corresponding Hox genes in the genomic DNA. As negative controls, the en1 nascent transcripts in B and C do not co-localize with those of Hox genes Ubx or abd-A. (**G**) Schematic representation of the predicted configuration of the Hox cluster in Parhyale. Previously identified genomic linkages are indicated with solid black lines, whereas linkages established by FISH are shown with dotted gray lines. The arcs connecting the green and red dots represent the linkages identified in D, E and F, respectively. The position of the Hox3 gene is still uncertain. Scale bars are 5 µm.

3I.3 Materials and methods (excerpted from **eLife** 2016;5:e20062)

Fluorescence in situ hybridization detection of Hox genes

Embryo fixation and in-situ hybridization was performed according to (Rehm et al., 2009). To enhance the nascent nuclear signal over mature cytoplasmic transcript, we used either early germband embryos (Stages 11 – 15) in which expression of *lab*, *Dfd*, and *Scr* are just starting (Serano et al., 2015), or probes that contain almost exclusively intron sequence (*Ubx, abd-A, Abd-B, and en1). Lab, Dfd*, and *Scr* probes are described in (Serano et al., 2015). Template for the intron-spanning probes were amplified using the following primers: en1-Intron1, AAGACACGACGAGCATCCTG and CTGTGTATGGCTACCCGTCC; Ubx-Intron1, GGTATGACAGCCGTCCAACA and AGAGTGCCAAGGATACCCGA; abd-A, CGATATACCCAGTCCGGTGC and TCATCAGCGAGGGCACAATT; Abd-B, GCTGCAGGATATCCACACGA and TGCAGTTGCCGCCATAGTAA.

A T7-adapter was appended to the 5' end of each reverse primer to enable direct transcription from PCR product. Probes were labeled with either Digoxigenin (DIG) or Dinitrophenol (DNP) conjugated UTPs, and visualized using sheep -DIG (Roche) and donkey -Sheep AlexaFluor 555 (Thermo Fischer Scientific), or Rabbit -DNP (Thermo Fischer Scientific) and Donkey -Rabbit AlexaFluor 488 (Jackson ImmunoResearch), respectively. Preparations were imaged on an LSM 780 scanning laser confocal (Zeiss), and processed using Volocity software (Perkin-Elmer).

The genome of the crustacean Parhyale hawaiensis, a model for animal development, regeneration, immunity and lignocellulose digestion

Damian Kao,[1†] Alvina G Lai,[1†] Evangelia Stamataki,[2†] Silvana Rosic,[3,4] Nikolaos Konstantinides,[5] Erin Jarvis,[6] Alessia Di Donfrancesco,[1] Natalia Pouchkina-Stancheva,[1] Marie Sémon,[5] Marco Grillo,[5] Heather Bruce,[6] Suyash Kumar,[2] Igor Siwanowicz,[2] Andy Le,[2] Andrew Lemire,[2] Michael B Eisen,[7] Cassandra Extavour,[8] William E Browne,[9] Carsten Wolff,[10] Michalis Averof,[5] Nipam H Patel,[6] Peter Sarkies,[3,4] Anastasios Pavlopoulos,[2] and Aziz Aboobaker1

[1]Department of Zoology, University of Oxford, Oxford, United Kingdom; [2]Howard Hughes Medical Institute, Janelia Farm Research Campus, Virginia, United States; [3]MRC Clinical Sciences Centre, Imperial College London, London, United Kingdom; [4]Clinical Sciences, Imperial College London, London, United Kingdom; 55Institut de Ge´ nomique Fonctionnelle de Lyon, Centre National de la Recherche Scientifique (CNRS) and E´ cole Normale Supe´ rieure de Lyon, Lyon, France; [6]Department of Molecular and Cell Biology, University of California, Berkeley, United States; [7]Molecular and Cell Biology, Howard Hughes Medical Institute, University of California, Berkeley, United States; [8]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, United States; [9]Department of Invertebrate Zoology, Smithsonian National Museum of Natural History, Washington, United States; [10]Vergleichende Zoologie, Institut fur Biologie, Humboldt-Universitat zu Berlin, Berlin, Germany

Selected sections from **eLife** 2016;5:e20062 DOI: 10.7554/eLife.20062

Abstract

The amphipod crustacean *Parhyale hawaiensis* is a blossoming model system for studies of developmental mechanisms and more recently regeneration. We have sequenced the genome allowing annotation of all key signaling pathways, transcription factors, and non-coding RNAs that will enhance ongoing functional studies. *Parhyale* is a member of the Malacostraca clade, which includes crustacean food crop species. We analysed the immunity related genes of *Parhyale* as an important comparative system for these species, where immunity related aquaculture problems have increased as farming has intensiĀied. We also find that *Parhyale* and other species within Multicrustacea contain the enzyme sets necessary to perform lignocellulose digestion ('wood eating'), suggesting this ability may predate the diversification of this lineage. Our data provide an essential resource for further development of *Parhyale* as an experimental model. The first malacostracan genome will underpin ongoing comparative work in food crop species and research investigating lignocellulose as an energy source.

eLife digest

The marine crustacean known as *Parhyale hawaiensis* is related to prawns, shrimps and crabs and is found at tropical coastlines around the world. This species has recently attracted scientific interest as a possible new model to study how animal embryos develop before birth and, because *Parhyale* can rapidly regrow lost limbs, how tissues and organs regenerate. Indeed, *Parhyale* has many characteristics that make it a good model organism, being small, fast-growing and easy to keep and care for in the laboratory.

Several research tools have already been developed to make it easier to study *Parhyale*. This includes the creation of a system for using the popular gene editing technology, CRISPR, in this animal. However, one critical resource that is available for most model organisms was missing; the complete sequence of all the genetic information of this crustacean, also known as its genome, was not available.

Kao, Lai, Stamataki et al. have now compiled the *Parhyale* genome – which is slightly larger than the human genome – and studied its genetics. Analysis revealed that *Parhyale* has genes that allow it to fully digest plant material. This is unusual because most animals that do this rely upon the help of bacteria. Kao, Lai, Stamataki et al. also identified genes that provide some of the first insights into the immune system of crustaceans, which protects these creatures from diseases.

Kao, Lai, Stamataki et al. have provided a resource and findings that could help to establish *Parhyale* as a popular model organism for studying several ideas in biology, including organ regeneration and embryonic development. Understanding how *Parhyale* digests plant matter, for example, could progress the biofuel industry towards efficient production of greener energy. Insights from its immune system could also be adapted to make farmed shrimp and prawns more resistant to infections, boosting seafood production.

3II.1 Introduction

Very few members of the Animal Kingdom hold the esteemed position of major model system for understanding living systems. Inventions in molecular and cellular biology increasingly facilitate the emergence of new experimental systems for developmental genetic studies. The morphological and ecological diversity of the phylum Arthropoda makes them an ideal group of animals for comparative studies encompassing embryology, adaptation of adult body plans and life history evolution (Akam, 2000; Budd and Telford, 2009; Peel et al., 2005; Scholtz and Wolff, 2013). While the most widely studied group are Hexapods, reĀlected by over a hundred sequencing projects available in the NCBI genome database, genomic data in the other three sub-phyla in Arthropoda are still relatively sparse.

Recent molecular and morphological studies have placed crustaceans along with hexapods into a pancrustacean clade (Figure 1A), revealing that crustaceans are paraphyletic (Mallatt

et al., 2004; Cook et al., 2005; Regier et al., 2005; Ertas et al., 2009; Richter, 2002). Previously, the only available fully sequenced crustacean genome was that of the water flea *Daphnia* which is a member of the Branchiopoda (Colbourne et al., 2011). A growing number of transcriptomes for larger phylogenetic analyses have led to differing hypotheses of the relationships of the major pancrustacean groups (Figure 1B) (Meusemann et al., 2010; Regier et al., 2010; Oakley et al., 2013; von Reumont et al., 2012). The genome of the amphipod crustacean *Parhyale hawaiensis* addresses the paucity of high quality non-hexapod genomes among the pancrustacean group, and will help to resolve relationships within this group as more genomes and complete proteomes become available (Rivarola-Duarte et al., 2014; Kenny et al., 2014). Crucially, genome sequence data is also necessary to further advance research in *Parhyale*, currently the most tractable crustacean model system. This is particularly true for the application of powerful functional genomic approaches, such as genome editing (Cong et al., 2013; Serano et al., 2015; Martin et al., 2015; Mali et al., 2013; Jinek et al., 2012; Gilles and Averof, 2014).

Parhyale is a member of the diverse Malacostraca clade with thousands of extant species including economically and nutritionally important groups such as shrimps, crabs, crayĀish and lobsters, as well as common garden animals like woodlice. They are found in all marine, fresh water, and higher humidity terrestrial environments. Apart from attracting research interest as an economically important food crop, this group of animals has been used to study developmental biology and the evolution of morphological diversity (for example with respect to Hox genes) (Martin et al., 2015; Averof and Patel, 1997; Liubicich et al., 2009; Pavlopoulos et al., 2009), stem cell biology (Konstantinides and Averof, 2014; Benton et al., 2014), innate immunity processes (Vazquez et al., 2009; Hauton, 2012) and recently the cellular mechanisms of regeneration (Konstantinides and Averof, 2014; Benton et al., 2014; Alwes et al., 2016). In addition, members of the Malacostraca, specifically both Amphipods and Isopods, are thought to be capable of 'wood eating' or lignocellulose digestion and to have microbiota-free digestive systems (King et al., 2010; Kern et al., 2013; Boyle and Mitchell, 1978; Zimmer et al., 2002).

The life history of *Parhyale* makes it a versatile model organism amenable to experimental manipulations (Figure 1C) (Wolff and Gerberding, 2015). Gravid females lay eggs every 2 weeks upon reaching sexual maturity and hundreds of eggs can be easily collected at all stages of embryogenesis. Embryogenesis takes about 10 days at 26°C and has been described in detail with an accurate staging system (Browne et al., 2005). Early embryos display an invariant cell lineage with each blastomere at the 8-cell stage contributing to a specific germ layer (Figure 1D) (Browne et al., 2005; Gerberding et al., 2002). Embryonic and post-embryonic stages are amenable to experimental manipulations and direct observation in vivo (Gerberding et al., 2002; Extavour, 2005; Rehm et al., 2009a, 2009b, 2009c, 2009d; Price et al., 2010; Alwes et al., 2011; Hannibal et al., 2012; Kontarakis and Pavlopoulos, 2014; Nast and Extavour, 2014; Chaw and Patel, 2012; Pavlopoulos and Averof, 2005). These can be combined with transgenic approaches (Pavlopoulos and Averof, 2005; Kontarakis et al., 2011; Kontarakis and Pavlopoulos, 2014; Pavlopoulos et al., 2009), RNA interference (RNAi) (Liubicich et al., 2009) and morpholino-mediated gene

knockdown (Ozhan-Kizil et al., 2009), and transgene-based lineage tracing (Konstantinides and Averof, 2014). Most recently the utility of the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) system for targeted genome editing has been elegantly demonstrated during the systematic study of *Parhyale* Hox genes (Martin et al., 2015; Serano et al., 2015). This arsenal of experimental tools (Table 1) has already established *Parhyale* as an attractive model system for biological research.

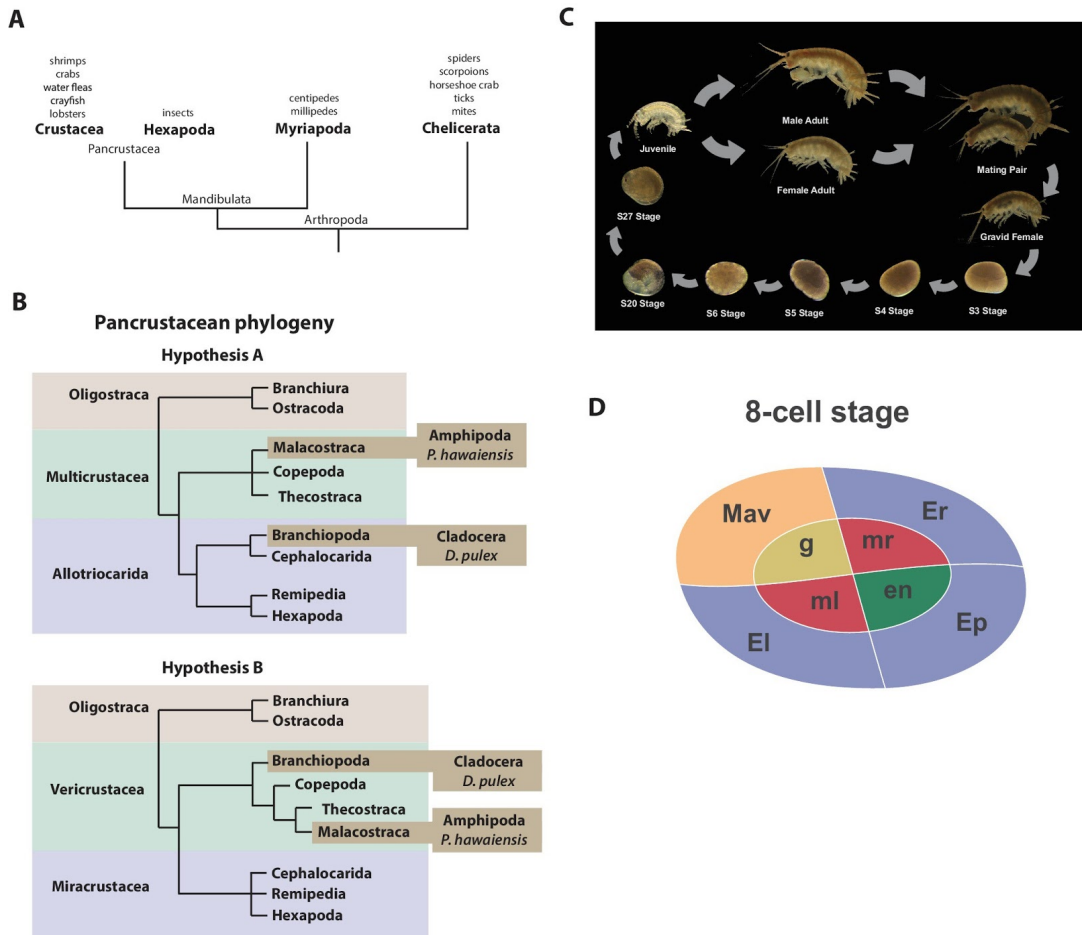


Figure 3.1: Introduction

(**A**) Phylogenetic relationship of Arthropods showing the Chelicerata as an outgroup to Mandibulata and the Pancrustacea clade which includes crustaceans and insects. Species listed for each clade have ongoing or complete genomes. Species include Crustacea: *Parhyale hawaiensis*, *D. pulex*; Hexapoda: *Drosophila melanogaster*, *Apis mellifera*, *Bombyx mori*, *Aedis aegypti*, *Tribolium castaneum*; Myriapoda: *Strigamia maritima*, *Trigoniulus corallines*; Chelicerata: *Ixodes scapularis*, *Tetranychus urticae*, *Mesobuthus martensii*, *Stegodyphus mimosarum*. (**B**) One of the unresolved issues concerns the placement of the Branchiopoda either together with the Cephalocarida, Remipedia and Hexapoda (Allotriocarida hypothesis A) or with the Copepoda, Thecostraca and Malacostraca (Vericrustacea hypothesis B). (**C**) Life cycle of *Parhyale* that takes about two months at 26C. *Parhyale* is a direct developer and a sexually dimorphic species. The fertilized egg undergoes stereotyped total cleavages and each blastomere becomes committed to a particular germ layer already at the 8-cell stage depicted in (**D**). The three macromeres Er, El, and Ep give rise to the anterior right, anterior left, and posterior ectoderm, respectively, while the fourth macromere Mav gives rise to the visceral mesoderm and anterior head somatic mesoderm. Among the 4 micromeres, the mr and ml micromeres give rise to the right and left somatic trunk mesoderm, en gives rise to the endoderm, and g gives rise to the germline.

Table 1: Experimental resources. Available experimental resources in *Parhyale* and corresponding references.

Experimental Resources	References
Embryological manipulations Cell microinjection, isolation, ablation	(Gerberding et al., 2002; Extavour, 2005; Price et al., 2010; Alwes et al., 2011; Hannibal et al., 2012; Rehm et al., 2009; Rehm et al., 2009; Kontarakis and Pavlopoulos, 2014; Nast and Extavour, 2014)
Gene expression studies In situ hybridization, antibody staining	(Rehm et al., 2009; Rehm et al., 2009)
Gene knock-down RNA interference, morpholinos	(Liubicich et al., 2009; Ozhan-Kizil et al., 2009)
Transgenesis Transposon-based, integrase-based	(Pavlopoulos and Averof, 2005; Kontarakis et al., 2011; Kontarakis and Pavlopoulos, 2014)
Gene trapping Exon/enhancer trapping, iTRAC (trap conversion)	(Kontarakis et al., 2011)
Gene misexpression Heat-inducible	(Pavlopoulos et al., 2009)
Gene knock-out CRISPR/Cas	(Martin et al., 2015)
Gene knock-in CRISPR/Cas homology-dependent or homology-independent	(Serano et al., 2015)
Live imaging Bright-field, confocal, light-sheet microscopy	(Alwes et al., 2011; Hannibal et al., 2012; Chaw and Patel, 2012; Alwes et al., 2016)

So far, work in *Parhyale* has been constrained by the lack of a reference genome and other standardized genome-wide resources. To address this limitation, we have sequenced, assembled and annotated the genome. At an estimated size of 3.6 Gb, this genome represents one of the largest animal genomes tackled to date. The large size has not been the only challenge of the *Parhyale* genome, that also exhibits some of the highest levels of sequence repetitiveness and polymorphism reported among published genomes. We provide information in our assembly regarding polymorphism to facilitate functional genomic approaches sensitive to levels of sequence similarity, particularly homology-dependent genome editing approaches. We analysed a number of key features of the genome as foundations for new areas of research in *Parhyale*, including innate immunity in crustaceans, lignocellulose digestion, non-coding RNA biology, and epigenetic control of the genome. Our data bring *Parhyale* to the forefront of developing model systems for a broad swathe of important bioscience research questions.

3II.2 Results and discussion (Selected Excerpts)

Genome assembly, annotation, and validation

The *Parhyale* genome contains 23 pairs (2n=46) of chromosomes (Figure 2) and with an estimated size of 3.6 Gb, it is currently the second largest reported arthropod genome after the locust genome (Parchem et al., 2010; Wang et al., 2014). Sequencing was performed on genomic DNA isolated from a single adult male taken from a line derived from a single female and expanded after two rounds of sib-mating.

. . .

One of the prime goals in sequencing the *Parhyale* genome was to achieve an assembly that could assist functional genetic and genomic approaches in this species. Different strategies have been employed to sequence highly heterozygous diploid genomes of non-model and wild-type samples (Kajitani et al., 2014).

Major signaling pathways and transcription factors in Parhyale

. . .

The ParaHox and *NK* gene clusters encode other *ANTP* class homeobox genes closely related to Hox genes (Brooke et al., 1998). In *Parhyale*, we found 2 caudal (*Cdx*) and 1 *Gsx* ParaHox genes. Compared to hexapods, we identified expansions in some NK-like genes, including 5 Bar homeobox genes (*BarH1/2*), 2 developing brain homeobox genes (*DBX*) and 6 muscle segment homeobox genes (*MSX/Drop*). Evidence from several bilaterian genomes suggests that *NK* genes are clustered together (Pollard and Holland, 2000; Jagla et al., 2001; Luke et al., 2003; Castro and Holland, 2003]. In the current assembly of the *Parhyale* genome, we identified an *NK2-3* gene and an *NK3* gene on the same scaffold (phaw_30.0004720) and the tandem duplication of an *NK2* gene on another scaffold (phaw_30.0004663). Within the *ANTP*class, we also observed 1 mesenchyme homeobox (*Meox*), 1 motor neuron homeobox (*MNX/Exex*) and 3 even-skipped homeobox (*Evx*) genes.

Non-coding RNAs and associated proteins in the Parhyale genome

Non-coding RNAs are a central, but still a relatively poorly understood part of eukaryotic genomes. In animal genomes, different classes of small RNAs are key for genome surveillance, host defense against viruses and parasitic elements in the genome, and regulation of gene expression through transcriptional, post-transcriptional and epigenetic control mechanisms (Castel and Martienssen, 2013; Aravin et al., 2001; Caplen et al., 2001; Brennecke et al., 2007; Gu et al., 2009; Lee et al., 2012; He and Hannon, 2004; Thomson et al., 2006; Filipowicz et al., 2008). The nature of these non-coding RNAs, as well as the proteins involved in their biogenesis and function, can vary between animals. For example, some nematodes have Piwi-interacting short RNAs (piRNAs), while others have replaced these by alternate small RNA based mechanisms to compensate for their loss (Sarkies et al., 2015).

. . .

Conserved linkages have also been observed between miRNAs and Hox genes in Bilateria (Enright et al., 2003a; Tanzer et al., 2005; Lemons and McGinnis, 2006; Stark et al., 2008; Shippy et al., 2008). For example, the phylogenetically conserved *mir-10* is present within both vertebrate and invertebrate Hox clusters between Hoxb4/*Dfd* and *Hoxb5/Scr* (Enright et al., 2003b). In the *Parhyale* genome and Hox BAC sequences, we found that *mir-10* is also located between *Dfd* and *Src* on BAC clone PA179-K23 and scaffold phaw_30.0001203 (Figure 14—figure supplement 2). However, we could not detect *mir-iab-4* near the *Ubx* and *AbdA* genes in *Parhyale*, the location where it is found in other arthropods/insects (Cumberledge et al., 1990).

Parhyale genome editing using homology-independent approaches

Parhyale has already emerged as a powerful model for developmental genetic research where the expression and function of genes can be studied in the context of stereotyped cellular processes and with a single-cell resolution. Several experimental approaches and standardized resources have been established to study coding and non-coding sequences (Table 1). These functional studies will be enhanced by the availability of the assembled and annotated genome presented here.

. . .

These results, together with the other recent applications of the CRISPR/Cas system to study Hox genes in *Parhyale* (Martin et al., 2015; Serano et al., 2015), demonstrate that the ability to manipulate the fertilized eggs together with the slow tempo of early cleavages can result in very high targeting frequencies and low levels of mosaicism for both knock-out and knock-in approaches. Considering the usefulness of the genome-wide resources described in this report, we anticipate that the *Parhyale* embryo will prove an extremely powerful system for fast and reliable G0 screens of gene expression and function..

3II.3 Conclusion (Selected Excerpts)

. . .

Parhyale has emerged since the early 2000's as an attractive animal model for developmental genetic and molecular cell biology research. It fulfills several desirable biological and technical requirements as an experimental model, including a relatively short life-cycle, year-round breeding under standardized laboratory conditions, availability of thousands of eggs for experimentation on a daily basis, and amenability to various embryological, cellular, molecular genetic and genomic approaches. In addition, *Parhyale* has stereotyped cell lineages and cell behaviors, a direct mode of development, a remarkable appendage diversity and the capacity to regenerate limbs post-embryonically. These qualities can be utilized to address fundamental long-standing questions in developmental biology, like cell fate specification, nervous system development, organ morphogenesis and regeneration (Stamataki and Pavlopoulos, 2016). Research on these topics will benefit enormously from the standardized genome-wide resources reported here. Forward and reverse genetic analyses using both unbiased screens and candidate gene approaches have already been devised successfully in *Parhyale* (Table 1). The availability of coding and non-coding sequences for all identified signaling pathway components, transcription factors and various classes of non-coding RNAs will dramatically accelerate the study of the expression and function of genes implicated in the aforementioned processes.

. . .

3II.4 Materials and methods (Selected Excerpts)

Raw genomic reads are deposited at NCBI with the project accession: PRJNA306836. All supplemental data including IPython notebook can be downloaded from this figshare link: https://āigshare.com/articles/supplemental_data_for_Parhyale_hawaniensis_genome/3498104 Alternatively, the IPython notebooks and associated scripts can also be viewed at the following github repository: https://github.com/damiankao/phaw_genome

. . .

Additional information (Author contributions)

DK, Devised assembly strategy, Assembled and analyzed the sequencing data, Annotated the genome, transcriptome and proteome, Performed orthology group analysis, Annotated small RNAs, Drafting and revising the article.

AGL, Analysed the genome including major signaling pathways, polymorphisms, immunity, lignocellulose digestion, epigenetic pathways, small RNA pathways and small RNAs, Cloning of DSCAM variants, Experimental confirmation of polymorphisms, Drafting and revising the article.

ES, Prepared the genomic libraries, performed CRISPR knock-out, performed CRISPR knock-in, Drafting and revising the article.

SR, Contributed bisulfite sequencing data and analysis of genome wide methylation.

NK, Contributed transcriptome data and transcriptome assembly.

EJ, Performed in situ hybridization detection of Hox genes, and interpreted data.

ADD, Contributed to confirmation of polymorphism and cloning of Ph-DSCAM variants.

NP-S, Contributed to confirmation of polymorphism and cloning of Ph-DSCAM variants.

MS, Contributed transcriptome data and transcriptome assembly.

MG, Contributed transcriptome data and transcriptome assembly.

HB, Contributed transcriptome data.

SK, Performed CRISPR knock-out.

IS, Performed Parhyale cuticle staining.

ALe, Performed CRISPR knock-in.

ALemi, Was consulted about sequencing strategy and helped with bioinformatics.

MBE, Contributed to RNAseq data production.

CE, Contributed to project planning.

WEB, Established the Chicago-F line.

CW, Performed karyotyping.

MA, Contributed transcriptome data and transcriptome assembly.

NHP, Performed in situ hybridization detection of Hox genes, and interpreted data. Contributed transcriptome data, Established the Chicago-F line.

PS, Conceived and designed bisulfite sequencing experiments, contributed bisultfite sequencing data and analysis of genome wide methylation.

AP, Conceived, designed and managed the project, Contributed to data acquisition and analysis, Drafting and revising the article.

AA, Devised assembly strategy, Contributed to data analysis, Conceived, designed and managed the project, Drafting and revising the article.

References (Part I)

Kmita M, Duboule D. Organizing axes in time and space; 25 years of colinear tinkering. Science. 2003;301:331–333. doi: 10.1126/science.1085753. [PubMed] [Cross Ref]

Kosman D, Mizutani CM, Lemons D, Cox WG, McGinnis W, Bier E. Multiplex detection of RNA expression in Drosophila embryos. Science. 2004;305:846 doi: 10.1126/science.1099247. [PubMed] [Cross Ref]

Rehm EJ, Hannibal RL, Chaw RC, Vargas-Vila MA, Patel NH. In Situ hybridization of labeled RNA probes to fixed Parhyale hawaiensis Embryos. Cold Spring Harbor Protocols. 2009b;2009:pdb.prot5130 doi: 10.1101/pdb.prot5130. [PubMed] [Cross Ref]

Ronshaugen M, Levine M. Visualization of trans-homolog enhancer-promoter interactions at the Abd-B Hox locus in the Drosophila embryo. Developmental Cell. 2004;7:925–932. doi: 10.1016/j.devcel.2004.11.001. [PubMed] [Cross Ref]

Serano JM, Martin A, Liubicich DM, Jarvis E, Bruce HS, La K, Browne WE, Grimwood J, Patel NH. Comprehensive analysis of Hox gene expression in the amphipod crustacean Parhyale hawaiensis. Developmental Biology. 2016;409:297–309. doi: 10.1016/j.ydbio.2015.10.029. [PubMed] [Cross Ref]

References (Selected Sections Part II)

Akam M. Arthropods: developmental diversity within a (super) phylum. PNAS. 2000;97:1–4. doi: 10.1073/pnas.97.9.4438. [PMC free article] [PubMed] [Cross Ref]

Alwes F, Enjolras C, Averof M. Live imaging reveals the progenitors and cell dynamics of limb regeneration. eLife. 2016;5:e19766 doi: 10.7554/eLife.19766. [PMC free article][PubMed] [Cross Ref]

Alwes F, Hinchen B, Extavour CG. Patterns of cell lineage, movement, and migration from germ layer speciĀicaɨon to gastrulation in the amphipod crustacean Parhyale hawaiensis. Developmental Biology. 2011;359:110–123. doi: 10.1016/j.ydbio.2011.07.029. [PubMed] [Cross Ref]

Averof M, Patel NH. Crustacean appendage evolution associated with changes in Hox gene expression. Nature. 1997;388:682–686. doi: 10.1038/41786. [PubMed] [Cross Ref]

Benton JL, Kery R, Li J, Noonin C, Söderhäll I, Beltz BS. Cells from the immune system generate adult-born neurons in crayfish. Developmental Cell. 2014;30:322–333. doi: 10.1016/j.devcel.2014.06.016. [PubMed] [Cross Ref]

Boyle PJ, Mitchell R. Absence of microorganisms in crustacean digestive tracts. Science. 1978;200:1157–1159. doi: 10.1126/science.200.4346.1157. [PubMed] [Cross Ref]

Browne WE, Price AL, Gerberding M, Patel NH. Stages of embryonic development in the amphipod crustacean, Parhyale hawaiensis. Genesis. 2005;42:124–149. doi: 10.1002/gene.20145. [PubMed] [Cross Ref]

Brooke NM, Garcia-Fernàndez J, Holland PW. The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. Nature. 1998;392:920–922. doi: 10.1038/31933. [PubMed] [Cross Ref]

Budd GE, Telford MJ. The origin and evolution of arthropods. Nature. 2009;457:812–817. doi: 10.1038/nature07890. [PubMed] [Cross Ref]

Castro LF, Holland PW. Chromosomal mapping of ANTP class homeobox genes in amphioxus: piecing together ancestral genomes. Evolution and Development. 2003;5:1–7. doi: 10.1046/j.1525-142X.2003.03052.x. [PubMed] [Cross Ref]

Chaw RC, Patel NH. Independent migration of cell populations in the early gastrulation of the amphipod crustacean Parhyale hawaiensis. Developmental Biology. 2012;371:94–109. doi: 10.1016/j.ydbio.2012.08.012. [PubMed] [Cross Ref]

Colbourne JK, Pfrender ME, Gilbert D, Thomas WK, Tucker A, Oakley TH, Tokishita S, Aerts A, Arnold GJ, Basu MK, Bauer DJ, Cáceres CE, Carmel L, Casola C, Choi JH, Detter JC, Dong Q, Dusheyko S, Eads BD, Fröhlich T, Geiler-Samerotte KA, Gerlach D, Hatcher P, Jogdeo S, Krijgsveld J, Kriventseva EV, Kültz D, Laforsch C, Lindquist E, Lopez J, Manak JR, Muller J, Pangilinan J, Patwardhan RP, Pitluck S, Pritham EJ, Rechtsteiner A, Rho M, Rogozin IB, Sakarya O, Salamov A, Schaack S, Shapiro H, Shiga Y, Skalitzky C, Smith Z, Souvorov A, Sung W, Tang Z, Tsuchiya D, Tu H, Vos H, Wang M, Wolf YI, Yamagata H, Yamada T, Ye Y, Shaw JR, Andrews J, Crease TJ, Tang H, Lucas SM, Robertson HM, Bork P, Koonin EV, Zdobnov EM, Grigoriev IV, Lynch M, Boore JL. The ecoresponsive genome of Daphnia pulex. Science. 2011;331:555–561. doi: 10.1126/science.1197761. [PMC free article] [PubMed] [Cross Ref]

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339:819–823. doi: 10.1126/science.1231143. [PMC free article] [PubMed][Cross Ref]

Cook CE, Yue Q, Akam M. Mitochondrial genomes suggest that hexapods and crustaceans are mutually paraphyletic. Proceedings of the Royal Society B: Biological Sciences. 2005;272:1295–1304. doi: 10.1098/rspb.2004.3042. [PMC free article][PubMed] [Cross Ref]

Ertas B, von Reumont BM, Wägele JW, Misof B, Burmester T. Hemocyanin suggests a close relationship of Remipedia and Hexapoda. Molecular Biology and Evolution. 2009;26:2711–2718. doi: 10.1093/molbev/msp186. [PubMed] [Cross Ref]

Extavour CG. The fate of isolated blastomeres with respect to germ cell formation in the amphipod crustacean Parhyale hawaiensis. Developmental Biology. 2005;277:387–402. doi: 10.1016/j.ydbio.2004.09.030. [PubMed] [Cross Ref]

Gerberding M, Browne WE, Patel NH. Cell lineage analysis of the amphipod crustacean Parhyale hawaiensis reveals an early restriction of cell fates. Development. 2002;129:5789–5801. doi: 10.1242/dev.00155. [PubMed] [Cross Ref]

Gilles AF, Averof M. Functional genetics for all: engineered nucleases, CRISPR and the gene editing revolution. EvoDevo. 2014;5:43–13. doi: 10.1186/2041-9139-5-43.[PMC free article] [PubMed] [Cross Ref]

Hannibal RL, Price AL, Patel NH. The functional relationship between ectodermal and mesodermal segmentation in the crustacean, Parhyale hawaiensis. Developmental Biology. 2012;361:427–438. doi: 10.1016/j.ydbio.2011.09.033. [PubMed] [Cross Ref]

Hauton C. The scope of the crustacean immune system for disease control. Journal of Invertebrate Pathology. 2012;110:251–260. doi: 10.1016/j.jip.2012.03.005. [PubMed][Cross Ref]

Jagla K, Bellard M, Frasch M. A cluster of Drosophila homeobox genes involved in mesoderm differentiation programs. BioEssays 2001;23:125–133. doi: 10.1002/1521-1878(200102)23:2<125::AID-BIES1019>3.0.CO;2-C. [PubMed] [Cross Ref]

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337:816–821. doi: 10.1126/science.1225829. [PubMed] [Cross Ref]

Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, Okuno M, Yabana M, Harada M, Nagayasu E, Maruyama H, Kohara Y, Fujiyama A, Hayashi T, Itoh T. Efficient de novo assembly of highly heterozygous genomes from

whole-genome shotgun short reads. Genome Research. 2014;24:1384–1395. doi: 10.1101/gr.170720.113.[PMC free article] [PubMed] [Cross Ref]

Kenny NJ, Sin YW, Shen X, Zhe Q, Wang W, Chan TF, Tobe SS, Shimeld SM, Chu KH, Hui JH. Genomic sequence and experimental tractability of a new decapod shrimp model, Neocaridina denticulata. Marine Drugs. 2014;12:1419–1437. doi: 10.3390/md12031419. [PMC free article] [PubMed] [Cross Ref]

Kern M, McGeehan JE, Streeter SD, Martin RN, Besser K, Elias L, Eborall W, Malyon GP, Payne CM, Himmel ME, Schnorr K, Beckham GT, Cragg SM, Bruce NC, McQueen-Mason SJ. Structural characterization of a unique marine animal family 7 cellobiohydrolase suggests a mechanism of cellulase salt tolerance. PNAS. 2013;110:10189–10194. doi: 10.1073/pnas.1301502110. [PMC free article] [PubMed][Cross Ref]

King AJ, Cragg SM, Li Y, Dymond J, Guille MJ, Bowles DJ, Bruce NC, Graham IA, McQueen-Mason SJ. Molecular insight into lignocellulose digestion by a marine isopod in the absence of gut microbes. PNAS. 2010;107:5345–5350. doi: 10.1073/pnas.0914228107. [PMC free article] [PubMed] [Cross Ref]

Konstantinides N, Averof M. A common cellular basis for muscle regeneration in arthropods and vertebrates. Science. 2014;343:788–791. doi: 10.1126/science.1243529.[PubMed] [Cross Ref]

Kontarakis Z, Pavlopoulos A, Kiupakis A, Konstantinides N, Douris V, Averof M. A versatile strategy for gene trapping and trap conversion in emerging model organisms. Development. 2011;138:2625–2630. doi: 10.1242/dev.066324. [PubMed] [Cross Ref]

Kontarakis Z, Pavlopoulos A. Transgenesis in non-model organisms: The case of Parhyale. Methods in Molecular Biology. 2014;1196:145–181. doi: 10.1007/978-1-4939-1242-1_10. [PubMed] [Cross Ref]

Liubicich DM, Serano JM, Pavlopoulos A, Kontarakis Z, Protas ME, Kwan E, Chatterjee S, Tran KD, Averof M, Patel NH. Knockdown of Parhyale Ultrabithorax recapitulates evolutionary changes in crustacean appendage morphology. PNAS. 2009;106:13892–13896. doi: 10.1073/pnas.0903105106. [PMC free article] [PubMed] [Cross Ref]

Luke GN, Castro LF, McLay K, Bird C, Coulson A, Holland PW. Dispersal of NK homeobox gene clusters in amphioxus and humans. PNAS. 2003;100:1–4. doi: 10.1073/pnas.0836141100. [PMC free article] [PubMed] [Cross Ref]

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas9. Science. 2013;339:823–826. doi: 10.1126/science.1232033. [PMC free article] [PubMed] [Cross Ref]

Mallatt JM, Garey JR, Shultz JW. Ecdysozoan phylogeny and Bayesian inference: first use of nearly complete 28S and 18S rRNA gene sequences to classify the arthropods and their kin. Molecular Phylogenetics and Evolution. 2004;31:178–191. doi: 10.1016/j.ympev.2003.07.013. [PubMed] [Cross Ref]

Martin A, Serano JM, Jarvis E, Bruce HS, Wang J, Ray S, Barker CA, O'Connell LC, Patel NH. CRISPR/Cas9 mutagenesis reveals versatile roles of Hox genes in crustacean limb specification and evolution. Current Biology. 2016;26:14–26. doi: 10.1016/j.cub.2015.11.021. [PubMed] [Cross Ref]

Meusemann K, von Reumont BM, Simon S, Roeding F, Strauss S, Kück P, Ebersberger I, Walzl M, Pass G, Breuers S, Achter V, von Haeseler A, Burmester T, Hadrys H, Wägele JW, Misof B. A phylogenomic approach to resolve the arthropod tree of life. Molecular Biology and Evolution. 2010;27:2451–2464. doi: 10.1093/molbev/msq130.[PubMed] [Cross Ref]

Nast AR, Extavour CG. Ablation of a single cell from eight-cell embryos of the amphipod crustacean Parhyale hawaiensis. Journal of Visualized Experiments. 2014;16:e51073 doi: 10.3791/51073. [PMC free article] [PubMed] [Cross Ref]

Oakley TH, Wolfe JM, Lindgren AR, Zaharoff AK. Phylotranscriptomics to bring the understudied into the fold: monophyletic ostracoda, fossil placement, and pancrustacean phylogeny. Molecular Biology and Evolution. 2013;30:215–233. doi: 10.1093/molbev/mss216. [PubMed] [Cross Ref]

Ozhan-Kizil G, Havemann J, Gerberding M. Germ cells in the crustacean Parhyale hawaiensis depend on Vasa protein for their maintenance but not for their formation. Developmental Biology. 2009;327:230–239. doi: 10.1016/j.ydbio.2008.10.028.[PubMed] [Cross Ref]

Parchem RJ, Poulin F, Stuart AB, Amemiya CT, Patel NH. BAC library for the amphipod crustacean, Parhyale hawaiensis. Genomics. 2010;95:261–267. doi: 10.1016/j.ygeno.2010.03.005. [PMC free article] [PubMed] [Cross Ref]

Pavlopoulos A, Averof M. Establishing genetic transformation for comparative developmental studies in the crustacean Parhyale hawaiensis. PNAS. 2005;102:7888–7893. doi: 10.1073/pnas.0501101102. [PMC free article] [PubMed] [Cross Ref]

Pavlopoulos A, Kontarakis Z, Liubicich DM, Serano JM, Akam M, Patel NH, Averof M. Probing the evolution of appendage specialization by Hox gene misexpression in an emerging model crustacean. PNAS. 2009;106:13897–13902. doi: 10.1073/pnas.0902804106. [PMC free article] [PubMed] [Cross Ref]

Peel AD, Chipman AD, Akam M. Arthropod segmentation: beyond the Drosophila paradigm. Nature Reviews. Genetics. 2005;6:905–916. doi: 10.1038/nrg1724. [PubMed][Cross Ref]

Pollard SL, Holland PW. Evidence for 14 homeobox gene clusters in human genome ancestry. Current Biology. 2000;10:1059–1062. doi: 10.1016/S0960-9822(00)00676-X.[PubMed] [Cross Ref]

Price AL, Modrell MS, Hannibal RL, Patel NH. Mesoderm and ectoderm lineages in the crustacean Parhyale hawaiensis display intra-germ layer compensation. Developmental Biology. 2010;341:256–266. doi: 10.1016/j.ydbio.2009.12.006. [PubMed] [Cross Ref]

Regier JC, Shultz JW, Kambic RE. Pancrustacean phylogeny: hexapods are terrestrial crustaceans and maxillopods are not monophyletic. Proceedings of the Royal Society B: Biological Sciences. 2005;272:395–401. doi: 10.1098/rspb.2004.2917.[PMC free article] [PubMed] [Cross Ref]

Regier JC, Shultz JW, Zwick A, Hussey A, Ball B, Wetzer R, Martin JW, Cunningham CW. Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. Nature. 2010;463:1079–1083. doi: 10.1038/nature08742. [PubMed][Cross Ref]

Rehm EJ, Hannibal RL, Chaw RC, Vargas-Vila MA, Patel NH. Antibody Staining of Parhyale hawaiensis Embryos. Cold Spring Harbor Protocols. 2009a;2009:pdb.prot5129 doi: 10.1101/pdb.prot5129. [PubMed] [Cross Ref]

Rehm EJ, Hannibal RL, Chaw RC, Vargas-Vila MA, Patel NH. In Situ hybridization of labeled RNA probes to fixed Parhyale hawaiensis Embryos. Cold Spring Harbor Protocols. 2009b;2009:pdb.prot5130 doi: 10.1101/pdb.prot5130. [PubMed] [Cross Ref]

Rehm EJ, Hannibal RL, Chaw RC, Vargas-Vila MA, Patel NH. Injection of Parhyale hawaiensis blastomeres with Ālluoꞁescently labeled tracers. Cold Spring Harbor Protocols. 2009c;2009:pdb.prot5128 doi: 10.1101/pdb.prot5128. [PubMed] [Cross Ref]

Rehm EJ, Hannibal RL, Chaw RC, Vargas-Vila MA, Patel NH. Fixation and Dissection of Parhyale hawaiensis Embryos. Cold Spring Harbor Protocols. 2009d;2009:pdb.prot5127 doi: 10.1101/pdb.prot5127. [PubMed] [Cross Ref]

Rehman A, Taishi P, Fang J, Majde JA, Krueger JM. The cloning of a rat peptidoglycan recognition protein (PGRP) and its induction in brain by sleep deprivation. Cytokine. 2001;13:8–17. doi: 10.1006/cyto.2000.0800. [PubMed] [Cross Ref]

Richter S. The Tetraconata concept: hexapod-crustacean relationships and the phylogeny of Crustacea. Organisms Diversity & Evolution. 2002;2:217–237. doi: 10.1078/1439-6092-00048. [Cross Ref]

Rivarola-Duarte L, Otto C, Jühling F, Schreiber S, Bedulina D, Jakob L, Gurkov A, Axenov-Gribanov D, Sahyoun AH, Lucassen M, Hackermüller J, Hoffmann S, Sartoris F, Pörtner H-O, Timofeyev M, Luckenbach T, Stadler PF. A first glimpse at the genome of the baikalian amphipod Eulimnogammarus verrucosus. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution. 2014;322:177–189. doi: 10.1002/jez.b.22560. [PubMed] [Cross Ref]

Scholtz G, Wolff C. Arthropod Biology and Evolution. 2013. Arthropod embryology: cleavage and germ band development; pp. 63–89.

Serano JM, Martin A, Liubicich DM, Jarvis E, Bruce HS, La K, Browne WE, Grimwood J, Patel NH. Comprehensive analysis of Hox gene expression in the amphipod crustacean Parhyale hawaiensis. Developmental Biology. 2016;409:297–309. doi: 10.1016/j.ydbio.2015.10.029. [PubMed] [Cross Ref]

Stamataki E, Pavlopoulos A. Non-insect crustacean models in developmental genetics including an encomium to Parhyale hawaiensis. Current Opinion in Genetics & Development. 2016;39:149–156. doi: 10.1016/j.gde.2016.07.004. [PubMed] [Cross Ref]

Vazquez L, Alpuche J, Maldonado G, Agundis C, Pereyra-Morales A, Zenteno E. Review: Immunity mechanisms in crustaceans. Innate Immunity. 2009;15:179–188. doi: 10.1177/1753425909102876. [PubMed] [Cross Ref]

von Reumont BM, Jenner RA, Wills MA, Dell'ampio E, Pass G, Ebersberger I, Meyer B, Koenemann S, Iliffe TM, Stamatakis A, Niehuis O, Meusemann K, Misof B. Pancrustacean phylogeny in the light of new phylogenomic data: support for Remipedia as the possible sister group of Hexapoda. Molecular Biology and Evolution. 2012;29:1031–1045. doi: 10.1093/molbev/msr270. [PubMed] [Cross Ref]

Wang X, Fang X, Yang P, Jiang X, Jiang F, Zhao D, Li B, Cui F, Wei J, Ma C, Wang Y, He J, Luo Y, Wang Z, Guo X, Guo W, Wang X, Zhang Y, Yang M, Hao S, Chen B, Ma Z, Yu D, Xiong Z, Zhu Y, Fan D, Han L, Wang B, Chen Y, Wang J, Yang L, Zhao W, Feng Y, Chen G, Lian J, Li Q, Huang Z, Yao X, Lv N, Zhang G, Li Y, Wang J, Wang J, Zhu B, Kang L. The locust genome provides insight into swarm formation and long-distance flight. Nature Communications. 2014;5:2957–2959. doi: 10.1038/ncomms3957.[PMC free article] [PubMed] [Cross Ref]

Wolff C, Gerberding M. "Crustacea": Comparative aspects of early development. Evolutionary Developmental Biology of Invertebrates. 2015;4:39–61. doi: 10.1007/978-3-7091-1853-5_2. [Cross Ref]

Zimmer M, Danko J, Pennings S, Danford A, Carefoot T. Cellulose digestion and phenol oxidation in coastal isopods (Crustacea: Isopoda) Marine Biology. 2002;140:1207–1213. doi: 10.1007/s00227-002-0800-2. [Cross Ref]

Chapter 4: Exploring the roles of individual Hox genes in the establishment of diverse appendage types using CRISPR/Cas9 knockout of individual Hox genes

Serano et al. (2016) (Chapter 2) laid the foundational framework for functional Hox studies in Parhyale by providing a detailed reference of the patterns of expression (both temporally and spatially) for each of the Parhyale Hox genes during embryonic development. The tight correlation revealed between unique combinations of modular/segmental Hox expression and distinct appendage identities suggest that a combinatorial "Hox code" may play a role in establishing the many diverse appendage types in *Parhyale* and other malacostracan crustaceans. It should be again emphasized that while each unique combination of Hox genes corresponded to a distinct limb type (and conversely that each distinct limb type expressed a unique combination of Hox genes), different *levels* of Hox expression (rather than spatial domain) was associated with minor differences among appendages that were considered to share the same type overall.

The following two chapters probes this "Hox code" in the specification of appendage identity during embryonic development — and in the establishment of diverse appendage types overall — through the systematic knockout of the different Hox genes individually (Chapter 4) and in tandem (Chapter 5), establishing the functional roles of the different Hox genes (Chapter 4) functioning as part of an entire complex (Chapter 5).

In Chapter 4, Martin et al. (2016) establishes and optimizes CRISPR/Cas9 mutagenesis in *Parhyale* for the first time and additionally co-localizes the expression of *Ubx* and *abd-A* at a cellular level. Chapter 5 extends the results of Martin et al. (2016) to the perspective of Hox genes acting as part of an entire complex by examining the "post-KO" Hox patterns in transformed embryos to observe co-regulatory interactions that act to define Hox expression domains as well as to make better predictions of combinatorial function in the establishment of limb identity. The establishment of an appendage identity by multiple Hox genes that is different from that established by either alone provides a mechanism for generating greater diversity in an animal that has more appendage types than Hox genes.

CRISPR/Cas9 Mutagenesis Reveals Ve. Roles of Hox Genes in Crustacean Limb Specification and Evolution

Arnaud Martin[1], Julia M. Serano[1], Erin Jarvis[2], Heather S. Bruce[1], Jennifer Wang[2], Shagnik Ray[1], Carryn A. Barker[1], Liam C. O'Connell[2], and Nipam H. Patel[1,2,]

1: Department of Molecular Cell Biology, University of California, Berkeley, California 94720-3200 USA.
2: Department of Integrative Biology, University of California, Berkeley, California 94720-3140 USA.

Published in **Current Biology** 26, 14–26

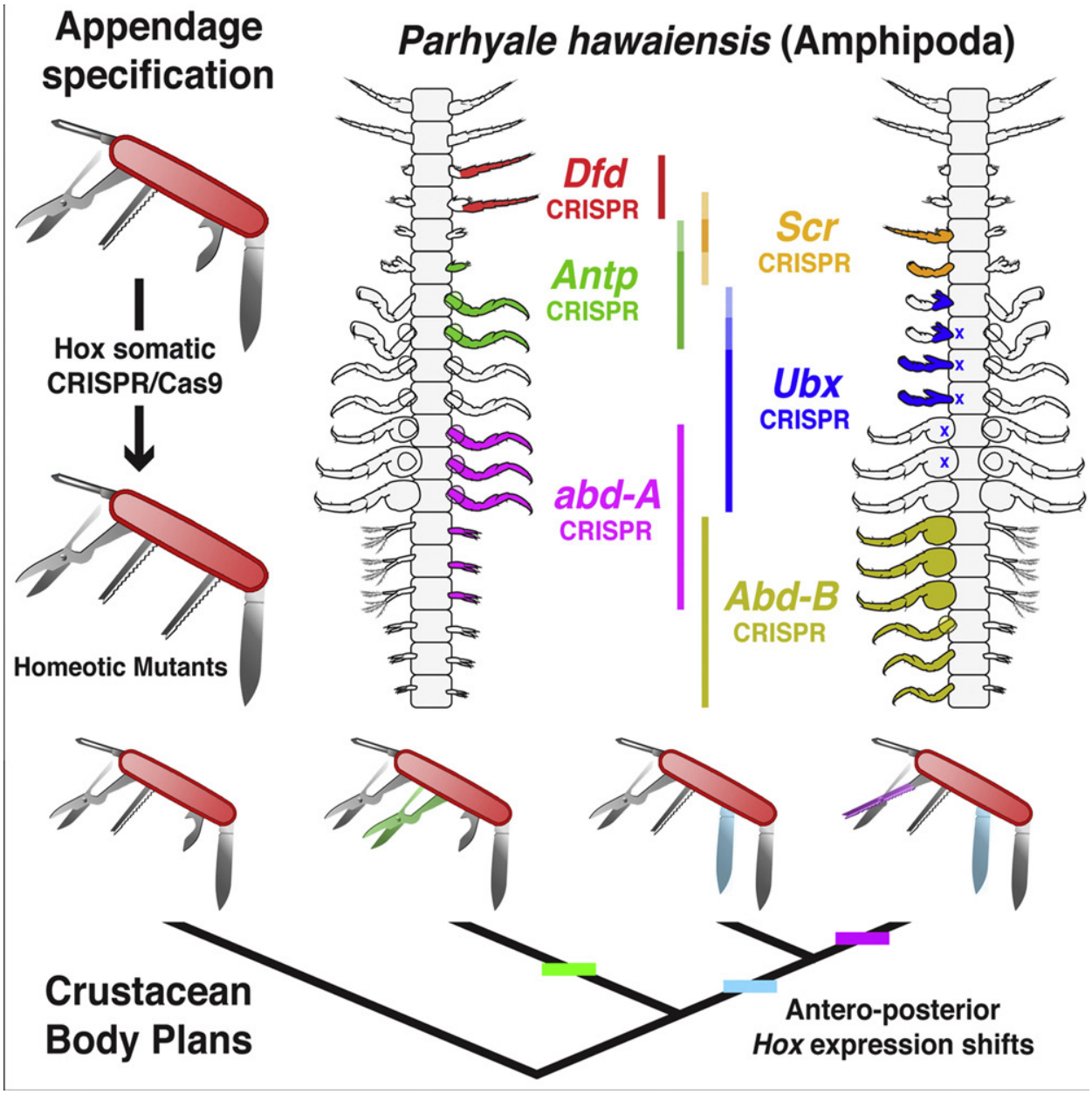
Highlights

- Amphipod crustaceans display a wide array of specialized limbs
- CRISPR mutagenesis and RNAi of Hox genes generate limb transformations
- Limb identity is specified by overlapping domains of Hox expression
- abd-A expression shifts created evolutionary diversification of the crustacean body

Summary

Crustaceans possess a diverse array of specialized limbs. Although shifts in Hox gene expression domains have been postulated to play a role in generating this limb diversity, little functional data have been provided to understand the precise roles of Hox genes during crustacean development. We used a combination of CRISPR/Cas9-targeted mutagenesis and RNAi knockdown to decipher the function of the six Hox genes expressed in the developing mouth and trunk of the amphipod *Parhyale hawaiensis*. These experimentally manipulated animals display specific and striking homeotic transformations. We found that *abdominal-A* (*abd-A*) and *Abdominal-B* (*Abd-B*) are required for proper posterior patterning, with knockout of *Abd-B* resulting in an animal with thoracic type legs along what would have been an abdomen, and *abd-A* disruption generating a simplified body plan characterized by a loss of specialization in both abdominal and thoracic appendages. In the thorax, *Ubx* is necessary for gill development and for repression of gnathal fate, and *Antp* dictates claw morphology. In the mouth, *Scr* and *Antp* confer the part-gnathal, part-thoracic hybrid identity of the maxilliped, and *Scr* and *Dfd* prevent antennal identity in posterior head segments. Our results allow us to define the role Hox genes play in specifying each appendage type in *Parhyale*, including the modular nature by which some appendages are patterned by Hox gene inputs. In addition, we define how changes in Hox gene expression have generated morphological differences between crustacean species. Finally, we also highlight the utility of CRISPR/Cas9-based somatic mutagenesis in emerging model organisms.

Graphical Abstract



4.1 Introduction

Arthropod appendages have diversified into a remarkable repertoire of specialized morphologies. Crustaceans of the Malacostraca class, such as crabs, lobsters, shrimps, or the emerging model organism *Parhyale hawaiensis*, provide remarkable illustrations of this principle [1, 2], as shown by the extensive morphological and functional diversity of limbs along their antero-posterior (AP) axis (Figure 4.1A). This extreme specialization provides a Swiss-army knife arrangement of appendages dedicated to perception (antennae), food processing and chewing (mouthparts), prehension (claws or "chelipeds"), walking (legs or "pereopods"), and propulsion (swimmerets or "pleopods"), and at the end of the *Parhyale* abdomen, forked shaped appendages (uropods) are used for anchoring.

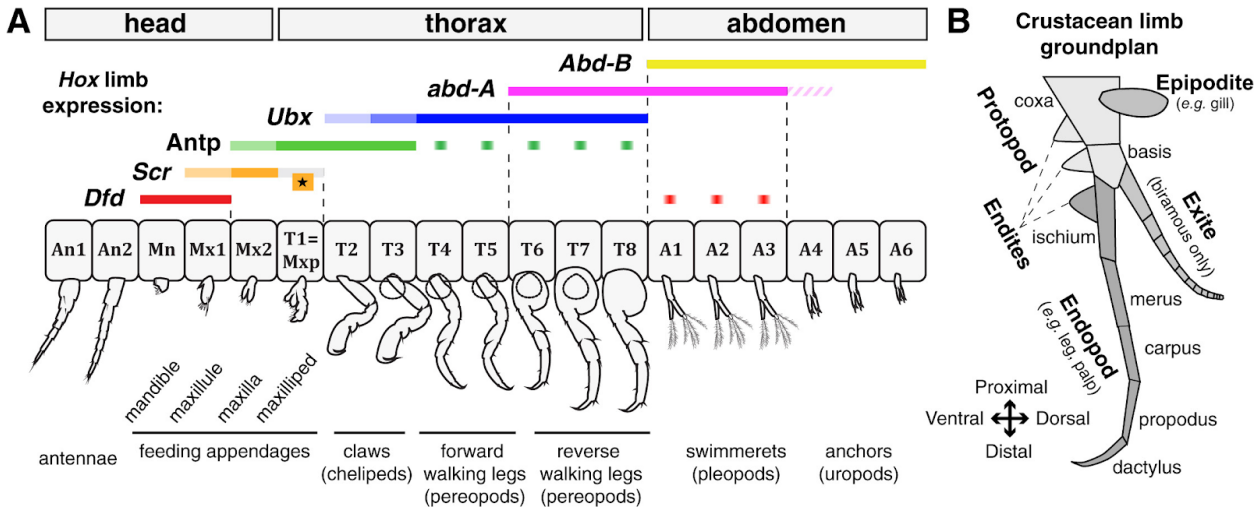


Figure 4.1. Hox Expression and the Crustacean Limb Body Plan

(A) Summary of Hox expression in *P. hawaiensis* in relationship to specialized segments (after [3]). Faded color bars depict weak expression domains. Green squares indicate neuronal expression of *Antp*. Due to post-translational processing of Antp transcripts [3], we show here the domain for *Antp* protein. Red squares indicate mesodermal expression of *Dfd* in the median section of pleopods. The star indicates late, appendage-specific expression of *Scr* in T1/Mxp. The hashed bar indicates transient, weak expression of *abd-A* in A4. (B) Schematic representation of the crustacean limb groundplan.

In spite of the diversity of forms they can take within a single individual, the limbs along the body axis of a crustacean or insect are serial homologs [4, 5]. Comparative anatomy and gene expression data have revealed that the proximo-distal (PD) limb axis is subdivided into two fundamental territories, a proximal protopod and a distal telopod [6–11]. In crustaceans, the protopod forms the base of this structure and is subdivided into two podomeres, the coxa and the basis (Figure 4.1B). The basis can be one-branched (uniramous, with an endopod) or two-branched (biramous, with both an endopod and an exite). All crustacean limb appendages are essentially variations on this common theme.

But how do appendages diverge from this basic organization and acquire a specific morphology based on their position along the body? Hox genes play an important role in establishing segmental identity along the AP axis of arthropods and other animals by regulating the transcription of downstream target genes [12, 13]. Furthermore, subtypes of thoracic and abdominal appendages vary in number and position between crustacean species, and comparative studies suggest that spatial shifts of Hox expression have facilitated such rearrangements by modulating a combinatorial code for limb identity [12, 14–20]. For example, comparative analysis of Ubx expression across crustacean species suggested that the Hox gene *Ubx* plays a role in defining the transition between feeding and locomotory type appendages in the anterior part of thorax [18]. Functional work in *Parhyale* supported this hypothesis: RNAi-based knockdowns transformed the T2 and T3 clawed appendages into a T1 type feeding limb [21], and misexpression of *Ubx* resulted in ectopic locomotory thoracic appendages in the head [22]. Of note, malacostracan T1 segments deviate from the thoracic leg-like archetype as they bear a maxilliped. Although the maxilliped is part of the T1 segment, it is integrated into the mouth apparatus and shows both gnathal and thoracic features [2]. The *Ubx* RNAi phenotypes (T2 and T3 to a T1 maxilliped) thus suggest that *Ubx* represses gnathal identity in the segments posterior to T1 in *Parhyale*, consistent with the observation that *Ubx* expression is restricted to non-maxilliped segments in every crustacean examined so far [16–18, 21].

Beyond *Ubx*, the potential roles of other Hox genes in determining the positional identity of crustacean limbs, and their evolutionary modification between species, remain unclear due to a lack of functional data. To fill this gap, we used CRISPR/Cas9-targeted mutagenesis and RNAi to systematically interfere with Hox function during the development of *P. hawaiensis* embryos. Gene knockouts of the six Hox genes expressed in the mouth and trunk generated homeotic shifts in limb features. These new results outline the combinatorial logic of a Hox code laying out the segmental identity of crustacean appendages.

4.2 Results

CRISPR/Cas9 Loss-of-Function Mutations in G0 Embryos

RNAi-based approaches have been used for gene expression knockdown during crustacean development [21, 23, 24], and although the approach achieves moderate knockdown of mRNA levels [21], the resulting intermediate phenotypes are still useful. As an alternative, we used CRISPR/Cas9 site-directed mutagenesis targeting the coding sequence of *P. hawaiensis* Hox genes and directly assessed effects in G0 embryos. Zygotic co-injections of Hox subgenomic RNA (sgRNA) and Cas9 mRNA or protein induced somatic insertion-deletion mutations, including null alleles at targeted sites, (Figures 2A, 2B, and S1). Injected animals can of course be mosaic and contain alleles generated by independent events, but in the case of *Parhyale*, we can use our detailed understanding of the early embryonic lineages to show that CRISPR/Cas9 targeting generates animals in which gene deletion has occurred in large domains, even easily generating unilaterally mutant individuals. To show this, we performed CRISPR injections into one of the two blastomeres

after the first zygotic cleavage. Given that this first division separates the left versus right sides for the majority of the body axis [25], we expected to see asymmetric effects. Figure 4.2C shows the result of such an experiment for *Antp* knockout and illustrates an embryo in which wild-type levels of Antp protein are detected in one half, but no protein is detected in the other half, indicating that both *Antp*alleles have been disrupted in all cells on one side of the embryo. Figure 4.2D shows an example where we targeted *Ubx* in one of two cells. In this case, the embryo shows wild-type expression of Ubx protein on one side and reduced levels on the other side. Of note for this embryo, the T2/3 limb primordia lacked one or two of the seven podomeres observed on the wild-type side, consistent with a transformation of these T2 and T3 appendages toward a T1 (maxilliped) appendage identity, similar to what was seen previously when Ubx levels were reduced by RNAi [21].

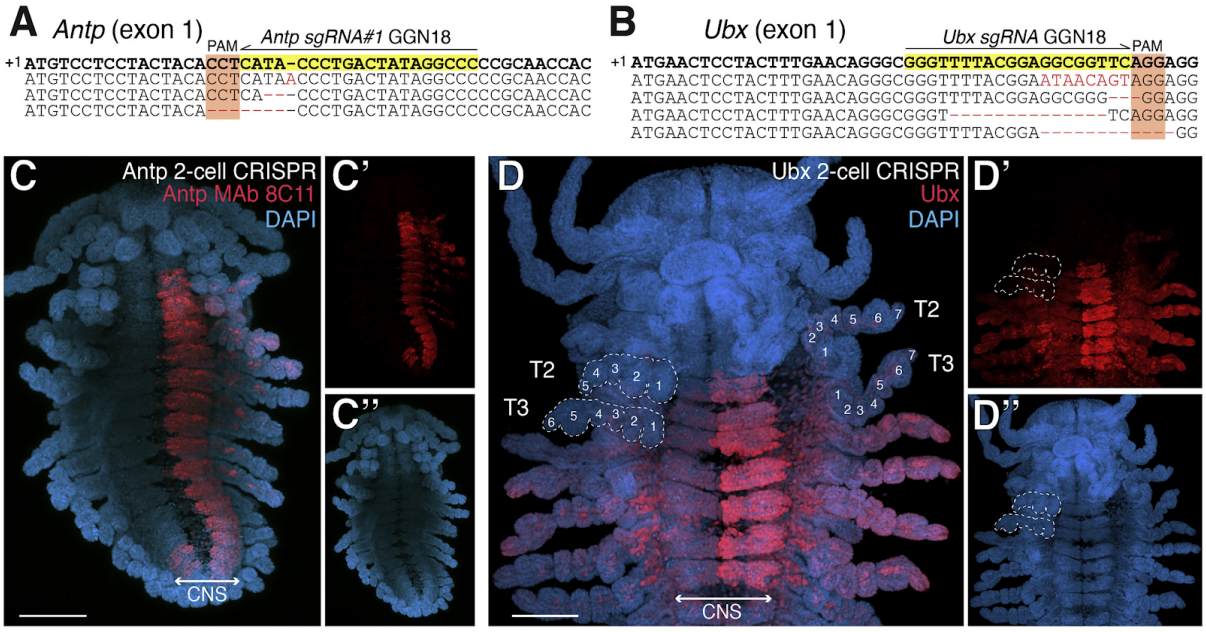


Figure 4.2. CRISPR Somatic Mutagenesis in *P. hawaiensis* Embryos

(A and B) Mutant alleles sequenced from single hatchlings around the NGG Protospacer Adjacent Motif (PAM) after zygotic CRISPR injections, targeting *Antp* (A) and *Ubx* (B). Recovered alleles show short, frameshift-inducing indels relative to the wild-type sequences (in bold). (C) *Antp* immunodetection in a stage 22 embryo after single-cell *Antp* CRISPR injection at the two-cell stage; unilateral expression suggests a complete knockout of *Antp* in the injected lineage. (D) Ubx immunolocalization in a stage-24 embryo, after single-cell CRISPR injection of Ubx sgRNA at the two-cell stage. Dotted lines contour T2 and T3 appendages with podomeres numbered. Expression is reduced on one side. Altered morphology of T2 and T3 on the side with reduced expression indicates the the transformation of T2 and T3 toward T1/Mxp. Scale bars, 100 μm. See also Figure 4S.1.

These results assaying expression after targeting suggest that CRISPR/Cas9 can be used as an efficient tool to generate somatic mutations interfering with gene function during the early embryonic development of P. hawaiensis and that the resultant mutant clones are large, most likely due to targeting soon after injection in the zygote. In this respect, it is useful to note that Parhyale embryos take 8 hr to go from the one-cell to eight-cell stage.

Given this success in knocking out expression, we carried out CRISPR-based loss-of-function experiments targeting the six Hox genes expressed in the mouth and trunk [3]—Deformed (*Dfd*), Sex comb reduced (*Scr*), *Antp*, *Ubx*, *abdominal-A* (*abd-A*), and *Abdominal-B* (*Abd-B*)—injecting the one-cell stage to limit mosaicism, and examined the effect of somatic mutagenesis on limb morphology. All but one sgRNA (abd-A sgRNA#2; penetrance = 12%) generated limb-specific mutant phenotypes in hatchlings at high efficiency with a penetrance ranging between 25% and 70% (Table S1).

Ubx Represses Mouth Features and Promotes Gill Development

As a proof of principle for generating phenotypes using CRISPR/Cas9, we first replicated the results of previous *Ubx* RNAi injections [21] (Figures 3A and 3B). We obtained embryos in which T2 and T3 were transformed toward T1 (the T1 appendage is a maxilliped and from now on is referred to as T1/Mxp). Importantly, T2/T3 retained a clawed morphology at their distal ends but lost the T2/T3-specific comb bristle, indicating a partial T2/T3-to-T1/Mxp homeosis, which was also what had been observed previously by RNAi. However, we also uncovered additional effects of *Ubx* loss-of-function in T4 and T5. As with transformed T2 and T3 appendages, T4 and T5 also acquired endites with multiple setae at their base, but with the addition of a claw-type morphology at their distal ends (Figures 3A–3C); thus, the transformed T4/T5 displayed aspects of both wild-type T1/Mxp and T2/T3 limbs. Finally, *Ubx* CRISPR resulted in the loss of the five pairs of gills, normally attached at the base of the T3–T7 segments (Figures 3A and 3D). This complements the effects of gain-of-function overexpression of *Ubx*, which induces ectopic gills [22]. Taken together, these results suggest that *Ubx* is necessary for the repression of gnathal identity in the base (proximal podomeres) of T2–T5, and required for gill development in T3–T7.

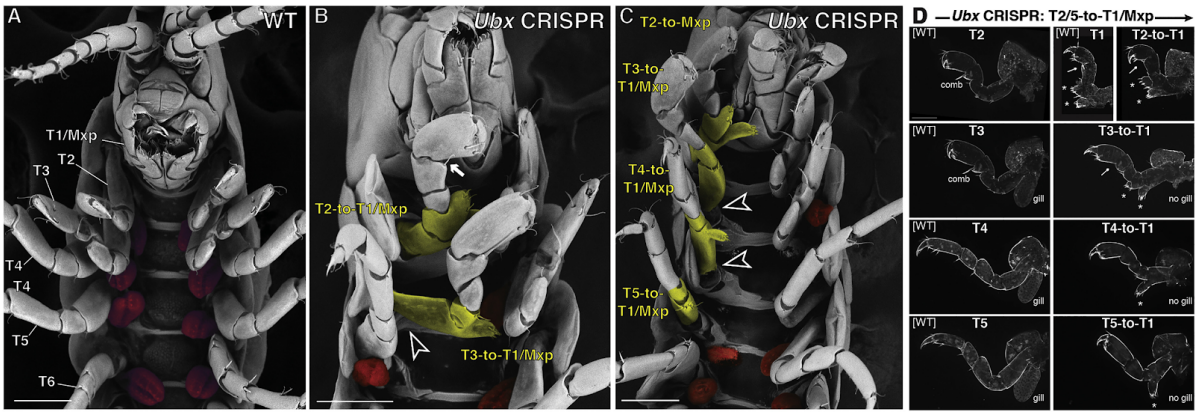


Figure 4.3. CRISPR Somatic Mutagenesis of *Ubx* Generates Thoracic Limb Anteriorizations and Gill Defects

(A) Ventral scanning electron microscopy (SEM) view of a wild-type hatchling showing thoracic gills (red). (B and C) SEM of an *Ubx* CRISPR mosaic mutants; these hatchling display modified T2–T5 proximal segments (yellow), sometimes displaying ectopic endites. Arrow, comb bristle absent; arrowheads, gills absent (red, wild-type gills). (D) Dark-field images of *Ubx* CRISPR limb homeoses. Arrow, comb bristle absent; asterisks, ectopic Mxp-like endites. Scale bars, 100 μm (A–C).

We interpret the more extreme phenotype seen with CRISPR/Cas9 (with two different sgRNAs) relative to RNAi as the difference between gene knockdown and knockout. The restriction of *Ubx* RNAi effects to T2–T3 suggests these segments are most sensitive to the reduction in the level of *Ubx* expression, and it is worth noting that *Ubx* RNA and protein are expressed at lower levels in these two segments than in T4–T8 in wild-type animals [21]. It is important, however, to consider that CRISPR/Cas9 mutagenesis is also expected to sometimes yield similar partial knockdown effects. In some cases, this may be because the mutant alleles that are generated retain some function or because of the tissue mosaicism inherent to our somatic analysis.

abd-A and *Abd-B* Organize the Specialization of Posterior Appendages

The posterior half of *P. hawaiensis* shows three pairs of reverse walking legs in T6–T8, three pairs of swimmerets in A1–A3, and three pairs of uropods in A4–A6 (Figure 4.4A). This anatomical parcellation is reĀlected at the molecular level by the expression of *abd-A* in the posterior legs and swimmerets (and weakly in the uropod of A4) and by *Abd-B* expression, which extends from the swimmerets to the uropods (Figures 1A and S2). It follows that the partially overlapping expression domains of two Hox genes creates three Hox states that correlate with morphological differences: (1) *abd-A* in T6–T8, (2) *abd-A* plus *Abd-B* in A1–T3, and (3) *Abd-B* in A4–T6. In the appendages of A1–A3, it appears that all ectodermal cells do co-express *abd-A* and *Abd-B* during limb development (Figure 4.S2). Here we tested the hypothesis that the posterior heteronomy of amphipods is specified by combinatorial Hox expression.

CRISPR somatic mutagenesis of *abd-A* validated this hypothesis and induced three notable limb modifications across its expression domain. First, posterior legs were transformed into anterior legs (T6/8-to-T4/5), as evidenced by their inverted orientation (for instance, with the dactyl pointing backward instead of forward) and by the absence of a large coxa, a characteristic of posterior legs (Figures 4A–4D). Second, the T8 segment acquired an ectopic gill, which is expected from a T8-to-T4/5 transformation (Figures 4B–4D). Third, swimmerets were transformed into uropods (A1/2/3-to-A4/6; Figures 4E–4G), and the A1–A3 abdominal body segments bearing them also transformed toward the A4/5/6 body segments in terms of size and shape, resulting in a severely contracted, narrow abdomen and an aberrant curvature of the body (Figure 4.4B). Interestingly, *abd-A* loss-of-function phenotypes can be seen as an anteriorization of T6–T8 and as a posteriorization of A1–A3. These results were replicated in CRISPR experiments that used an sgRNA targeting the second exon of *abd-A* (Table S1; Figure 4.S1), thus ruling out off-target effects on limb morphology. In comparison, two independent expression knockdown experiments failed to recreate T6–T8 transformations but succeeded in replicating the effects of *abd-A* CRISPR in the abdomen (Figures 4H and 4I). Specifically, the injection of siRNA and two transgenic lines expressing *abd-A* hairpin RNAs under the control of a heatshock promoter [26] all resulted in aberrant swimmerets resembling uropods (A1–A3 transformed toward A4/5/6), characterized by detached pairs of appendages, a curved basis, and a failure to develop propulsive setae. This suggests again the importance of Hox expression levels—in this case, A1–A3 are more sensitive to lowering *abd-A* expression than are T6–T8.

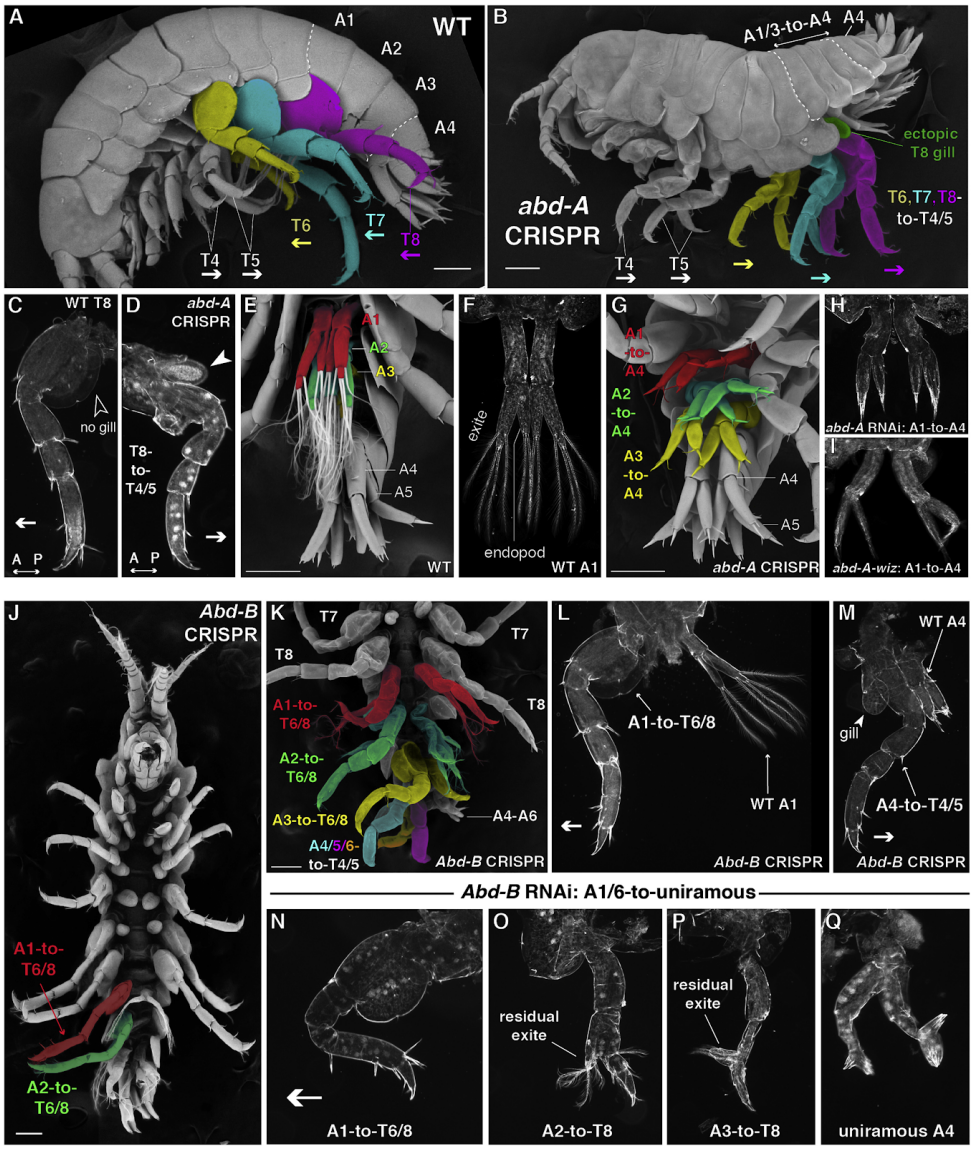


Figure 4.4. *abd-A* and *Abd-B* Pattern Functional Subdivisions in Thoracic and Abdominal Appendages

(A and B) Lateral SEM views of wild-type (A) and *abd-A* CRISPR-injected (B) *Parhyale* hatchlings; *abd-A* CRISPR results in leg homonomy, with anteriorization of the reverse-walking morphology in T6–T8 and ectopic gills in normally gill-less T8. Mutant abdomens curl upward due to A1–A3 posteriorization. (C and D) Dark-field images of dissected wild-type (C) and *abd-A* mutant (D) T8, with reversed polarity in the antero-posterior (AP) axis. (E and F) Ventral views of wild-type pleopods (A1–A3 swimmerets), characterized by a biramous morphology terminated by long setae. (G) Ventral view of an *abd-A* CRISPR somatic mutant. The A1–A3 pleopods are disjointed and acquire a more posterior, uropod-like morphology lacking long setae. (H) Pleopod-to-uropod transformation obtained after zygotic injection of *abd-A* siRNA. (I) Pleopod-to-uropod transformation obtained in the *hsp70-abd-A-wiz* transgenic line. (J and K) Ventral views of *Abd-B* CRISPR mosaic mutant hatchlings. Abdominal segments show gradual transitions toward a leg-like, uniramous morphology, including complete A1/3-to-T8 and A4/6-to-T4/5 transformations. (L) Dorsal view of the flattened appendages of the A1 segment of an *Abd-B* CRISPR mosaic mutant with a unilateral transformation of a pleopod into a posterior leg with large coxa (A1-to-T8). (M) Dorsal view of an *Abd-B* CRISPR mosaic mutant with a unilateral transformation of a uropod into a gilled anterior leg (A4-to-T4/5). (N–Q) Variable partial transformations of the biramous A1–A4 limb toward a uniramous morphology after *Abd-B* RNAi zygotic injections, including complete transformations into reverse-walking (T8-like) legs (N). Effects similar to those in (Q) were obtained for A5–A6. See also Figure 4.S2.

In accordance with its pan-abdominal expression, *Abd-B* CRISPR transformed both swimmerets and uropods into walking legs (Figures 4J–4M), culminating in a densely packed array of legs as seen across one side of the entire abdomen in Figure 4.4K. The induced legs displayed a large (T8-like) coxa in the A1–A3 segments and a narrow coxa characteristic of anterior legs (T4/5-like) in the A4–A6 segments. We deduce that disruption of *Abd-B* transforms A1–A3 toward T6/7/8 and A4–A6 toward T4/5 in *abd-A*-positive and *abd-A*-weak/negative domains, respectively. In extreme cases, transformed legs in the *abd-A*-weak A4 segment showed an ectopic gill reminiscent of the *abd-A* CRISPR T8-to-T4/5 transformation (Figure 4.4M). It will be interesting to test whether *Ubx* expression extends posteriorly into the abdomen upon *Abd-B* knockout, which could explain these gill acquisitions.

Abd-B RNAi injections showed a similar, but less severe effect: both swimmerets and uropods underwent a biramous-to-uniramous transition, with most swimming segments taking a walking-leg morphology (Figures 4N–4Q). Although these incomplete transformations are most likely due to a limitation of the knockdown approach, *Abd-B* RNAi succeeded in forming fully differentiated reverse walking legs in the first abdominal segment (A1-to-T8). Thus, several lines of evidence show that *Abd-B* is necessary for the maintenance of abdominal limb identity and promotes biramous morphology.

Antp Functions in Claw Specification

Amphipod thoracic legs are subdivided into three types—prehensile claw-like chelipeds (T2–T3), forward-walking legs (T4–T5), and reverse-walking legs (T6–T8). While *abd-A* directs the differentiation of reverse- versus forward-walking legs, *Ubx* disruption did not explain the genetic demarcation between claws and walking legs in the anterior thorax. This functional subdivision could depend on Antp, which is expressed in clawed segments, but not in developing walking legs [3]. *Antp* CRISPR validated this hypothesis and yielded cheliped-to-forward-walking-leg transformations (T2/3-to-T4/5), as evidenced by the narrow shape of the propodus segment and by the loss of T2/3-specific comb bristles involved in grooming (Figures 5A–5D). Modified chelipeds failed to acquire normal segmentation and retained a fused ischium-merus (Figure 4.5C). In wild-type animals, gills are only observed in T3–T7 segments, but remarkably, T2 transformed limbs also displayed an ectopic gill indicative of a more posterior specification (T2-to-T4/5). This anatomical gain provides additional evidence for a homeotic effect of *Antp* somatic mutagenesis on the entire appendage. As these results were obtained for two distinct sgRNAs of comparable penetrance, we have ruled out off-target effects of *Antp* CRISPR on limb morphology and conclude that *Antp* is required for the specification of chelipeds.

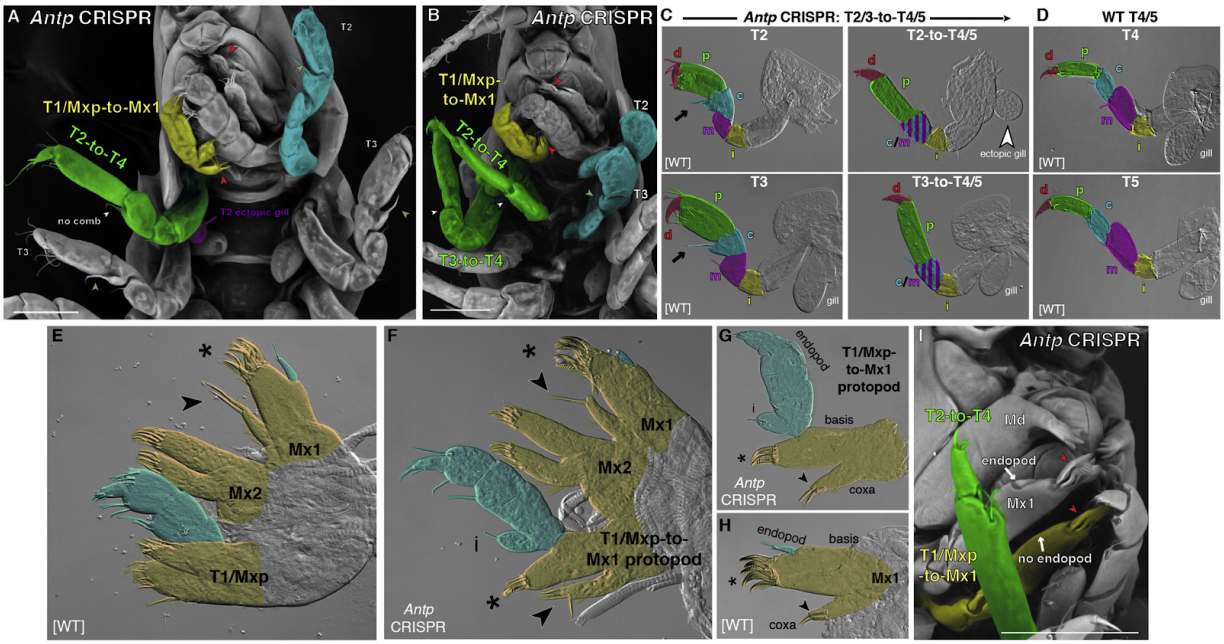


Figure 4.5. *Antp* Is Required for Limb Specialization in the Anterior Thorax

(A) Ventral SEM view of a unilateral mutant hatchling obtained by *Antp* CRISPR. The mouth is dislocated due to an incomplete T1/Mxp-to-Mx1 transformation (yellow), bearing Mx1-specific setal teeth (red arrowheads). The transformed T2 limb (green) lacks a normal claw morphology and the T2/3 specific comb bristle (white arrowheads, absent; green, wild-type) and shows an ectopic gill (purple). (B) Another example of a unilateral *Antp* CRISPR mutant, notably showing a T2/3-to-walking-leg transformation (green). (C) Differential interference contrast (DIC) imaging of T2/3 limbs transformed by *Antp* CRISPR, with endopod podomeres false colored. Arrow, comb bristle (absent in mutants); d, dactylus; p, propodus; c, carpus; m, merus; i, ischium. (D) Wild-type T4 and T5 forward-walking legs. (E–H) Detailed morphology of the maxillary apparatus in wild-types (E and H) and *Antp* CRISPR mutants (F and G). Blue, endopods; yellow, protopods; asterisk, Mx1-basis-specific setal teeth; arrowhead, Mx1-coxa-specific endite. (I) Unilateral *Antp* CRISPR mutant showing a T1/Mxp-to-Mx1 transformation with a complete ablation of the endopod. Scale bars, 100 μm (A, B, and I).

Complementary Effects of *Antp* and *Scr* in Maxillipeds

In contrast with the thorax and abdomen, arthropod mouthparts generally show a sequential heteronomy where all consecutive segments are distinct and differ from each other, without repetition. In the next two sections, we explore how sequential expression of Hox genes might explain the differentiation of the amphipod mouth apparatus. In addition to its effects on chelipeds, *Antp* CRISPR also resulted in visible defects in maxillipeds. *Antp* mutant jaws showed T1/Mxp-to-Mx1 transformations, as revealed by the acquisition of Mx1-specific serrated setal teeth on the basis endite and by a narrow coxal endite, topped by two long simple setae (Figures 5A–5B and 5E–5G). In contrast, wild-type T1/Mxp endites both resemble the Mx2 condition (Figure 4.5H). Disruption of *Antp* also showed graded effects of the T1/Mxp endopods, the more distal part of the limb. In the milder forms, the T1/Mxp endopod regressed into a palp of bulging aspect, due to an abnormally narrow attachment site on the basis article (Figure 4.5G). These transformed limbs display

three endites instead of two, due to the maintenance of the ischium and basis endites with an ectopic and prominent, Mx1-like coxal endite. In more extreme forms, the endopod was missing from the T1/Mxp-to-Mx1 transformed limb (Figure 4.5I). Altogether, these results underline the dual role of *Antp* in maxilliped development, as it selects the identity of the proximal domain while also being required for palp growth in the distal domain.

Scr expression (Mx1-T1/Mxp) overlaps with *Antp* (Mx2-T3) in the mouth, and hatchlings that were injected with *Scr* CRISPR showed a mild to severe disorganization of the jaw due to an imperfect interlocking of the modified mouthparts. Upon closer inspection, *Scr* mosaic mutants revealed maxilliped-to-cheliped transformations (T1/Mxp-to-T2/3), with the distal palp acquiring both a T2/3-specific comb bristle and the morphology of a prehensile claw, characterized by an enlarged propodus and an opposing dactyl (Figures 6A–6C). In the proximal domain, T1/Mxp endites regressed upon *Scr* loss of function, consistent with a conversion of this appendage toward a thoracic identity.

In summary, both the CRISPR phenotypes of both *Antp* and *Scr*mutants reveal a modular, composite organization of the maxilliped, with dual effects on the proximal and distal domains. *Scr* functions as a determinant of the gnathal identity of the protopod, and *Antp* is necessary for preventing Mx1-like morphology in this limb domain. Conversely, in the T1/Mxp endopod, *Scr* inhibits the posterior claw-like morphology of the palp, and *Antp* is required for endopod presence. The antagonistic roles of these genes may thus explain the hybrid nature of maxillipeds, by conferring a combination of thoracic (presence of an endopodal extension) and gnathal (sensory endites and clawless palp) features.

Conserved Functions of *Scr* and *Dfd* in Mouth Patterning

The mandible (Mn), maxillule (Mx1), and maxillae (Mx2) are consecutive mouth appendages involved in food processing along with the more posterior T1 maxilliped. Mn or Mx1 palps are common among other amphipods, but in *P. hawaiensis*, a residual palp is apparent on Mx1 only and the three appendages thus appear to repress endopod development. CRISPR-induced mutagenesis of both *Scr* and *Dfd* revealed that these genes control different aspects of the regional identity of these segments on the PD axis.

Wild-type Mx2 have two lobes with simple setae (the coxa and basis endites). Although this appendage is thus devoid of an endopod, both RNAi- and CRISPR-based loss of function of *Scr*activated endopodal growth (Figures 6D and 6E). *Scr* CRISPR individuals showed a gradual series of Mx2 modifications, starting with the presence of an ectopic endopod and the acquisition of serrated setae characteristic of the Mx1 segment on the endites (Mx2-to-Mx1). In the most extremes cases, the endites regressed, and an antenna-like endopod protruded from the side of the mouth (Mx2-to-An). *Scr* RNAi resulted in less dramatic, but still striking, Mx2 phenotypes with the formation of ectopic but incomplete endopods, and the acquisition of an additional endite-bearing ischium—a condition that exists in the maxilliped of *Parhyale* and other amphipods [2]. This transformation of Mx2 to T1/Mxp (Figure 4.6D) is consistent with the phenotype seen when low levels of *Ubx*

misexpression cause a reduction in the levels of *Scr* [22]. *Scr* RNAi had no effect on the Mx2 proximal domain or on Mxp, suggesting a lower expressivity than *Scr* CRISPR.

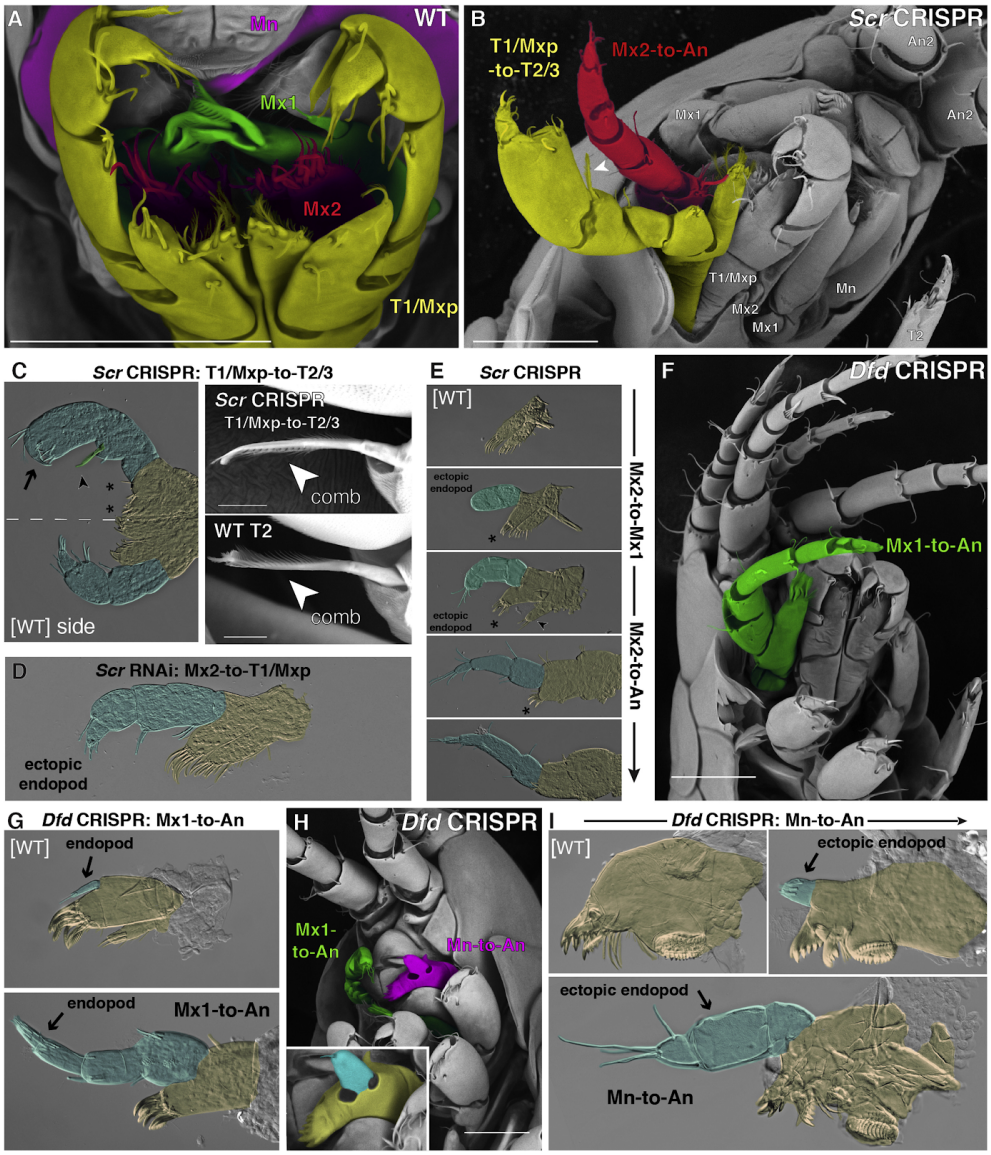


Figure 4.6. *Scr* and *Dfd* Maintain Gnathal Features in Mouth Appendages

(A) Ventral SEM view of a wild-type mouth apparatus. (B) Mouth identity defects in a unilateral *Scr* CRISPR mutant. Arrowhead, T2/3-specific comb bristle. (C) Ventral view of a dissected, unilaterally transformed pair of *Scr* CRISPR maxillipeds. Arrow, T2/3-like claw morphology of the dactylus/propodus; arrowhead, T2/3-specific comb bristle; asterisk, regressed T1/Mxp endites. (D) Effect of *Scr* RNAi on Mx2, with growth of an ectopic endopod (blue). (E) Spectrum of effects of *Scr* CRISPR on Mx2, with growth of an ectopic endopod (blue), acquisition of Mx1 protopod identity (asterisk, Mx1-specific setal teeth on basis endite; arrowhead, Mx1-specific coxal endite), and in most marked phenotypes, acquisition of an antenna-like morphology. (F and G) Transformation of the Mx1-endopod into an antennal morphology (green) in *Dfd* CRISPR mosaic mutant hatchlings. (H and I) Acquisition of an ectopic antenna-like endopod (blue) in the Mn appendage after *Dfd* CRISPR. Scale bars, 100 μm (A–B, F, and H) and 10 μm (C).

Like *Scr* in Mx2, *Dfd* CRISPR promoted endopodal development in Mn and Mx1 (Figures 6F–6I). In *Dfd* mutants, the Mx1 vestigial endopod developed into a segmented antenna, whereas the proximal domain retained an Mx1 identity (partial Mx1-to-An). *Dfd* CRISPR also induced the formation of an ectopic endopod on Mn, resulting in a dislocated mandible protruding from the mouth apparatus or culminating in the formation of a short and segmented antennal primordium (Mn-to-An).

Taken together, these results show that *Scr*/*Dfd* loss-of-function experiments both induce antenna-like appendages in the mouth. Similar phenotypes have been observed in homologous segments upon *Scr/Dfd* knockdown in hemipterans and coleopterans, suggesting an evolutionarily conserved role in the maintenance of gnathal identity between insects and crustaceans [27–31]. In the proximal domain of mouth appendages, *Scr* in particular patterns the setulation of Mx2/T1 endites, whereas *Dfd* prevents spurious Mx1 morphology in the Mx2 segment. We conclude that in addition to an ancestral function in the distal repression of antennal fate, the sequential expression of *Scr* and *Dfd* in the mouth also contributes to the heteronomy of this body region via modular effects along the limb PD axis.

4.3 Discussion

Using CRISPR for Somatic Analyses of Gene Function

The recent development and apparent universality of CRISPR/Cas9 genome editing [32] allowed us to analyze the function of the six Hox genes expressed in the mouth and trunk of *P. hawaiensis*, an emerging model organism (Figures 7A and 7B). We used zygotic injections to generate DNA lesions in the soma, without attempting stable germline transformation. Here we discuss this strategy and the extent to which it could foster discovery in analogous experimental systems.

Cost

We generated ready-to-inject samples in 2–3 days and at low cost (less than $80 per target in reagents).

Penetrance

CRISPR/Cas9 somatic loss-of-function experiments generated homeotic phenotypes at high frequency for eight out of the nine sgRNAs that were assessed (Table S1). The lower efficiency of *abd-A sgRNA#2* (12% penetrance) may be explained by the fact it was the only sgRNA targeting a short second exon (Figure 4.S1), which could be subject to splicing.

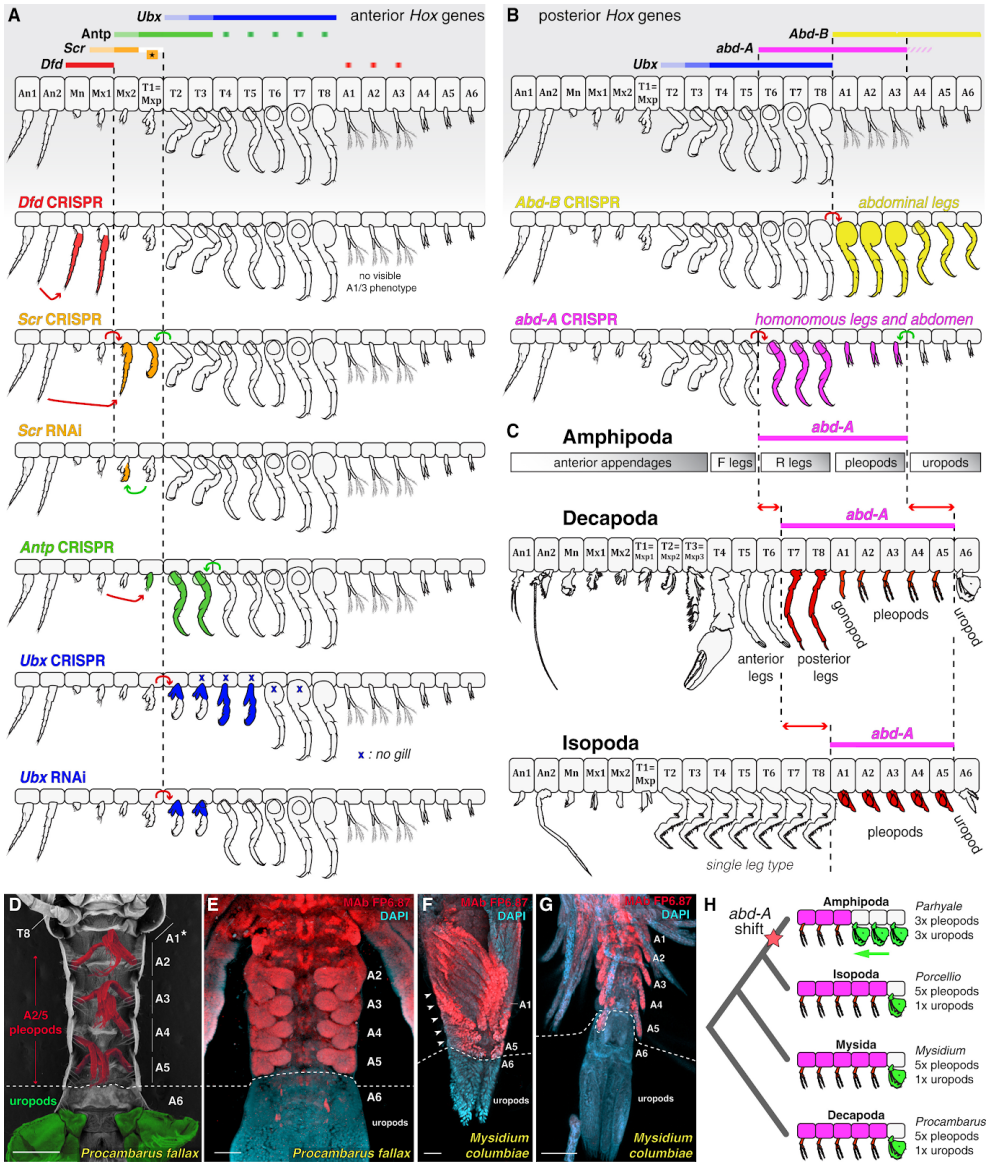


Figure 4.7. The Control of Crustacean Limb Identity by Hox Genes
(A and B) Summary of all known Hox loss-of-function phenotypes in *P. hawaiensis*. Arrows indicate the directionality of the homeosis (red, anteriorization; green, posteriorization); dotted lines indicate gills (C) AP shifts in *abd-A* expression recapitulate the evolution of limb-type subdivision in both thoracic legs and abdominal appendages (this study) [16, 17, 26]. Notice that *abd-A* loss of function triggers homeotic shift of opposite directions on each side of its expression domain in *P. hawaiensis* (B). F legs, forward-walking legs; R legs, reverse-walking legs. (D) Ventral SEM view of the abdominal appendages of a decapod crayfish hatchling, with only the A6 segment bearing a uropod. Asterisk, A1 appears to be limbless in hatchlings, but males develop a gonopod (modified pleopod involved in reproduction) at the juvenile stages. (E) FP6.87 staining of a crayfish embryonic abdomen (red); staining is absent from A6 and uniform in A2–A5 limb primordial. (F and G) FP6.87 staining of the embryonic abdomen of a mysid shrimp at successive stages (red); staining is absent from the uropod-bearing A6 segment and uniform in A1–A5 pleopod primordia. (H) Modification of abdominal limb distribution in Amphipoda (tree topology after [33]). Scale bars, 500 μm (D and E) and 100 μm (F and G). See also Table S1.

Expressivity

In all our comparisons (*Scr*, *Ubx*, *abd-A*, and *Abd-B*), CRISPR showed more marked effects than siRNA injections, with an increased degree of transformation. These limitations of RNAi are most likely due to incomplete mRNA knockdown in *Parhyale* [21], although clearly the combination of CRISPR and RNAi data was useful in revealing the relative sensitivity of different segments to Hox gene perturbation.

Reproducibility and Target Specificity

Although it would be difficult to assess the target specificity of CRISPR in our model system, we sought to test the reproducibility of limb transformation phenotypes using non-overlapping sgRNAs (for *Dfd*, *Scr*, *Antp*, and *abd-A*). In these four cases, mutant phenotypes were equivalent regardless of the 19–20 bp nucleotides targeted. RNAi phenotypes obtained for *Ubx*, *abd-A*, and *Abd-B* were also consistent with the effects of CRISPR mutagenesis in these genes. These results provide independent replications and rule out off-target effects of CRISPR/Cas9 somatic mutagenesis on limb morphology.

Mosaicism

A caveat of somatic mutagenesis is linked to the random occurrence of DNA cleavage in post-zygotic stages [34]. We have seen that CRISPR injections have the potential to generate bi-allelic knockouts that spread to large sections of the injected individual by clonal inheritance (Figure 4.2C). For any given transformed animal, the distribution of mutant cells and their respective allelic dosage are unknown (Figure 4.2D). That said, the resulting mosaicism can be advantageous for several reasons. First, unilateral mutant phenotypes can be directly compared to a wild-type state within the same animal, providing an internal control. Second, for pleiotropic genes involved in several processes across development, mosaicism may increase the rate of surviving "escapers" by randomizing the distribution of mutant clones. Last, mosaicism can generate phenotypic series that are biologically informative. In our case, this was true for *Scr* CRISPR, in which intermediate (Mx2-to-Mx1) and severe homeosis (Mx2-to-An) suggested two distinct functions of this gene in the maxilla segment.

Overall, we encourage the use of CRISPR/Cas9 somatic mutagenesis for the rapid analysis of gene function in emerging model organisms with injectable eggs, complementing the already widespread use of RNAi. This should notably facilitate the systematic study of Hox gene function across a broad sample of arthropods, drawing the promise of an extended understanding of segmental and serial homolog evolution. For instance, CRISPR has been successfully carried out in the branchiopod *Daphnia magna* [35]. Hox mutagenesis in this species would extend existing gene expression analyses [36] and could yield important comparative insights into the macroevolution of crustaceans.

Hox Expression Shifts and the Evolution of the Crustacean Trunk

CRISPR mosaic mutants reveal that *abd-A* and *Abd-B* expression domains determine segmental identity in the thorax and abdomen. This combinatorial model sheds light into the evolution of abdominal appendages. In four malacostracans, *Procambarus*, *Porcellio*, *Mysidium*, and *Mysidopsis*, *abd-A* is expressed in the first five abdominal segments (containing pleopods), but not in the uropods of A6, the sixth and final abdominal segment [4, 16, 17] (summarized in Figure 4.7C). To replicate these previous results, we profiled the distribution of the Ubx/abd-A proteins using the cross-reactive FP6.87 monoclonal antibody [37, 38] in *Procambarus fallax* (Decapoda) and *Mysidium columbiae* (Mysida) embryos. As previously reported, we found that the A6 segment of both decapods and mysids lacked abd-A, correlating with the presence of a pair of uropods rather than pleopods on A6 (Figures 7D–7F). In other words, abd-A-weak abdominal segments are always associated with a uropod identity. The three pleopod plus three uropod segment arrangement is unique to amphipods and may have been caused by an amphipod-specific loss of abd-A expression in A4–A5 (Figure 4.7H).

CRISPR somatic knockouts approximate this evolutionary scenario, as the experimental disruption of *abd-A* resulted in pleopod-to-uropod transformations, which validates a functional link between *abd-A* deployment and pleopod/uropod ratio. The posteriorization of the A1–A3 domain also suggests that *abd-A* works in conjunction with the overlapping expression of *Abd-B* in this region to establish pleopod identity. This is a new exception to the "posterior prevalence rule" [39], which would have predicted an absence of *abd-A* function in cells co-expressing the more posterior Hox gene *Abd-B*. Both comparative and functional data thus suggest that spatial shifts of *abd-A* deployment modulate the number of *abd-A*-negative uropods, explaining divergent arrangements of abdominal appendages in crustaceans.

Amphipods are also characterized by the presence of two types of legs ("amphi-poda," *gr.* "different foot"), in contrast with isopods, which possess a single type of walking-leg morphology. In our amphipod model organism, *abd-A* mutagenesis replaced the reverse-walking legs with additional forward-walking legs, resulting in an isopod-like conĀiguration. Accordingly, *abd-A* is not expressed in the legs of an isopod [16], suggesting that disruption of *abd-A* in the amphipod thorax effectively recapitulated the isopod state (Figure 4.4.7H). Given the central role that Hox genes play in determining arthropod segment identity, evolutionary shifts in Hox expression domains may provide a recurring strategy to generate diverse arrangements of specialized limb types [12, 40].

Hox Functions in the Modular Evolution of Maxillipeds

Our model for how shifts in *abd-A* and *Antp* expression have accompanied morphological evolution of the crustacean body plan is similar to the proposed role of *Ubx* in generating diversity in the number of crustacean maxillipeds [18]. Crustaceans exhibit anywhere from zero to three pairs of maxillipeds, and the number of maxillipeds correlates with the position of the anterior boundary of *Ubx* expression; in other words, appendages of the

anterior thorax that lack *Ubx* expression become maxillipeds, whereas those that express *Ubx* become claws or legs [18, 21, 22].

Although *Ubx* represses gnathal identity the thorax, it does not explain the "chimeric"—both gnathal and thoracic—identity of the maxilliped [41], a composite identity that relies on the selector activities of more anterior Hox genes. Indeed, we found that interfering with Hox gene functions triggered modular effects on feeding appendages, with endite-bearing articles requiring *Scr*, the maxilliped endopod requiring *Antp*, and either *Scr* or *Dfd*repressing antenna-like endopods in maxillae (Mx1) and maxillules (Mx2). The compartmented functions of these consecutive genes may contribute to the robust establishment of differentiated morphologies in adjacent mouth segments, and they also shed light on the composite nature of maxillipeds. Indeed, *Scr*and *Antp* show dual functions that are complementary in each section of T1/Mxp. In the proximal section, *Scr* is necessary for the growth of endites while *Antp* provides positional identity. The Mxp distal domain shows a reverse pattern, with a requirement of *Antp*for palp growth, while *Scr* provides positional identity. The ability of Hox genes to perform different functions along the PD axis has been linked to the Hox co-factors *Homothorax* (*Hth*) and *Extradenticle* (*Exd*) in insects [31, 42–51]. Because *Exd* and *Hth* expression mark proximal limb domains in the crustacean limb [6, 7, 10, 11], Exd/Hth/Hox protein interactions could explain the differential effects of genes such as *Antp*, *Scr*, and *Dfd* in the proximal versus distal domains of the *Parhyale* feeding segments. Combinations of Hox genes and proximal co-factors may thus form a molecular canvas for the modular evolution of crustacean limbs, as reĀlecɨed by the chimeric organization of crustacean maxillipeds.

Arthropod appendages are often used as an example of dramatic evolutionary diversification of form and function, and there are numerous examples of correlations between gene expression and differing morphologies between species. In some cases, these correlations have included expression changes in Hox genes, and other studies have implicated evolutionary shifts in Hox gene targets. Our functional studies can now point to the shifting domains of Hox genes as part of the basis for crustacean body-plan evolution, but there is no doubt that this will only be part of the explanation.

4.4 Experimental Procedures

Detailed experimental procedures are presented in the Supplemental Experimental Procedures.

CRISPR/Cas9 Somatic Mutagenesis

Preparation of *Cas9* mRNA followed a published procedure [52], with the exception that we used the *pCasX* plasmid linearized with *Acc65I* as a template for T7 transcription [53]. For sgRNA design, we used the ZiFiT Targeter webtool [54] to scan for GGN17–18[NGG] motifs in Hox gene open reading frames and generated oligonucleotides integrating target specific GGN17–18 into a tracrRNA sequence (Supplemental Experimental Procedures). PCR

assembly of the DNA templates, in vitro transcription, and purification of sgRNAs followed a previously published protocol [52]. *Cas9* mRNA and sgRNA were mixed in a 1:2 molar ratio, purified, suspended in water, and stored at −80°C until injection. Injection mixes based on Cas9 protein consisted of aqueous re-suspensions of 333 ng/μl recombinant Cas9-NLS protein (PNA Bio, catalog number CP01), 200 ng/μl sgRNA, and 0.05% phenol red dye for injection visualization.

RNA Interference

Stealth siRNA duplexes were designed using the BLOCK-IT RNAi Designer tool (Thermo Fischer Scientific) in non-conserved coding regions of the *Parhyale Scr*, *abd-A*, and *Abd-B* transcripts [3]. The *abd-A-wiz* construct was generated by directional cloning of 544nt *abd-A* fragments into a modified version of the *pWIZ* vector [26]. The resulting construct, consisting of the *abd-A* fragment in opposite orientations on either side of the *Drosophila white* intron, was then placed downstream of the *Parhyale hsp70* promoter within a *pMinos* transformation vector, and germline transformation was carried out as previously described [55, 56]. Embryos from females that carried the *hsp70-abd-A-wiz* transgene were isolated and subjected to daily, 1-hr-long heat shocks beginning around stage 14 (i.e., just before the onset of endogenous *abd-A* expression) and continuing until hatching. Two independent transgenic lines were established, and both yielded similar results.

Injections and Imaging

Embryo injection followed a published protocol [57]. For CRISPR somatic mutagenesis, approximately 40–60 picoliters of 400–600 ng/μl *Cas9* mRNA/sgRNA mixture were injected into one-cell embryos. For RNAi, approximately 20–40 picoliters of 200 μM siRNAs were injected into one-cell embryos or both cells of two-cell embryos. For examination of limb phenotypes, *P. hawaiensis*hatchlings were fixed for 2 hr in 3.7% formaldehyde, and appendages were removed individually, mounted in 70% glycerol, and visualized with dark-field and DIC optics. For SEM, *P. hawaiensis* and *P. fallax* hatchlings were fixed for 2 hr in 3.7% formaldehyde, dehydrated via an ethanol series prior to critical point drying, and examined on a Hitachi TM-1000. *M. columbiae*specimens used for SEM were obtained from a stock of adults that had been fixed and stored in MeOH. SEM images were false-colored using the “Darken” and “Soft Light” layer functions of Adobe Photoshop.

Author Contributions

N.H.P., A.M., and J.M.S. conceived the project. A.M., E.J., J.W., H.S.B., and S.R. performed CRISPR/Cas9 mutagenesis and analysis. J.M.S. and C.A.B. carried out RNAi. E.J. and L.C.O. performed antibody and in situ analysis. A.M., N.H.P., and J.M.S. wrote the manuscript with input from all co-authors.

Acknowledgments

We thank Ira Blitz and Arul Subramanian at University of California, Irvine for sharing the *pCasX* plasmid, the personnel of the Electron Microscope Lab at University of California, Berkeley for assistance with SEM, and three reviewers for insightful comments on the manuscript. This work was supported by NSF grant IOS-1257379 to N.H.P.

References

1. Brusca, R. C., and Brusca, G. J. (2003). Invertebrates, Second, (Sunderland, MA: Sinauer Associates, Inc.).

2. Schram, F. R. (1986). Crustacea, (USA: Oxford University Press).

3. Panganiban, G., Sebring, A., Nagy, L., and Carroll, S. B. (1995). The development of crustacean limbs and the evolution of arthropods. Science *270*, 1363 – 1366.

4. Williams, T. A. (2003). The evolution and development of crustacean limbs: an analysis of limb homologies. In Evolutionary Developmental Biology of Crustacea, Crustacean Issues. (Boca Raton, FL: CRC Press), p. 220.

5. Giorgianni, M. W., and Patel, N. H. (2004). Patterning of the branched head appendages in *Schistocerca americana* and *Tribolium castaneum*. Evol. Dev. *6*, 402–410.

6. Gonzalez-Crespo, S., and Morata, G. (1996). Genetic evidence for the subdivision of the arthropod limb into coxopodite and telopodite. Development *122*, 3921–3928.

7. Jockusch, E., Williams, T., and Nagy, L. (2004). The evolution of patterning of serially homologous appendages in insects. Dev. Genes Evol. *214*, 324–338.

8. Kojima, T. (2004). The mechanism of *Drosophila* leg development along the proximodistal axis. Dev. Growth Differ. *46*, 115–129.

9. Prpic, N.-M., and Telford, M. (2008). Expression of *homothorax* and *extradenticle* mRNA in the legs of the crustacean *Parhyale hawaiensis*: evidence for a reversal of gene expression regulation in the pancrustacean lineage. Dev. Genes Evol. *218*, 333–339.

10. Williams, T. A., Nulsen, C., and Nagy, L. M. (2002). A Complex Role for Distal-less in Crustacean Appendage Development. Dev. Biol. *241*, 302–312.

11. Hughes, C. L., and Kaufman, T. C. (2002). Hox genes and the evolution of the arthropod body plan. Evol Dev *4*, 459 – 499.

12. Pearson, J. C., Lemons, D., and McGinnis, W. (2005). Modulating Hox gene functions during animal body patterning. Nat Rev Genet *6*, 893–904.

13. Abzhanov, A., and Kaufman, T. C. (1999). Novel regulation of the homeotic gene *Scr* associated with a crustacean leg-to-maxilliped appendage transformation. Development *126*, 1121–1128.

14. Abzhanov, A., and Kaufman, T. C. (1999). Homeotic genes and the arthropod head: Expression patterns of the *labial*, *proboscipedia*, and *Deformed* genes in crustaceans and insects. Proc. Natl. Acad. Sci. *96*, 10224–10229.

15. Abzhanov, A., and Kaufman, T. C. (2000). Crustacean (malacostracan) Hox genes and the evolution of the arthropod trunk. Development *127*, 2239–2249.

16. Abzhanov, A., and Kaufman, T. C. (2000). Embryonic expression patterns of the Hox genes of the crayĀish *Procambarus clarkii* (Crustacea, Decapoda). Evol. Dev. *2*, 271–283.

17. Averof, M., and Patel, N. H. (1997). Crustacean appendage evolution associated with changes in Hox gene expression. Nature *388*, 682–686.

18. Deutsch, J. S., and Mouchel-Vielh, E. (2003). Hox genes and the crustacean body plan. BioEssays *25*, 878–887.

19. Schram, F. R., and Koenemann, S. (2001). Developmental genetics and arthropod evolution: part I, on legs. Evol. Dev. *3*, 343–354.

20. Liubicich, D. M. et al. (2009). Knockdown of Parhyale Ultrabithorax recapitulates evolutionary changes in crustacean appendage morphology. Proc. Natl. Acad. Sci. *106*, 13892–13896.

21. Pavlopoulos, A., Kontarakis, Z., Liubicich, D. M., Serano, J. M., Akam, M., Patel, N. H., and Averof, M. (2009). Probing the evolution of appendage specialization by Hox gene misexpression in an emerging model crustacean. Proc. Natl. Acad. Sci. *106*, 13897–13902.

22. Kato, Y., Shiga, Y., Kobayashi, K., Tokishita, S., Yamagata, H., Iguchi, T., and Watanabe, H. (2011). Development of an RNA interference method in the cladoceran crustacean *Daphnia magna.* Dev. Genes Evol. *220*, 337–345.

23. Copf, T., Rabet, N., and Averof, M. (2006). Knockdown of spalt function by RNAi causes de-repression of Hox genes and homeotic transformations in the crustacean *Artemia franciscana*. Dev. Biol. *298*, 87–94.

24. Gerberding, M., Browne, W. E., and Patel, N. H. (2002). Cell lineage analysis of the amphipod crustacean *Parhyale hawaiensis* reveals an early restriction of cell fates. Development *129*, 5789–5801.

25. Serano, J. M., Martin, A., Liubicich, D. M., Jarvis, E., Bruce, H. S., La, K., Browne, W. E., Grimwood, J., and Patel, N. H. (in press). Comprehensive analysis of Hox gene expression in the amphipod crustacean *Parhyale hawaiensis*. Dev Biol.

26. Lee, Y. S., and Carthew, R. W. (2003). Making a better RNAi vector for *Drosophila*: use of intron spacers. Methods *30*, 322–329.

27. Angelini, D. R., and Kaufman, T. C. (2005). Functional analyses in the milkweed bug *Oncopeltus fasciatus* (Hemiptera) support a role for Wnt signaling in body segmentation but not appendage development. Dev Biol *283*, 409 – 423.

28. Brown, S. J., Shippy, T. D., Beeman, R. W., and Denell, R. E. (2002). *Tribolium* Hox genes repress antennal development in the gnathos and trunk. Mol. Phylogenet. Evol. *24*, 384–387.

29. DeCamillis, M. A., Lewis, D. L., Brown, S. J., Beeman, R. W., and Denell, R. E. (2001). Interactions of the *Tribolium Sex combs reduced* and *proboscipedia* Orthologs in Embryonic Labial Development. Genetics *159*, 1643–1648.

30. Hrycaj, S., Chesebro, J., and Popadić, A. (2010). Functional analysis of Scr during embryonic and post-embryonic development in the cockroach, Periplaneta americana. Spec. Sect. Morphog. *341*, 324–334.

31. Smith, F. W., and Jockusch, E. L. (2014). Hox genes require homothorax and extradenticle for body wall identity specification but not for appendage identity specification during metamorphosis of *Tribolium castaneum*. Dev. Biol. *395*, 182–197.

32. Gilles, A., and Averof, M. (2014). Functional genetics for all: engineered nucleases, CRISPR and the gene editing revolution. EvoDevo *5*, 43.

33. Yen, S.-T., Zhang, M., Deng, J. M., Usman, S. J., Smith, C. N., Parker-Thornburg, J., Swinton, P. G., Martin, J. F., and Behringer, R. R. (2014). Somatic mosaicism and allele complexity induced by CRISPR/Cas9 RNA injections in mouse zygotes. Dev. Biol. *393*, 3–9.

34. Nakanishi, T., Kato, Y., Matsuura, T., and Watanabe, H. (2014). CRISPR/Cas-Mediated Targeted Mutagenesis in *Daphnia magna*. PLoS ONE *9*, e98363.

35. Shiga, Y., Sagawa, K., Takai, R., Sakaguchi, H., Yamagata, H., and Hayashi, S. (2006). Transcriptional readthrough of Hox genes Ubx and Antp and their divergent post-transcriptional control during crustacean evolution. Evol. Dev. *8*, 407–414.

36. Averof, M., and Akam, M. (1995). Hox genes and the diversiĀicaȚon of insect and crustacean body plans. Nature *376*, 420–423.

37. Kelsh, R., Weinzierl, R. O., White, R. A., and Akam, M. (1994). Homeotic gene expression in the locust Schistocerca: an antibody that detects conserved epitopes in Ultrabithorax and abdominal-A proteins. Dev Genet *15*, 19 – 31.

38. Durston, A. J. (2012). Global posterior prevalence is unique to vertebrates: A dance to the music of time? Dev. Dyn. *241*, 1799–1807.

39. Akam, M. (1998). Hox genes, homeosis and the evolution of segment identity: no need for hopeless monsters. Int J Dev Biol *42*, 445–451.

40. Averof, M., Pavlopoulos, A., and Kontarakis, Z. (2010). Evolution of new appendage types by gradual changes in Hox gene expression – the case of crustacean maxillipeds. Paleodiversity *3*, 141–145.

41. Abzhanov, A., Holtzman, S., and Kaufman, T. C. (2001). The *Drosophila* proboscis is specified by two Hox genes, *proboscipedia* and *Sex combs reduced*, via repression of leg and antennal appendage genes. Development *128*, 2803–2814.

42. Angelini, D. R., and Kaufman, T. C. (2004). Functional analyses in the hemipteran *Oncopeltus fasciatus* reveal conserved and derived aspects of appendage patterning in insects. Dev. Biol. *271*, 306–321.

43. Ebner, A., Cabernard, C., Affolter, M., and Merabet, S. (2005). Recognition of distinct target sites by a unique Labial/Extradenticle/Homothorax complex. Development *132*, 1591–1600.

44. Mann, R. S., Lelli, K. M., and Joshi, R. (2009). Chapter 3 Hox Specificity: Unique Roles for Cofactors and Collaborators. In Current Topics in Developmental Biology, Olivier Pourquié, ed. (Academic Press), pp. 63–101.

45. Merabet, S., and Hudry, B. (2013). Hox transcriptional specificity despite a single class of cofactors: Are Āleȣible interaction modes the key? BioEssays *35*, 88–92.

46. Mito, T., Ronco, M., Uda, T., Nakamura, T., Ohuchi, H., and Noji, S. (2008). Divergent and conserved roles of extradenticle in body segmentation and appendage formation, respectively, in the cricket *Gryllus bimaculatus*. Dev. Biol. *313*, 67–79.

47. Rivas, M. L., Espinosa-Vázquez, J. M., Sambrani, N., Greig, S., Merabet, S., Graba, Y., and Castelli-Gair Hombría, J. (2013). Antagonism Versus Cooperativity with TALE Cofactors at the Base of the Functional DiversiĀicaȣon of Hox Protein Function. PLoS Genet *9*, e1003252.

48. Ronco, M., Uda, T., Mito, T., Minelli, A., Noji, S., and Klingler, M. (2008). Antenna and all gnathal appendages are similarly transformed by homothorax knock-down in the cricket *Gryllus bimaculatus*. Dev. Biol. *313*, 80–92.

49. Sambrani, N., Hudry, B., Maurel-Zaffran, C., Zouaz, A., Mishra, R., Merabet, S., and Graba, Y. (2013). Distinct Molecular Strategies for Hox-Mediated Limb Suppression in *Drosophila*: From Cooperativity to Dispensability/Antagonism in TALE Partnership. PLoS Genet *9*, e1003307.

50. Slattery, M. et al. (2011). Cofactor Binding Evokes Latent Differences in DNA Binding Specificity between Hox Proteins. Cell *147*, 1270–1282.

51. Bassett, A., and Liu, J.-L. (2014). CRISPR/Cas9 mediated genome engineering in Drosophila. Methods *69*, 128–136.

52. Blitz, I. L., Biesinger, J., Xie, X., and Cho, K. W. Y. (2013). Biallelic genome modification in F0 Xenopus tropicalis embryos using the CRISPR/Cas system. Genesis *51*, 827–834.

53. Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J.-R. J., and Joung, J. K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotech *31*, 227–229.

54. Kontarakis, Z., and Pavlopoulos, A. (2014). Transgenesis in Non-model Organisms: The Case of *Parhyale*. In Hox Genes, Methods in Molecular Biology., Y. Graba, R. Rezsohazy, eds. (Springer New York), pp. 145–181.

55. Pavlopoulos, A., and Averof, M. (2005). Establishing genetic transformation for comparative developmental studies in the crustacean *Parhyale hawaiensis*. Proc. Natl. Acad. Sci. *102*, 7888–7893.

56. Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A., and Patel, N. H. (2009). Injection of *Parhyale hawaiensis* Blastomeres with Fluorescently Labeled Tracers. Cold Spring Harb. Protoc. *2009*, pdb.prot5128.

57. Spears, T., DeBry, R. W., Abele, L. G., Chodyla, K., and Boyko, C. B. (2005). Peracarid monophyly and interordinal phylogeny inferred from nuclear small-subunit ribosomal DNA sequences (Crustacea: Malacostraca: Peracarida). Proc. Biol. Soc. Wash. *118*, 117–157.

Supplemental Information

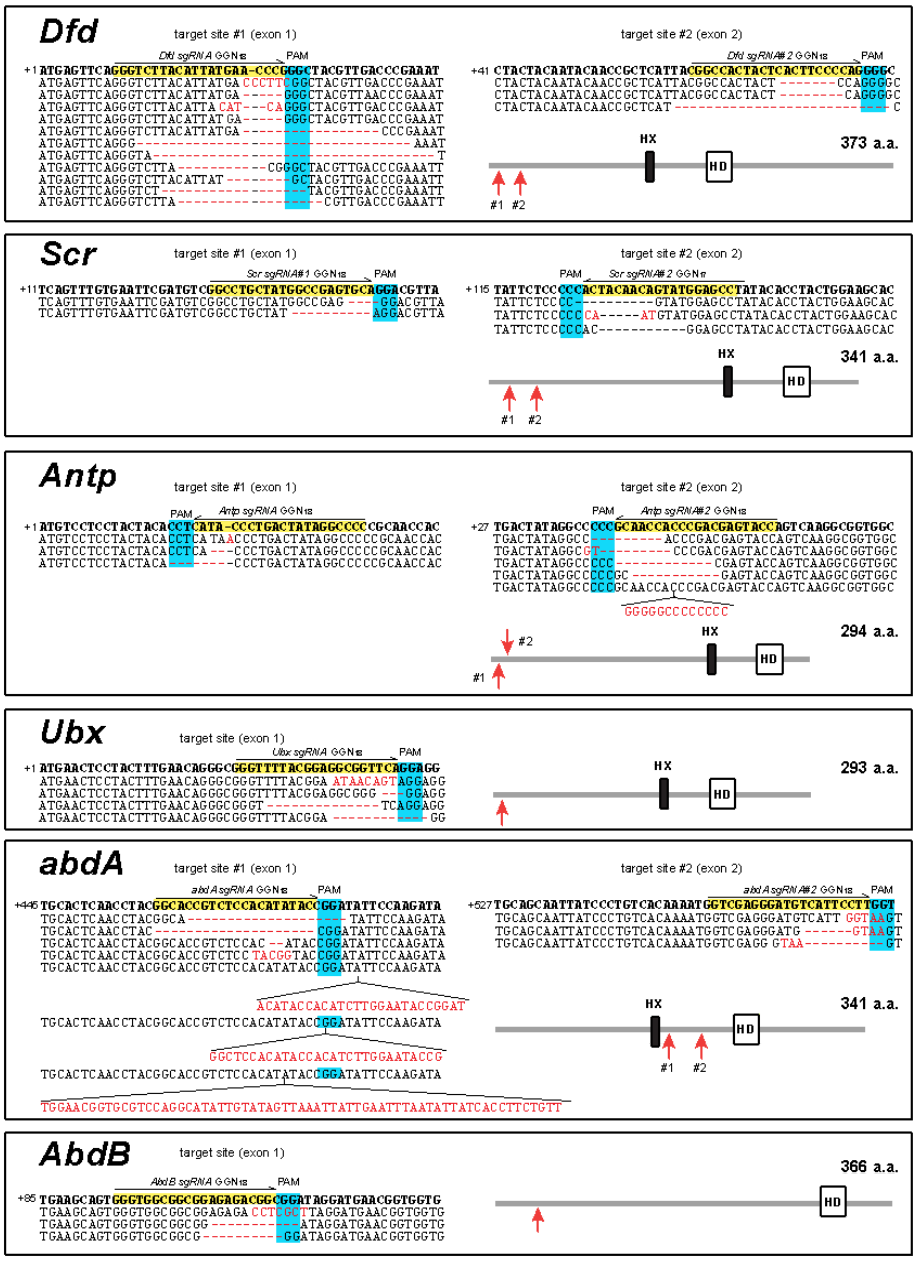


Figure S1. Detection of CRISPR-induced Hox mutant alleles in G0 individuals – related to Figure 4.2. Sequences at target sites obtained from single whole hatchlings after injection of a single CRISPR sgRNA (in yellow) reveal multiple DNA lesions around the [NGG] protospacer adjacent motif (PAM, in blue), including insertions or deletions expected to result in coding frame shifts (indels in red; reference wild-type allele in bold). Sequence coordinates indicate the position of the alignments relative to the start codon. All the sequences obtained outside the range of the provided alignments were identical to the wild-type reference (in bold). Red arrows: target sites on the protein sequence (grey bar); HX: hexapeptide motif; HD: Homeodomain.

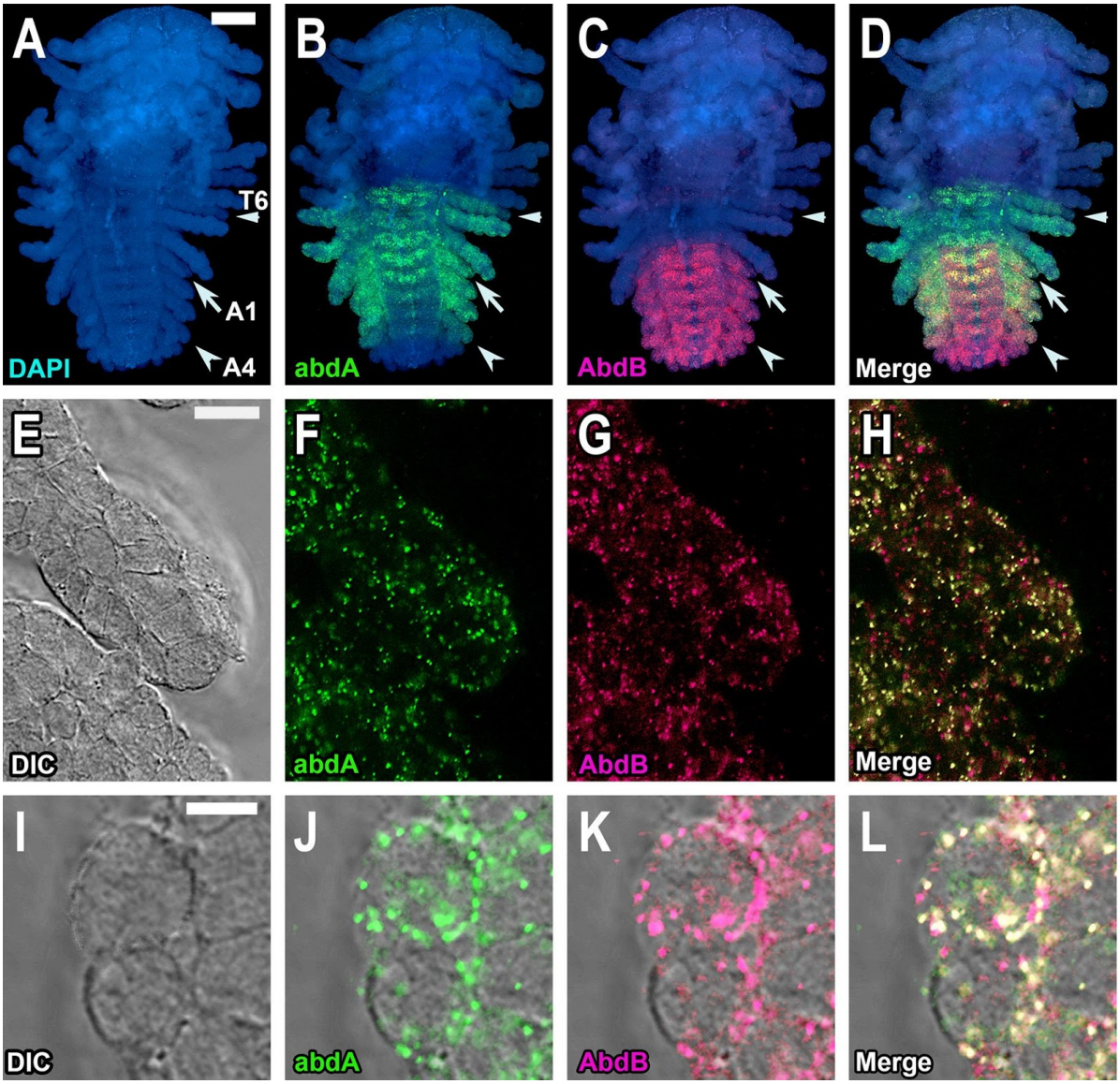


Figure S4.2. *abd-A / Abd-B* double *in situ* hybridization – related to Figure 4.4.
Stage 22 *Parhyale* embryo showing the distribution of *abd-A* (green) and *Abd-B* (red) transcripts. **A-D**. View of entire embryo (extended focus). **E-H**. Image of the of the A2 limb (2 µm optical z-section). **I-L**. Image of two ectodermal cells in the A2 limb (2 µm optical z-section). **A**. DAPI staining. **E and I**. DIC image. Merge in **H** does not include the DIC image. **A-D**. *abd-A* transcripts extend from T6 to A3 and weakly in A4. *Abd-B* transcripts extend from A1 to the end of the abdomen. **E-F**. Throughout their area of overlap in the limbs of A1-A4, both *abd-A* and *Abd-B* transcripts appear to be in all ectodermal cells, as shown here in the A2 limb. **I-L**. Image of two ectodermal cells in A2 limb clearly showing that both genes are expressed in both cells. Triangle, T6 limb; arrow, A1 limb; arrowhead, A4 limb. Scale bars 100µm **A-D**, 15µm **E-H**, 5µm **I-L**.

Supplemental Table

Target gene	Target specific reagent	Cas9 mRNA	Cas9 protein	No. injected	No. hatched	Survival rate	No. phenotype	Penetrance	Minimum criterion
Deformed (*Dfd*)	***Dfd* CRISPR *sgRNA#1***	X		139	92	66%	64	**70%**	Md or Mx1 endopod protrusion
	Dfd* CRISPR *sgRNA#2	X		104	50	48%	29	**58%**	
Sex combs reduced (*Scr*)	***Scr* CRISPR *sgRNA#1***	X		144	88	61%	47	**53%**	Mx2 endopod protrusion
	Scr* CRISPR *sgRNA#2		X	195	79	41%	36	**46%**	
	PhScr_72 siRNA			not recorded	84	-	0	0%	
	PhScr_123 siRNA			not recorded	29	-	10	34%	
	PhScr_297 siRNA			not recorded	104	-	6	6%	
	PhScr pool (3 siRNAs)			not recorded	143	-	26	18%	
Antennapedia (*Antp*)	***Antp* CRISPR *sgRNA#1***	X		276	207	75%	128	**62%**	Mouth dislocation / asymmetric Mxp
	Antp* CRISPR *sgRNA#2	X		132	82	62%	31	**38%**	
Ultrabithorax (*Ubx*)	***Ubx* CRISPR *sgRNA***	X		130	90	69%	24	**27%**	Short T2 with endites
	Ubx* CRISPR *sgRNA		X	271	192	71%	67	**35%**	
abdominal-A (*abdA*)	***abdA* CRISPR *sgRNA#1***	X		181	110	61%	47	**43%**	Deformed pleopod lacking setae
	abdA* CRISPR *sgRNA#2	X		158	110	70%	13	**12%**	
	PhabdA pool (4 siRNAs)			135	54	40%	6	11%	
	PhabdA-86 siRNA			130	95	73%	9	9%	
	PhabdA-306 siRNA			59	43	73%	11	26%	
	PhabdA-358 siRNA			46	17	37%	4	24%	
	PhabdA-459 siRNA			46	37	80%	7	19%	
Abdominal-B (*AbdB*)	***AbdB* CRISPR *sgRNA***	X		251	133	53%	59	**44%**	Uniramous abdominal appendage
	PhAbdB pool (4 siRNAs)			45	34	76%	5	15%	
	PhAbdB-277 siRNA			81	68	84%	12	18%	
	PhAbdB-357 siRNA			80	65	81%	7	11%	
	PhAbdB-520 siRNA			35	26	74%	0	0%	
	PhAbdB-649 siRNA			55	46	84%	2	4%	

Table S4.1. Summary of CRISPR and RNAi zygotic injection results in *P. hawaiensis* – related to Figure 4.7.
Note: *Scr siRNA* experiments had a survival rate similar to other injections.

Supplemental Experimental Procedures

Animals

Parhyale used in the experiments were from the inbred Chicago-F lab strain, which was established as an iso-female line approximately ĀifŁen years ago. Females of the parthenogenetic *Procambarus fallax* f. *virginalis* ("marbled crayĀisłł" or "marmorkrebs") originated from a commercial supplier and were maintained in artiĀicial pond water. *Mysidium colombiae* embryos were collected and fixed in Belize in 1994.

Preparation of Cas9 mRNA

We used the *pCasX* plasmid as a template for T7 *in vitro* transcription of *Cas9* mRNA [S1]. This vector contains a T7 promoter, the coding sequence of a human-codon optimized version of *Streptococcus pyogenes*, a C-terminal nuclear localization signal, and SV40 sequences for transcript stabilization [S1]. This vector was linearized with *Acc65I*, puriĀied by salt/EtOH precipitation, and resuspended in H_2O at a concentration of 500ng/µL. This linear template was then used for in vitro transcription using the mMessage mMachine T7 Kit (Thermo Fischer ScientiĀic), poly-adenylated, puriĀied using the RNeasy mini Kit (Qiagen), and frozen as 10µg/µL aliquots. A published protocol provides step-by-step guidelines for this procedure [S2].

CRISPR sgRNA design

Hox gene transcript sequences [S3] served as a template for the design of CRISPR sgRNAs. We used the ZiFiT Targeter webtool [S4] to scan for $GGN_{17\text{-}18}$[NGG] motifs in Āirstor second exons. The 5'GG dinucleotide terminates the T7 promoter and may maximize transcript yield, but it has been suggested that T7-transcribed sgRNAs with sequence-specific nucleotides are also viable [S5]. The resulting $GGN_{17\text{-}18}$ were synthetized by PCR assembly with the common *CRISPR sgR* primer [S2]. Incorporating the $GGN_{17\text{-}18}$ target-specific sequences for each sgRNA (underlined below), the oligonucleotides used in this study were:

***CRISPR F_Dfd#1* (for *Dfd* CRISPR *sgRNA#1*):**
5'GAAATTAATACGACTCACTATAGGGTCTTACATTATGAACCCGTTTTAGAGCTAGAAATAGC3'

***CRISPR F_Dfd#2* (for *Dfd* CRISPR *sgRNA#2*):**
5'GAAATTAATACGACTCACTATAGGCCACTACTCACTTCCCCAGTTTTAGAGCTAGAAATAGC3'

***CRISPR F_Scr#1* (for *Scr* CRISPR *sgRNA#1*):**
5'GAAATTAATACGACTCACTATAGGCCTGCTATGGCCGAGTGCGTTTTAGAGCTAGAAATAGC3'

***CRISPR F_Scr#2* (for *Scr* CRISPR *sgRNA#2*):**
5'GAAATTAATACGACTCACTATAGGCTCCATACTGTTGTAGTGTTTTAGAGCTAGAAATAGC3'

***CRISPR F_Antp#1* (for *Antp* CRISPR *sgRNA#1*):**
5'GAAATTAATACGACTCACTATAGGGCCTATAGTCAGGGTATGGTTTTAGAGCTAGAAATAGC3'

***CRISPR F_Antp#2* (for *Antp* CRISPR *sgRNA#2*):**
5'GAAATTAATACGACTCACTATAGGTACTCGTCGGGTGGTTGCGTTTTAGAGCTAGAAATAGC3'

***CRISPR F_Ubx#1* (for *Ubx* CRISPR *sgRNA#1*):**
5'GAAATTAATACGACTCACTATAGGGTTTTACGGAGGCGGTTCGTTTTAGAGCTAGAAATAGC3'

***CRISPR F_Ubx#2* (for *Ubx* CRISPR *sgRNA#2*):**
5'GAAATTAATACGACTCACTATAGGAAATGGATTACGGCGACGGTTTTAGAGCTAGAAATAGC3'

***CRISPR F_abd-A#1* (for *abd-A* CRISPR *sgRNA#1*):**
5'GAAATTAATACGACTCACTATAGGCACCGTCTCCACATATACGTTTTAGAGCTAGAAATAGC3'

***CRISPR F_abd-A#2* (for *abd-A* CRISPR *sgRNA#2*):**
5'GAAATTAATACGACTCACTATAGGTCGAGGGATGTCATTCCTGTTTTAGAGCTAGAAATAGC3'

***CRISPR F_Abd-B* (for *Abd-B* CRISPR *sgRNA*):**
5'GAAATTAATACGACTCACTATAGGGTGGCGGCGGAGAGACGGGTTTTAGAGCTAGAAATAGC3'

***CRISPR sgR* (universal):**
5'AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTA
GCTCTAAAAC3'

CRISPR injection mixes

PCR concatenates of the CRISPR F / CRISPR sgR oligonucleotides were assembled with the Phusion High-Fidelity DNA polymerase (New England Biolabs), puriĀied with the PCR puriĀicationkit (Qiagen), and used for *in vitro* transcription using the MEGAshortscript T7 Transcription (Thermo Fischer ScientiĀic).The resulting short RNA products were puriĀied with the MEGAclear Transcription Clean-Up kits (Thermo Fischer ScientiĀic),mixed with *Cas9* mRNA were mixed in a 1:2 molar ratio (10µg *Cas9* mRNA: 0.5µg sgRNA), puriĀiedby salt/EtOH precipitation, and suspended in water. For protein-based injections of Cas9, we used mixes containing 333ng/µL of commercial Cas9 (PNABio, catalog number CP01) re-suspended in water, 200ng/µL sgRNA, and Phenol Red at a 0.05% Āinalconcentration for injection visualization.

RNA interference

Stealth™ siRNA duplexes were designed using the BLOCK-IT™ RNAi Designer tool (Thermo Fischer ScientiĀic).We targeted non-conserved coding regions of the *Parhyale Scr*, *abd-A* and *Abd-B* transcripts that were shared by all previously isolated splicing isoforms [S3]. siRNAs were resuspended in water and mixed with Phenol Red (0.05% Āinal concentration). The siRNA sequences (forward strand) are as follows:

***PhScr_72*:** 5'CGCAGAUUAUUAUGGCGCUGCUAUG3'
***PhScr_123*:** 5'CCCACUACAACAGUAUGGAGCCUAU3'
***PhScr_297*:** 5'UCGACUUCACCAACCCUCAUCUUCU3'
Phabd-A-80: 5'GCCAGAACAGAAGUAUGUAUCCAUA3'
***Phabd-A-306*:** 5'CACUAAUCCUUUGAAUCCGUGCACU3'
***Phabd-A-358*:** 5'CCGGAUAUUCCAAGAUACCCUUGGA3'
***Phabd-A-459*:** 5'CAUAGAGAACCAAUGGCGUGGAUUA3'
***PhAbd-B-277*:** 5'GGUGUAAUUCGUCAUUCCCAUUCAA3'
***PhAbd-B-357*:** 5'CCAGUACAGUCAAGCCUUCUGCAGA3'
***PhAbd-B-520*:** 5'GCUGCCCAAACUUGGUGCAAUUAUG3'
***PhAbd-B-649*:** 5'CCACACGAUGGAUAUCUCAGGAACU3'

The *abd-A-wiz* construct was generated by directional cloning of 544nt *abd-A* fragments into a modiĀiedversion of the *pWIZ* vector [S6]. The resulting construct, consisting of the *abd-A* fragment in opposite orientations on either side of the *Drosophila white* intron, was then placed downstream of the *Parhyale hsp70* promoter within a *pMinos* transformation vector, and germline transformation was carried out as previously described [S7, S8]. Embryos from females that carried the *hsp70-abd-A-wiz* transgene were isolated and subjected to daily, one-hour-long heat shocks beginning around stage 14 (*i.e.* just before the onset of endogenous *abd-A* expression) and continuing until hatching. Two independent transgenic lines were established, and both yielded similar results.

Injections and imaging

We micro-injected approximately 40-60 picoliters of CRISPR injection mixes or 20-40 picoliters of 200µM siRNAs in zygotes or two-cell embryos using a published procedure [S9]. To examine limb phenotypes, *Parhyale* hatchlings were Āixed in 3.7% formaldehyde, appendages were removed individually, mounted in 70% glycerol, viewed with darkĀield and Nomarski optics on a Zeiss Axio Imager.A1 microscope, and photographed using ProgRes C14Plus camera and software (Jenoptik). For scanning electron microscopy (SEM), *Parhyale* and *Procambarus* hatchlings were Āixed 2hrs in 3.7% formaldehyde, dehydrated via an ethanol series, and stored in 90% ethanol. Afterwards, they were washed several times in 100% ethanol prior to critical point drying, and examined on a Hitachi TM-1000 microscope. SEM images were false-colored using the "Darken" and "Soft Light" layer functions of Adobe Photoshop.

Mutant allele sequencing

We extracted the DNA of single hatchlings that showed a CRISPR-induced phenotype using the DNeasy Blood and Tissue Kit (Qiagen), performed PCR ampliĀicationof the targeted sites, cloned the resulting product using the pGEM-T Easy Vector system, and sequenced 16 positive clones per individual with the M13R universal primer. All individuals yielded redundant sequences consisting of a mix of wild-type and mutant alleles. The primers used for PCR amplifications were:

***PhDfd_exon1F*:** 5'CTGAAGTGCACGCAAAAGAG3'
***PhDfd_exon1R*:** 5'GCACTGTGCAAGTTGAGGAC3'
***PhScr_exon1F*:** 5'CGTGCAATGTTTTGAAGCAA3'
***PhScr_exon1R*:** 5'AAGGCTGTGCTTCCAGTAGG3'
***PhAntp_exon1F*:** 5'TCCAATCCAGCCGTGAGTTA3'
***PhAntp_exon1R*:** 5'CATGCAAGTAGGAGCTGCTG3'
***PhUbx_exon1F*:** 5'GTGCATAGTGGTGATTCAGTGT3'
***PhUbx_exon1R*:** 5'GTTGGTGGTGTTGTTGGTGT3'
***Phabd-A_exon1F*:** 5'CGTCTGCGGATAGTGACAAA3'
***Phabd-A_exon1R*:** 5'GTGATGGACATCCAAGGGTA3'
***Phabd-A_exon2F*:** 5'ACAAGCAGGTTTTGGGAGTG3'
***Phabd-A_exon2R*:** 5'ATTGGGTGAGTGCAAAGCTG3'
***PhAbd-B_exon1F*:** 5'GTCCTCCTGCTCCTCGATC3'
***PhAbd-B_exon1R*:** 5'CTCTTCGTATACAGCGCCAC3'

Immunofluorescence and fluorescent in situ hybridization

Antibody staining of *Parhyale*, *Procambarus* and *Mysidium* were performed as described previously [S10–S12]. The following antibodies were used: monoclonal mouse anti-Antp antibody 8C11 (1:30-1:50 dilution) [S13], polyclonal rat anti-Ubx (1:4000-1:6000) [S14], and monoclonal mouse anti-Ubx/abd-A FP6.87 (1:30-1:50) [S15]. Primary antibodies were

detected with AlexaFluor 555 conjugated anti-mouse and anti-rat secondary antibodies (Thermo Fischer Scientific) (1:500).

In-situ probes for *abd-A* and *Abd-B* are as described in [S3] and were produced with either Digoxigenin (DIG)- (Roche) or Dinitrophenol (DNP)- (Perkin-Elmer) conjugated UTPs. Hybridization was performed as described in [S16]. Transcript visualization was achieved using a combination of sheep anti-DIG (Roche) and rabbit anti-DNP (Thermo Fischer ScientiĀic) primary antibodies, followed by donkey anti-Sheep AlexaFluor 555 (Thermo Fischer ScientiĀic) and Donkey anti-Rabbit AlexaFluor 488 or 647 (Jackson ImmunoResearch) secondary antibodies [S17]. Embryos were imaged on an LSM 700 or LSM 780 laser scanning confocal microscope (Zeiss) and processed using Volocity software (Perkin-Elmer).

Supplemental References

S1. Blitz, I. L., Biesinger, J., Xie, X., and Cho, K. W. Y. (2013). Biallelic genome modification in F0 *Xenopus tropicalis* embryos using the CRISPR/Cas system. Genesis *51*, 827–834.

S2. Bassett, A., and Liu, J.-L. (2014). CRISPR/Cas9 mediated genome engineering in Drosophila. Methods *69*, 128–136.

S3. Serano, J. M., Martin, A., Liubicich, D. M., Jarvis, E., Bruce, H. S., La, K., Browne, W. E., Grimwood, J., and Patel, N. H. (in press). Comprehensive analysis of Hox gene expression in the amphipod crustacean *Parhyale hawaiensis*. Dev Biol.

S4. Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J.-R. J., and Joung, J. K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotech *31*, 227–229.

S5. Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., and Lim, W. A. (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell *152*, 1173–1183.

S6. Lee, Y. S., and Carthew, R. W. (2003). Making a better RNAi vector for *Drosophila*: use of intron spacers. Methods *30*, 322–329.

S7. Kontarakis, Z., and Pavlopoulos, A. (2014). Transgenesis in Non-model Organisms: The Case of *Parhyale*. In Hox Genes, Methods in Molecular Biology., Y. Graba, R. Rezsohazy, eds. (Springer New York), pp. 145–181.

S8. Pavlopoulos, A., and Averof, M. (2005). Establishing genetic transformation for comparative developmental studies in the crustacean *Parhyale hawaiensis*. Proc. Natl. Acad. Sci. *102*, 7888–7893.

S9. Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A., and Patel, N. H. (2009). Injection of *Parhyale hawaiensis* Blastomeres with Fluorescently Labeled Tracers. Cold Spring Harb. Protoc. *2009*, pdb.prot5128.

S10. Patel, N. H. (1994). Imaging neuronal subsets and other cell types in whole-mount *Drosophila* embryos and larvae using antibody probes. Methods Cell Biol. *44*, 445–487.

S11. Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A., and Patel, N. H. (2009). Fixation and Dissection of *Parhyale hawaiensis* Embryos. Cold Spring Harb. Protoc. *2009*, pdb.prot5127.

S12. Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A., and Patel, N. H. (2009). Antibody Staining of *Parhyale hawaiensis* Embryos. Cold Spring Harb. Protoc. *2009*, pdb.prot5129.

S13. Condie, J. M., Mustard, J. A., and Brower, D. L. (1991). Generation of anti-Antennapedia monoclonal antibodies and Antennapedia protein expression in imaginal discs. Dros Info Serv *70*, 52 – 54.

S14. Liubicich, D. M. et al. (2009). Knockdown of Parhyale Ultrabithorax recapitulates evolutionary changes in crustacean appendage morphology. Proc. Natl. Acad. Sci. *106*, 13892–13896.

S15. Kelsh, R., Weinzierl, R. O., White, R. A., and Akam, M. (1994). Homeotic gene expression in the locust *Schistocerca*: an antibody that detects conserved epitopes in Ultrabithorax and abdominal-A proteins. Dev Genet *15*, 19 – 31.

S16. Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A., and Patel, N. H. (2009). In situ hybridization of labeled RNA probes to fixed *Parhyale hawaiensis* embryos. Cold Spring Harb Protoc *2009*, pdb.prot5130.

S17. Ronshaugen, M., and Levine, M. (2004). Visualization of trans-homolog enhancer-promoter interactions at the Abd-B Hox locus in the *Drosophila* embryo. Dev. Cell *7*, 925–932.

Chapter 5. Exploring the interactions among Hox genes in the establishment of diverse appendage types using CRISPR/Cas9 knockout of multiple Hox genes simultaneously

Combinatorial interactions of Hox genes establish appendage diversity of the amphipod crustacean *Parhyale hawaiensis*

The following article is intended to be published with E. Jarvis as the primary author

Keywords

Ultrabithorax, abdominal-A, Abdominal-B, Hox Combinatorial model, posterior prevalence, CRISPR/Cas9 mutagenesis, Appendage specification

Summary Statement

Contrasting the canonical model of posterior prevalence, *Parhyale* limbs are specified by a combinatorial "Hox code". Changes to neighboring Hox domains after single, double, and triple Hox gene knockout reveal cross-regulatory interactions.

Abstract

Homeotic transformations are a striking demonstration of how regional identity in animals is encoded in their genomes. The mutation of a single Hox gene has the capacity to mediate transformation of appendage identity in arthropods. These transformations exhibit a pattern that has led to three models of how Hox genes shape morphology. The interactions between Hox genes in patterning has only been described in limited number of organisms. Here we focus on the interactions of three Hox gene, *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*), and *Abdominal-B* (*Abd-B*) in patterning thoracic and abdominal appendages in the crustacean *Parhyale hawaiensis*. The patterning function of these genes in *Drosophila* is the basis for the, posterior prevalence model of Hox function that has proven to be predictive for understanding many Hox gene interactions that shape *Drosophila* development. Studies of Hox function in *Parhyale*, however, have revealed two striking phenotypes which differed markedly from what this model would predict. Knockout of *abd-A* and *Abd-B* result in

homeotic transformations that are not consistent. We probe the logic of the *Parhyale* transformation by using CRISPR/Cas9 to systematically examine all combinations of *Ubx*, *abd-A*, and *Abd-B* loss of function in *Parhyale*, and conduct comprehensive post-knockout expression analysis of other Hox genes to identify cross-regulatory interactions to establish a model of how patterns of Hox expression that underlie the transformed limbs respond to Hox mutation. From this, we demonstrate that in *Parhyale* these genes function combinatorially, not dominantly, to specify limb identity. We find that *abd-A* has largely broken free from the rules of posterior prevalence. These results provide insight into how not only *Parhyale*, but Crustacea in general can specify such a diversity of limb identities, and how crustaceans may have been able to generate an especially rich pattern of body plan variation among species.

5.1 Introduction

The Hox family of homeodomain-encoding transcription factors specify regional identity along the head to tail axis in nearly all metazoan animals. Their role in regionalization of animal bodies is dramatically demonstrated by the striking phenotypes arising from homeotic mutations where structures in one part of the body are transformed into the likeness of another. In the fruit fly *Drosophila melanogaster* altering the function of a single Hox gene, such as in the *Antennapedia* gain-of-function (GOF) mutation results in near perfect transformation of antennae into legs or halteres to wings in the case of *Ultrabithorax* loss-of-function (LOF) mutants (McGinnis and Krumlauf 1992; Gehring et al. 1994; Graba et al. 1997; Mann and Morata 2000).

Hox genes are clustered in the genome and have a conserved organization of paralogs that is mirrored by their expression along the anterior-posterior (AP) axis. Hox gene expression domains have distinct anterior and posterior boundaries and, in general, the posterior expression of one Hox gene partially overlaps the anterior expression of the neighboring and more posteriorly expressed Hox gene in the cluster. This ordered regional expression of Hox genes is widely conserved and acts to specify different morphologies through the activation and repression of distinct sets of downstream targets (cite). This regionalization is reinforced by the refinement of Hox expression boundaries by neighboring Hox genes through cross-regulatory interactions and potentially through functional competition in the regions where the expression domains of multiple Hox genes overlap (Levine and Harding 1989; Hughes and Kaufman 2002).

The extensive conservation of Hox expression and regulatory interactions have led to a number of useful generalizations regarding the consequences of Hox mutation to be made from a limited number of well-characterized species (e.g., *Tribolium* and *Drosophila* (Denell 2008; Reviewed in Hughes and Kaufman 2002)). One key principle has been a clear AP directionality in the logic of morphological transformations. In arthropods, numerous studies of multiple Hox genes have found that LOF of a posteriorly expressed Hox gene results in homeotic transformation to the next-most anterior identity while GOF mutations

result in the opposite anterior to posterior transformation (Reviewed by Akam 1998). This phenomenon has been termed posterior prevalence and is related to the concept of phenotypic suppression where Hox genes more posteriorly expressed are phenotypically dominant to more anteriorly expressed Hox genes (Noro et al. 2011; Macías and Morata 1996).

The potential for variation in the expression and function of Hox genes to explain the evolution of change in animal body plans has made them an attractive developmental system to study in diverse organisms. Increasingly, new technologies have enabled the expansion of the scope of these investigations to include a much greater diversity of species and with much greater detail. One organism prime for expanding our current body of knowledge is the amphipod crustacean *Parhyale hawaiensis*, which boasts an abundance of diverse appendage types, relevant phylogenetic positioning, and a rapidly expanding collection of tools and genetic resources (Browne et al. 2005; Kao et al. 2016; Rehm et al. 2009). A complete analysis of Hox expression in *Parhyale* has revealed that each distinct appendage type expresses a unique set of Hox genes (Serano et al. 2016), suggesting that a unique "Hox code" may be responsible for each of the specific limb morphologies. The function of each Hox gene in the establishment of appendage identity was explored by examining the patterns of limb morphologies that resulted from the individual knockout (KO) of each using CRISPR/Cas9 (Martin et al. 2016). This study revealed a number of appendage transformations that appeared to violate the canonical model of posterior prevalence.

Of particular note, KO of *Abd-B*, which is expressed exclusively and throughout the *Parhyale* abdomen, resulted in a non-contiguous "skip" in the transformation of the posterior set of abdominal appendages (A4-A6) from short/spiky, biramous uropods toward a uniramous leg identity most similar to the forward-facing walking legs of the anterior thorax (T4-T5) (Figure 1B). Additionally, knockout of abd-A led to both anterior *and posterior* shifts of identity with the transformation of A1 - A3 swimming appendages to A4 - A6 uropods (Figure 1C). Although these patterns of transformation are not predicted by classic invertebrate Hox logic, they can be logically explained using a combinatorial "Hox code" model where the combined expression of multiple Hox genes would have combinatorial influence on downstream targets rather than the more posterior Hox gene phenotypically dominating the establishment of limb identity (Reviewed in Kmita and Duboule 2003). Similar patterns of homeotic transformations more consistent with a Hox code model than with posterior prevalence have been described. For example, mutation of the *Drosophila* Hox gene *Sex combs reduced* (*Scr*) in the Antennapedia Complex (ANTP-C) results in a bi-directional transformation similar in polarity to that observed in *Parhyale abd-A* KO (Duncan and Kaufman 1975). However, the non-linear skip in transformation observed in *Parhyale Abd-B* KO is, to the best of our knowledge, unprecedented and warrants further investigation.

In this study, we investigated these non-canonical transformations and the potential combinatorial nature of *Parhyale* Hox function using CRISPR/Cas9 to generate G_0 mutants

for each double KO permutation of the posterior Hox genes *Ubx, abd-A* and *Abd-B* (*Abd-B* + *abd-A* 2xKO, *abd-A* + *Ubx* 2xKO, and *Abd-B* + *Ubx* 2xKO) as well as the triple KO of all three (*Abd-B + abd-A + Ubx* 3xKO). In addition to characterizing the resulting phenotypic transformations, we performed comprehensive post-knockout expression analyses of Hox genes to establish the consequences of Hox KO underpinning the transformed identities and to identify possible cross-regulatory interactions among *Parhyale* Hox genes that might contribute to the observed appendage transformations.

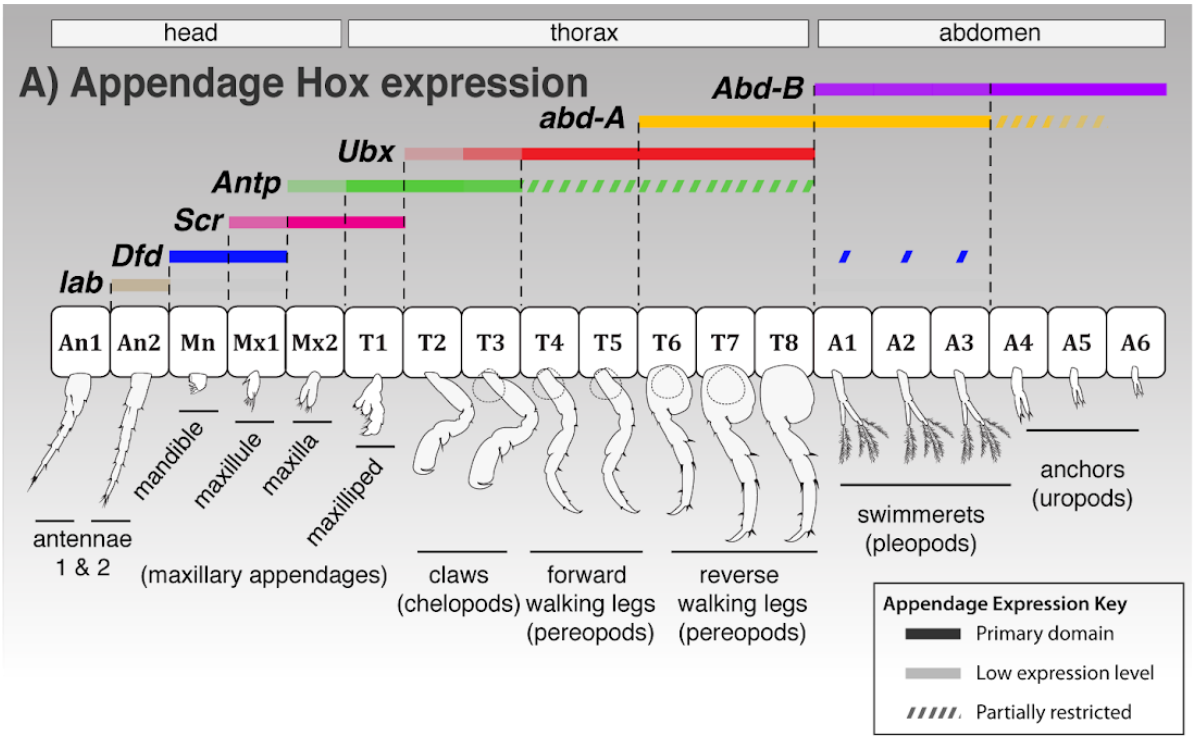


Fig 5.1. (A) Schematic representation of Hox gene expression in wildtype *Parhyale* appendages
Solid lines represent Hox expression throughout the appendage. Diagonally dashed lines represent the restricted expression of *abd-A* (anteriorly-restricted domain in A4 - A5 pleopods) and *Antp* (T4 - T8 legs do not express *Antp* in the superficial ectoderm). *Dfd* triangles in A1 - A3 represent a small patch of its expression in pleopods. Faded color represents weaker expression. Expression not shown for Hox genes *pb* and *Hox3*, which have very limited domains spanning An2 and Mn, and do not include the entire appendage.For pb and Hox3, expression occurs within very limited domains that does not include the entire appendage. Note that each unique combination of Hox genes corresponds to a distinct appendage type: An1 (no Hox expression), An2 (*lab*, not shown), Mn (*Dfd*), Mx1 (*Dfd + Scr*), Mx2 (*Scr +* low levels of *Antp*), T1 Mxp (*Scr* + *Antp*), T2 - T3 Claws (*Ubx* + *Antp*), T4 - T5 walking legs (*Ubx*), T6 - T8 jumping legs (*Ubx + abd-A)*, A1 - A3 pleopods (*abd-A + Abd-B*), and A4 - A6 uropods (*Abd-B*). Thoracic legs and abdominal appendages are patterned by the posterior Hox genes *Ubx*, *abd-A*, and *Abd-B*, corresponding to the *Drosophila* Bithorax complex. The expression domains of *Ubx* (expressed throughout the thorax excluding the first head-fused segment/maxillary appendage T1) and *Abd-B* (expressed throughout the abdominal) remain distinct from one another, while the expression domain of *abd-A* partially overlaps both, spanning the posterior thorax (where its anterior boundary delineates the T4 - T5 walking legs from the T6 - T8 jumping legs in the *Ubx*-expressing thorax) through anterior abdomen (where its posterior boundary delineates the A1 - A3 pleopods from the A4 - A6 uropods in the *Abd-B*-expressing abdomen.

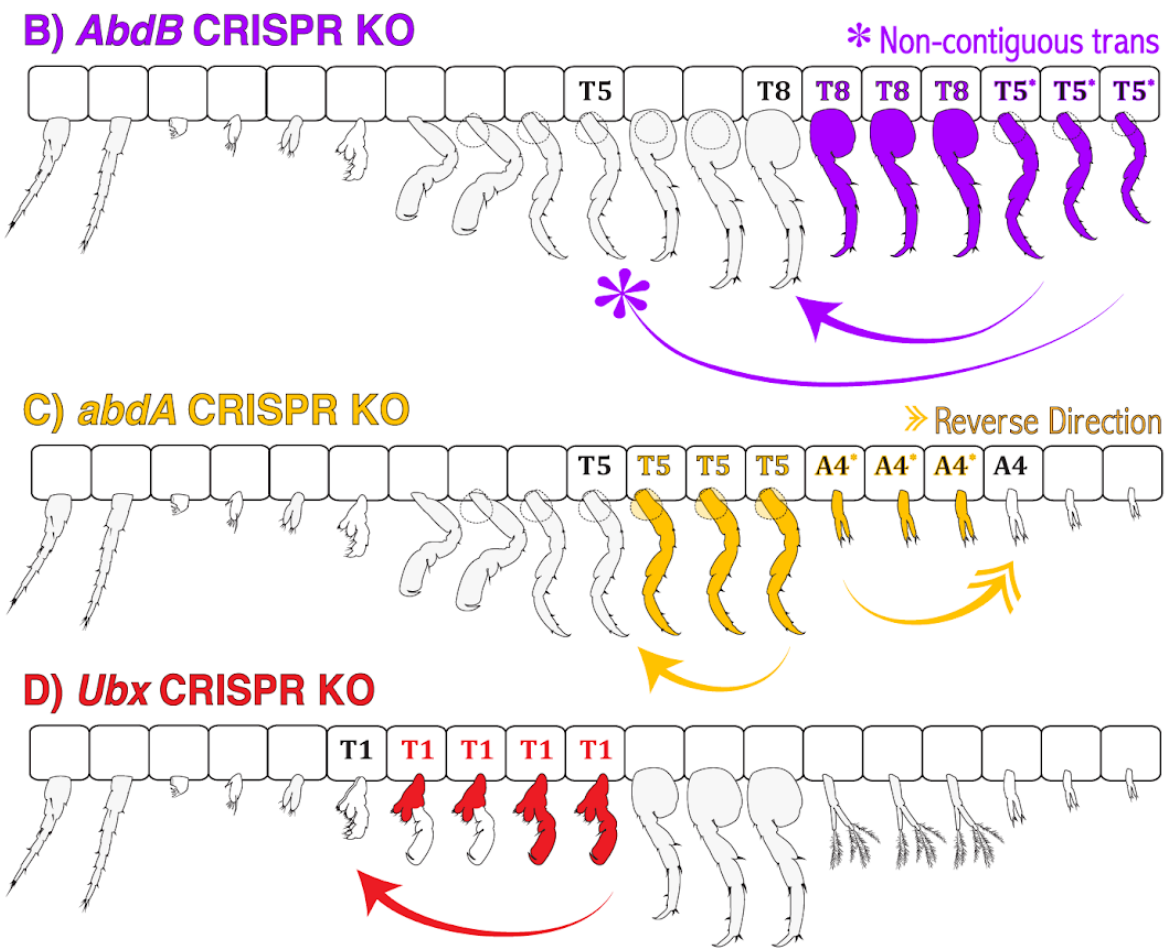


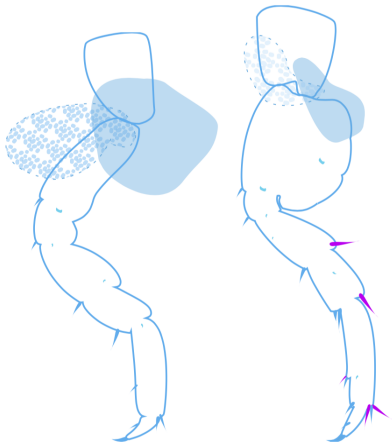
Fig 5.1 (B - D) Transformation of limb morphologies by Hox gene KO.
Homeotic transformation by CRISPR/Cas9 Hox gene knockout (adapted from Martin et al. 2016). Arrows show direction of transformation, double arrow indicates unusual transformation. Color overlay of appendage represents region of appendage transformation. **(B) *Abd-B* knockout** results in the transformation of the A1 - A3 pleopods to the T6 - T8 jumping leg phenotype that is immediately anterior as well as the non-linear skip in transformed identity of the A4 - A6 uropods towards that of the non-adjacent walking legs of T4 - T5. **(C) *abd-A* knockout** In the *Ubx* expressing thorax, *abd-A* knockout results in the transformation of the *abd-A*-expressing T6 - T8 jumping legs to the adjacently anterior T4 - T6 walking legs that do not express *abd-A*. The *abd-A*-expressing A1 - A3 swimming appendages are transformed to the adjacently posterior A4 - A6 uropods that fall outside the strong *abd-A* expression domain within the *Abd-B*-expressing abdomen. **(D) *Ubx* knockout** results in the linear transformation of the thoracic legs T2 - T5 toward the maxillary identity of T1 that is adjacently anterior. Note that while all appendages expressing either *abd-A* or *Abd-B* are affected by the respective Hox gene knockout, only *Ubx* appendages not also co-expressing the more posterior *abd-A* are transformed upon *Ubx* knockout, as is consistent with the more traditional Hox models.

Results

The 'Hox code' of thoracic and abdominal limb identities

Non-linear transformation of abdominal appendages corresponds with abd-A *expression in* Abd-B *KO*

In *Abd-B* KO, A4 - A6 uropods are non-linearly transformed toward T4 - T5 walking legs rather than to the T6 - T8 jumping legs (which the A1 - A3 appendages are transformed to) as would be predicted using classic Hox logic. One possible explanation for this unusual transformation is a combinatorial Hox code model. In this scenario, *abd-A* would function together with either *Ubx* or *Abd-B* in their respective domains of overlap to establish a phenotypic identity unique from either domain alone. As previous studies did not establish the consequences of *Abd-B* KO on the expression of other Hox genes in the *Abd-B* domain in its absence, we performed RNA *in-situ* hybridization (*abd-A*) and immunofluorescence (UBX) to characterize the expression patterns underlying the transformed leg identities in *Abd-B* KO embryos.[1]



In the *Abd-B* expressing abdomen of wildtype *Parhyale*, *abd-A* is strongly expressed in the A1–A3 segments and primordial swimming appendages (pleopods), which are uniform in size and virtually indistinguishable from one another. The posterior edge of *abd-A* expression marks a boundary between these and the posterior A4 – A6 segments and uropods, which mostly lack the expression of *abd-A* with the exception of weak expression in the anterior compartment of the A4 segment and appendage, and a similar but even weaker anterior domain in A5 (Serano et al. 2016). Despite low-levels of *abd-A* transcript in their respective anterior compartments, A4 and A5 appendages are distinctly uropod in identity. The main phenotypic difference is in their size, where the width of the A4 segment and corresponding size of uropod is significantly greater than those of A5, which are in turn greater than those of A6. In *Abd-B* KO animals, the A1 - A3 appendages are transformed uniformly and distinctly towards a reverse jumping leg identity, whereas the transformed A4 - A6 appendages are exhibit graded intermediate qualities of both forward walking and reverse jumping legs (Figure 5.2i-ii).

[1] Schematic is forward walking leg and reverse jumping leg. Note that the forward walking leg is shown in its reverse orientation to more easily compare to the reverse jumping leg

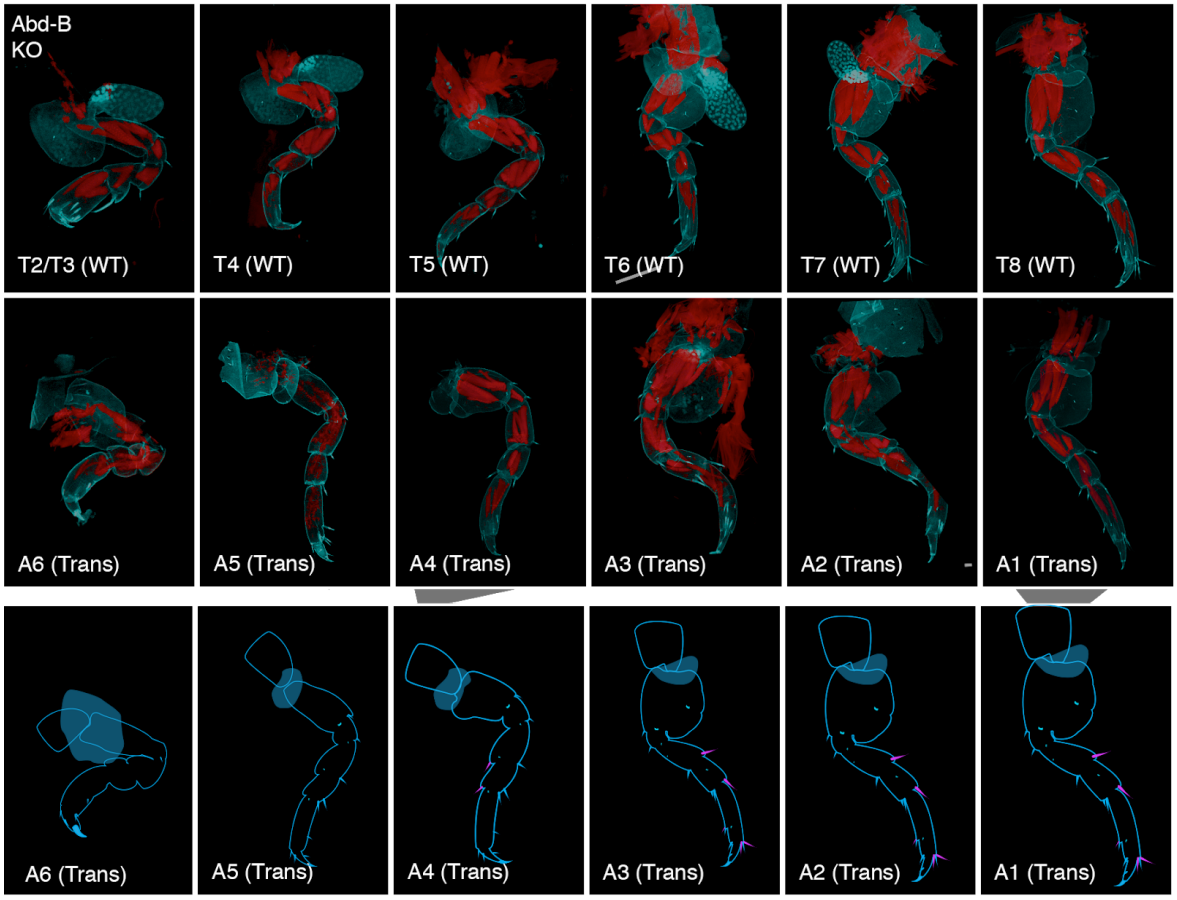


Figure 5.2i-ii Intermediate jumping leg → walking leg phenotypes
Panel 5.2i shows the transformed and schematized limb phenotypes of a representative *Abd-B* KO hatchling. (D-F) *Abd-B* KO gives rise to a striking and apparently complete homeotic transformation of the A1 - A3 set of abdominal appendages (which strongly express *abd-A* throughout) from feathery, biramous swimming limbs to the anteriorly adjacent uniramous, backward-facing jumping legs of T6 - T8 (which also strongly express *abd-A* throughout). Panel 5.2ii The A4 - A6 set of abdominal appendages (which express low levels of *abd-A* in an anteriorly restricted domain) are transformed to a phenotype most resembling the T4 - T5 walking legs (which express no *abd-A*) but with intermediate jumping leg properties.

To investigate *abd-A* expression in the absence of *Abd-B*, we performed RNA *in situ* hybridization on *Abd-B* KO embryos alongside age-matched uninjected control embryos from the same collection of broods. The strong expression of *abd-A* remained largely unchanged in the transformed A1 - A3 appendages and segments of *Abd-B* KO embryos compared to the uninjected wildtype controls (Figure 5.2). The relatively weak *abd-A* expression observed in the wildtype A4 appendages, however, is clearly increased in *Abd-B* KO to near the levels observed in A1 – A3 but remained limited to the anterior portion of the limb. Strikingly, this intermediate level of *abd-A* expression strongly corresponds with the intermediate nature of the respective appendage transformation that has attributes of

both forward-facing walking legs (which lack *abd-A* in the wildtype *Parhyale* thorax) and reverse-facing jumping legs (which strongly expresses *abd-A* throughout). Likewise, the low anterior expression of *abd-A* in A5 appendages is also increased in *Abd-B* KO mutants, though to a lesser degree than in A4. Strong *abd-A* expression is limited to distal, but not the most distal segments of the primordial appendage. Correspondingly, the transformed A5 appendages also feature intermediate qualities from both thoracic leg types, though are further transformed towards the forward walking leg identity compared to the A4 transformations. The segment and appendages of A6 remain *abd-A*-less and exhibit no qualities distinctive to the reverse jumping leg identity (Figure 5.2A).

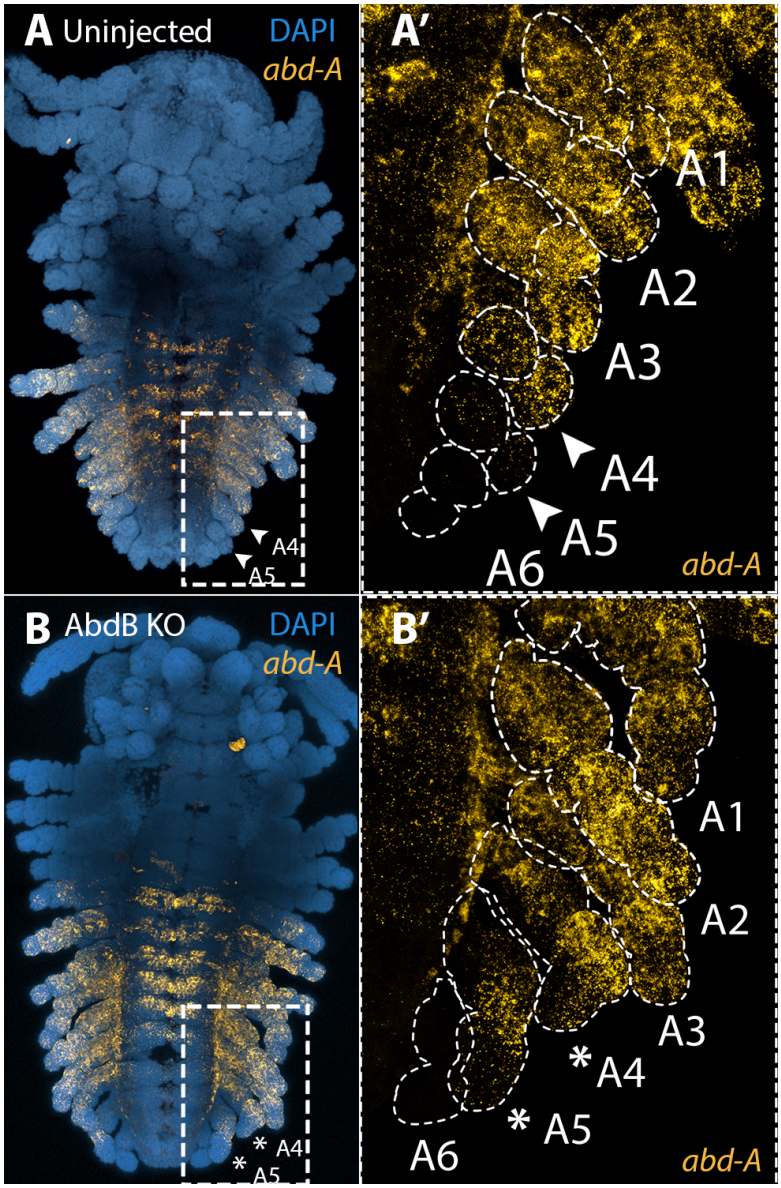


Fig 5.2. *abd-A* expression correlates with limb size and intermediate walking/jumping leg features

(A & A') *abd-A* expression in wildtype embryo. Anterior boundary of *abd-A* delineates the distinct identities of the T4/T5 forward walking legs from the T6-T8 reverse jumping legs in the thorax. In the abdomen, *abd-A* is strongly expressed throughout the A1-A3 pleopods, and in an anteriorly-restricted domain in A4 and to a lesser degree in A5 (white arrowheads), and is absent from A6. **(B) *abd-A* expression in *Abd-B* KO embryos**. A1-A6 abdominal limbs show clear transformation from bifurcated appendages to a uniramous legs. A1 - A3 exhibit all distinct podomeres, while A4 and A5 are clearly uniramous (asterisk), they are relatively shorter and exhibit fewer distinct podomeres. The levels of *abd-A* in A4 and A5 appear to be slightly increased relative the wildtype control, though their spatial restriction to the anterior domain of the limb was maintained. A6 remains absent of visible *abd-A* expression.

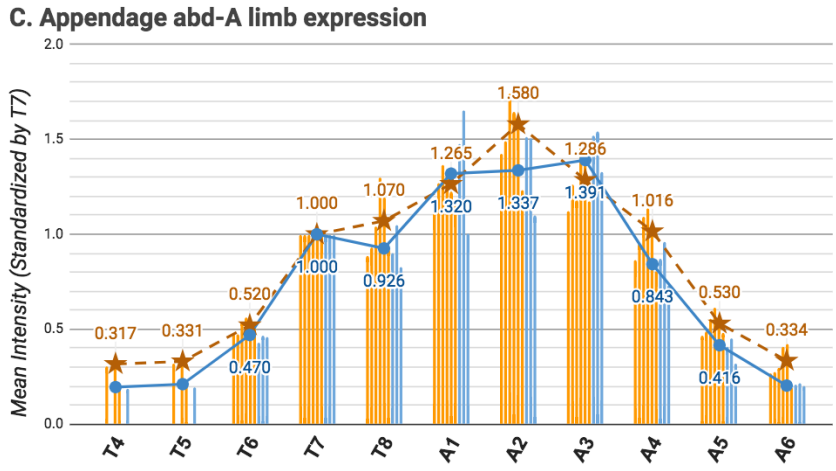


Fig 5.2 (C) Quantitative *abd-A* analysis of the mean average limb intensity for wildtype and *Abd-B* KO embryos.

The smaller but equivalently shaped T6 limb shows a distinctly lower level of *abd-A* relative relative T7 and T8, which are of equivalent size and intensity. A1 - A3 in both wildtype and *Abd-B* KOs are of similar intensity that is distinctly higher than that of T7 and T8. As was qualitatively suggested, A4 - A6 appendages show a graded drop in intensity, with the overall mean levels of the *Abd-B* KOs slightly higher than wildtype.

The intermediate identity of A4 and A5 walking legs in *Abd-B* KO hatchlings has no clear corollary in wildtype animals and likely results from the expression of anteriorly-restricted *abd-A* in overlap with *Ubx* and not *Abd-B* as it is within the wildtype abdomen. To test this further, we generated *abd-A* + *Abd-B* double KO animals and examined their limb morphology. Consistent with the prediction that *abd-A* establishes the specific qualities of reverse jumping legs, simultaneous KO of *Abd-B* and *abd-A* homonymizes all T4 - A6 appendages of the posterior thorax and abdomen to the forward walking leg identity (Figure 5.2D).

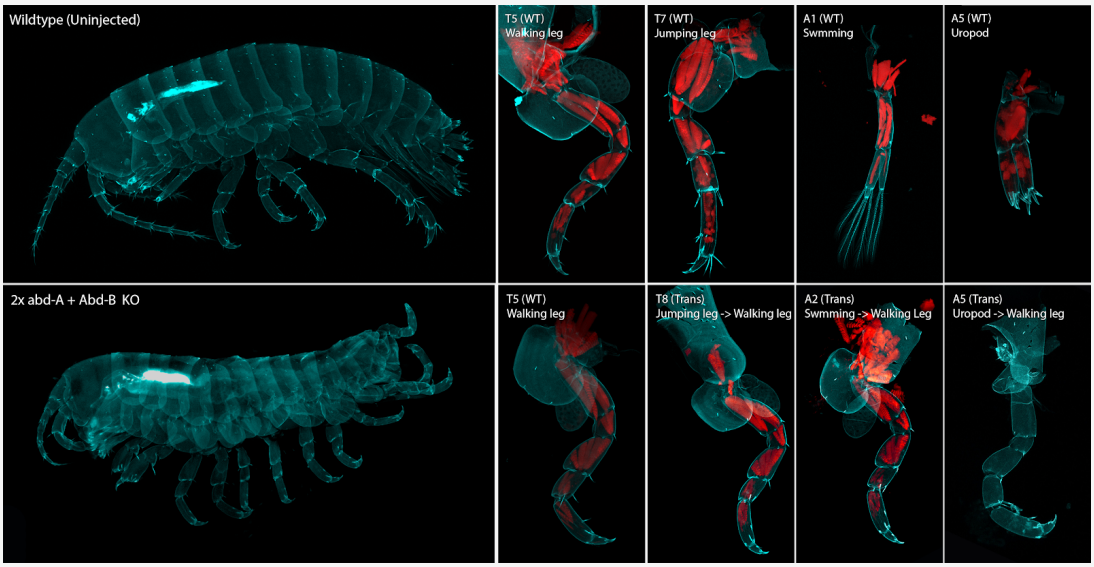


Fig 5.2D
Simultaneous KO of *Abd-B* and *abd-A* results in the removal of all jumping leg-specific properties, thus homonomyzing T4 - A6 appendages to forward walking legs.

Abd-B *KO enables posterior expansion of* Ubx *expression into the thoracically-transformed abdomen*

The above result indicates *abd-A* as a key determinant of the reverse jumping vs. forward walking leg identity in the *Abd-B* KO transformed abdomen as well as in the thorax, where the principle Hox gene associated with the wildtype walking/jumping leg identities is *Ubx*. In *Parhyale*, *Ubx* is expressed strongly throughout the thoracic legs (T4 - T8) and at lower levels in the claws (T2 - T3), and is absent from wildtype abdominal appendages (Serano et al. 2016) (Figure 5.3). To further test that the identity of abdominal appendages transformed toward thoracic walking or jumping leg are in adherence to a putative "Hox code", we examined *Abd-B* KO embryos for ectopic expression of *Ubx* in the abdomen. True to code, we find that *Abd-B* KO embryos exhibit a posterior expansion of *Ubx* throughout the abdominal segments and transformed appendages. Moreover, individuals with a mix of wildtype and transformed appendages exhibit corresponding mosaicism of abdominal *Ubx* expression (Figure 5.3).

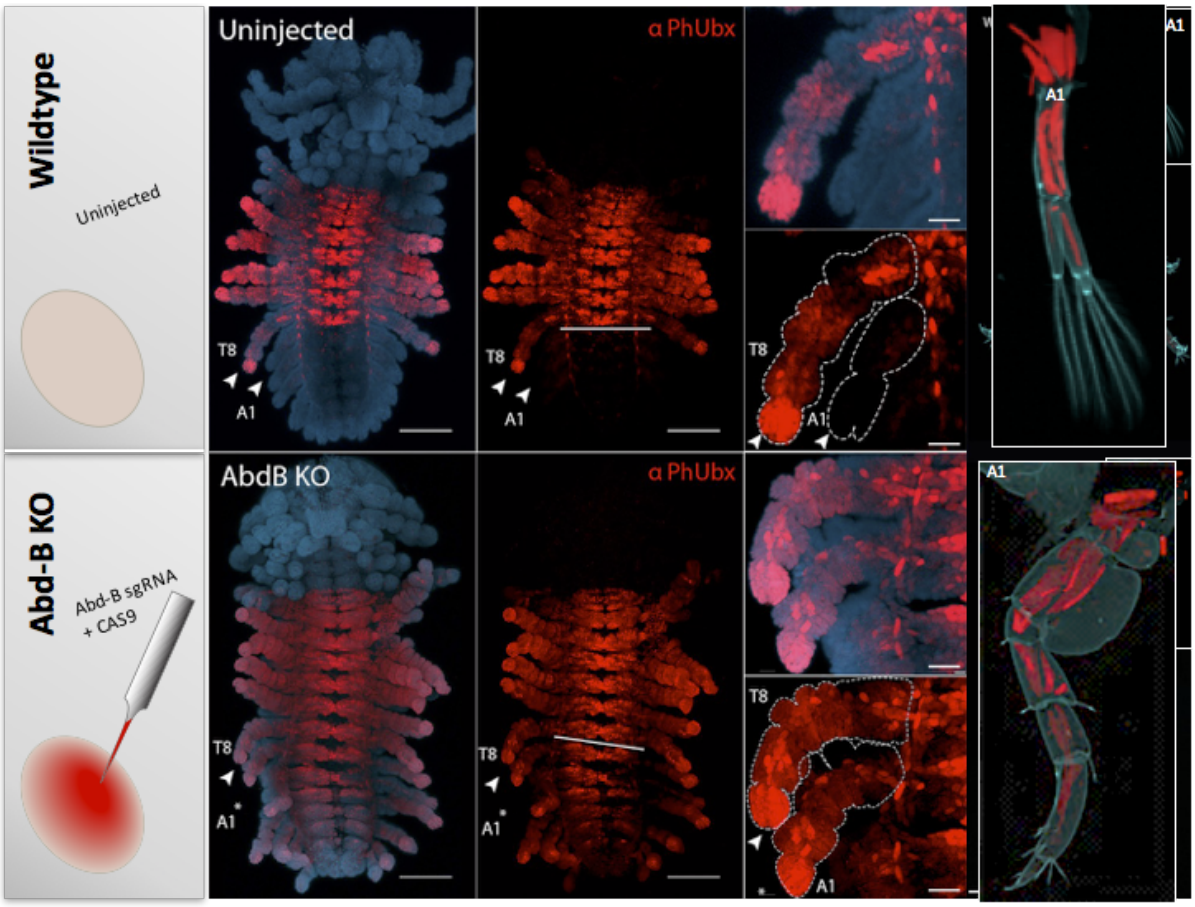


Fig. 5.3. *Abd-B* KO results in the posterior expansion of the *Ubx* boundary.
(A-D') Stage 22 *Parhyale* embryos stained immunofluorescently for *Ph*-UBX (Red) and counterstained with DAPI (blue). Arrowhead marks wildtype appendages; Asterisk denotes transformed appendage. In wildtype embryos (A-B') UBX is expressed at high levels throughout the T4-T8 thoracic legs and at lower

levels in the T2 - T3 claws. Although there appears to be some intermittent expression of *Ubx* in select abdominal neurons/neuromeres, *Ubx* is absent from abdominal appendages. The thoracic leg vs. abdominal appendage phenotypes are clearly distinguishable in embryos at this stage, as is highlighted by the nascent T8 leg and A1 pleopod in (B'). Walking and jumping legs are uniramous with 7 podomeres (leg segments) while pleopods and uropods are biramous, with two rami (an endopod and exopod) emerging from a single protopod. (C-D') *Abd-B* KO embryos show clear transformation of their abdominal appendages to a thoracic leg identity (D' highlights uniramous, multi-podomeres of transformed A1). The posterior boundary of *Ubx* is very clearly expanded throughout the abdominal segments and appendages in the absence *Abd-B*.

Antp *and* Dfd *expression in the transformed abdominal legs of* Abd-B *KOs correspond with wildtype leg expression patterns*

To complete an expression survey of all Hox genes expressed in wildtype legs and/or abdominal appendages in the absence of *Abd-B*, we examined the expression patterns of *Dfd* (*in-situ* hybridization) and ANTP (immunofluorescence) in *Abd-B* KO embryos to determine if shifts in expression of other Hox genes may be associated with thoracic legs and abdominal appendage transformations in *Parhyale*. We find that the non-ectodermal expression pattern of *Antp* observed in wildtype thoracic legs (but not abdominal appendages) is extended into the transformed walking and jumping legs of the abdomen (Figure 3). *Dfd*, which is expressed in the distal tips of wildtype A1 - A3 swimming appendages (but not thoracic legs) is correspondingly absent from the transformed abdominal legs (Figure 3). Whether these patterns serve a functional purpose or simply molecular markers following the phenotype needs to be explored further.

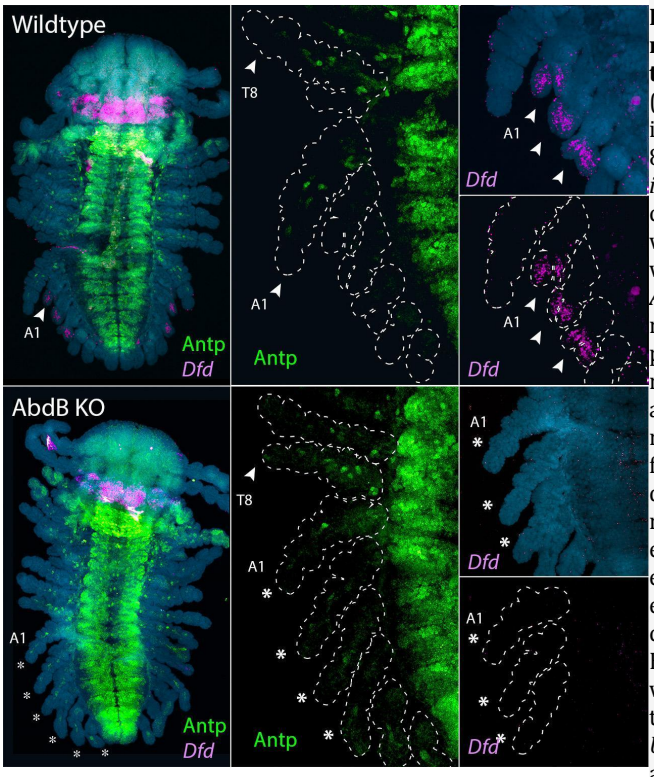


Fig 5.4 Wildtype *Antp* & *Dfd* leg patterns recapitulated by abdominal appendages transformed to legs in *Abd-B* KO embryos
(A-F') Triple stain of stage 22 embryos by immunoĀluorescence (cross-reactive monoclonal ANTP 8C11, green and *Ph-UBX*, not shown) and Āluorescent *in-situ* hybridization (*Dfd*, fuschia); embryo counterstained with DAPI (cyan). Arrowheads point to wildtype appendages, transformed appendages labeled with asterisks. (A-C') Wildtype expression patterns of *ANTP* and *Dfd*. ANTP is expressed in the T1 - T3 maxilliped and claws, and in a limited, non-ectodermal pattern in the remaining T3 - T8 thoracic legs. *ANTP* is not expressed in the abdominal appendages (B). While anterior Hox gene *Dfd* is expressed primarily in the maxillary appendages Mn and Mx1 (A), this study focuses on its discrete expression outside its primary domain of expression in the not quite distal-most tips in nascent wildtype pleopods (C - C'). (D-F') In *Abd-B* KO embryos, the same non-ectodermal *Antp* patterning exhibited by T4 - T8 thoracic legs is ectopically expressed in the A1 - A6 abdominal appendages that are clearly transformed to a thoracic-like leg identity (D - E). Furthermore, the A1 - A3 expression of *Dfd* in wildtype pleopods (but not thoracic legs) is absent from the transformed legs (D, F-F'). As is shown in Figure 3, *Ubx* is expressed throughout the *Abd-B* KO transformed abdomen of this specific embryo (not shown).

The above analyses encompass all Hox genes expressed in the wildtype thorax and abdomen of *Parhyale* and demonstrate that the Hox expression underlying each of the limbs transformed by *Abd-B* KO corresponds to the unique combination of Hox genes expressed by the respective wildtype limb each is transformed toward.

Cross-regulatory interactions among Hox genes Scr, Ubx, abd-A, and Abd-B explain non-canonical phenotypic appendage transformations

The above observation of interactions among *Abd-B*, *abd-A*, and *Ubx* in the specification of limb identities strongly suggests that combinatorial interactions on target genes, weather additive or competitive, control aspects of appendage morphology in *Parhyale*. In order to further explore such interactions, we systematically examined the transformed limb morphologies for the remaining *Ubx, abd-A*, and *Abd-B* double KO combinations as well as the triple KO of all three. Each set of transformations was accompanied by a post-KO expression analysis to uncover the shifts in expression domains of other Hox genes that may be regulated by cross-regulatory interactions and to examine the underlying combinations of Hox genes putatively responsible for each transformed identity.

Simultaneous KO of Abd-B *and* Ubx *transforms the posterior-most appendages toward a maxilliped fate*

The posterior expansion of *Ubx* into the abdomen of *Abd-B* KO embryos, whether direct or indirect, establishes the cross-regulatory repression of *Ubx* by *Abd-B*. The correlation between the transformed abdominal legs and their ectopic expression of *Ubx* suggests that *Ubx* has a functional role in establishing the walking/jumping leg identities in the transformed abdomen. To validate *Ubx* function (as opposed to an alternative explanation that the walking leg identity is simply the default identity in absence of *Abd-B*), we targeted the simultaneous KO of *Abd-B* and *Ubx* to eliminate the role ectopic expression of *Ubx* may have in the transformed abdominal appendage identity. Strikingly, the double KO of *Abd-B* + *Ubx* results in the posterior-most abdominal appendages taking on maxillary features of the anterior-most thoracic identity, the shorter, branching gnathal T1 maxilliped (Mxp) (Figure 5.5A-C; G-I). In some cases, the A6 appendage underwent complete transformation to a maxilliped identity (Fig. 5.5C). A1 - A3 appendages, however, continued to exhibit the same reverse jumping leg transformation as with *Abd-B* KO alone.

Previous studies functionally demonstrate the role of *Ubx* in selecting leg vs. maxillary identities at its anterior boundary (Martin et al. 2016; Liubicich et al. 2009; Pavlopoulos et al. 2009). This study is the first to our knowledge to demonstrate a maxillary transformation at the posterior border of *Ubx*. Consistent with the single KO of *Ubx*, double *Ubx* + *Abd-B* KO transforms the T2/T3 claws and T4/T5 walking legs, but not T6 – T8 jumping legs, towards the shorter, branching gnathal phenotype of the T1 maxilliped (Figure 1D). While previous experiments achieved only partial transformation towards a maxilliped identity in T4/T5 walking limbs, the simultaneous use of multiple *Ubx* guides

allowed us to achieve complete maxilliped transformation of all T2 – T5 appendages (see Table 1).

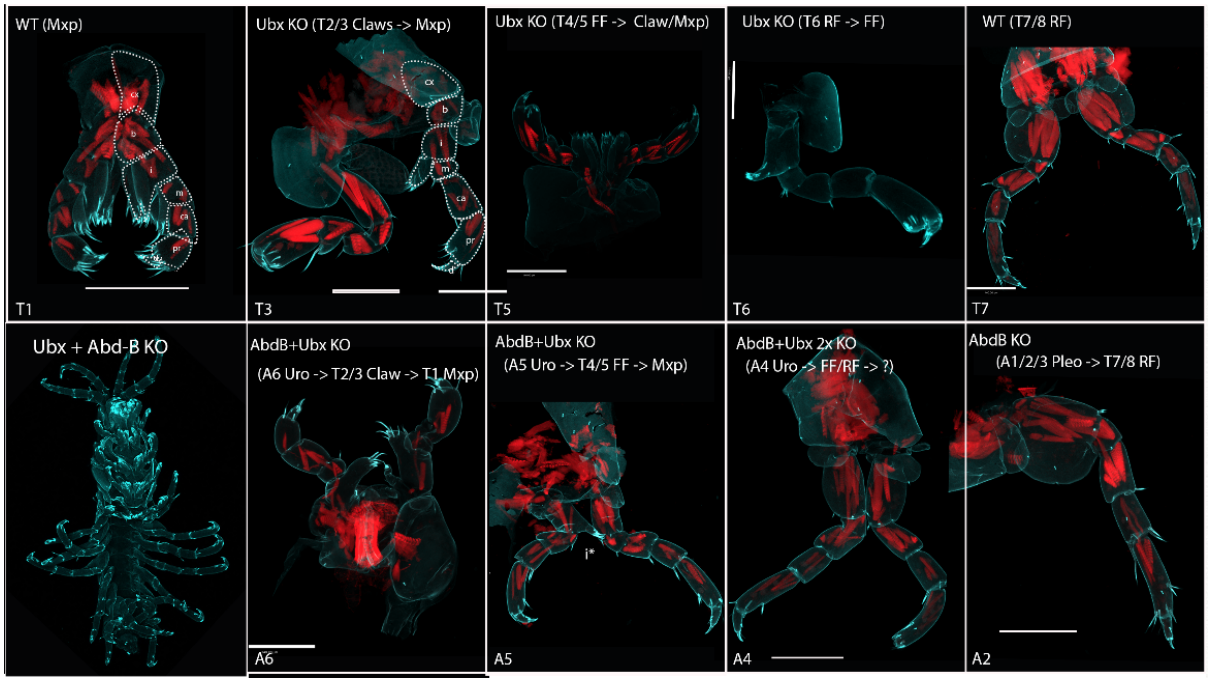
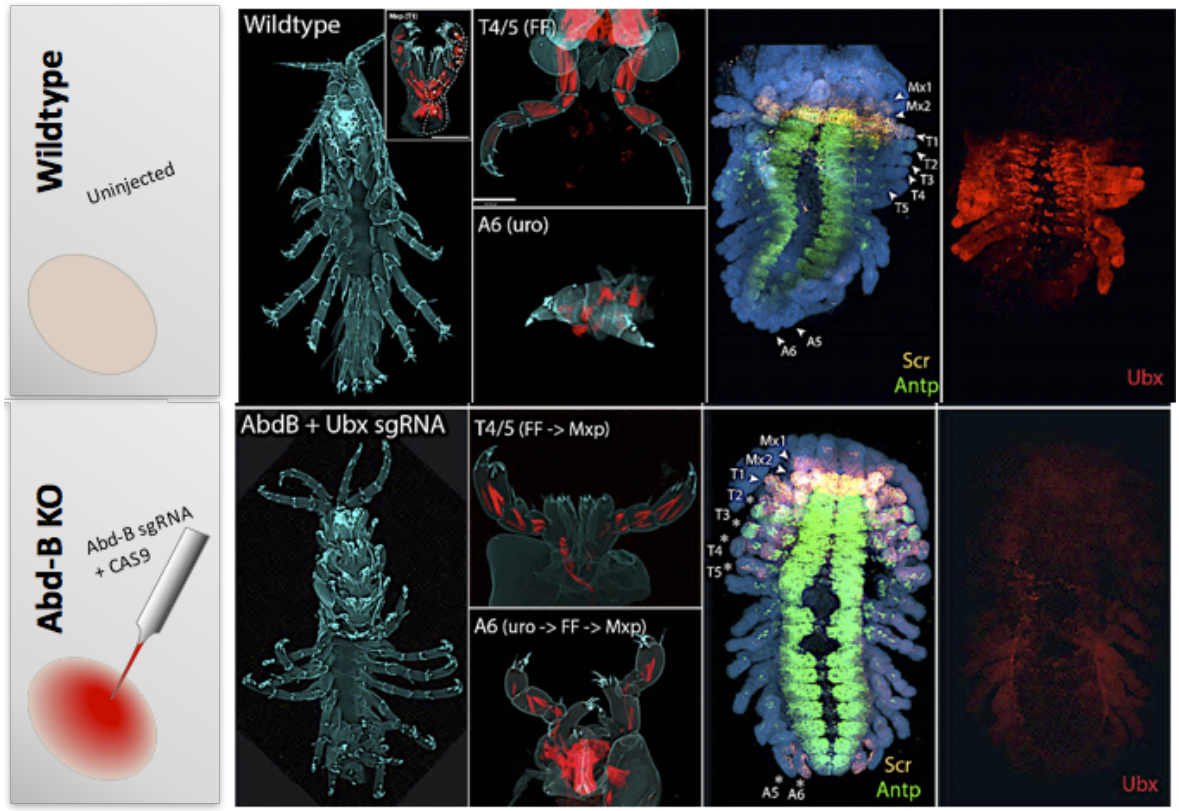


Fig 5.5 Panel 1 show appendages of wildtype (A-C) vs. *Ubx* + *Abd-B* 2x KO (G-I) hatchlings. *Ubx* + *Abd-B* 2x KO transforms the posterior-most walking leg(s) (the A6 appendage, which lacks *abd-A*) towards a maxilliped fate (note that single KO of *Ubx* transformants show no transformation in the A6 uropod identity). The *abd-A* expressing appendages T6 – T8 (reverse-facing jumping legs) and A1 – A3 (transformed to reverse-facing jumping legs by single *Abd-B* knockout) exhibit the jumping leg phenotype exhibited in *Abd-B* KO alone. The intermediate walking-jumping legs exhibited by A4 and A5 in the single KO of *Abd-B* (which have lowered levels of anteriorly-restricted *abd-A*) continue to exhibit their respective *Abd-B* KO identities as well. As is exhibited by the single KO of *Ubx*, the anterior thoracic appendages T2/T3 (claws) and T4/T5 (forward walking legs) are also transformed towards the maxilliped identity. When paired with the simultaneous use of an additional *Ubx* guide (Table ?), our results yield T4/T5 maxilliped transformations that are much more complete than has previously been demonstrated.

Double KO of Abd-B and Ubx induces ectopic expression of Scr

In *Parhyale*, the maxilliped identity is associated with the expression of Scr and Antp, both of which appear to be required for its complete identity. (Martin et al. 2016) functionally demonstrated that the loss of *Scr* leads to the corresponding loss of the maxillary branches and an overall lengthening of the primary endopod. We performed a post-knockout expression analysis of *Scr* and *Antp* on double *Ubx*+ *Abd-B* KO embryos and clearly observe ectopic expression of *Scr*, but not *Antp*, in the transformed limb primordia of the anterior thorax and posterior abdomen (Figure 5). This result demonstrates the cross-regulatory repression of *Scr* by the more posterior Hox genes *Ubx* and *Abd-B* as well as the continued association of the maxilliped identity with *Scr* no matter its location.



5.5 Panel 2 show the corresponding embryonic expression of *Scr, Antp*, and UBX in wildtype (D-F) vs. *Ubx* + *Abd-B* KO (J-L) animals. Multi-label fluorescent staining panels show the wildtype and post-knockout expression of *Scr* (*in-situ* hybridization) and ANTP and UBX *(immunofluorescence)* show that the transformation towards a maxilliped identity is paired with loss of *Ubx* (relative its ectopic expansion upon *Abd-B* knockout) and gain of *Scr*. 40x zoom shows that the Scr expression in A6 is expressed fully throughout the entire developing limb, whereas its expression in A5---which expresses low-levels of *abd-A* in its anterior compartment---appears to be limited to the posterior compartment.
(One more panel?) The triple KO of *Ubx + abd-A+ Abd-B* results in the ectopic expression of *Scr* throughout thoracic and abdominal appendages in the absence of these posterior-most genes, creating a fully "foot-in-mouth" monster with maxillipeds on all thoracic and abdominal segment (T1 - A6). Likewise, the 2x KO of *Ubx + abd-A* (M-O) shows that the loss of *abd-A* along with *Ubx* allows the posterior boundary of *Scr* to expand throughout the entire thorax, transforming all thoracic appendages toward a maxilliped fate, but stopping short of the *Abd-B* expressing abdomen. as well as abd-A 2x KO of *Ubx + abd-A* transforms all thoracic appendages

Triple Ubx + abd-A + Abd-B KO transforms all appendages of the thorax and abdomen to the maxilliped fate

The ectopic *Scr* expression and the associated transformation to maxilliped identity is only observed in regions of the thorax and abdomen that lack the expression of *abd-A*. This suggests that the expression of *Scr* and the associated maxilliped identity may also be repressed by *abd-A*. To test this cross-regulatory interaction and examine *abd-A* function in these processes, we carried out the simultaneous triple knockout of *Ubx* + *Abd-B* + *abd-A*. Strikingly, but not surprisingly, *Ubx* + *abd-A* + *Abd-B* KO hatchlings exhibit the maxilliped identity for all thoracic and abdominal appendages (Figure 6). In order to examine the complete set of *Ubx*, *abd-A*, and *Abd-B* KO combinations, we perform the simultaneous KO of *Ubx* and *abd-A*. As would be expected, all thoracic limbs are transformed to maxillipeds.

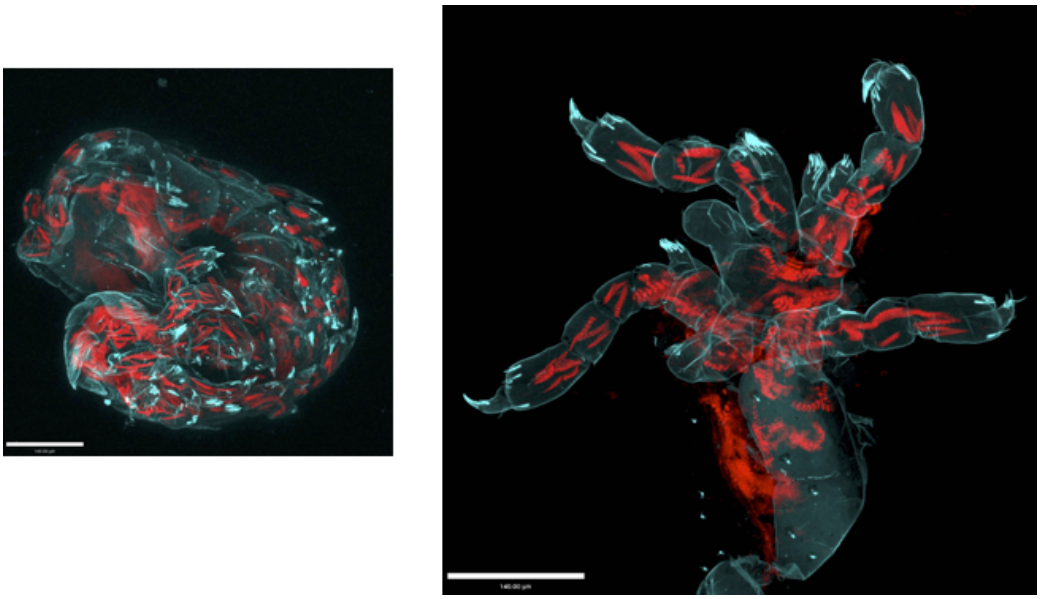


Fig 5.6: 3x Ubx + *abd-A* + *Abd-B* KO
(A) All thoracic and abdominal appendages are transformed towards the maxilliped phenotype (B) Dissection of T8 and A1 appendages (a jumping leg and swimming leg, respectively, in wildtype animals)

Additional Data

abd-A KO results in the more uniform expression of Abd-B in the Parhyale abdomen

While Abd-B is expressed throughout the abdomen, it is expressed more strongly in the A4 - A6 segments and appendages than in A1 - A3. We show here that siRNA knockdown and CRISPR/Cas9 knockout of *abd-A* results in a more even expression of Abd-B in Parhyale embryos (Supplementary Figure S5.2). As expected, immunofluorescent staining of Antp and Ubx in abd-A knockdown/knockout embryos reveals no obvious changes to their expression (data not shown).

Discussion

There is a striking order apparent in the body plans of segmented animals such as arthropods. Comparative study of body plan diversity among arthropod taxa reveals regularities in the identities of these segments and the transitions between these identities, particularly in the appendages. One clear example is the variation and interchange of maxilliped and thoracic walking/swimming leg number in crustaceans (Averof and Patel 1997). The revelation that mutation of a single gene could completely transform the identity of appendages that arose from studies in *Drosophila* has shaped the mechanistic and evolutionarily explanations underlying these transitions. One key component of this has been the posterior prevalence model of Hox function, which is a very simple rule that makes clear predictions. Its simplicity is attractive; if such a simple model can (seemingly) explain much of Hox function, there must be a simple underlying mechanism. This is akin to other early ideas about pattern formation such as those of Alan Turing, who suggested that a very simple rule, the diffusion-reaction system, would explain many diverse patterns found in nature such as spots and stripes seen in animals. When the expression of many early patterning genes in *Drosophila* were found to be in very regular striped patterns, and that these patterns were regulated in part by a transcription factor expressed in a gradient, it was hypothesized by many that a single genetic element would use something simple like Turing had suggested to form the seven stripes. The molecular dissection of the even skipped gene revealed that it was not at all that simple, however, and in fact numerous diverse separable genetic elements--each modular and functioning independently--were responsible for individual stripes. Once again, the striking order that seems apparent in a biological system is more likely the result of a modular and highly contingent evolutionary process.

Here we have examined the cross-regulatory interactions and combinatorial effects of the Hox genes *Scr*, *Ubx*, *abd-A*, and *Abd-B* in establishing appendage identities in the *Parhyale* thorax and abdomen. We find that a simple posterior prevalence model is not sufficient to account for the observed transformations and that rather a "Hox code" model best explains the role of Hox genes in the formation of *Parhyale's* diverse range of appendage types**.** The initial motivation for this study was the observation of two atypical appendage transformations first observed by Martin et al. (2016) following the single KO of *abd-A* or *Abd-B* in *Parhyale*. As the transformations violated the posterior prevalence rules, we explored the possibility that a combinatorial "Hox code" model of appendage specification might explain these transformations. To do this, we have examined the appendage transformations for all single, double, and triple KO combinations of the posterior Hox genes *Ph-Ubx*, *Ph-abd-A*, and *Ph-Abd-B* and determined the patterns of Hox expression associated with each mutant background. We uncovered a number of cross-regulatory interactions that overall explain the pattern of appendage identity in mutant and wild type animals. The Hox expression patterns underlying each specific transformed appendage type correspond with morphologically similar wildtype appendages and can be explained by altered expression domains of Hox genes affected by the respective Hox gene KO(s).

Furthermore, appendage transformations that exhibited novel intermediate identities have patterns of Hox expression distinct from any wildtype pattern. Interestingly, the novel combinations of features in these intermediate transformed appendages reflect aspects of the morphologies associated with wild type Hox codes.

Combinatorial effects of *abd-A*

The proposal that the diversification of appendage types can be achieved by combinatorial Hox interaction such as we show for *abd-A* with either *Ubx* or *Abd-B* is not without precedent. Expression data from several other malacostracan crustacean species including *Procambarus fallax* (crayfish), *Mysidium columbiae* (mysid), and *Porcellio scaber* (woodlouse) show that the distinction and relative numbering of limb subtype in both the thorax and abdomen correspond to their respective overlap with *abd-A* (Abzhanov and Kaufman 2000; Martin et al. 2016). This suggests that shifts in the boundaries of *abd-A* expression may have provided a modular mechanism for macroevolutionary change in crustaceans. Although *Ubx* and *abd-A* suppress the formation of limbs in most insects, a number of basal hexapods retain their abdominal appendages. In the collembolan *Orchesella cincta* (springtail), combined expression of *Oc-Ubx* and *Oc-abd-A* specify formation of a stabilizing appendage that is different from the *Oc-Ubx* specified springtube or the *Oc-abd-A* leaping organ. (Konopova and Akam 2014). siRNA knockdown of *Oc-Ubx* results in the transformation of the A3 stabilizing appendage toward the more posterior leaping organ of A4 that expresses *Oc-abd-A* alone. Rather than following the expected prediction of posterior prevalence, this outcome parallels the abd-A KO transformation in the abdomen Parhyale, supporting the combinatorial specification of limb type by *abd-A* and *Ubx*.

In this study we show that suppression of *Ubx* expression by *Abd-B* establishes its posterior boundary. The KO of *Abd-B* permits a posterior expansion of *Ubx* throughout the abdomen. This expansion corresponds with a transformation of abdominal appendage types toward wild type identities normally associated with two other Hox codes: A1 - A3 → T6 - T8 (*Ubx* + *abd-A*) and A4 - A6 → T4 - T5 (*Ubx*). We find that the levels of anteriorly-restricted *abd-A* expression in A4 - A5 appendages is increased in the *Abd-B* KO. This increase is responsible for the corresponding intermediacy of the transformed A4 - A6 walking legs, which take on some jumping leg properties associated with wild type appendages normally expressing *abd-A* at higher levels. This is confirmed by the double *Abd-B* + *abd-A* KO where appendages develop exclusively as forward walking legs on all T4 - A6 segments. Consistent with this model, Martin et al. (2016) found that the single KO of *abd-A* results in the transformation of A1 - A3 swimming appendages that normally express both *abd-A* and *Abd-B* to the more posterior identity of the A4 - A6 uropods that express *Abd-B* alone.

Scr (Mxp development) and *Ubx* (thoracic leg development) are suppressed by posteriorly expressed genes *except abd-A*

To investigate the possible interaction with *Ubx*, we performed the simultaneous KO of *Abd-B* and *Ubx.* We found that the reverse jumping legs present on T6 - A3 in the *Abd-B KO* are largely unaffected. This indicates that *abd-A* specifies a thoracic jumping leg identity independent of its co-expression with *Ubx.* The phenotypic dominance of *abd-A* and the absence of interaction with *Ubx* is in accordance with the posterior prevalence model of Hox gene function. Nonetheless, this further bolsters the evidence that the A1 - A3 wildtype pleopod identity is specified by the combined function of *abd-A* and *Abd-B* and not by *Abd-B*, which when alone specifies wild type A4 - A6 uropods, and not by *abd-A*, which when alone specifies wildtype T6 – T8 and transformed A1 - A3 jumping legs in *Ubx + Abd-B* KO.

Despite the lack of transformation of the jumping legs described above, the loss of *Ubx* results in a near complete transformation of the T2 - T5 as well as A6 appendages to maxillipeds in the absence of *abd-A* and/or *Abd-B*. Furthermore the intermediate transformation of A4 and A5 may be explained as a consequence of the phenotypic dominance of the limited domain of *abd-A* expression. The transformation of abdominal appendages to a maxilliped fate by the double KO of *Ubx* and *Abd-B* suggests that *Ubx* is important for walking legs more globally. The maxilliped identity is established largely by *Scr* and is suppressed by *Ubx*, and maxillipeds can in a way be thought of as modified walking legs with jaw-like properties. Previous studies examining expression data from from multiple crustacean species correlated the macroevolutionary gain (and loss) of maxillipeds with the retraction (or expansion) of the anterior expression boundary of *Ubx* (Averof and Patel 1997). Although there are a number of LOF (Martin et al. 2016; Liubicich et al. 2009) and GOF (Pavlopoulos et al. 2009) studies in *Parhyale* that recapitulate the function of *Ubx* in selecting legs vs. maxillary appendages at its anterior boundary, this is, to the best of our knowledge, the only study to date to demonstrate a maxillary transformation in the posterior.

Although numerous functional overexpression and loss of function experiments support a role for *Scr* in maxilliped development (Pavlopoulos et al. 2009; Abzhanov and Kaufman 2000) these studies do not fully explore changes in the expression of Hox genes. Furthermore, while *Antp* KO leads to a transformation towards a more anterior maxillary appendage, the maxilla (Mx2), *Scr* KO leads to the loss of excites and lengthening of endopod, and an overall identity that is walking leg-like in nature. (Martin et al. 2016) data suggest that *Scr*, whose posterior boundary ends with the anterior boundary of Ubx, is important for designating a maxilliped vs. walking leg identity. Our finding reveal for the first time the ectopic expression of *Scr* in those appendages transformed toward the maxilliped fate in *Ubx* + *Abd-B* KO, including in A6 (Figure 5.6).

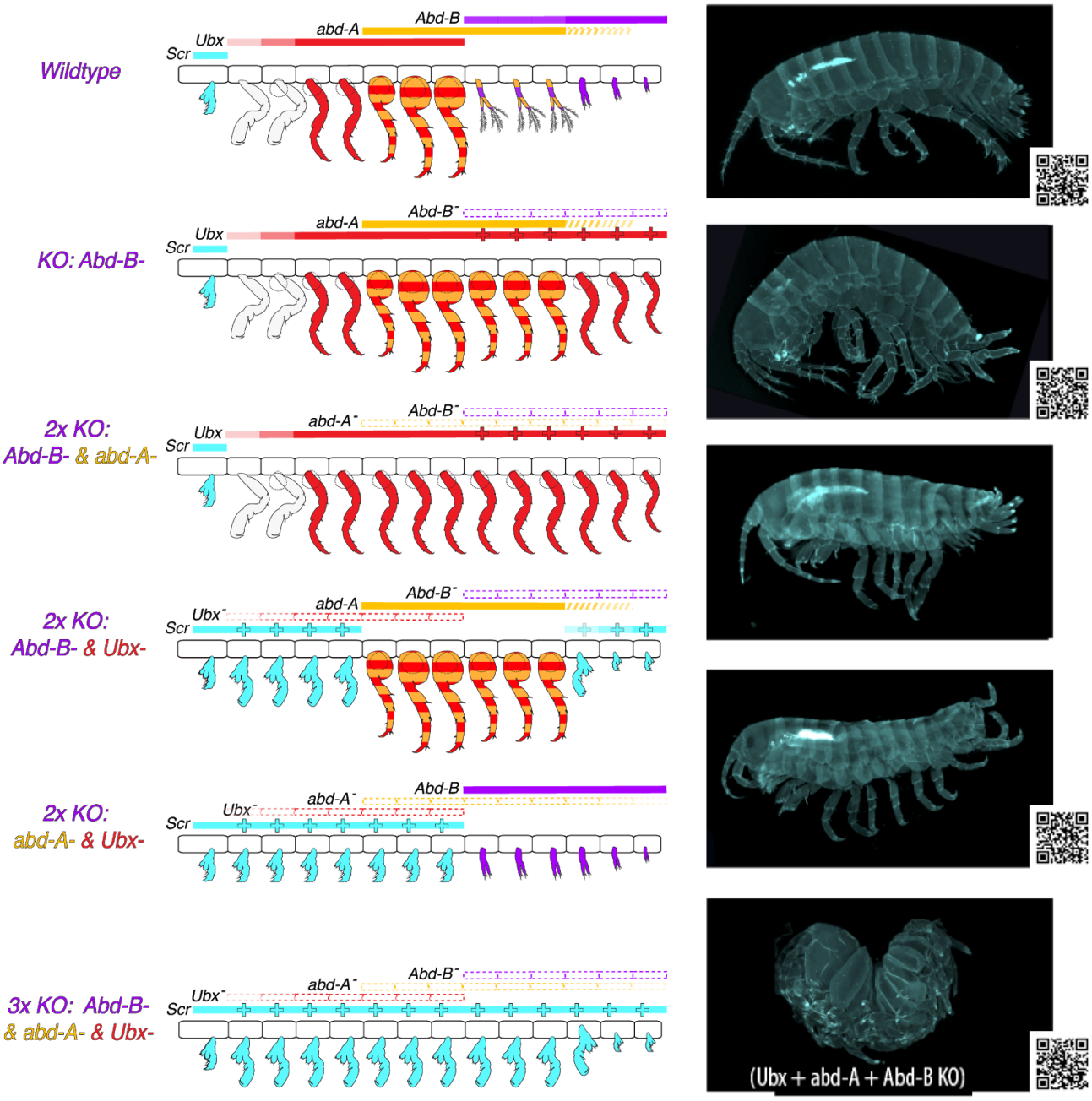


Fig. 5.7. Dynamics of Hox gene cross regulatory repression and co-functional modulation in defining diverse limb phenotypes.

Concluding

Hox genes have sparked considerable attention since their discovery, primarily due to the striking morphological transformations that result from their overexpression and loss demonstrating the importance of even a single Hox gene in establishing regional identity. Embryos are patterned not by a series of Hox genes in isolation, however. Axial regionalization is contingent to some degree upon the cross-regulatory interactions among Hox genes. These can occur on many levels, including pre-translational refinement of one another's expression boundaries and post-translational selection of identity in regions where multiple Hox domains overlap (for example, through additive determinants and functional competition for targets and cofactors). In addition, Hox genes may synergistically or competitively co-regulate downstream pathways, where the genetic interactions among Hox genes modulate phenotype by fine-tuning the initial establishment of one another's boundaries through cross-regulatory interactions.

By studying post-transformational Hox expression in *Parhyale* Hox KOs, we have revealed numerous cross-regulatory interactions and have garnered a fuller picture of the Hox expression patterns underlying the transformed identities, thus allowing us to make a more informed model of limb designation and to provide greater insight into Hox specification of regional identities. We have shown that posterior prevalence is not sufficient to predict and understand the appendage transformations in *Parhyale* and that the modular 'Hox code' expressed in individual segments functions to establish segmental identity. We believe this study demonstrates that interactions among Hox genes provide a mechanism for the diversification of appendage morphologies and the modularization of integrated appendage arrangements over evolutionary time.

Materials and Methods

CRISPR/Cas9 Somatic Mutagenesis

Recombinant Cas9-NLS protein (QB3, UC Berkeley) was combined with purified single-guide RNA (sgRNA) to form a ribonucleoprotein complex (RNP) and delivered directly by microinjection (Rehm et al. 2009). Details of RNP assembly and *Parhyale*-specific genome editing techniques developed by our lab are visually described in (Farboud et al. 2018). CRISPR guides *Abd-B*#1, *abd-A*#1, and *Ubx*#1 are described in (Martin et al. 2016). Guides new to this study were designed using Geneious software (Biomatters, version R9 (Kearse et al. 2012)) and generated synthetically by Synthego (Redwood City, CA). We found ordering synthetic RNA to be a significant time-waver and cost-effective, and, in our hands, the editing efficiency and survival rates of Synthego sgRNA were equivalent to or superior to that generated by *in vitro* transcription. See Table 1 for guide sequences, rates of survival and efficiency, and reference to published *Parhyale* Hox gene sequences.

Approximately 40-60 picoliter of the final RNP injection mix (4 - 8 uM sgRNA (150-200 ng/uL) + 2 uM Cas9 (333ng/uL) + 0.05% phenol red for visualization) was microinjected into one-cell (or both blastomeres of two-cell) embryos following the published protocol (Rehm et al. 2009). Injected embryos, along with batch/age-matched uninjected controls, were cultured at 26°C (12h day-night cycles) to hatching (for phenotypic and genetic analysis) or sacrificed midway through embryogenesis for downstream embryonic Hox expression analyses.

Staining and Imaging

The expression patterns for multiple Hox genes were visualized in parallel by *in-situ* hybridization (*dfd, Scr, abd-A,* and *Abd-B*) and/or immunofluorescence (UBX and ANTP) following modified versions of the published protocols (Rehm et al. 2009; Rehm et al. 2009). *In-situ* probes are described in (Serano et al. 2016) and labeled with either Digoxigenin (DIG) or Dinitrophenol (DNP) conjugated UTPs. When performed in parallel, Hox-specific and anti-hapten primaries were added concurrently after hybridization. *In-situ* hybridization reactions were visualized either fluorescently with AlexaFluor secondaries in parallel with immunofluorescence or enzymatically using FastRed (Sigma F4648).. Antibodies are as follows. Hox-specific primaries: polyclonal Rat-anti-*Ph*-UBX (1:5000) (Liubicich et al. 2009), cross-reactive anti-ANTP at 1:40 (DSHB monoclonal, 8C11), and cross-reactive anti-UBX/ABD-A) at 1:20 (DSHB monoclonal FP6.87, Kelsh et al.(1994)). Anti-hapten primaries: sheep αDIG (Roche) and/or Rabbit anti-DNP (Thermo Fisher) or αDIG-AP. Secondary antibodies were used at 1:500 and included donkey-anti-rabbit AlexaFluor-488 and/or donkey-anti-rat AlexaFluor-647 (Jackson ImmunoResearch) and/or donkey anti-Sheep AlexaFluor-555 (ThermoFisher Scientific). *In-situ* antibody scheme modified from (Ronshaugen and Levine 2004).

Expression panels were performed on G_0 CRISPR/Cas9 KO *Parhyale* embryos along with age-matched uninjected control from the same brood batches. Embryos intended for downstream expression analyses were cultured for approximately 115 - 125h (S21/S22) for *in-situ* hybridization and 120 - 132h (Stage 21) for immunofluorescence. We found these age ranges optimal for balancing embryonic development that has advanced enough for the fate of the nascent limbs to become distinguishable but before the deposition of cuticle in later stages becomes problematic. Embryonic staging is based on (Browne et al. 2005). Embryos were dissected out of their membranes in 3.2% paraformaldehyde (PFA) in Āilter sterilized seawater or in a 9:1:1 ratio of PEM buffer, 10xPBS, and 32% PFA following published methods (Rehm et al. 2009) and fixed for a total of 20 minutes (immunoĀluorescence) or 40 minutes (*in-situ* hybridization). A combination of immunoĀluorescence (UBX and ANTP) and/or *in-situ* hybridization (*dfd, Scr, abd-A,* and *Abd-B*) was performed against multiple Hox genes in parallel

Parhyale hatchlings were sacrificed at 0 - 3 days post hatching and fixed in 3.2% paraformaldehyde as described above for approximately 30 minutes. Fixed hatchlings were

placed in glycerol in a glass bottom dish to image the overall arrangement of appendages. Individual appendages were then carefully removed using sharpened tungsten needles.

Parhyale are direct developers with the appendage morphology of hatchlings reflect the adult morphology. Hatchlings appendages and stained embryos were mounted with 70% glycerol on glass slides, and imaged using a LSM 780 scanning laser confocal (Zeiss) using 10x (entire hatchlings) 20x (individual hatchling legs and entire embryos) or 40x magnification (embryo appendages). LSM files were processed using Volocity software (Perkin-Elmer) and individual channels overlaid in Photoshop (Adobe). Schematic representations made in Illustrator (Adobe)

Acknowledgements

Matt Ronshaugen for his invaluable edits and improvements to the manuscripts. Panda Kreidler for animal care and tool making. Kevin Chung for finishing my experiments when I was unable. Synthego (Redwood City, CA) for technical assistance and hand-delivered sgRNA guides.

Competing Interests

The authors declare no competing interests.

Author Contributions

NHP and EJ conceived of the project. EJ carried out staining experiments, microscopy, and analysis. EJ and KC performed CRISPR/Cas9 mutagenesis. EJ wrote the manuscript with input and edits from NHP.

Chapter 6: Concluding Remarks

Each and every one of these—from a tiny fruit fly to you—begins life as a single cell that must develop just so in order to become its particular complex form. Deviations during development can lead to shocking malformations and congenital defects, but it is these "glitches" that provide the fodder for the evolution of diverse new forms and are most informative for understanding the genes, mechanisms, and pathways that specify one form over another. The study of developmental anomalies (both natural and induced) led to the surprising discovery that it is a relatively small set of highly conserved genes that regulates the successive specification of embryonic development in all sorts of disparate critters, allowing us to gain substantial insight into our own development by studying another's.

Among these regulatory genes are the Hox genes, which designate regional identities along the anteroposterior (head-to-tail) axis of animals, where they are expressed in distinct but overlapping domains. Striking homeotic mutations demonstrate the significance of individual Hox genes in specifying regional attributes of animals, yet it is only when taken in aggregate do the most remarkable properties of Hox genes emerge — namely, that of collinearity and posterior prevalence. The foundational studies of these phenomena were performed in *Drosophila melanogaster*, a foundational model that has taken us far in our understanding in genetics, but one that is relatively derived and lacking in appendage diversity. Despite the sweeping "truthiness" of the models founded in *Drosophila*, our understanding of the extent to which these models hold true are limited by a lack of comprehensive studies characterizing the interactions among Hox genes across a multitude of representative species.

My dissertation investigates the functional role of Hox genes in specifying and integrating the diversity of appendages in the amphipod crustacean *Parhyale hawaiensis. Parhyale* holds a much greater diversity of appendages and segments as does *Drosophila*, and unlike *Drosophila*, which must undergo a second round of development (metamorphosis) to transform into its adult form, *Parhyale* are "direct developers" and develop directly from an embryo to an immature adult form, as is the case in human development. Contrary to the model of posterior prevalence, my experiments show combinatorial function among Hox genes in the specification of appendage identities, and that the induced patterns of Hox expression underlying the transformed identities reflect that of the wildtype "Hox codes" of the respective limb identities. These results demonstrate the capacity of shifting Hox domains to readily influence the expression boundaries of neighboring Hox genes, and promote a model of functional co-modulation for the specification of the highly diverse limb phenotypes of *Parhyale*, which would provide a mechanism for the increased generation of unique appendage types.

Why should we care about how an arthropod develops its appendages? As shown with the popularity of Aesop's fables, questions such as "How did the butterĀły get its eyespots" and "How did the lobster evolve three pairs of jaw-like maxillipeds" the story of the evolution

and development of animal patterns and forms is fascinating. But that's not the end of the motivation to these studies. Studying induced developmental anomalies is also very useful.

So important is the function of the core set of developmental regulatory genes that Hox genes are a part of that they are highly conserved among disparate organisms due to the weight of their significance. Congenital defects are the number one cause of infant mortality and and a whopping 1 in 3 hospital beds are taken due to malformations; by expanding the depth and diversity of Hox studies, we may greatly inform our understanding of human development. Inducing developmental anomalies is incredibly useful for understanding development, and show us very clearly how changes to the expression of just a few regulatory genes could be an important mechanism in the evolution of morphological form.

This thesis details one of the most (if not the most) comprehensive set of studies to date of the roles, functions, and interactions of Hox genes in a non-model organism and in a crustacean. The advancement of CRISPR/Cas9 editing allows us to target specific genes relatively cheaply and easily, allowing us to perform experiments in a much greater range of animals and to a depth of comprehensiveness not feasible even a short few years ago. It would be very informative to continue the Hox gene interactions portion of this dissertation using the more anterior Hox genes of the Antp-Complex, where the role seems more complex and the adherence to canonical models less straightforward. Furthermore, I would very much like to continue looking into how integrated a role the Hox genes have in specifying the nervous system and musculature that underlie the ectodermal transformations exhibited here.

All in all, these "seawater monsters" show us more about our development, our evolution, and ourselves.

References from Non-Published Sections

Abzhanov, A. and Kaufman, T. C. (1999). Novel regulation of the homeotic gene Scr associated with a crustacean leg-to-maxilliped appendage transformation. *Development* **126**, 1121–1128.

Abzhanov, A. and Kaufman, T. C. (2000). Crustacean (malacostracan) Hox genes and the evolution of the arthropod trunk. *Development* **127**, 2239–2249.

Akam, M. (1989). Hox and HOM: Homologous gene clusters in insects and vertebrates. *Cell* **57**, 347–349.

Akam, M. (1998). Hox genes: from master genes to micromanagers. *Curr. Biol.* **8**, R676–8.

Akam, M. (2000). Developmental Genetics and the Diversity of Animal Form: Hox Genes in Arthropods. In *The Biology of Biodiversity* (ed. Kato, M.), pp. 195–208. Tokyo: Springer Japan.

Akam, M., Averof, M., Castelli-Gair, J., Dawes, R., Falciani, F. and Ferrier, D. (1994). The evolving role of Hox genes in arthropods. *Dev. Suppl.* 209–215.

Averof, M. and Patel, N. H. (1997a). Crustacean appendage evolution associated with changes in Hox gene expression. *Nature* **388**, 682–686.

Averof, M. and Patel, N. H. (1997b). Crustacean appendage evolution associated with changes in Hox gene expression. *Nature* **388**, 682–686.

Averof, M. and Patel, N. H. (1997c). Crustacean appendage evolution associated with changes in Hox gene expression. *Nature* **388**, 682–686.

Ball, P. (2011). *Shapes: Nature's Patterns: a Tapestry in Three Parts*. Oxford University Press.

Bateson, W. (1894). *Materials for the study of variation treated with especial regard to discontinuity in the origin of species. By William Bateson*.

Brena, C., Liu, P. Z., Minelli, A. and Kaufman, T. C. (2005). Abd-B expression in Porcellio scaber Latreille, 1804 (Isopoda: Crustacea): conserved pattern versus novel roles in development and evolution. *Evol. Dev.* **7**, 42–50.

Browne, W. E., Price, A. L., Gerberding, M. and Patel, N. H. (2005). Stages of embryonic development in the amphipod crustacean, Parhyale hawaiensis. *Genesis* **42**, 124–149.

Brunet, T. and King, N. (2017). The Origin of Animal Multicellularity and Cell Differentiation. *Dev. Cell* **43**, 124–140.

Carrasco, A. E., McGinnis, W., Gehring, W. J. and De Robertis, E. M. (1984). Cloning of an X. laevis gene expressed during early embryogenesis coding for a peptide region

homologous to Drosophila homeotic genes. *Cell* **37**, 409–414.

Corsetti, M. T., Briata, P., Sanseverino, L., Daga, A., Airoldi, I., Simeone, A., Palmisano, G., Angelini, C., Boncinelli, E. and Corte, G. (1992). Differential DNA binding properties of three human homeodomain proteins. *Nucleic Acids Res.* **20**, 4465–4472.

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection: Or the Preservation of Favored Races in the Struggle for Life*. W. B. Conkey.

Deglincerti, A., Croft, G. F., Pietila, L. N., Zernicka-Goetz, M., Siggia, E. D. and Brivanlou, A. H. (2016). Self-organization of the in vitro attached human embryo. *Nature* **533**, 251–254.

De Robertis, E. M. (2006). Spemann's organizer and self-regulation in amphibian embryos. *Nat. Rev. Mol. Cell Biol.* **7**, 296–302.

Dickinson, D. J., Nelson, W. J. and Weis, W. I. (2011). A polarized epithelium organized by beta- and alpha-catenin predates cadherin and metazoan origins. *Science* **331**, 1336–1339.

Dobzhansky, T. and Dobzhansky, T. G. (1982). *Genetics and the Origin of Species*. Columbia University Press.

Duncan, I. W. and Kaufman, T. C. (1975). Cytogenic analysis of chromosome 3 in Drosophila melanogaster: mapping of the proximal portion of the right arm. *Genetics* **80**, 733–752.

Eldredge, N. and Gould, S. J. (1972). Punctuated Equilibria: An Alternative to Phyletic Gradualism. In *Models in Paleobiology* (ed. Thomas, J. M.), pp. 82–115. Freeman, Cooper and Company, San Francisco.

Farboud, B., Jarvis, E., Roth, T. L., Shin, J., Corn, J. E., Marson, A., Meyer, B. J., Patel, N. H. and Hochstrasser, M. L. (2018). Enhanced Genome Editing with Cas9 Ribonucleoprotein in Diverse Cells and Organisms. *J. Vis. Exp.*

Galant, R. and Carroll, S. B. (2002). Evolution of a transcriptional repression domain in an insect Hox protein. *Nature* **415**, 910–913.

Gehring, W. J., Affolter, M. and Bürglin, T. (1994). Homeodomain proteins. *Annu. Rev. Biochem.* **63**, 487–526.

Gerhart, J. (1999). Pieter Nieuwkoop's contributions to the understanding of meso-endoderm induction and neural induction in chordate development. *Int. J. Dev. Biol.* **43**, 605–613.

Gillham, N. W. (2001). Evolution by jumps: Francis Galton and William Bateson and the mechanism of evolutionary change. *Genetics* **159**, 1383–1392.

Goldschmidt, R. (1933). SOME ASPECTS OF EVOLUTION. *Science* **78**, 539–547.

Goldschmidt, R. (1940). *The Material Basis of Evolution*. Yale University Press.

Gompel, N., Prud'homme, B., Wittkopp, P. J., Kassner, V. A. and Carroll, S. B. (2005). Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in Drosophila. *Nature* **433**, 481–487.

Gould, S. J. (1977). The return of hopeful monsters. *Nat. Hist.* **86**, 22–30.

Gould, S. J. (1990). *Wonderful Life: The Burgess Shale and the Nature of History*. W. W. Norton & Company.

Graba, Y., Aragnol, D. and Pradel, J. (1997a). Drosophila Hox complex downstream targets and the function of homeotic genes. *Bioessays* **19**, 379–388.

Graba, Y., Aragnol, D. and Pradel, J. (1997b). Drosophila Hox complex downstream targets and the function of homeotic genes. *Bioessays* **19**, 379–388.

Graba, Y., Aragnol, D. and Pradel, J. (1997c). Drosophila Hox complex downstream targets and the function of homeotic genes. *Bioessays* **19**, 379–388.

Harding, K., Wedeen, C., McGinnis, W. and Levine, M. (1985). Spatially regulated expression of homeotic genes in Drosophila. *Science* **229**, 1236–1242.

Hughes, C. L. and Kaufman, T. C. (2002). Hox genes and the evolution of the arthropod body plan. *Evol. Dev.* **4**, 459–499.

Huxley, J. (1942). *Evolution, the Modern Synthesis*. G. Allen & Unwin Limited.

Janssen, R. and Budd, G. E. (2010). Gene expression suggests conserved aspects of Hox gene regulation in arthropods and provides additional support for monophyletic Myriapoda. *Evodevo* **1**, 4.

Johannsen, W. (2014). The genotype conception of heredity. 1911. *Int. J. Epidemiol.* **43**, 989–1000.

Kao, D., Lai, A. G., Stamataki, E., Rosic, S., Konstantinides, N., Jarvis, E., Di Donfrancesco, A., Pouchkina-Stancheva, N., Semon, M., Grillo, M., et al. (2016a). The genome of the crustacean Parhyale hawaiensis, a model for animal development, regeneration, immunity and lignocellulose digestion. *Elife* **5**,.

Kao, D., Lai, A. G., Stamataki, E., Rosic, S., Konstantinides, N., Jarvis, E., Di Donfrancesco, A., Pouchkina-Stancheva, N., Semon, M., Grillo, M., et al. (2016b). The genome of the crustacean a model for animal development, regeneration, immunity and lignocellulose digestion. *Elife* **5**,.

Kaufman, T. C. (1978). Cytogenetic Analysis of Chromosome 3 in DROSOPHILA

MELANOGASTER: Isolation and Characterization of Four New Alleles of the Proboscipedia (pb) Locus. *Genetics* **90**, 579–596.

Kaufman, T. C., Lewis, R. and Wakimoto, B. (1980). Cytogenetic Analysis of Chromosome 3 in DROSOPHILA MELANOGASTER: The Homoeotic Gene Complex in Polytene Chromosome Interval 84a-B. *Genetics* **94**, 115–133.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649.

King, N. (2004). The unicellular ancestry of animal development. *Dev. Cell* **7**, 313–325.

Kmita, M. and Duboule, D. (2003). Organizing axes in time and space; 25 years of colinear tinkering. *Science* **301**, 331–333.

Konopova, B. and Akam, M. (2014). The Hox genes Ultrabithorax and abdominal-A specify three different types of abdominal appendage in the springtail Orchesella cincta (Collembola). *Evodevo* **5**, 2.

Krumlauf, R. (1994a). Hox genes in vertebrate development. *Cell* **78**, 191–201.

Krumlauf, R. (1994b). Hox genes in vertebrate development. *Cell* **78**, 191–201.

Lawrence, P. A. and Morata, G. (1994). Homeobox genes: their function in Drosophila segmentation and pattern formation. *Cell* **78**, 181–189.

Levine, M. S. and Harding, K. W. (1989). Drosophila: the zygotic contribution. In *Genes and Embryos* (ed. Glover, D. M.) and Hames, B. D.), pp. 39–94. Oxford; New York: IRL Press.

Lewis, E. B. (1951). PSEUDOALLELISM AND GENE EVOLUTION. *Cold Spring Harb. Symp. Quant. Biol.* **16**, 159–174.

Lewis, E. B. (1978a). A gene complex controlling segmentation in Drosophila. *Nature* **276**, 565.

Lewis, E. B. (1978b). A gene complex controlling segmentation in Drosophila. *Nature* **276**, 565–570.

Liu, Y. (2008). A new perspective on Darwin's Pangenesis. *Biol. Rev. Camb. Philos. Soc.* **83**, 141–149.

Liubicich, D. M., Serano, J. M., Pavlopoulos, A., Kontarakis, Z., Protas, M. E., Kwan, E., Chatterjee, S., Tran, K. D., Averof, M. and Patel, N. H. (2009). Knockdown of Parhyale Ultrabithorax recapitulates evolutionary changes in crustacean appendage morphology. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13892–13896.

Loxdale, H. D. (1991). *J. Anim. Ecol.* **60**, 375–376.

Lutz, B., Lu, H. C., Eichele, G., Miller, D. and Kaufman, T. C. (1996). Rescue of Drosophila labial null mutant by the chicken ortholog Hoxb-1 demonstrates that the function of Hox genes is phylogenetically conserved. *Genes Dev.* **10**, 176–184.

Macías, A. and Morata, G. (1996). Functional hierarchy and phenotypic suppression among Drosophila homeotic genes: the labial and empty spiracles genes. *EMBO J.* **15**, 334–343.

Malicki, J., Schughart, K. and McGinnis, W. (1990). Mouse Hox-2.2 specifies thoracic segmental identity in Drosophila embryos and larvae. *Cell* **63**, 961–967.

Malicki, J., Bogarad, L. D., Martin, M. M., Ruddle, F. H. and McGinnis, W. (1993). Functional analysis of the mouse homeobox gene HoxB9 in Drosophila development. *Mech. Dev.* **42**, 139–150.

Maloof, A. C., Porter, S. M., Moore, J. L., Dudas, F. O., Bowring, S. A., Higgins, J. A., Fike, D. A. and Eddy, M. P. (2010). The earliest Cambrian record of animals and ocean geochemical change. *Geol. Soc. Am. Bull.* **122**, 1731–1774.

Mann, R. S. and Morata, G. (2000). The developmental and molecular biology of genes that subdivide the body of Drosophila. *Annu. Rev. Cell Dev. Biol.* **16**, 243–271.

Margulis, L. (1970). *Origin of eukaryotic cells: Evidence and research implications for a theory of the origin and evolution of microbial, plant and animal cells on the precambrian Earth*. Yale University Press.

Martin, A., Serano, J. M., Jarvis, E., Bruce, H. S., Wang, J., Ray, S., Barker, C. A., O'Connell, L. C. and Patel, N. H. (2016a). CRISPR/Cas9 Mutagenesis Reveals Versatile Roles of Hox Genes in Crustacean Limb Specification and Evolution. *Curr. Biol.* **26**, 14–26.

Martin, A., Serano, J. M., Jarvis, E., Bruce, H. S., Wang, J., Ray, S., Barker, C. A., O'Connell, L. C. and Patel, N. H. (2016b). CRISPR/Cas9 Mutagenesis Reveals Versatile Roles of Hox Genes in Crustacean Limb Specification and Evolution. *Curr. Biol.* **26**, 14–26.

Martin, A., Serano, J. M., Jarvis, E., Bruce, H. S., Wang, J., Ray, S., Barker, C. A., O'Connell, L. C. and Patel, N. H. (2016c). CRISPR/Cas9 Mutagenesis Reveals Versatile Roles of Hox Genes in Crustacean Limb Specification and Evolution. *Curr. Biol.* **26**, 14–26.

Martin, A., Serano, J. M., Jarvis, E., Bruce, H. S., Wang, J., Ray, S., Barker, C. A., O'Connell, L. C. and Patel, N. H. (2016d). CRISPR/Cas9 Mutagenesis Reveals Versatile Roles of Hox Genes in Crustacean Limb Specification and Evolution. *Curr. Biol.* **26**, 14–26.

Martindale, M. Q. and Henry, J. Q. (1998). The Development of Radial and Biradial Symmetry: The Evolution of Bilaterality. *Am. Zool.* **38**, 672–684.

McGinnis, W. and Krumlauf, R. (1992a). Homeobox genes and axial patterning. *Cell* **68**,

283–302.

McGinnis, W. and Krumlauf, R. (1992b). Homeobox genes and axial patterning. *Cell* **68**, 283–302.

McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A. and Gehring, W. J. (1984). A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. *Nature* **308**, 428.

McGinnis, N., Kuziora, M. A. and McGinnis, W. (1990). Human Hox-4.2 and Drosophila deformed encode similar regulatory specificities in Drosophila embryos and larvae. *Cell* **63**, 969–976.

Meinhardt, H. (2001). Organizer and axes formation as a self-organizing process. *Int. J. Dev. Biol.* **45**, 177–188.

Mendel, G. (1865). Experiments in plant hybridization (1865). *Verhandlungen des naturforschenden Vereins Brünn*.

Minelli, A. (2008). The entangled phylogeny of the Bilateria. In *Perspectives in Animal Phylogeny and Evolution*, Oxford: Oxford University Press.

Morata, G. and Kerridge, S. (1981). Sequential functions of the bithorax complex of Drosophila. *Nature* **290**, 778–781.

Morgan, T. H. (1915). *The Mechanism of Mendelian Heredity*. H. Holt.

Morgan, T. H. (1916). *A Critique of the Theory of Evolution*. Princeton University Press.

Noro, B., Lelli, K., Sun, L. and Mann, R. S. (2011). Competition for cofactor-dependent DNA binding underlies Hox phenotypic suppression. *Genes Dev.* **25**, 2327–2332.

Nüsslein-Volhard, C. and Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. *Nature* **287**, 795–801.

Palmer, A. R. (2004). Symmetry breaking and the evolution of development. *Science* **306**, 828–833.

Parchem, R. J., Poulin, F., Stuart, A. B., Amemiya, C. T. and Patel, N. H. (2010). BAC library for the amphipod crustacean, Parhyale hawaiensis. *Genomics* **95**, 261–267.

Pavlopoulos, A., Kontarakis, Z., Liubicich, D. M., Serano, J. M., Akam, M., Patel, N. H. and Averof, M. (2009). Probing the evolution of appendage specialization by Hox gene misexpression in an emerging model crustacean. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13897–13902.

Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A. and Patel, N. H. (2009a). The Crustacean Parhyale hawaiensis: A New Model for Arthropod Development. *Cold Spring*

Harb. Protoc. **2009**, db.emo114–pdb.emo114.

Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A. and Patel, N. H. (2009b). In situ hybridization of labeled RNA probes to fixed Parhyale hawaiensis embryos. *Cold Spring Harb. Protoc.* **2009**, db.prot5130.

Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A. and Patel, N. H. (2009c). The crustacean Parhyale hawaiensis: a new model for arthropod development. *Cold Spring Harb. Protoc.* **2009**, db.emo114.

Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A. and Patel, N. H. (2009d). Antibody staining of Parhyale hawaiensis embryos. *Cold Spring Harb. Protoc.* **2009**, db.prot5129.

Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A. and Patel, N. H. (2009e). In situ hybridization of labeled RNA probes to fixed Parhyale hawaiensis embryos. *Cold Spring Harb. Protoc.* **2009**, db.prot5130.

Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A. and Patel, N. H. (2009f). Fixation and dissection of Parhyale hawaiensis embryos. *Cold Spring Harb. Protoc.* **2009**, db.prot5127.

Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A. and Patel, N. H. (2009g). Injection of Parhyale hawaiensis Blastomeres with Fluorescently Labeled Tracers. *Cold Spring Harb. Protoc.* **2009**, db.prot5128.

Rivera-Pomar, R. and Jäckle, H. (1996). From gradients to stripes in Drosophila embryogenesis: filling in the gaps. *Trends Genet.* **12**, 478–483.

Ronshaugen, M. and Levine, M. (2004). Visualization of trans-homolog enhancer-promoter interactions at the Abd-B Hox locus in the Drosophila embryo. *Dev. Cell* **7**, 925–932.

Ronshaugen, M., McGinnis, N. and McGinnis, W. (2002). Hox protein mutation and macroevolution of the insect body plan. *Nature* **415**, 914–917.

Scott, M. P. and Weiner, A. J. (1984). Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of Drosophila. *Proc. Natl. Acad. Sci. U. S. A.* **81**, 4115–4119.

Scott, M. P., Weiner, A. J., Hazelrigg, T. I., Polisky, B. A., Pirrotta, V., Scalenghe, F. and Kaufman, T. C. (1983). The molecular organization of the Antennapedia locus of Drosophila. *Cell* **35**, 763–776.

Serano, J. M., Martin, A., Liubicich, D. M., Jarvis, E., Bruce, H. S., La, K., Browne, W. E., Grimwood, J. and Patel, N. H. (2016a). Comprehensive analysis of Hox gene expression in the amphipod crustacean Parhyale hawaiensis. *Dev. Biol.* **409**, 297–309.

Serano, J. M., Martin, A., Liubicich, D. M., Jarvis, E., Bruce, H. S., La, K., Browne, W. E., Grimwood, J. and Patel, N. H. (2016b). Comprehensive analysis of Hox gene expression in the amphipod crustacean Parhyale hawaiensis. *Dev. Biol.* **409**, 297–309.

Shiga, Y., Sagawa, K., Takai, R., Sakaguchi, H., Yamagata, H. and Hayashi, S. (2006). Transcriptional readthrough of Hox genes Ubx and Antp and their divergent post-transcriptional control during crustacean evolution. *Evol. Dev.* **8**, 407–414.

Spemann, H. and Mangold, H. (1924). über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. *Archiv für mikroskopische Anatomie und Entwicklungsmechanik* **100**, 599–638.

Spemann, H. and Mangold, H. (2001). Induction of embryonic primordia by implantation of organizers from a different species. 1923. *Int. J. Dev. Biol.* **45**, 13–38.

St Johnston, D. and Nüsslein-Volhard, C. (1992). The origin of pattern and polarity in the Drosophila embryo. *Cell* **68**, 201–219.

Turing, A. M. (1952). The chemical basis of morphogenesis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **237**, 37–72.

Usami, Y. (2006). Theoretical study on the body form and swimming pattern of Anomalocaris based on hydrodynamic simulation. *J. Theor. Biol.* **238**, 11–17.

Wanscher, J. H. (1975). The history of Wilhelm Johannsen's genetical terms and concepts from the period 1903 to 1926. *Centaurus* **19**, 125–147.

Watson, J. D. and Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **171**, 737.

Wedeen, C., Harding, K. and Levine, M. (1986). Spatial regulation of Antennapedia and bithorax gene expression by the Polycomb locus in Drosophila. *Cell* **44**, 739–748.

Whittington, H. B. (1975). The enigmatic animal Opabinia regalis, middle Cambrian, Burgess Shale, British Columbia. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **271**, 1–43.

Whittington, H. B. and Briggs, D. E. G. (1985). The Largest Cambrian Animal, Anomalocaris, Burgess Shale, British Columbia. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **309**, 569–609.

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. *J. Theor. Biol.* **25**, 1–47.

Zhao, J. J., Lazzarini, R. A. and Pick, L. (1993). The mouse Hox-1.3 gene is functionally equivalent to the Drosophila Sex combs reduced gene. *Genes Dev.* **7**, 343–354.

All non-data images that lack an image citation are in the Public Domain and are free to alter and re-publish and without directly crediting the original source.