UNIVERSITÀ DI TRENTO | Department of Cellular, Computational and Integrative Biology | CIBIO

Circular RNAs as Potential Biomarkers in Huntington's Disease Pathogenesis

Miguel Pellegrini[1], Jessica Michela Döring[1], Guendalina Bergonzoni[1], Federica Perrone[2], Giulia Cardamone[3], Alan Monziani[1], Francesca Di Leva[1], Michele Arnoldi[1], Sabrina Maffi[2], Simone Migliore[2], Ludovica Busi[5], Rosanna Asselta[3], Erik Dassi[4], Ferdinando Squitieri[2] and Marta Biagioli[1]

[1]NeuroEpigenetics laboratory, Department of Cellular, Computational and Integrative Biology, University of Trento, Trento, Italy

[2]Huntington and Rare Diseases Unit, CSS-Mendel at IRCCS Casa Sollievo della Sofferenza, Rome, Italy

[3]Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, and Istituto Clinico Humanitas IRCCS, Rozzano, Milan, Italy

[4]RNA Regulatory Networks laboratory, Department of Cellular, Computational and Integrative Biology, University of Trento, Trento, Italy

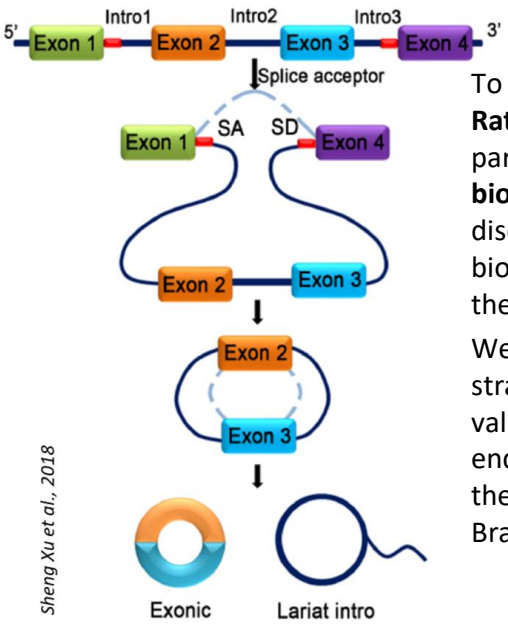
[5] Italian League for Research on Huntington (LIRH) Foundation, Rome, Italy

Twitter ORCID NeuroEpigenetics Lab NeuroEpigenetics Lab @CIBIO

This work was supported by: Huntington's Disease Society of America

UNIVERSITÀ DI TRENTO
Department of Cellular, Computational and Integrative Biology
CIBIO

1. Introduction & Aim



Background

To date, HD patients are classified by the **Unified Huntington Disease Rating Scale** (**UHDRS**), which relies on non-objectively measurable parameters, thus there is an increasing need to develop and validate **biomarkers in accessible biofluids** (such as blood) to either follow disease progression and/or to predict treatment's outcome. Such biomarkers will also be essential in assessing the effectiveness of new therapeutic treatments.

We focused on circular RNAs (circRNAs), covalently-bound, single-stranded RNA circles, obtained by backsplicing [1]. CircRNAs could be valuable biomarkers since (i) they are more stable and resistant to endonucleases than their linear counterparts; (ii) they are abundant in the Central Nervous System (CNS) [2]; (iii) they can easily cross the Blood-Brain Barrier (BBB) [3].

Aim

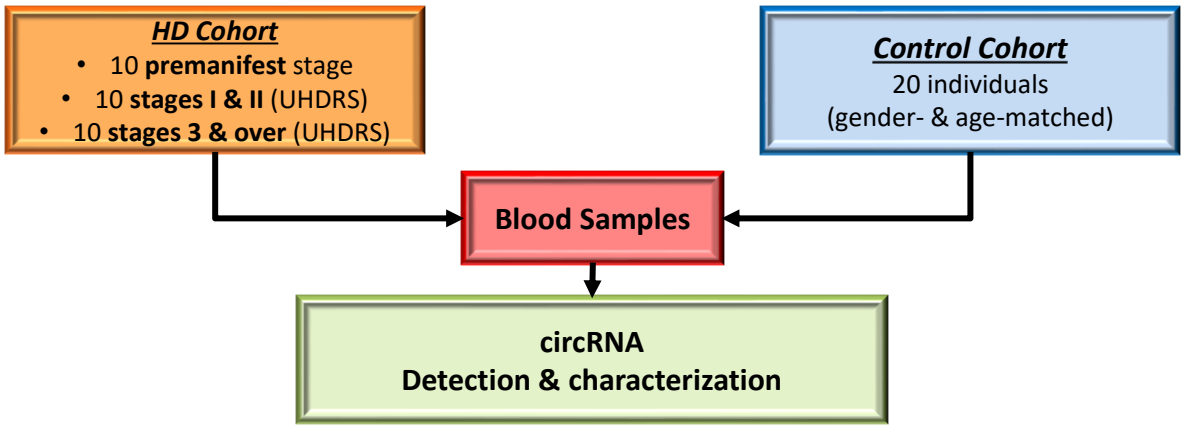
Here, for the first time in HD, we aim to focus, detect, and characterize blood-derived circRNAs. We investigate their potential as new biomarkers that can forecast the onset of motor symptoms as well as consistently follow the later stages of HD disease stage and length of the CAG repeat.

Blood Samples

50 blood samples were collected, divided in 2 cohorts: a **HD Cohort**, composed of **30 individuals** affected by Huntington at different stages of the disease; a **Control Cohort**, comprising of **20 healthy, age and sex-matched individuals**. The **Blood Samples** were provided Prof. **Ferdinando Squitieri** (Huntington and Rare Diseases Unit in Rome).

This protocol was **approved by the appointed ethical committee** and all **those involved signed** a form of **informed consent.**

Fig. 1



Results 1. Sequencing quality control

We first examined the number of read counts for all samples, specifically the reads mapping to unique loci in the genome, as they are indicative of the quality of the library: 2 samples of the asymptomatic stage (HD26 and HD37) had a number of reads outside of the 'Mean ± 2SD' range, and were removed from downstream analyses (Fig. 2A). We also examined the chimeric reads, used to detect and quantify circRNAs (Fig. 2B).

Fig. 2A

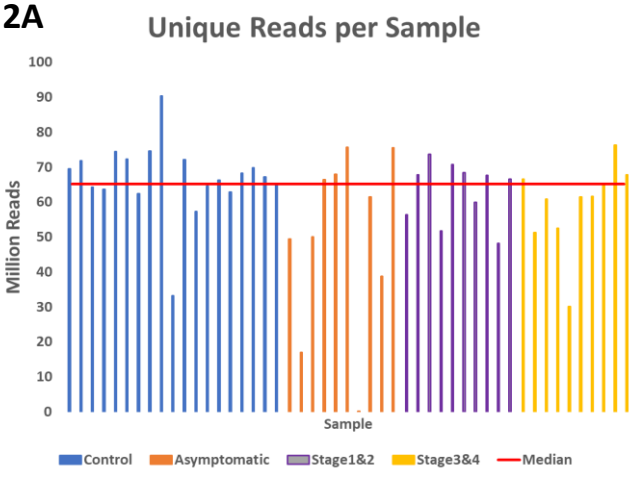
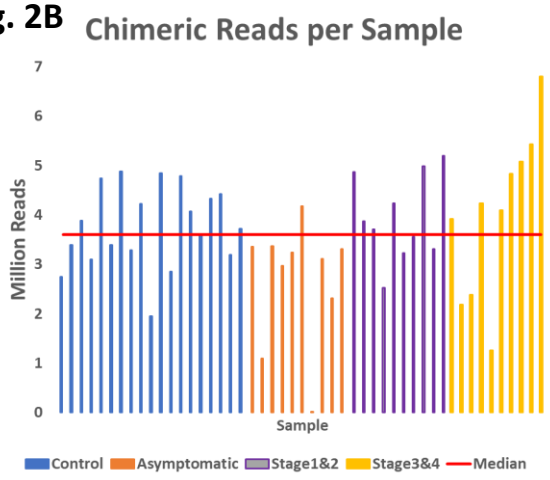


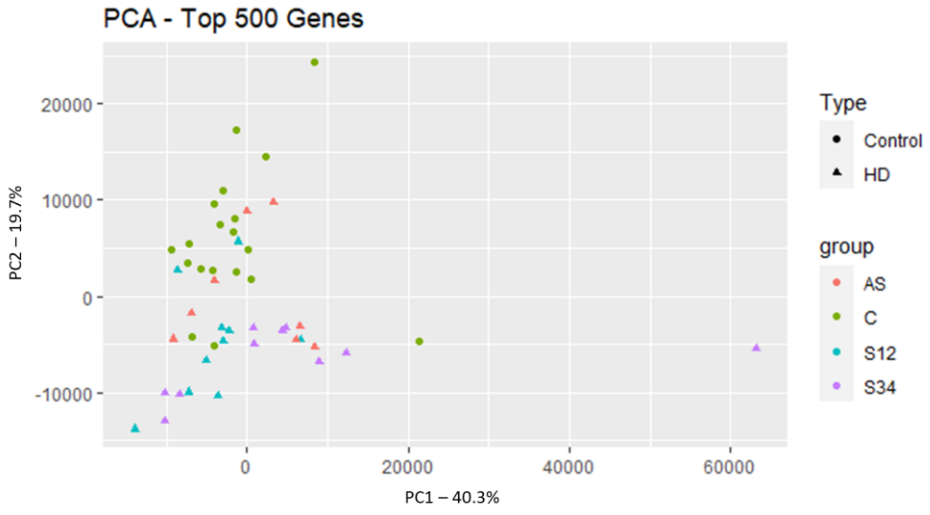
Fig. 2B



Principal Component Analysis

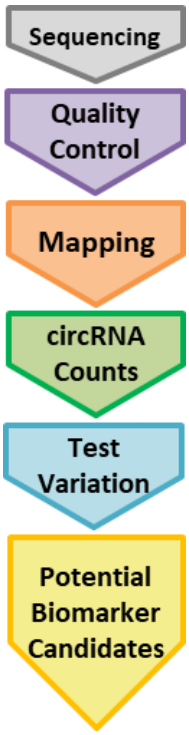
Using **STAR-generated gene counts**, we performed PCA analysis, using the top 500 genes with highest variance across samples (Fig. 3). A clear separation between Controls (green dots) and the HD group (orange, cyan and purple triangles).

Fig. 3



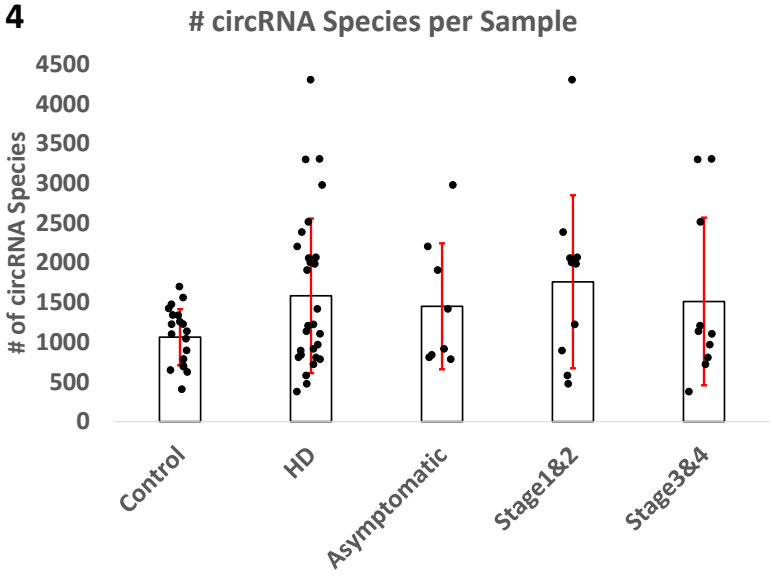
UNIVERSITÀ DI TRENTO
Department of Cellular, Computational and Integrative Biology
CIBIO

Results 2. CircRNA expression is dysregulated in HD patients



By applying the bionformatic pipeline depicted in the diagram [4-6], we evaluated the overall number of different species of circRNA present in each of the 2 main conditions (CTRL vs HD). Only circRNAs with a number of counts per sample ≥ 3 in either group were considered (Fig. 4). **As shown in the figure, the control group has a lower average number of expressed circRNA species when compared to the patient group (whether considered as a whole – 'HD' bar – or stage by stage: asymptomatic, Stage 1&2, Stage 3&4).**

Fig. 4

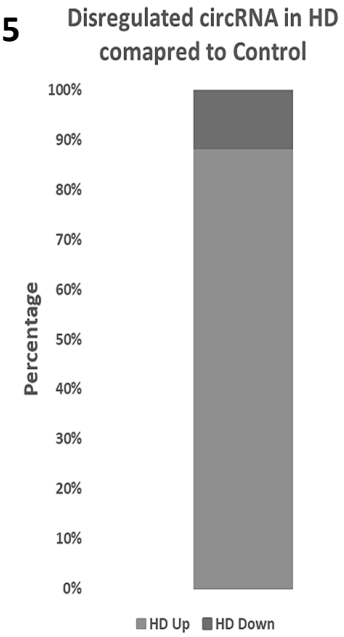


Most circRNAs are upregulated in HD patients

Among all the detected circRNA, 88.3% were up-regulated in HD, while the remaining were down-regulated (Fig. 5).

	# circRNA	%
HD Up	730	88.3
HD Down	97	11.7

Fig. 5



Results 3. Selection of viable circRNAs candidates

CircTest: Control vs HD group differentially expressed circRNA

With **CircTest** we were able to identify **35 circRNAs** whose expression was enriched in HD patients (as a whole) compared to controls (Fig. 6A). **GO and KEGG** pathway analysis of the obtained list, revealed that these genes are predominantly involved to **ubiquitination-related** pathways (Fig. 6B).

Fig. 6A

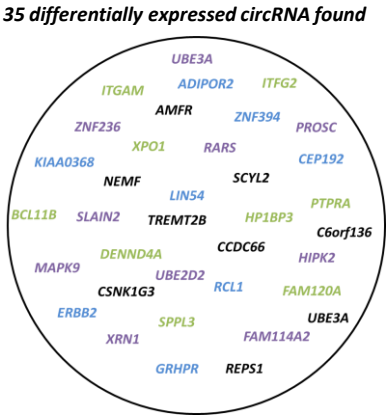
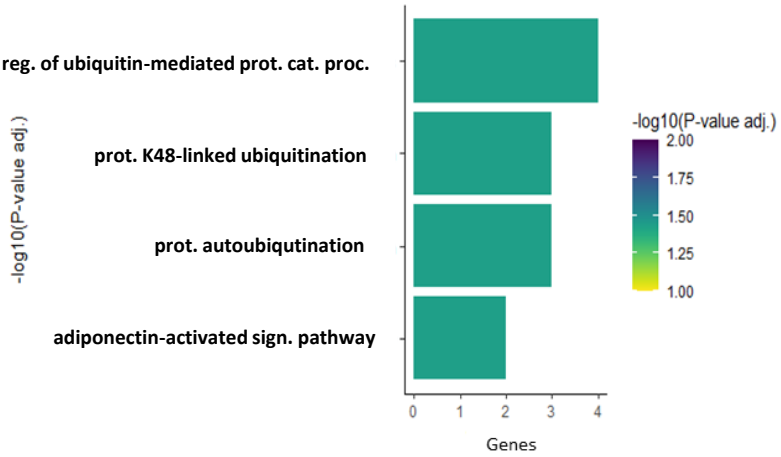
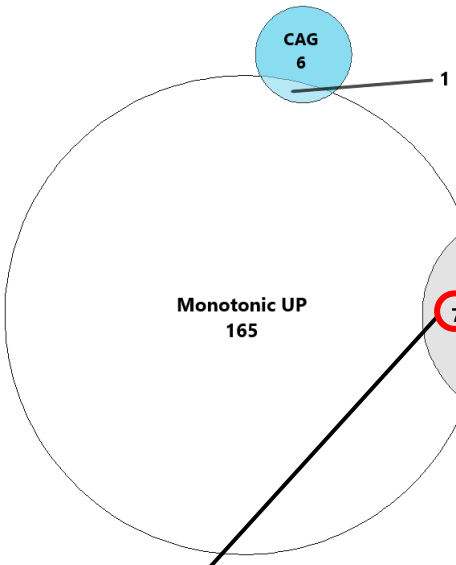


Fig. 6B





Chr	Start	End	Gene	circRNA ID
3	170145423	170149244	*PHC3*	hsa_circ_0001360

Chr	Start	End	Gene	circRNA ID
X	101041316	101042312	*TRMT2B*	hsa_circ_0001937
16	56385918	56389375	*AMFR*	hsa_circ_0000704
9	93471140	93498886	*FAM120A*	hsa_circ_0001875
15	25650607	25657118	*UBE3A*	hsa_circ_0000586
8	37623043	37623873	*PROSC*	hsa_circ_0001788
5	167915606	167921655	*RARS*	hsa_circ_0001550

Fig. 7

To identify a suitable list of circRNAs candidates, we intersected the results from 3 analytical approaches :

1) circRNAs differentially expressed in HD;

2) circRNAs whose expression increases with disease progression;

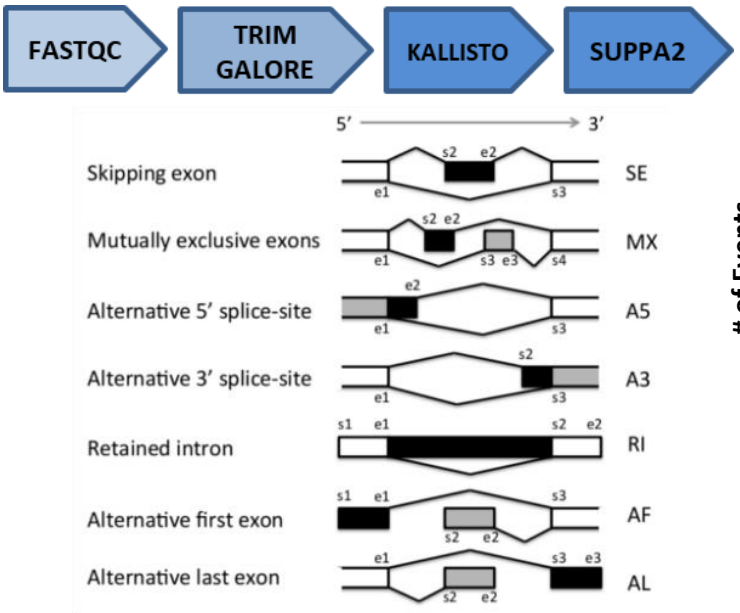
3) circRNA correlated to the CAG length.

7 circRNA candidates were identified that satisfied conditions 1) and 2), while 1 satisfied conditions 2) and 3) (Fig. 7).

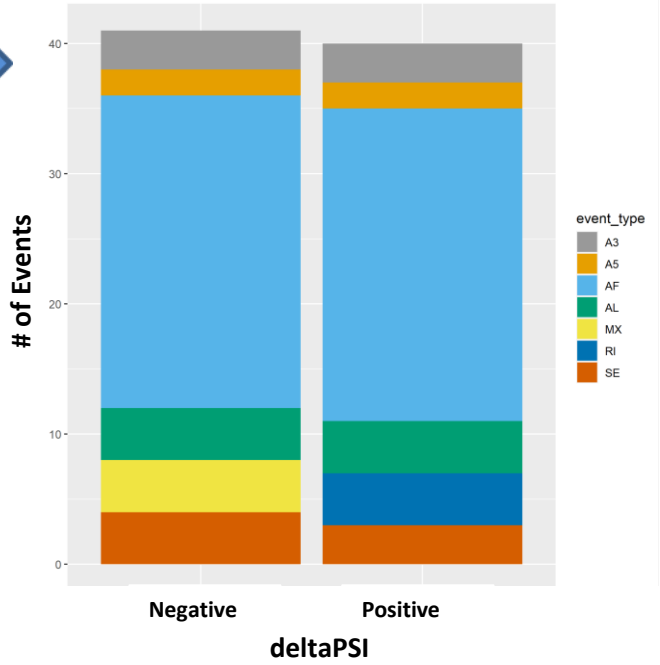
UNIVERSITÀ DI TRENTO
Department of Cellular, Computational and Integrative Biology
CIBIO

Results 4. Alternative Splicing Analysis in HD Blood Samples

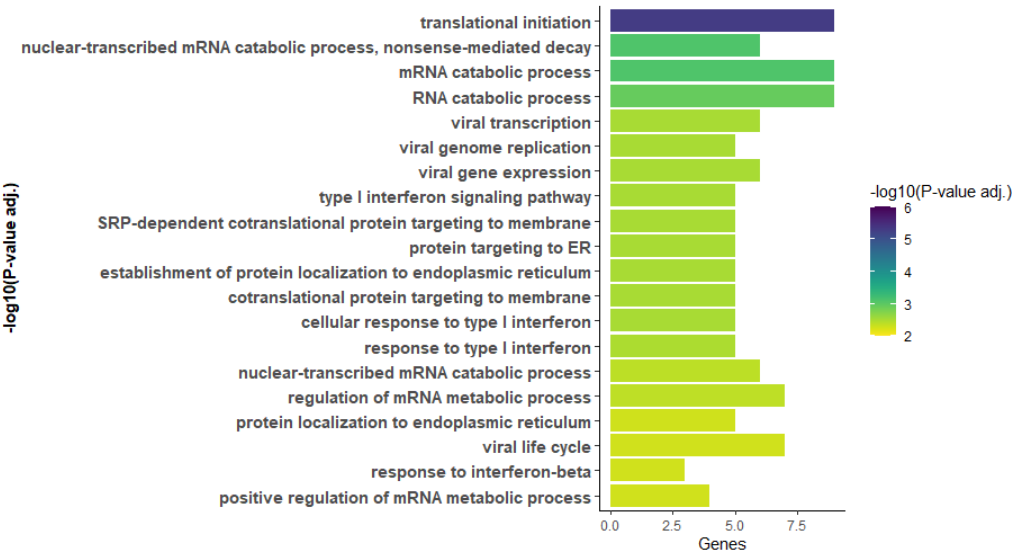
We investigated **Alternative Splicing events** (**AS**) that may differ between the HD and Control groups. Transcripts were first quantified with **Kallisto** [7] and then the difference in 'Percentage of Spliced-In' (dPSI) was calculated with **SUPPA2** [8,9]. **81 splicing events** reached significance (**P-value < 0.05**) with a **dPSI above 10%** in the HD group compared to the control. Even if only few events show a significant difference, it is still possible that one or more of these may be of particular relevance for the disease, thus further analyses concerning this topic are required.



(Adapted from: *Alamancos G.P. et al.*, 2015)



GO and KEGG pathway analysis for the genes related to the identified events: here we report the 20 most significant. Most of the reported terms are related to **RNA processing**, **Interferon response** and **protein localization**. RNA processing is a well established process that is disrupted by mut*HTT*: alterations in the functioning of this pathways might lead to alterations in splicing and, possibly, to backsplicing.



UNIVERSITÀ DI TRENTO
Department of Cellular, Computational and Integrative Biology
CIBIO

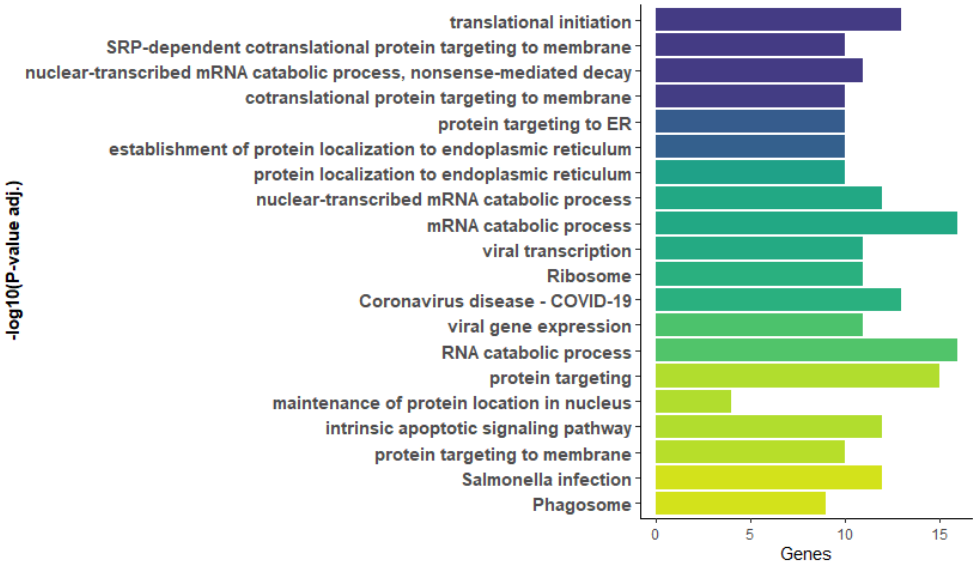
Results 5. Gene expression in HD Blood Samples

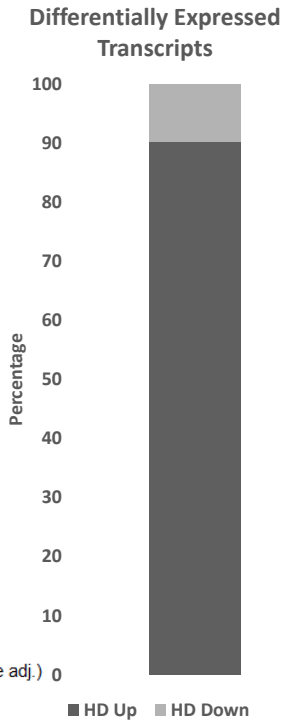
Most Differentially Expressed Transcripts are Up-regulated in HD

FASTQC → TRIM GALORE → KALLISTO → DESeq2

Transcript counts obtained with Kallisto [10] were used as input for DESeq2 [9]. After the removal of Globin and rRNA transcripts, 287 transcripts resulted differentially expressed (adjusted P-value < 0.05, |logFC| > 0.5), of which >90% are enriched in HD.

Some of the most prominent processes that emerged from the **GO and KEGG** pathway analysis for these genes include: '**intracellular transport', 'RNA processing'**.

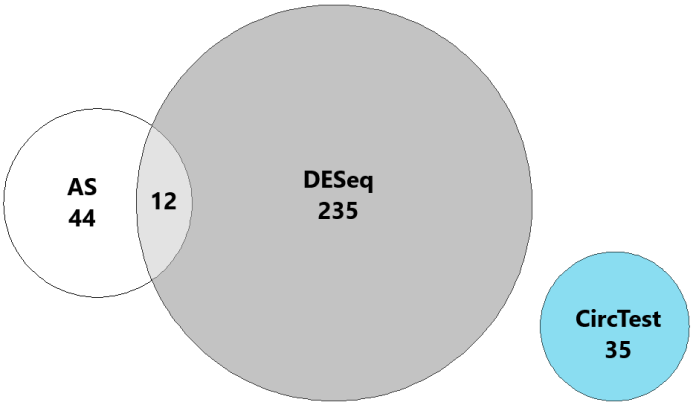




	DE	%
HD Up	259	90.2
HD Down	28	9.8

Intersection between AS events, DE transcripts and circRNA

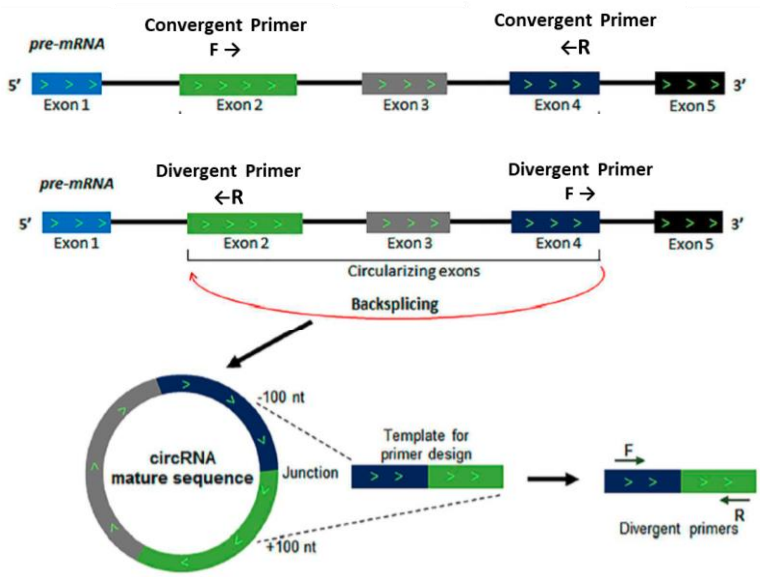
We intersected the gene lists obtained from AS, DESeq and CircTest analyses to verify whether the observed chnages in As and circRNAs expression could be linked to differential expression. As evidenced by the Venn diagrams, the overlaps are really minimal, suggesting that mut*HTT* is affecting the AS and backsplicing processes independently of gene expression changes.



Future Directions

Validation of the circRNA candidates identified by RNA-seq via **RT-qPCR assay to detect and quantify** these possible new **biomarkers** and

1. **Segregate** HD patients based on the **disease stage.**



(Image adapted from Panda A. C. and Gorospe M*, "Detection and Analysis of Circular RNAs by RT-PCR", 2018).*

Specific circRNA amplification will be achieved through the use of Divergent Primers.

2. **Evaluate reproducibility** of our findings in a **different cohort** of patients;
3. **Comparison** between **newly identified** biomarkers and the **pre-existing** ones, specifically **how they change in relation to mutHTT levels**;
4. **Evaluate** if the candidate behave similarly **in the Cerebro Spinal Fluid (CSF)** as they do in **blood**.

References

1. Xu, Sheng et al., ***"A comprehensive review of circRNA: from purification and identification to disease marker potential."*** *PeerJ* vol. 6 e5503. 24 Aug. 2018, doi: 10.7717/peerj.5503;

2. Szabo, Linda et al., **"Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development."** *Genome biology* vol. 16,1 126. 16 Jun. 2015, doi: 10.1186/s13059-015-0690-5;

3. Lim, Ryan G. et al., ***"Huntington's Disease iPSC-Derived Brain Microvascular Endothelial Cells Reveal WNT-Mediated Angiogenic and Blood-Brain Barrier Deficits"***, *Cell Reports*, Volume 19, Issue 7, 1365 – 1377, doi: https://doi.org/10.1016/j.celrep.2017.04.021;

4. Giulia Cardamone et al., "***The Characterization of GSDMB Splicing and Backsplicing Profiles Identifies Novel Isoforms and a Circular RNA That Are Dysregulated in Multiple Sclerosis****", Int. J. Mol. Sci.* 2017, 18, 576; doi: 10.3390/ijms18030576;

5. Elvezia Maria Paraboschi et al., "***Interpreting Non-coding Genetic Variation in Multiple Sclerosis Genome-Wide Associated Regions***", *Frontiers in Genetics* 2018, 9, 647; doi: 10.3389/fgene.2018.00647;

6. Jun Cheng et al., "***Specific identification and quantification of circular RNAs from sequencing data"***, *Bioinformatics*, Volume 32, Issue 7, 1 April 2016, Pages 1094-1096, doi: https://doi.org/10.1093/bioinformatics/btv656.

7. Bray N. et al., ***"Near-optimal probabilistic RNA-seq quantification"***, *Nature Biotechnology* 34, 525-527(2016), doi:10.1038/nbt.3519;

8. Alamancos G.P. et al., "***Leveraging transcript quantification for fast computation of alternative splicing profiles"****. RNA*. 2015 Sep;21(9):1521-31;

9. Trincado J.L. et al., ***"SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions."****, Genome Biol.,* 2018 Mar 23;19(1):40;

10. Love M.I. et al., ***"Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2."*** Genome Biology, 2014, 15, 550. doi: 10.1186/s13059-014-0550-8