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

Sablok, Gaurav

2019-06-26

Sablok , G , He , X , Miranto , M , Peltomaa , E , Flores , J , Sleith , R , Karol , K , Delwiche , C , Bell , N , Paulin , L , Poczai , P & Hyvönen , J 2019 , ' Gametophytic transcriptomics of an early embryophyte (Blasia pusilla) and comparative landscape of gametophyte evolution ' , International Symposium on Frontiers in Molecular Science , Lisbon , Portugal , 26/06/2019 - 28/06/2019 .

---

http://hdl.handle.net/10138/304307

---

other

ISFMS 2019
3rd International Symposium on Frontiers in Molecular Science
RNA Regulatory Networks
26 – 28 June 2019, Lisbon, Portugal

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

Gaurav Sablok[1,2], Xiaolan He[1], Mari Miranto[1], Elina Peltomaa[3], Jorge R. Flores[1], Robin Sleith[4], Kenneth Karol[4], Charles Delwiche[5], Neil Bell[6], Lars Paulin[7], Peter Poczai[1], Jaakko Hyvonen[1,2]*
[1]Finnish Museum of Natural History, University of Helsinki, Finland; [2]OEB, Faculty of Biology and Environmental Sciences, University of Helsinki, Finland; [3]ECOENV, Faculty of Biology and Environmental Sciences, University of Helsinki, Finland; [4]NewYork Botanical Garden, USA; [5]University of Maryland USA; [6]Royal Botanic Garden, Edinburgh, UK; [7]Insitute of Biotechnology, University of Helsinki, Finland.
***Email:jaakko.hyvonen@helsinki.fi**

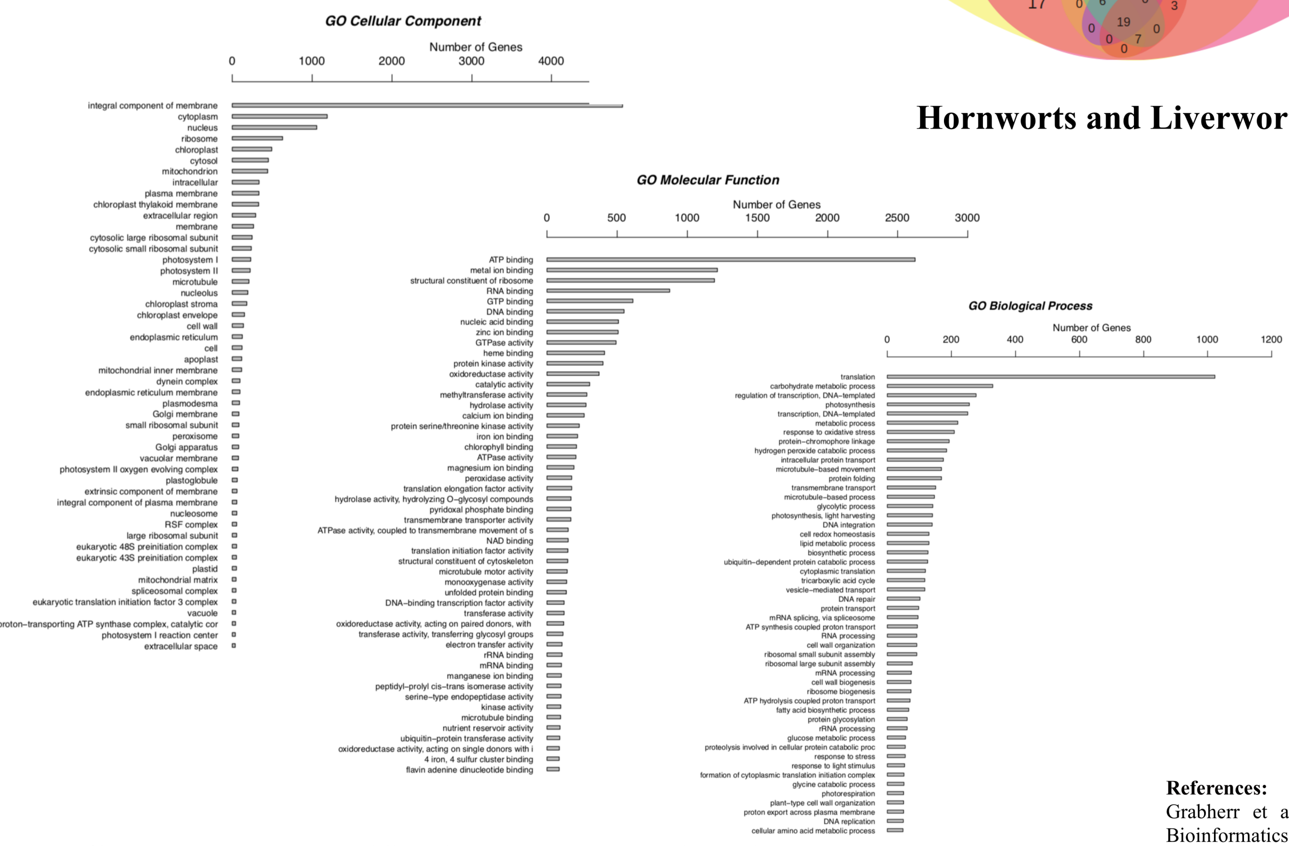
Introduction:

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 unraveling 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* phylo-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 Biotop RNA Extraction kit (ISOLATE II RNA Plant Kit (BIO-52076) and were subsequently used for library preparation and sequencing using Illumina Next-Seq at Institute of Biotechnology, University of Helsinki, Finland. In total, 555.696062 M with an (87.086811%) over Q30 were further used for the transcriptome reconstruction using the Trinity assembler version 2.6.4. Gametophyte transcriptome of the representative species were downloaded from 1KP (http://www.onekp.com) and were compared for the orthologus transcripts using ProteinOrtho (Lechner et al. 2011). Venn diagram for the shared orthologus proteins across the gametophytes were plotted using the VennPainter (Lin et al. 2016). Annotations for the assembled transcriptome were done using the eggnog-mapper with virNOG as the backand database (http://eggnogdb.embl.de/#/app/emapper). 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 accessed using DOGMA (Dohmen et al. 2016) and BUSCO (Waterhouse et al. 2017) using embryophytes as a lineage.

Species	Family	
Philonotis fontana	Bartramiaceae	Mosses
Plagiomnium insigne	Mielichhoferiaceae	Mosses
Racomitrium elongatum	Grimmiaceae	Mosses
Racomitrium varium (A)	Grimmiaceae	Mosses
Dicranum scoparium	Dicranaceae	Mosses
Leucobryum albidum	Leucobryaceae	Mosses
Stereodon subimponens	Pylaisiaceae	Mosses
Orthotrichum lyellii	Orthotrichaceae	Mosses
Polytrichum commune	Polytrichaceae	Mosses
Syntrichia princeps	Pottiaceae	Mosses
Sphagnum lescurii	Sphagnaceae	Mosses
Takakia lepidozioides	Takakiaceae	Mosses
Timmia austriaca	Timmiaceae	Mosses
Paraphymatoceros hallii	Notothyladaceae	Hornworts
Phaeoceros carolinianus	Notothyladaceae	Hornworts
Nothoceros aenigmaticus	Dendrocerotaceae	Hornworts
Nothoceros vincentianus (A)	Dendrocerotaceae	Hornworts
Megaceros flagellaris	Dendrocerotaceae	Hornworts
Bazzania trilobata	Lepidoziaceae	Liverworts
Pellia neesiana	Pelliaceae	Liverworts
Porella navicularis	Porellaceae	Liverworts



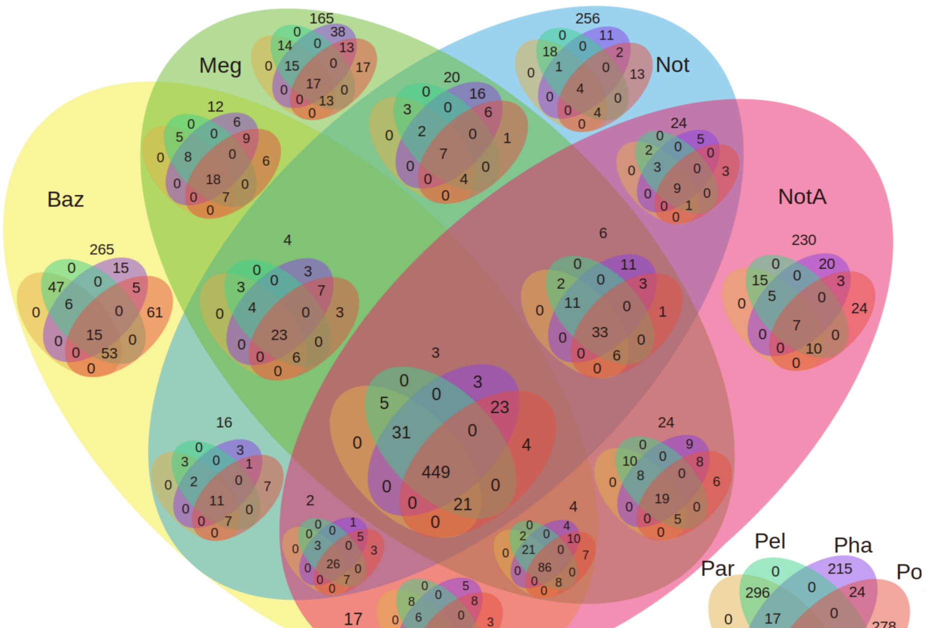
Gene ontology classification at cellular, biological and molecular functions.

Blasia pusilla gametophyte transcriptome summary statistics

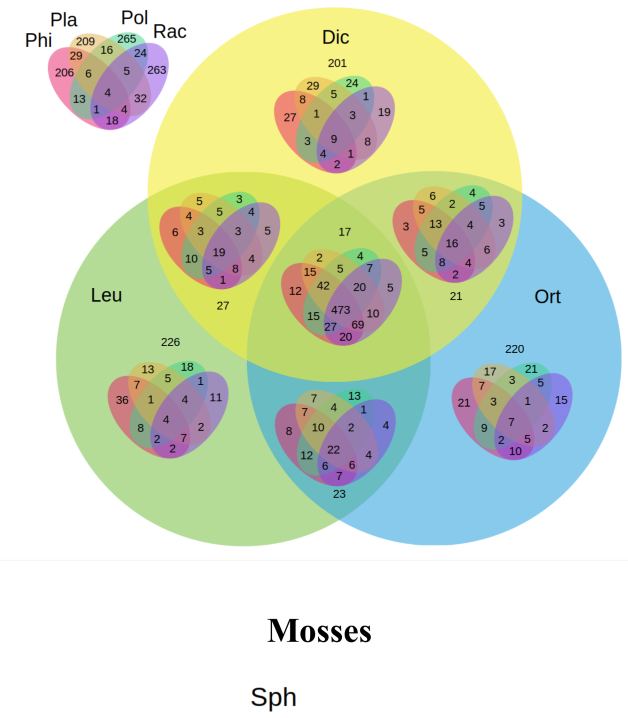
Blasia pusilla assembled transcripts	130871
Blasia pusilla Supertranscripts	80880
Percent GC	46.55
* *Blasia pusilla* **Transcriptome statistics based on ALL transcript contigs**	
Contig N50:	3191bp
Median contig length	653bp
Average contig	1533.29bp
* *Blasia pusilla* **Transcriptome statistics based on longest isoform**	
Contig N50	2334bp
Median contig length	395bp
Average contig	921.32bp
Number of CDSpredicted using GeneMark-ST (Transcriptome)	60812
Number of CDS predicted using GeneMark-ST (Supertranscripts)	24336
Number of transcripts associated virNOG:	39499
Number of supertrancripts associated virNOG:	11696
Number of transcripts mapped to KEGG:	22996
Number of supertranscripts mapped to KEGG:	7199

Domain assessment of transcriptome completeness: 59.28% completeness representing 77.25% (CDAsize 1), 55.39% (CDAsize2)
BUSCO assessment of transcriptome completeness: 44% completeness representing 635 (CB), 610 (CSB)

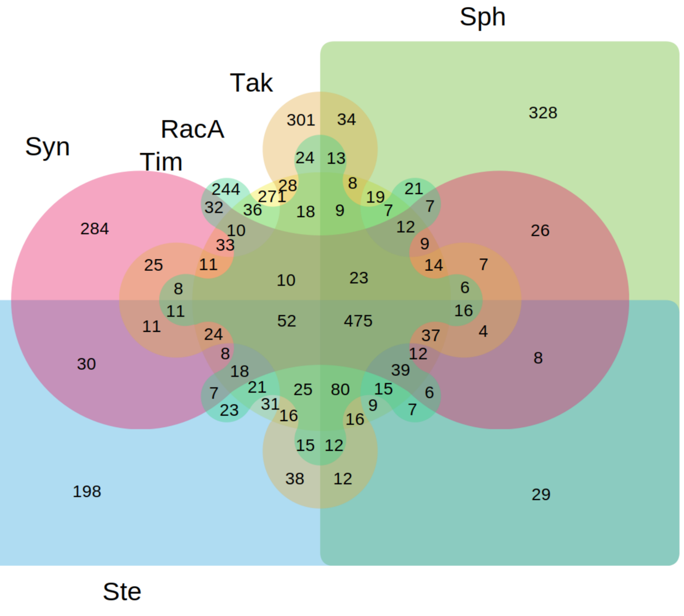
Venn diagram showing the distribution of the shared *Blasia pusilla* transcripts across the mosses, hornworts and liverworts 1KP gametophyte transcriptome. * The first three letter defines the species in the table with (A) added to same genus name.



Hornworts and Liverworts



Mosses



References:
Grabherr et al. Nat Biotechnol. 2011; 29:644-652; Lin et al. PLoS One. 2016;11(4): e0154315.; Lechner et al. BMC Bioinformatics. 2011; 12:124. Huerta-Cepas et al. Nucleic Acids Res. 2019;47(D1): D309–D314. Dohmen et al. Bioinformatics, 2016; 32:2577-81; Waterhouse et al. Mol Biol Evol. 2017; 35:543–548.

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