

Genomic Resources for the Pacific Geoduck (*Panopea generosa*)

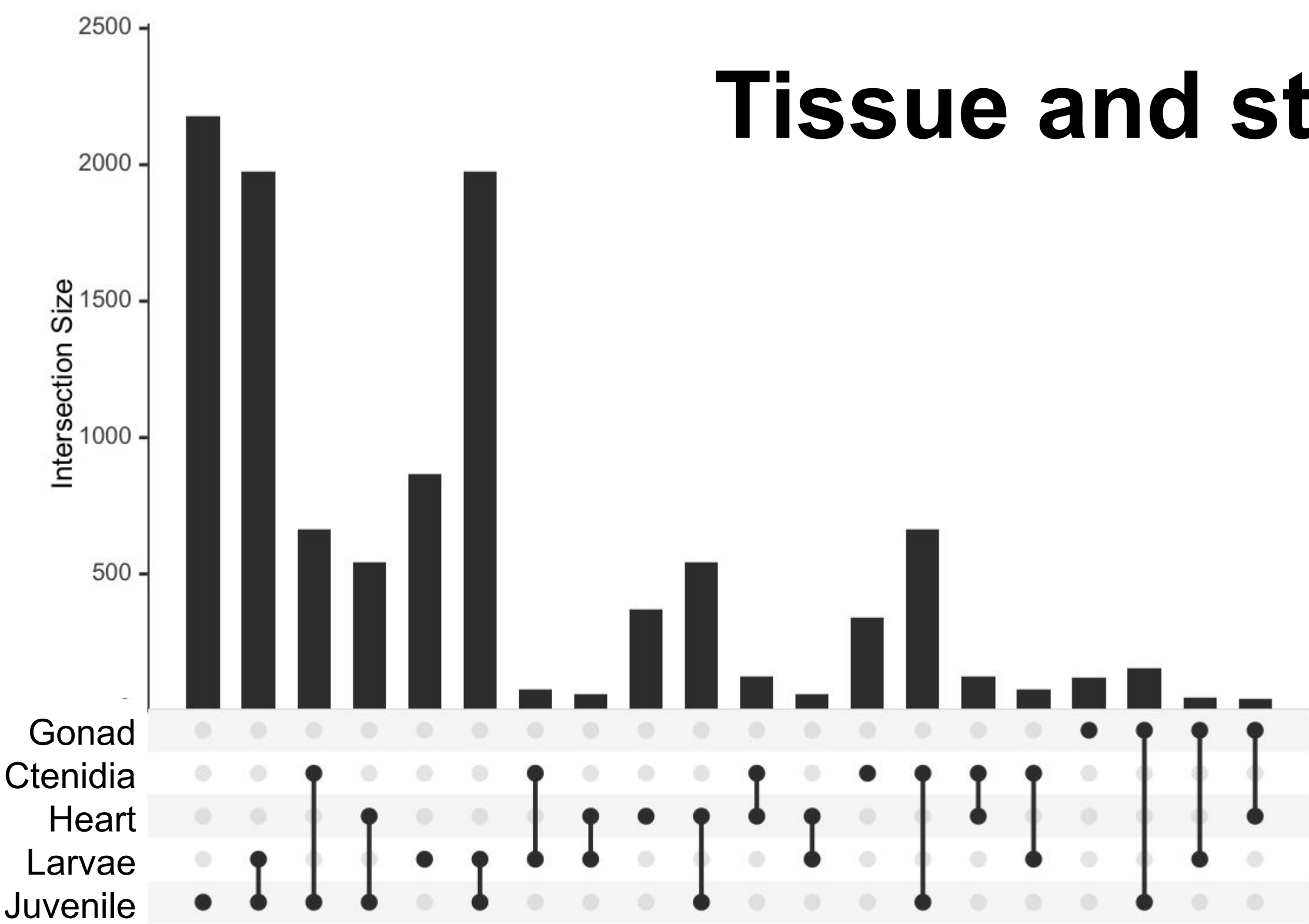
Olivia Cattau¹, Sam White¹, Hollie Putnam², Shelly Trigg¹, Mackenzie Gaverty³, Brent Vadopalas¹, Steven Roberts¹

¹School of Aquatic and Fishery Sciences, University of Washington. ²University of Rhode Island. ³Northwest Fisheries Science Center, NOAA.

Summary

- The Pacific Geoduck genome has been sequenced and assembled
- Deep coverage tissue specific transcriptomes are available
- Single-cell RNA-seq efforts provide new insight into developmental trajectories

Tissue and stage specific transcriptomes

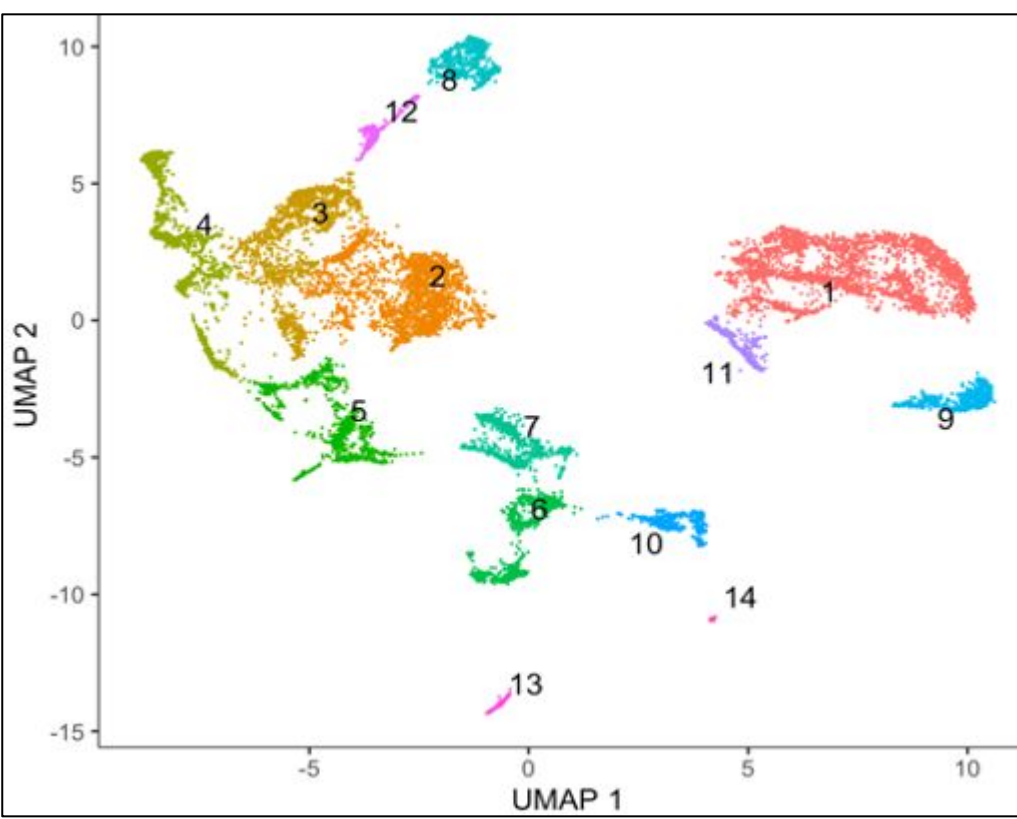


| Gene-set Descriptor | Number of Genes |
|--------------------------|-----------------|
| Ctenidia Unique | 340 |
| Gonad Unique | 119 |
| Heart Unique | 371 |
| Juvenile Unique | 2180 |
| Larvae Unique | 868 |
| Express in all libraries | 11294 |

Comparative Genomics

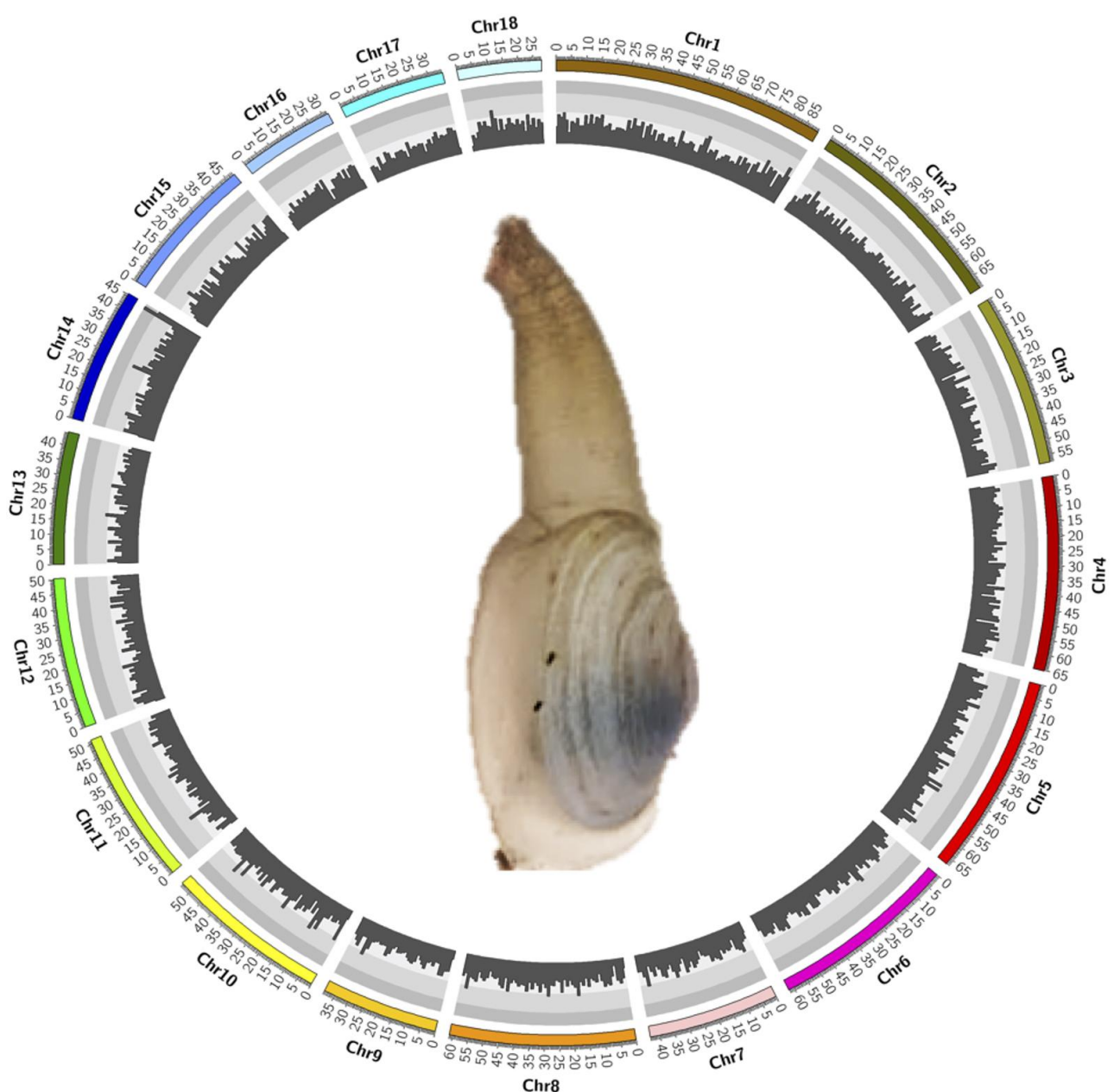
| Species | GenBank Assembly Accession | Genome size (Mbp) | Chromosomes | Number of Genes | GC content (%) | Reference |
|-------------------------|----------------------------|-------------------|-------------|-----------------|----------------|-----------------------------|
| <i>P. generosa</i> | GCA_902825435.1 | 942 | 18 | 34,947 | 33.78 | (Putnam et al. 2023) |
| <i>R. philippinarum</i> | GCA_009026015.1 | 1230 | 19 | 27,652 | 32.26 | (Mun et al. 2017) |
| <i>C. virginica</i> | GCA_002022765.4 | 685 | 11 | 38,826 | 34.82 | (Gómez-Chiarri et al. 2015) |
| <i>C. gigas</i> | GCA_902806645.1 | 648 | 10 | 38,296 | 33.50 | (Zhang et al. 2012) |

Single-Cell RNA-seq



Single-cell sequencing has been carried out on 17 hpf embryos with distinct cell clusters identified based on expression (above). Cell group 4 patterns suggest association with gonad development. Top marker genes for group 4 are listed and corresponding expression in tissue libraries (below).

| gene_id | gene_ID | ctenidia | heart | gonad |
|----------------|---|----------|-------|-------|
| PGEN_00g199430 | G2/mitotic-specific cyclin-B | 28 | 0 | 56 |
| PGEN_00g070070 | inner centromere protein | 3 | 0 | 2 |
| PGEN_00g089980 | G2/mitotic-specific cyclin-A | 6 | 0 | 9 |
| | histone-lysine N-methyltransferase | | | |
| PGEN_00g259100 | NSD2 | 7 | 4 | 10 |
| PGEN_00g054930 | sperm-specific protein PHI-2B/PHI-3 | 0 | 0 | 50 |
| PGEN_00g068740 | dentin sialoprophosphoprotein-like | 2 | 21 | 3 |
| PGEN_00g078980 | homeobox protein CDX-1 | 0 | 0 | 3 |
| PGEN_00g330580 | lymphocyte-specific helicase | 2 | 0 | 2 |
| | putative UDP-GlcNAc:betaGal | | | |
| PGEN_00g120520 | beta-1,3-N-acetylglucosaminyltransferase LOC100288842 | 2 | 0 | 2 |
| | metal-response element-binding | | | |
| PGEN_00g281640 | transcription factor 2-like | 3 | 1 | 1 |



Geoduck genome representation with chromosome name on the outermost track, followed by chromosome position in kilobase pairs (kbp), color bars indicating each chromosome, and gene density (genes/kbp) on the innermost track.

Support

Access to data, papers, and poster
github.com/RobertsLab/poster-geoduck-NSA-2023



Genome Statistics

| | | |
|----------------|--------------------------------|-------------|
| Genome | Total scaffold length (bp) | 942,353,201 |
| | Scaffold N50 (bp) | 57,743,597 |
| | GC content, N excluded (%) | 33.78 |
| Genes | Number of genes | 34,947 |
| | Mean gene length (bp) | 201 |
| | Mean coding region length (bp) | 201 |
| Exons | Mean number of exons per gene | 6.78 |
| | Mean length (bp) | 201 |
| | Total length (Mbp) | 47 |
| Introns | Genes with introns (%) | 72.04 |
| | Mean length (bp) | 2185 |
| | Total length (Mbp) | 338 |