Alternative splicing: An overlooked mechanism contributing to local adaptation?

Sarah Salisbury[1], Maria Lisette Delgado Aquije[1], and Anne Dalziel[2]

[1]Dalhousie University
[2]Saint Mary's University

August 10, 2021

Molecular Ecology News and Views - Perspective on Jacobs and Elmer (2021) "Alternative splicing and gene expression play contrasting roles in the parallel phenotypic evolution of a salmonid fish"

Alternative splicing: An overlooked mechanism contributing to local adaptation?

Sarah J. Salisbury [1] , M. Lisette Delgado[1] , and Anne C. Dalziel [2,3]

[1] Department of Biology, 1355 Oxford Street, Dalhousie University, Halifax, NS, Canada, B3H 4R2[2] Department of Biology, 923 Robie Street, Saint Mary's University, Halifax, Nova Scotia, Canada, B3H 3C3[3] Correspondence: anne.dalziel@smu.ca

Identifying the molecular mechanisms contributing to fitness-associated phenotypic variation in natural populations is a major goal of molecular ecology. However, the multiple regulatory steps between genotype and phenotype mean that many potential regulatory mechanisms can evolve to influence trait divergence. To date, the role of transcriptional regulation in local adaptation has received the most focus, as we can readily measure mRNA quantity and have a reasonable grasp of how variation in the expression of many protein-coding genes might influence phenotype. Thus, studying the evolution of protein coding gene mRNA abundance in candidate tissues has led to some successes in detecting the molecular mechanisms underlying local adaptation (reviewed by Hill et al. 2021). However, the contribution of differential splicing of precursor mRNA (pre-mRNA) to adaptive differentiation, as well as the loci controlling this variation, remains largely unexplored in wild populations. In their "From the Cover"' article in this issue of Molecular Ecology, Jacobs and Elmer (2021) re-analzye muscle RNA sequencing (RNA-seq) data to quantify the relative contributions of variation in mRNA quantity (differentially expressed "DE" genes) and splice variant identity (differentially spliced "DS" genes) to parallel ecotypic divergence of wild "benthic" and "pelagic" Arctic charr (*Salvelinus alpinus*). They found little overlap in the identity and biological functions of DE and DS genes, suggesting that these two regulatory mechanisms act on different cellular traits to complementarily alter organismal phenotype. Furthermore, many DE and DS genes could be mapped to *cis*-acting QTL, arguing that some of this regulatory divergence is genetically based. DE and DS genes were also more likely to be "hub genes" than their non-divergent counterparts, hinting that this regulatory variation may have a variety of meaningful phenotypic effects. The comparison of three independently evolved pairs of benthic and pelagic charr uncovered greater than expected parallelism in both expression and splicing between ecotypes across different lakes, supporting a role for these molecular phenotypes in adaptive divergence. Overall, Jacobs and Elmer's (2021) findings highlight the importance of DS as a potential mechanism underlying local adaptation and provide a framework for others hoping

to make the most of their RNA-seq data.

Species that have repeatedly evolved comparable phenotypes in response to similar selective pressures provide a unique opportunity to investigate the mechanisms and predictability of evolution (Elmer and Meyer 2011). The great phenotypic diversity and recent, repeated post-glacial evolution of Arctic charr ecotypes in many locations across this species' Holarctic distribution makes them a good model for linking genotypes to ecologically-relevant phenotypes (Klemetsen 2010). Two commonly occurring ecotypes are the "benthic" form that feeds in the littoral zone and the "pelagic" form that feeds on plankton in the limnetic zone. Living in these different habitats has led to adaptive divergence in feeding morphology and swimming activity (Klemetsen 2010).

In prior work, Jacobs et al. (2020) found that the evolution of this ecotype-specific feeding morphology is associated with significant parallelism in both genetic and gene expression divergence. However, as in most transcriptomic studies, the potential contributions of splicing was not directly assessed (Fig. 1). In contrast to differential gene expression, which results in a quantitative difference in the number of mRNA transcripts (a balance between transcription and degradation), alternative splicing leads to a qualitative difference in the transcriptome, based on which exons are included in the final mRNA transcript prior to translation (Fig.1). If alternatively spliced transcripts vary in their stability or function, splice variation can also influence the quantity of specific protein isoforms. A major goal of Jacobs and Elmer (2021) was to test for ecotype-associated variation in mRNA splicing and contrast this form of regulatory variation to divergence in mRNA quantity. To accomplish this, they employed a novel suite of analyses on RNA-seq data collected from white muscle samples (a functionally-relevant tissue given swimming differences between ecotypes) of wild benthic and pelagic fish in three Scottish lakes (Awe, Tay, Dughaill; Jacobs et al., 2020).

Differential gene expression between ecotypes was re-assessed using Deseq2; this analysis is based on a gene-level count, so is not entirely independent of splicing variation. Differential splicing was assessed using two techniques: differential exon usage (using Dexseq) and differential intron excision (using Leafcutter; Li et al. 2018). These analyses revealed that DE genes differed from those which were DS and were associated with different GO and regulatory processes (Fig.2a). However, both DE and DS genes were enriched for GO terms related to DNA regulation. In addition, DE and DS genes both occupied more central positions in gene co-expression networks than non-differentiated genes (Fig. 2b), suggesting increased pleiotropy. DS, but not DE, genes were also associated with more GO terms, than non-differentiated genes, suggesting that DS genes may have greater pleiotropy than DE genes. These data match findings from tissue-specific splicing QTLs (sQTLs) in humans, which were found to be more pleiotropic and have a stronger effect on phenotype than expression QTLs (eQTLs) (Garrido-Martín et al. 2021). Thus, studies of sQTL may be a critical addition to our toolkit used to find the genetic basis of adaptation.

As these tissue samples were collected from wild fish populations living in diverse environments, the relative contributions of genetically based variation versus pure phenotypic plasticity to the observed transcriptional variation is unclear. Given Arctic charr's known high capacity for plasticity (Klemetsen 2010), Jacobs and Elmer (2021) tested if transcriptomic variation has evolved by conducting QTL mapping of*cis*- variation associated with differential expression (eQTL) and splicing (sQTL). Despite limited power (n=4 per ecotype per lake), they were able to find many such loci, suggesting that at least some of the observed molecular divergence has a genetic basis. *cis*- eQTL and sQTL mapping also found that expression and splicing variation across ecotypes and lakes were controlled by distinct loci (Fig. 2c), but few of these loci were outliers. Since *trans* -eQTL/sQTL , which could not be located in this study, are predicted to co-regulate the expression and splicing of multiple genes (reviewed by Hill et al. 2015), further studies are needed to conclusively test the extent to which these molecular phenotypes are controlled by different loci.

When studying the mechanisms of adaptation, a key question is the level of repeatability, and thus evolutionary predictability, we see in populations independently adapting to a given environment. Parallelism can indicate that phenotypes of interest are the result of adaptation, and is predicted to increase when selective pressures are similar, populations have high levels of shared genetic variation, and there are strong genetic, developmental or physiological constraints (reviewed by Rosenblum et al. 2014). Both alternative

splicing and expression showed significant parallelism among ecotypes (Fig. 2d). While little is known about the constraints on ecotypic adaptation, the effects of shared genetic variation are likely at play, as all lakes were recently colonized by a common glacial lineage (Atlantic lineage) after Scotland's deglaciation ˜10 - 15 000 years ago. Interestingly, Jacobs and Elmer (2021) found that in some instances where DS genes were identical across lakes, the specific way these genes were differentially spliced was non-identical across lakes. This finding matches prior work suggesting that the extent of convergence or parallelism may increase with increasing levels of biological organization, due to the hierarchical nature of biological traits. For example, many different mutations could lead to convergent differences in mRNA expression (Jacobs et al. 2020). It will be interesting to see if non-parallel transcriptional regulatory changes found in replicate charr populations might lead to convergent changes in swimming muscle size or function among ecotypes, as found in benthic and limnetic Lake Whitefish (*Coregonus clupeaformis*) populations (Dalziel et al. 2017).

In sum, the work of Jacobs and Elmer (2021) suggests that post-transcriptional processes may regulate adaptive differentiation and therefore deserve further attention as potential mechanisms contributing to ecological divergence. Additionally, this study provides a blueprint for those interested in measuring variation in pre-mRNA splicing in their own RNA-seq datasets; Leafcutter's ability to investigate differential intron excision from just a genomic reference, without the need for an annotated genome, will be especially useful for molecular ecologists studying non-genomic model organisms (Li et al. 2018).

References

Dalziel, A.C., Laporte, M., Rougeux, C., Guderley, H., & Bernatchez, L. (2017) Convergence in organ size but not energy metabolism enzyme activities among wild Lake Whitefish (*Coregonus clupeaformis*) species pairs. *Molecular Ecology* , 26, 225–244.

de Klerk, E., & 't Hoen, P.A.C. (2015). Alternative mRNA transcription, processing, and translation: insights from RNA sequencing. *Trends in Genetics* , 31 (3), 98-110.

Elmer, K. R., & Meyer, A. (2011). Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends in Ecology & Evolution* , 26(6), 298-306.

Garrido-Martín, D., Borsari, B., Calvo, M. et al. (2021). Identification and analysis of splicing quantitative trait loci across multiple tissues in the human genome. *Nature Communications* , 12, 727.

Hill, M.S., Zande, P.V., & Wittkopp, P.J., (2021) Molecular and evolutionary processes generating variation in gene expression.*Nature Reviews Genetics* . 22 (4), 203-215.

Jacobs, A., Carruthers, M., Yurchenko, A., Gordeeva, N. V., Alekseyev, S. S., Hooker, O., ... & Elmer, K. R. (2020). Parallelism in eco-morphology and gene expression despite variable evolutionary and genomic backgrounds in a Holarctic fish. *PLoS Genetics* , 16(4), e1008658.

Jacobs, A., & Elmer, K. R. (2021). Alternative splicing and gene expression play contrasting roles in the parallel phenotypic evolution of a salmonid fish. *Molecular Ecology* . 29, XXX-XXX.

Klemetsen, A. (2010). The charr problem revisited: exceptional phenotypic plasticity promotes ecological speciation in postglacial lakes. *Freshwater Reviews* , *3* (1), 49-74.

Lee, Y., & Rio, D. C. (2015). Mechanisms and regulation of alternative Pre-mRNA splicing. *Annual Review of Biochemistry* , 84, 291–323.

Li, Y. I., Knowles, D. A., Humphrey, J., Barbeira, A. N., Dickinson, S. P., Im, H. K., & Pritchard, J. K. (2018). Annotation-free quantification of RNA splicing using LeafCutter. *Nature Genetics* , 50(1), 151-158.

Rosenblum, E.G., Parent, C. E., & Brandt, E.E. (2014). The Molecular Basis of Phenotypic Convergence. *Annual Review of Ecology, Evolution, and Systematics* , 45, 203-226.