Gametophytic transcriptomics of an early embryophyte (Blasia pusilla) and comparative landscape of gametophyte evolution

Gaurav Sabhok1,2, Xiaolun He1, Mari Miranto1, Elna Peltomaa1, Jorge R. Flores1, Robin Sleth1, Kenneth Karol1, Charles Delwiche3, Neil Bell5 Lars Paulin1, Jarkko Hyvonen1,2

1Finnish Museum of Natural History, University of Helsinki, Finland; 2OEB, Faculty of Biology and Environmental Sciences, University of Helsinki, Finland; 3ECOFEN, Faculty of Biology and Environmental Sciences, University of Helsinki, Finland; 4NewYork Botanical Garden, USA; 5University of Maryland USA; Royal Botanic Garden, Edinburgh, UK, Institute of Botany, University of Helsinki, Finland.

*Email: jarkko.hyvonen@helsinki.fi

Introduction:
Transcriptomics plays a central role in understanding the genetic underpinning of post-transcriptional regulation. In recent years, transcriptomics has not only acted as a model for unravelling post-transcriptional regulatory events but has also provided a basis for species-level comparative transcriptomics. In the absence of genomic data, transcriptomics not only allows for de-novo phyto-transcriptomics but also enables the detection of specific genes, which are specific to certain lineages or clades. We present the first reference gametophytic transcriptome of Blasia pusilla sequenced using Illumina Next-Seq with the aim of unravelling the gametophytic transcriptomics of an early embryophyte. We also highlight the role of transcriptomics in identifying early genes and possible phylogenetic pressures on ancestral gene evolution across embryophytes (land plants).

Material and Methods:
In-vitro grown gametophytic cultures of B. pusilla on MS medium were used for RNA extractions using the HiTrap RNA Extraction kit (ISTATE II RNA Plant Kit (BIO–52076)) and were subsequently used for library preparation and sequencing using Illumina Next-Seq at the Biotechnology Center of the University of Helsinki, Finland. In total, 555,690,062 reads with an (87.08%) were further used for the transcriptome reconstruction using the Trinity assembler version 2.6.4. Gametophytic transcriptome of the representative species were downloaded from KEGG (http://www.kenp.org) and were compared for the orthologous transcripts using ProteinOrtho (Lechner et al. 2011). Venn diagram for the shared orthologous proteins across the gametophytes were plotted using the VennMarker (Lin et al. 2016). Annotations for the assemblered transcriptome were done using the eggnoG mapper with v1NOC as the back-end database (http://eggnoGdb.embl.de/#!/app/eggnomapper). Additionally, annotations were done by performing the BLASTx searches (1E-5) against plant species UNIPROT databases available from (https://www.uniprot.org). Transcriptome completeness was assessed using DOGS (Dohmen et al. 2016) and BUSCO (Waterhouse et al. 2017) using embryophytes as a lineage.

Venn diagram showing the distribution of the shared Blasia pusilla transcripts across the mosses, hornworts and liverworts 1KP gametophyte transcriptome. * The first three letter defines the species in the table (A) added to same genus name.

Gene ontology classification at cellular, biological and molecular functions.

References:

Acknowledgements: Academy of Finland for their grant 295595 to RJ & XR, and to CSC - IT Centre for Science for providing computing environment.

Hornworts and Liverworts

Mosses

Sph

Gene Ontology Process

Cellular Component

Biological Process

References