**Message Box**

Audience:

* Committee members
  + Jon: principal investigator of the EAGER EecSeq project, mentor of Jacob Green, and professor of biological science in the evolution and marine biology specialization focusing on a seascape genomics approach to anthropogenic stress on the life history and connectivity of marine invertebrate populations
  + Marta: professor of fisheries, animal, and veterinary sciences researching the prevention and management of infectious disease in shellfish and finfish from a multidisciplinary approach
  + Dina: USDA research focusing on evaluation of wild and commercial oyster stocks and developing new approaches and resources for aquaculture genomics
  + Bryan: principal investigator of the science education and society research program focusing on teaching and learning of science within the social context

The Issue:

* Developing expressed exome sequencing method to assess genes under selection in any organism

The Problem:

* Sampling enough individuals in a study to assess genotype, phenotype, environment interactions
* Sampling the correct region of the genome that leads to phenotypic change
* The research budget of single PhD project or grant projects rarely supports the full development of producing well-annotated and complete genomes, transcriptomes, and proteomes
* Current reduced representation target sequencing tools require robust genomic resources and fail to capture variation in transcriptomic profiles that are divergent from the reference

The Solution:

* Transform the proof-of-concept EecSeq project into a readily available method to sequencing the exome of any organism.
* Optimize molecular lab protocol for probe length, probe diversity, and library hybridization
* Compare and validate EecSeq with traditional target capture methods, RNASeq, and WGS
* Develop a bioinformatic pipeline to analyze EecSeq sequencing data accessible to any evolutionary biologist

The Benefit:

* EecSeq probes are created specifically for the study of interest removing reference bias from other target capture approaches
* EecSeq probes are designed to reflect specific stressor studies or unique populations and applied across other populations or life history
* Provide the background information allowing researchers to make an informed decision about what type of next generation sequencing approach they could take to best balance research budget and experimental design
* Design a robust and tested molecular protocol that will guide probe creation, synthesis of capture pool libraries and hybridization of probes to target regions
* Create a fully develop bioinformatic pipeline that i

The So What:

**Contents**

**Title of the Study**

Development of a tool to sequence the exome of any organism rapidly and cost-effectively

**Statement of the Problem**

Next generation sequencing (NGS) techniques, like whole genome sequencing and RNA sequencing, have greatly increased our capacity to explore issues of genetic conservation and adaptation, but are costly and time-intensive to accurately assess regional population diversity and adaptation. Target capture sequencing allow us to reduce cost and selectively enrich only the specific expressed exons, but this design requires a robust and well-annotated genome. We are developing a novel sequencing technique called Expressed Exome Capture Sequencing (EecSeq) to sequence the exome of any organism rapidly and cost-effectively. Our research will transform a proof of concept into a readily available approach by improving and optimizing the molecular laboratory workflow, comparing and validating EecSeq against other NGS approaches, and developing a reproducible and open-source bioinformatic pipeline to analyze EecSeq sequencing data accessible to any evolutionary biologist. This project will provide researchers with the information, justification, and protocols needed to implement EecSeq for their own questions.

**Justification for and Significance of the Study**

Paragraph 1: Next generation sequencing techniques have greatly increased our capacity to explore issues of genetic conservation and adaptation.

Paragraph 2: Genomic technologies like whole genome sequencing and RNAseq are costly and time-intensive to accurately assess even regional population diversity.

Paragraph 3: Reduced representation sequencing methods such as RADseq are relatively easy to implement and have high throughput. But due to the way RADseq samples the genome scientists are unable to explore genes under selection.

Paragraph 4: Targeted capture approaches are an alternative that selectively enrich specific regions of the genome with designed genetic probes. But this approach requires

Paragraph 5: What is EecSeq and how does it provides a solution to the pitfalls of other NGS approaches?

Paragraph 6: What are the benefits and importance of EecSeq as a sequencing technology to the broader field of biology?

Paragraph 7: Why use the oyster as an example of non-model species for the implementation and testing of EecSeq

Paragraph 8: Research Statement with objective and broad hypothesis

**Methodology or Procedures**

**Research Design**

Objective 1: Developing an open source bioinformatic pipeline to analyze EecSeq data

* Objective 1a: Design and implement a de novo assembly method using only captured exome reads
  + Source of data:
    - Hypoxia experiment of oyster adults
    - Coastal acidification and sewage effluent exposed oyster larvae
    - Oyster genome
    - Note: all data obtained from objective 2 and 3 will be process through this pipeline
  + Methods
    - Read quality control
      * Fastp
      * Multiqc
      * Normalization
      * Unique read filtering
    - De novo Assembly gDNA capture reads
      * Oases
      * Transabyss
      * Trinity
    - Post-processing clustering
      * CD-hit
    - Assessment
      * Assembly statistics
        + Transrate
        + rnaQuast
        + trinity stats
      * BUSCO scores
      * Mapping read quality to de novo assembly
      * Mapping read quality to reference assembly
      * Exome capture efficiency
      * Exome capture specificity
* Objective 1b: Design and implement a hybrid assembly method utilizing cDNA probe reads and captured exome reads
  + Source of data:
    - Hypoxia experiment of oyster adults
    - Coastal acidification and sewage effluent exposed oyster larvae
    - Oyster genome
    - Note: all data obtained from objective 2 and 3 will be process through this pipeline
  + Methods
    - Read quality control
      * Fastp
      * Multiqc
      * Normalization
      * Unique read filtering
    - De novo cDNA Assembly
      * Oases
      * Transabyss
      * Trinity
    - De novo Assembly gDNA capture reads
      * Oases
      * Transabyss
      * Trinity
    - Post-processing clustering
      * CD-hit
    - Assessment
      * Assembly statistics
        + Transrate
        + rnaQuast
        + trinity stats
      * BUSCO scores
      * Mapping read quality to de novo assembly
      * Mapping read quality to reference assembly
      * Exome capture efficiency
      * Exome capture specificity
    - Final Assembly
* Objective 1c: Contrast and compare de novo and reference-based bioinformatic methods across all experimental datasets

Objective 2: Optimize molecular lab protocol for probe length, probe diversity, and library hybridization

* Objective 2a: Quantify the role of capture probe and capture pool insert length in exome capture and efficiency
  + Source of data
    - NB oyster project 48 individuals, 10 from 3 locations and 9 from 2 locations in Narragansett Bay
    - Extract DNA/RNA from 10 individuals 2 from each population
  + Methods
    - Molecular
      * DNA/RNA extraction
      * mRNA library synthesis
        + alter mRNA fragment lengths to 150, 250, 350 bp with mRNA hyper prep kit protocol
      * Probe Synthesis
      * gDNA library synthesis
        + alter gDNA to 150 bp and 350 bp via sonication
      * Hybridization
      * Sequencing
    - Assessment
      * Quick
        + Read filtering
        + Map reads to oyster genome
        + Mapping quality
        + Exome capture specificity
        + Exome capture efficiency
      * Long:
        + De novo pipeline
        + Hybrid pipeline
* Objective 2b: Determine the optimal level of probe and capture pool diversity to assess drop out of divergent alleles
  + Source of data
    - NB oyster project 48 individuals, 10 from 3 locations and 9 from 2 locations in Narragansett Bay
    - Extract DNA/RNA from 48 individuals, max number from each population
  + Methods
    - Molecular
      * DNA/RNA extraction
      * mRNA library synthesis
        + create pools 8, 16, 32, 48 of mRNA samples
      * Probe Synthesis
      * gDNA library synthesis
        + create pools 8, 16, 32, 48 of gRNA samples
      * Hybridization
      * Sequencing
    - Assessment
      * Quick
        + Read filtering
        + Map reads to oyster genome
        + Mapping quality
        + Exome capture specificity
        + Exome capture efficiency
      * Long:
        + De novo pipeline
        + Hybrid pipeline
* Objective 2c: Optimize capture probe to capture pool insert libraries ratio for efficient hybridization and capture
  + Source of data
    - NB oyster project 48 individuals, 10 from 3 locations and 9 from 2 locations in Narragansett Bay
    - Recommendations from Objective 1 and 2 will be integrated into Objective 3
  + Methods
    - Molecular
      * DNA/RNA extraction
      * mRNA library synthesis
        + create pools 8, 16, 32, 48 of mRNA samples
      * Probe Synthesis
      * gDNA library synthesis
        + create pools 8, 16, 32, 48 of gRNA samples
      * Hybridization
      * Sequencing
    - Assessment
      * Quick
        + Read filtering
        + Map reads to oyster genome
        + Mapping quality
        + Exome capture specificity
        + Exome capture efficiency
      * Long:
        + De novo pipeline
        + Hybrid pipeline

Objective 3: Compare and validate EecSeq with traditional target capture methods, RNASeq, and WGS

* Objective 3.1: Compare commercial exome capture probes designed from the Eastern Oyster Reference genome to EecSeq probes using identical samples
  + Source of Data
    - Commercial [probes](https://arborbiosci.com/genomics/targeted-sequencing/mybaits/?gclid=CjwKCAjwur-SBhB6EiwA5sKtjkVmgTTEy39b66Q9RcwIxwii1-9xw5U3Qg1XGZyc6P5acBPdQooY_hoCXaUQAvD_BwE) design from the eastern oyster genome
    - EecSeq probes created from
      * NB oyster project
      * Hypoxia
      * CASE
      * Data from Dina
      * 6 individuals, 3 replicate captures
  + Methods
    - See Objective 2 results
* Objective 3.2: Validate genotyping accuracy and test ability to target genes of interest by utilizing refence individuals form the Eastern Oyster genome project
  + **Need to rethink this experimental design**
* Objective 3.3: Compare the accuracy of pooled EecSeq to pooled RNA seq data for allele frequency estimation using replicate pools of known genotyped individuals
  + **Need to rethink this experimental design**

**General characteristics about the study population**

[Add information about oyster RI oyster populations and the oyster genome here]

**Location or setting calendar events**

[Narragansett Bay and oyster populations; calendar events for study]

**Sampling Design and Procedures**

[Overall design from oyster sampling, dissection, genetic material acquisition, EecSeq molecular protocol, sequencing, bioinformatic analysis]

**Description of tool used for collecting and analyzing data**

[Describe EecSeq molecular kits and reagents? Bioinformatic tools?]

**Definition of the most important terms and concepts**

[List of concepts and terms]

**Data processing procedures and procedures of data analysis as appropriate**

[EecSeq protocol/Bioinformatic pipeline]

**Literature Cited**

1. Abrouk M, Athiyannan N, Müller T, Pailles Y, Stritt C, Roulin AC, Chu C, Liu S, Morita T, Handa H, et al. 2021. Population genomics and haplotype analysis in spelt and bread wheat identifies a gene regulating glume color. Commun. Biol. 4:1–11. doi:10.1038/s42003-021-01908-6.
2. Alex Buerkle C, Gompert Z. 2013. Population genomics based on low coverage sequencing: How low should we go? Mol. Ecol. 22:3028–3035. doi:10.1111/mec.12105.
3. Alexander MA, Scott JD, Friedland KD, Mills KE, Nye JA, Pershing AJ, Thomas AC. 2018. Projected sea surface temperatures over the 21st century: Changes in the mean, variability and extremes for large marine ecosystem regions of Northern Oceans. Elem Sci Anth 6:9. doi:10.1525/elementa.191.
4. Ali OA, O’Rourke SM, Amish SJ, Meek MH, Luikart G, Jeffres C, Miller MR. 2016. Rad capture (Rapture): Flexible and efficient sequence-based genotyping. Genetics 202:389–400. doi:10.1534/genetics.115.183665.
5. Allendorf FW, Hohenlohe PA, Luikart G. 2010. Genomics and the future of conservation genetics. Nat. Rev. Genet. 11:697–709. doi:10.1038/nrg2844.
6. Almany GR, Planes S, Thorrold SR, Berumen ML, Bode M, Saenz-Agudelo P, Bonin MC, Frisch AJ, Harrison HB, Messmer V, et al. 2017. Larval fish dispersal in a coral-reef seascape. Nat. Ecol. Evol. 1:1–7. doi:10.1038/s41559-017-0148.
7. Anand S, Mangano E, Barizzone N, Bordoni R, Sorosina M, Clarelli F, Corrado L, Boneschi FM, D’Alfonso S, De Bellis G. 2016. Next generation sequencing of pooled samples: Guideline for variants’ filtering. Sci. Rep. 6:1–9. doi:10.1038/srep33735.
8. Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS One 3:1–7. doi:10.1371/journal.pone.0003376.
9. Baruzzo G, Hayer KE, Kim EJ, DI Camillo B, Fitzgerald GA, Grant GR. 2017. Simulation-based comprehensive benchmarking of RNA-seq aligners. Nat. Methods 14:135–139. doi:10.1038/nmeth.4106.
10. Benestan L, Quinn BK, Maaroufi H, Laporte M, Clark FK, Greenwood SJ, Rochette R, Bernatchez L. 2016. Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (Homarus americanus). Mol. Ecol. 25:5073–5092. doi:10.1111/mec.13811.
11. Bernatchez S, Xuereb A, Laporte M, Benestan L, Steeves R, Laflamme M, Bernatchez L, Mallet MA. 2019. Seascape genomics of eastern oyster (Crassostrea virginica) along the Atlantic coast of Canada. Evol. Appl. 12:587–609. doi:10.1111/eva.12741.
12. Bernatchez S, Xuereb A, Laporte M, Benestan L, Steeves R, Laflamme M, Bernatchez L, Mallet MA. 2019. Seascape genomics of eastern oyster (Crassostrea virginica) along the Atlantic coast of Canada. Evol. Appl. 12:587–609. doi:10.1111/eva.12741.
13. Bi K, Vanderpool D, Singhal S, Linderoth T, Moritz C, Good JM. 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. BMC Genomics 13. doi:10.1186/1471-2164-13-403.
14. Booker TR, Yeaman S, Whitlock MC. 2020. Variation in recombination rate affects detection of outliers in genome scans under neutrality. Mol. Ecol. 29:4274–4279. doi:10.1111/mec.15501.
15. Booker TR, Yeaman S, Whitlock MC. 2020. Variation in recombination rate affects detection of outliers in genome scans under neutrality. Mol. Ecol. 29:4274–4279. doi:10.1111/mec.15501.
16. Brennan RS, Garrett AD, Huber KE, Hargarten H, Pespeni MH. 2019. Rare genetic variation and balanced polymorphisms are important for survival in global change conditions. Proc. R. Soc. B Biol. Sci. 286. doi:10.1098/rspb.2019.0943.
17. Brind’Amour A, Mahévas S, Legendre P, Bellanger L. 2018. Application of Moran Eigenvector Maps (MEM) to irregular sampling designs. Spat. Stat. 26:56–68. doi:10.1016/j.spasta.2018.05.004.
18. Capblancq T, Forester BR. 2021. Redundancy analysis: A Swiss Army Knife for landscape genomics. Methods Ecol. Evol. 12:2298–2309. doi:10.1111/2041-210X.13722.
19. Caron DA, Alexander H, Allen AE, Archibald JM, Armbrust EV, Bachy C, Bell CJ, Bharti A, Dyhrman ST, Guida SM, et al. 2017. Probing the evolution, ecology and physiology of marine protists using transcriptomics. Nat. Rev. Microbiol. 15:6–20. doi:10.1038/nrmicro.2016.160.
20. Chapman RW, Mancia A, Beal M, Veloso A, Rathburn C, Blair A, Holland AF, Warr GW, Didinato G, Sokolova IM, et al. 2011. The transcriptomic responses of the eastern oyster, Crassostrea virginica, to environmental conditions. Mol. Ecol. 20:1431–1449. doi:10.1111/j.1365-294X.2011.05018.x.
21. Chapman RW, Mancia A, Beal M, Veloso A, Rathburn C, Blair A, Sanger D, Holland AF, Warr GW, Didonato G. 2009. A transcriptomic analysis of land-use impacts on the oyster, Crassostrea virginica, in the South Atlantic bight. Mol. Ecol. 18:2415–2425. doi:10.1111/j.1365-294X.2009.04194.x.
22. Cherry KM, Qian L. 2018. Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. Nature 559:370–388. doi:10.1038/s41586-018-0289-6.
23. Christmas MJ, Biffin E, Breed MF, Lowe AJ. 2016. Finding needles in a genomic haystack: targeted capture identifies clear signatures of selection in a nonmodel plant species. Mol. Ecol. 25:4216–4233. doi:10.1111/mec.13750.
24. Combs M, Puckett EE, Richardson J, Mims D, Munshi-South J. 2018. Spatial population genomics of the brown rat (Rattus norvegicus) in New York City. Mol. Ecol. 27:83–98. doi:10.1111/mec.14437.
25. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, Szcześniak MW, Gaffney DJ, Elo LL, Zhang X, et al. 2016. A survey of best practices for RNA-seq data analysis. Genome Biol. 17:1–19. doi:10.1186/s13059-016-0881-8.
26. Coutelier M, Hammer MB, Stevanin G, Monin ML, Davoine CS, Mochel F, Labauge P, Ewenczyk C, Ding J, Gibbs JR, et al. 2018. Efficacy of exome-targeted capture sequencing to detect mutations in known cerebellar ataxia genes. JAMA Neurol. 75:591–599. doi:10.1001/jamaneurol.2017.5121.
27. Cowley MJ, Liu YC, Oliver KL, Carvill G, Myers CT, Gayevskiy V, Delatycki M, Vlaskamp DRM, Zhu Y, Mefford H, et al. 2019. Reanalysis and optimisation of bioinformatic pipelines is critical for mutation detection. Hum. Mutat. 40:374–379. doi:10.1002/humu.23699.
28. Davey JL, Blaxter MW. 2010. RADseq: Next-generation population genetics. Brief. Funct. Genomics 9:416–423. doi:10.1093/bfgp/elq031.
29. Davey JW, Cezard T, Fuentes-Utrilla P, Eland C, Gharbi K, Blaxter ML. 2013. Special features of RAD Sequencing data: Implications for genotyping. Mol. Ecol. 22:3151–3164. doi:10.1111/mec.12084.
30. De Mita S, Thuillet AC, Gay L, Ahmadi N, Manel S, Ronfort J, Vigouroux Y. 2013. Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. Mol. Ecol. 22:1383–1399. doi:10.1111/mec.12182.
31. De Oliveira AL, Wollesen T, Kristof A, Scherholz M, Redl E, Todt C, Bleidorn C, Wanninger A. 2016. Comparative transcriptomics enlarges the toolkit of known developmental genes in mollusks. BMC Genomics 17. doi:10.1186/s12864-016-3080-9.
32. De Villemereuil P, Frichot É, Bazin É, François O, Gaggiotti OE. 2014. Genome scan methods against more complex models: When and how much should we trust them? Mol. Ecol. 23:2006–2019. doi:10.1111/mec.12705.
33. Dolby GA, Ellingson RA, Findley LT, Jacobs DK. 2018. How sea level change mediates genetic divergence in coastal species across regions with varying tectonic and sediment processes. Mol. Ecol. 27:994–1011. doi:10.1111/mec.14487.
34. Dong C, Zhang L, Chen Z, Xia C, Gu Y, Wang J, Li D, Xie Z, Zhang Q, Zhang X, et al. 2020. Combining a New Exome Capture Panel With an Effective varBScore Algorithm Accelerates BSA-Based Gene Cloning in Wheat. Front. Plant Sci. 11:1–12. doi:10.3389/fpls.2020.01249.
35. Dumbauld BR, Ruesink JL, Rumrill SS. 2009. The ecological role of bivalve shellfish aquaculture in the estuarine environment: A review with application to oyster and clam culture in West Coast (USA) estuaries. Aquaculture 290:196–223. doi:10.1016/j.aquaculture.2009.02.033.
36. Epelboin Y, Quintric L, Guévélou E, Boudry P, Pichereau V, Corporeau C. 2016. The kinome of Pacific oyster Crassostrea gigas, its expression during development and in response to environmental factors. PLoS One 11:1–24. doi:10.1371/journal.pone.0155435.
37. Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Syst. Biol. 61:717–726. doi:10.1093/sysbio/sys004.
38. Flanagan SP, Jones AG. 2018. Substantial differences in bias between single-digest and double-digest RAD-seq libraries: A case study. Mol. Ecol. Resour. 18:264–280. doi:10.1111/1755-0998.12734.
39. Gasc C, Peyretaillade E, Peyret P. 2016. Sequence capture by hybridization to explore modern and ancient genomic diversity in model and nonmodel organisms. Nucleic Acids Res. 44:4504–4518. doi:10.1093/nar/gkw309.
40. Geniza M, Jaiswal P. 2017. Tools for building de novo transcriptome assembly. Curr. Plant Biol. 11–12:41–45. doi:10.1016/j.cpb.2017.12.004.
41. Gibert P, Debat V, Ghalambor CK. 2019. Phenotypic plasticity, global change, and the speed of adaptive evolution. Curr. Opin. Insect Sci. 35:34–40. doi:10.1016/j.cois.2019.06.007.
42. Glover N, Dessimoz C, Ebersberger I, Forslund SK, Gabaldón T, Huerta-Cepas J, Martin MJ, Muffato M, Patricio M, Pereira C, et al. 2019. Advances and Applications in the Quest for Orthologs. Mol. Biol. Evol. 36:2157–2164. doi:10.1093/molbev/msz150.
43. Golestan Hashemi FS, Razi Ismail M, Rafii Yusop M, Golestan Hashemi MS, Nadimi Shahraki MH, Rastegari H, Miah G, Aslani F. 2018. Intelligent mining of large-scale bio-data: Bioinformatics applications. Biotechnol. Biotechnol. Equip. 32:10–29. doi:10.1080/13102818.2017.1364977.
44. Grover CE, Salmon A, Wendel JF. 2012. Targeted sequence capture as a powerful tool for evolutionary analysis. Am. J. Bot. 99:312–319. doi:10.3732/ajb.1100323.
45. Hart MW, Marko PB. 2010. It’s about time: Divergence, demography, and the evolution of developmental modes in marine invertebrates. Integr. Comp. Biol. 50:643–661. doi:10.1093/icb/icq068.
46. Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, Poss ML, Reed LK, Storfer A, Whitlock MC. 2016. Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. Am. Nat. 188:379–397. doi:10.1086/688018.
47. Hodel RGJ, Chen S, Payton AC, McDaniel SF, Soltis P, Soltis DE. 2017. Adding loci improves phylogeographic resolution in red mangroves despite increased missing data: Comparing microsatellites and RAD-Seq and investigating loci filtering. Sci. Rep. 7:1–14. doi:10.1038/s41598-017-16810-7.
48. Hoffberg SL, Kieran TJ, Catchen JM, Devault A, Faircloth BC, Mauricio R, Glenn TC. 2016. RADcap: sequence capture of dual-digest RADseq libraries with identifiable duplicates and reduced missing data. Mol. Ecol. Resour. 16:1264–1278. doi:10.1111/1755-0998.12566.
49. Hölzer M, Marz M. 2019. De novo transcriptome assembly: A comprehensive cross-species comparison of short-read RNA-Seq assemblers. Gigascience 8:1–16. doi:10.1093/gigascience/giz039.
50. Honaas LA, Wafula EK, Wickett NJ, Der JP, Zhang Y, Edger PP, Altman NS, Chris Pires J, Leebens-Mack JH, DePamphilis CW. 2016. Selecting superior de novo transcriptome assemblies: Lessons learned by leveraging the best plant genome. PLoS One 11:1–42. doi:10.1371/journal.pone.0146062.
51. Jones MR, Good JM. 2016. Targeted capture in evolutionary and ecological genomics. Mol. Ecol. 25:185–202. doi:10.1111/mec.13304.
52. Kersey PJ, Allen JE, Allot A, Barba M, Boddu S, Bolt BJ, Carvalho-Silva D, Christensen M, Davis P, Grabmueller C, et al. 2018. Ensembl Genomes 2018: An integrated omics infrastructure for non-vertebrate species. Nucleic Acids Res. 46:D802–D808. doi:10.1093/nar/gkx1011.
53. LaCava MEF, Aikens EO, Megna LC, Randolph G, Hubbard C, Buerkle CA. 2020. Accuracy of de novo assembly of DNA sequences from double-digest libraries varies substantially among software. Mol. Ecol. Resour. 20:360–370. doi:10.1111/1755-0998.13108.
54. Lal MM, Southgate PC, Jerry DR, Bosserelle C, Zenger KR. 2017. Swept away: Ocean currents and seascape features influence genetic structure across the 18,000 Km Indo-Pacific distribution of a marine invertebrate, the black-lip pearl oyster Pinctada margaritifera. BMC Genomics 18:1–21. doi:10.1186/s12864-016-3410-y.
55. Langa J, Estonba A, Conklin D. 2020. EXFI: Exon and splice graph prediction without a reference genome. Ecol. Evol. 10:8880–8893. doi:10.1002/ece3.6587.
56. Lelieveld SH, Veltman JA, Gilissen C. 2016. Novel bioinformatic developments for exome sequencing. Hum. Genet. 135:603–614. doi:10.1007/s00439-016-1658-6.
57. Li B, Fillmore N, Bai Y, Collins M, Thomson JA, Stewart R, Dewey CN. 2014. Evaluation of de novo transcriptome assemblies from RNA-Seq data. Genome Biol. 15:1–21. doi:10.1186/s13059-014-0553-5.
58. Liao X, Li M, Zou Y, Wu FX, Yi-Pan, Wang J. 2019. Current challenges and solutions of de novo assembly. Quant. Biol. 7:90–109. doi:10.1007/s40484-019-0166-9.
59. Lim HC, Braun MJ. 2016. High-throughput SNP genotyping of historical and modern samples of five bird species via sequence capture of ultraconserved elements. Mol. Ecol. Resour. 16:1204–1223. doi:10.1111/1755-0998.12568.
60. Linck EB, Hanna ZR, Sellas A, Dumbacher JP. 2017. Evaluating hybridization capture with RAD probes as a tool for museum genomics with historical bird specimens. Ecol. Evol. 7:4755–4767. doi:10.1002/ece3.3065.
61. Lindegren M, Holt BG, MacKenzie BR, Rahbek C. 2018. A global mismatch in the protection of multiple marine biodiversity components and ecosystem services. Sci. Rep. 8:1–8. doi:10.1038/s41598-018-22419-1.
62. Liu J, Li G, Chang Z, Yu T, Liu B, McMullen R, Chen P, Huang X. 2016. BinPacker: Packing-Based De Novo Transcriptome Assembly from RNA-seq Data. PLoS Comput. Biol. 12:1–15. doi:10.1371/journal.pcbi.1004772.
63. Lotterhos KE. 2019. The effect of neutral recombination variation on genome scans for selection. G3 Genes, Genomes, Genet. 9:1851–1867. doi:10.1534/g3.119.400088.
64. Lotterhos KE, Card DC, Schaal SM, Wang L, Collins C, Verity B. 2017. Composite measures of selection can improve the signal-to-noise ratio in genome scans. Methods Ecol. Evol. 8:717–727. doi:10.1111/2041-210X.12774.
65. Lotterhos KE, Whitlock MC. 2015. The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. Mol. Ecol. 24:1031–1046. doi:10.1111/mec.13100.
66. Lotterhos KE, Whitlock MC. 2014. Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. Mol. Ecol. 23:2178–2192. doi:10.1111/mec.12725.
67. Lowry DB, Hoban S, Kelley JL, Lotterhos KE, Reed LK, Antolin MF, Storfer A. 2017. Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. Mol. Ecol. Resour. 17:142–152. doi:10.1111/1755-0998.12635.
68. Maisano Delser P, Corrigan S, Hale M, Li C, Veuille M, Planes S, Naylor G, Mona S. 2016. Population genomics of C. melanopterus using target gene capture data: Demographic inferences and conservation perspectives. Sci. Rep. 6:1–12. doi:10.1038/srep33753.
69. Marko PB, Hart MW. 2011. The complex analytical landscape of gene flow inference. Trends Ecol. Evol. 26:448–456. doi:10.1016/j.tree.2011.05.007.
70. Matz M V. 2018. Fantastic Beasts and How To Sequence Them: Ecological Genomics for Obscure Model Organisms. Trends Genet. 34:121–132. doi:10.1016/j.tig.2017.11.002.
71. McKinney GJ, Larson WA, Seeb LW, Seeb JE. 2017. RADseq provides unprecedented insights into molecular ecology and evolutionary genetics: comment on Breaking RAD by Lowry et al. (2016). Mol. Ecol. Resour. 17:356–361. doi:10.1111/1755-0998.12649.
72. Micheletti SJ, Matala AR, Matala AP, Narum SR. 2018. Landscape features along migratory routes influence adaptive genomic variation in anadromous steelhead (Oncorhynchus mykiss). Mol. Ecol. 27:128–145. doi:10.1111/mec.14407.
73. Milbury CA, Lee JC, Cannone JJ, Gaffney PM, Gutell RR. 2010. Fragmentation of the large subunit ribosomal RNA gene in oyster mitochondrial genomes. BMC Genomics 11:1–17. doi:10.1186/1471-2164-11-485.
74. Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Res. 17:240–248. doi:10.1101/gr.5681207.
75. O’Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS. 2018. These aren’t the loci you’e looking for: Principles of effective SNP filtering for molecular ecologists. Mol. Ecol. 27:3193–3206. doi:10.1111/mec.14792.
76. O’Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS. 2018. These aren’t the loci you’e looking for: Principles of effective SNP filtering for molecular ecologists. Mol. Ecol. 27:3193–3206. doi:10.1111/mec.14792.
77. Patnaik BB, Wang TH, Kang SW, Hwang HJ, Park SY, Park EB, Chung JM, Song DK, Kim C, Kim S, et al. 2016. Sequencing, de novo assembly, and annotation of the transcriptome of the endangered freshwater pearl bivalve, Cristaria plicata, provides novel insights into functional genes and marker discovery. PLoS One 11:1–28. doi:10.1371/journal.pone.0148622.
78. Payne A, Holmes N, Clarke T, Munro R, Debebe B, Loose M. 2020. Nanopore adaptive sequencing for mixed samples, whole exome capture and targeted panels. bioRxiv:1–28.
79. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat. Protoc. 11:1650–1667. doi:10.1038/nprot.2016.095.
80. Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One 7. doi:10.1371/journal.pone.0037135.
81. Portnoy DS, Puritz JB, Hollenbeck CM, Gelsleichter J, Chapman D, Gold JR. 2015. Selection and sex-biased dispersal in a coastal shark: The influence of philopatry on adaptive variation. Mol. Ecol. 24:5877–5885. doi:10.1111/mec.13441.
82. Puritz JB, Gold JR, Portnoy DS. 2016. Fine-scale partitioning of genomic variation among recruits in an exploited fishery: Causes and consequences. Sci. Rep. 6:1–6. doi:10.1038/srep36095.
83. Puritz JB, Keever CC, Addison JA, Barbosa SS, Byrne M, Hart MW, Grosberg RK, Toonen RJ. 2017. Life-history predicts past and present population connectivity in two sympatric sea stars. Ecol. Evol. 7:3916–3930. doi:10.1002/ece3.2938.
84. Puritz JB, Lotterhos KE. 2018. Expressed exome capture sequencing: A method for cost-effective exome sequencing for all organisms. Mol. Ecol. Resour. 18:1209–1222. doi:10.1111/1755-0998.12905.
85. Puritz JB, Matz M V., Toonen RJ, Weber JN, Bolnick DI, Bird CE. 2014. Demystifying the RAD fad. Mol. Ecol. 23:5937–5942. doi:10.1111/mec.12965.
86. Puritz JB, Toonen RJ. 2011. Coastal pollution limits pelagic larval dispersal. Nat. Commun. 2. doi:10.1038/ncomms1238.
87. Quattrini AM, Faircloth BC, Dueñas LF, Bridge TCL, Brugler MR, Calixto-Botía IF, DeLeo DM, Forêt S, Herrera S, Lee SMY, et al. 2018. Universal target-enrichment baits for anthozoan (Cnidaria) phylogenomics: New approaches to long-standing problems. Mol. Ecol. Resour. 18:281–295. doi:10.1111/1755-0998.12736.
88. Rellstab C, Dauphin B, Zoller S, Brodbeck S, Gugerli F. 2019. Using transcriptome sequencing and pooled exome capture to study local adaptation in the giga-genome of Pinus cembra. Mol. Ecol. Resour. 19:536–551. doi:10.1111/1755-0998.12986.
89. RI Coastal Resources Management Council. 2014. Shellfish management plan. RI Gen. Laws.Richmond JQ, Backlin AR, Galst-Cavalcante C, O’Brien JW, Fisher RN. 2018. Loss of dendritic connectivity in southern California’s urban riverscape facilitates decline of an endemic freshwater fish. Mol. Ecol. 27:369–386. doi:10.1111/mec.14445.
90. Russell PH, Johnson RL, Ananthan S, Harnke B, Carlson NE. 2018. A large-scale analysis of bioinformatics code on GitHub. PLoS One 13:1–35. doi:10.1371/journal.pone.0205898.
91. Schiebelhut LM, Puritz JB, Dawson MN. 2018. Decimation by sea star wasting disease and rapid genetic change in a keystone species, Pisaster ochraceus. Proc. Natl. Acad. Sci. U. S. A. 115:7069–7074. doi:10.1073/pnas.1800285115.
92. Schlötterer C, Tobler R, Kofler R, Nolte V. 2014. Sequencing pools of individuals-mining genome-wide polymorphism data without big funding. Nat. Rev. Genet. 15:749–763. doi:10.1038/nrg3803.
93. Schmid S, Genevest R, Gobet E, Suchan T, Sperisen C, Tinner W, Alvarez N. 2017. HyRAD-X, a versatile method combining exome capture and RAD sequencing to extract genomic information from ancient DNA. Methods Ecol. Evol. 8:1374–1388. doi:10.1111/2041-210X.12785.
94. Schmitz-Abe K, Li Q, Rosen SM, Nori N, Madden JA, Genetti CA, Wojcik MH, Ponnaluri S, Gubbels CS, Picker JD, et al. 2019. Unique bioinformatic approach and comprehensive reanalysis improve diagnostic yield of clinical exomes. Eur. J. Hum. Genet. 27:1398–1405. doi:10.1038/s41431-019-0401-x.
95. Schweizer RM, VonHoldt BM, Harrigan R, Knowles JC, Musiani M, Coltman D, Novembre J, Wayne RK. 2016. Genetic subdivision and candidate genes under selection in North American grey wolves. Mol. Ecol. 25:380–402. doi:10.1111/mec.13364.
96. Selkoe KA, D’Aloia CC, Crandall ED, Iacchei M, Liggins L, Puritz JB, Von Der Heyden S, Toonen RJ. 2016. A decade of seascape genetics: Contributions to basic and applied marine connectivity. Mar. Ecol. Prog. Ser. 554:1–19. doi:10.3354/meps11792.
97. Shafer ABA, Peart CR, Tusso S, Maayan I, Brelsford A, Wheat CW, Wolf JBW. 2017. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. Methods Ecol. Evol. 8:907–917. doi:10.1111/2041-210X.12700.
98. Shih BB, Angus T, Barnett MMW, Chen S, Summers KKM, Klein K, Faulkner GJG, Saini HKH, Watson M, Dongen S Van, et al. 2018. Visualisation and analysis of RNA-Seq assembly graphs. bioRxiv:1–14. doi:10.1101/409573.
99. Silliman K. 2019. Population structure, genetic connectivity, and adaptation in the Olympia oyster (Ostrea lurida) along the west coast of North America. Evol. Appl. 12:923–939. doi:10.1111/eva.12766.
100. Smith BT, Harvey MG, Faircloth BC, Glenn TC, Brumfield RT. 2014. Target capture and massively parallel sequencing of ultraconserved elements for comparative studies at shallow evolutionary time scales. Syst. Biol. 63:83–95. doi:10.1093/sysbio/syt061.
101. Smukowski CS, Noor MAF. 2011. Recombination rate variation in closely related species. Heredity (Edinb). 107:496–508. doi:10.1038/hdy.2011.44.Sohn J Il, Nam JW. 2018. The present and future of de novo whole-genome assembly. Brief. Bioinform. 19:23–40. doi:10.1093/bib/bbw096.
102. Suchan T, Pitteloud C, Gerasimova NS, Kostikova A, Schmid S, Arrigo N, Pajkovic M, Ronikier M, Alvarez N. 2016. Hybridization capture using RAD probes (hyRAD), a new tool for performing genomic analyses on collection specimens. PLoS One 11:1–22. doi:10.1371/journal.pone.0151651.
103. Syring J V., Tennessen JA, Jennings TN, Wegrzyn J, Scelfo-Dalbey C, Cronn R. 2016. Targeted capture sequencing in whitebark pine reveals range-wide demographic and adaptive patterns despite challenges of a large, repetitive genome. Front. Plant Sci. 7:1–15. doi:10.3389/fpls.2016.00484.
104. Therkildsen NO, Palumbi SR. 2017. Practical low-coverage genome wide sequencing of hundreds of individually barcoded samples for population and evolutionary genomics in nonmodel species. Mol. Ecol. Resour. 17:194–208. doi:10.1111/1755-0998.12593.
105. Tian S, Yan H, Klee EW, Kalmbach M, Slager SL. 2018. Comparative analysis of de novo assemblers for variation discovery in personal genomes. Brief. Bioinform. 19:893–904. doi:10.1093/bib/bbx037.
106. Verity R, Collins C, Card DC, Schaal SM, Wang L, Lotterhos KE. 2017. minotaur: A platform for the analysis and visualization of multivariate results from genome scans with R Shiny. Mol. Ecol. Resour. 17:33–43. doi:10.1111/1755-0998.12579.
107. Via S, Lande R. 2016. Genotype-Environment Interaction and the Evolution of Phenotypic Plasticity Author ( s ): Sara Via and Russell Lande Published by : Society for the Study of Evolution Stable URL : http://www.jstor.org/stable/2408649 Accessed : 01-03-2016 03 : 35 UTC Your . Evolution (N. Y). 39:505–522.
108. Wang J, Chen K, Ren Q, Zhang Y, Liu J, Wang G, Liu A, Li Y, Liu G, Luo J, et al. 2021. Systematic Comparison of the Performances of De Novo Genome Assemblers for Oxford Nanopore Technology Reads From Piroplasm. Front. Cell. Infect. Microbiol. 11:1–9. doi:10.3389/fcimb.2021.696669.
109. Wang S, Meyer E, Mckay JK, Matz M V. 2012. 2b-RAD: A simple and flexible method for genome-wide genotyping. Nat. Methods 9:808–810. doi:10.1038/nmeth.2023.
110. Wang S, Gribskov M. 2017. Comprehensive evaluation of de novo transcriptome assembly programs and their effects on differential gene expression analysis. Bioinformatics 33:327–333. doi:10.1093/bioinformatics/btw625.
111. Waterhouse RM, Seppey M, Simao FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva E V., Zdobnov EM. 2018. BUSCO applications from quality assessments to gene prediction and phylogenomics. Mol. Biol. Evol. 35:543–548. doi:10.1093/molbev/msx319.
112. Willis SC, Hollenbeck CM, Puritz JB, Portnoy DS. 2022. Genetic recruitment patterns are patchy and spatiotemporally unpredictable in a deep-water snapper (Lutjanus vivanus) sampled in fished and protected areas of western Puerto Rico. Conserv. Genet. doi:10.1007/s10592-021-01426-2.
113. Xuereb A, Benestan L, Normandeau É, Daigle RM, Curtis JMR, Bernatchez L, Fortin MJ. 2018. Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine invertebrate (Parastichopus californicus). Mol. Ecol. 27:2347–2364. doi:10.1111/mec.14589.
114. Xuereb A, D’Aloia CC, Andrello M, Bernatchez L, Fortin MJ. 2021. Incorporating putatively neutral and adaptive genomic data into marine conservation planning. Conserv. Biol. 35:909–920. doi:10.1111/cobi.13609.
115. Yahav T, Privman E. 2019. A comparative analysis of methods for de novo assembly of hymenopteran genomes using either haploid or diploid samples. Sci. Rep. 9:1–10. doi:10.1038/s41598-019-42795-6.