Stage-Specific Disease Genes in the Cortex and Striatum

John C. Obenauer[1], Jian Chen[2], Jeffrey S. Aaronson[2], Jim Rosinski[2], Jonathan Bard[2]

[1]Rancho BioSciences LLC, San Diego, CA; [2]CHDI Management/CHDI Foundation, Princeton, NJ

Abstract

Extensive RNA-seq and proteomics experimental data from mouse models of Huntington's disease have been made publicly available (Alexandrov V et al., 2016), and comprehensive lists of the differentially expressed disease genes and proteins in these experiments have been deposited in HDinHD.org (Langfelder P et al., 2016). We examined the cortex and striatum RNA and protein results in these lists across 17 disease stages, where we defined the disease stages as Q50 6 months (6M), Q50 10M, Q80 2M, Q80 6M, Q80 10M, Q92 2M, etc., ending with Q175 10M (defined as, Qn knock-in (KI) allelic-series). We observed stage-specific gene expression in the cortex proteins, where each disease stage can be characterized by a block of proteins that first becomes significant at that stage. Many of the proteins in the earlier disease stages are no longer significant in the Q175 mice, and conversely, some of the Q175 disease proteins are not significant in any earlier stage. Out of 570 total cortex disease proteins, the highest percentage of them is present in the Q50 mice (276 proteins, 48% of total), while only 10% (58) are present in the Q175 mice. The cortex proteins first appearing at each disease stage are enriched in different GO terms, with regulation of membrane potential being significant in the earliest Q50 6M stage, positive regulation of MAP kinase activity being significant in the intermediate Q111 6M stage, and locomotory behavior being significant in the late Q175 6M stage. We went on to test the cortex RNA, striatum proteins, and striatum RNA, and similarly found blocks of stage-specific disease genes or proteins in each case. The striatum RNA and proteins included relatively stable cores of disease genes at the later stages (Q111 and higher, ages 6M and 10M), making possible the Str266R and Str115P disease signatures on HDinHD.org, while consistent disease signatures, were not detected in the more diverse cortex RNA and proteins. We tentatively conclude that HD's effects or mechanisms change more during disease progression in the cortex than in the striatum in the Qn KI allelic-series models.

Objective

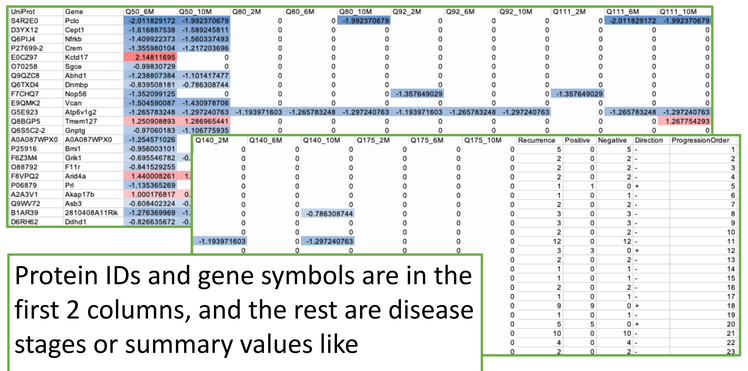
Can mRNAs or proteins be identified that are expressed at distinct stages of HD disease progression in the cortex and striatum?

Methods

HDinHD.org contains extensive differential expression results for all the RNA-seq and proteomics data in the mouse allelic series experiments. All the significant genes and proteins for the cortex and striatum full series experiments (GSE65770 for cortex RNA, PXD005485 for cortex proteins, GSE65774 for striatum RNA, and PXD006302 for striatum proteins) were downloaded. The significant genes and proteins were organized into a table in order of disease progression. Disease progression was defined by increasing Q length, and by age within each Q length: Q50 6M (6 months), Q50 10M, Q80 2M, Q80 6M, Q80 10M, Q92 2M, Q92 6M, Q92 10M, Q111 2M, Q111 6M, Q111 10M, Q140 2M, Q140 6M, Q140 10M, Q175 2M, Q175 6M, and Q175 10M. Each of these 17 Q-length and age combinations was considered a disease stage. For proteins, UniProt IDs that are assigned to the same gene symbol were considered the same protein in this analysis.

Results

Cortical Protein Analyses: Disease Progression Table
Significant protein changes as log10 fold change
(Red = increased in disease, blue = decreased)

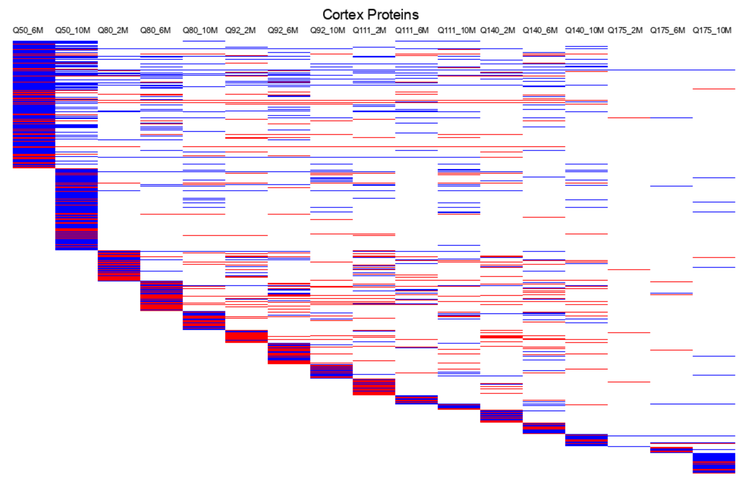


UniProt	Gene	Q50_6M	Q50_10M	Q80_2M	Q80_6M	Q80_10M	Q92_2M	Q92_6M	Q92_10M	Q111_2M	Q111_6M	Q111_10M
S4R2E0	Pclo	-2.011829172	-1.992370679	0	0	-1.992370679	0	0	0	0	-2.011829172	-1.992370679
D3YX12	Cept1	-1.616887538	-1.589245811	0	0	0	0	0	0	0	0	0
Q8PIJ4	Nfkb	-1.409922373	-1.560337493	0	0	0	0	0	0	0	0	0
P27699-2	Cnrn	-1.355980104	-1.217203696	0	0	0	0	0	0	0	0	0
E0CZ97	Kctd17	2.14811695	0	0	0	0	0	0	0	0	0	0
O70258	Sgce	-0.99830729	0	0	0	0	0	0	0	0	0	0
Q9QZC8	Abhd1	-1.238807384	-1.101417477	0	0	0	0	0	0	0	0	0
Q6TXD4	Dnmbp	-0.839508181	-0.786308744	0	0	0	0	0	0	0	0	0
F7CHQ7	Nop56	-1.352099125	0	0	0	0	-1.357649029	0	0	-1.357649029	0	0
E9QMK2	Vcan	-1.504590087	-1.430978706	0	0	0	0	0	0	0	0	0
G5E923	Atp6v1g2	-1.265783248	-1.297240763	-1.193971603	-1.265783248	-1.297240763	-1.193971603	-1.265783248	-1.297240763	0	-1.265783248	-1.297240763
Q8BGP5	Tmem127	1.250908893	1.286965441	0	0	0	0	0	0	0	0	1.267754293
Q8S5C2-2	Gngtg	-0.97060183	-1.106775935	0	0	0	0	0	0	0	0	0
A0A087WPX0	A0A087WPX0	-1.254571026										
P25916	Bmi1	-0.956003101										
F8ZM4	Grik1	-0.695546782	-0									
O88792	F11r	-0.841529255										
F8VPQ2	Arid4a	1.440008261	1.									
P06879	Prl	-1.135365269										
A2A3V1	Akap17b	1.000176817	0.									
Q9WV72	Aab3	-0.608402324	-0									
B1AR39	2810408A11Rik	-1.276369969	-1									
D6RH62	Ddhd1	-0.826635672	-0									

Q140_2M	Q140_6M	Q140_10M	Q175_2M	Q175_6M	Q175_10M	Recurrence	Positive	Negative	Direction	ProgressionOrder
0	0	0	0	0	0	5	0	5	-	1
0	0	0	0	0	0	2	0	2	-	2
0	0	0	0	0	0	2	0	2	-	3
0	0	0	0	0	0	2	0	2	-	4
0	0	0	0	0	0	1	1	0	+	5
0	0	0	0	0	0	1	0	1	-	6
0	0	0	0	0	0	2	0	2	-	7
0	0	0	0	0	0	3	0	3	-	8
0	0	0	0	0	0	3	0	3	-	9
0	0	0	0	0	0	2	0	2	-	10
0	0	0	0	0	0	12	0	12	-	11
0	0	0	0	0	0	3	3	0	+	12
0	0	0	0	0	0	2	0	2	-	13
0	0	0	0	0	0	1	1	0	+	14
0	0	0	0	0	0	1	0	1	-	15
0	0	0	0	0	0	2	0	2	-	16
0	0	0	0	0	0	1	0	1	-	17
0	0	0	0	0	0	9	9	0	+	18
0	0	0	0	0	0	1	0	1	-	19
0	0	0	0	0	0	5	5	0	+	20
0	0	0	0	0	0	10	0	10	-	21
0	0	0	0	0	0	4	0	4	-	22
0	0	0	0	0	0	2	0	2	-	23

Protein IDs and gene symbols are in the first 2 columns, and the rest are disease stages or summary values like recurrence.

The full table contains 570 rows of unique proteins that were significant in at least one of the disease stages. The plot below is a compact view of the table, in which each row is a single line with the same red and blue coloring as in the table. The plot shows blocks of proteins that first become significant at a specific disease stage (defined by Q length and age). Some proteins like Htt are always significant, but others first appear as late as the Q175 Q length. Additionally, the number of proteins that are significant are greater in the lower Q (e.g. Q50/80) compared to the higher Q models (e.g. Q175).



The 570 significant proteins are distributed among the Q lengths as shown in the table below. The numbers add up to more than 570 because some proteins are significant in multipe Q lengths. Surprisingly, the Q50 model had the most disease proteins (48%), and the Q175 model had the least (10%).

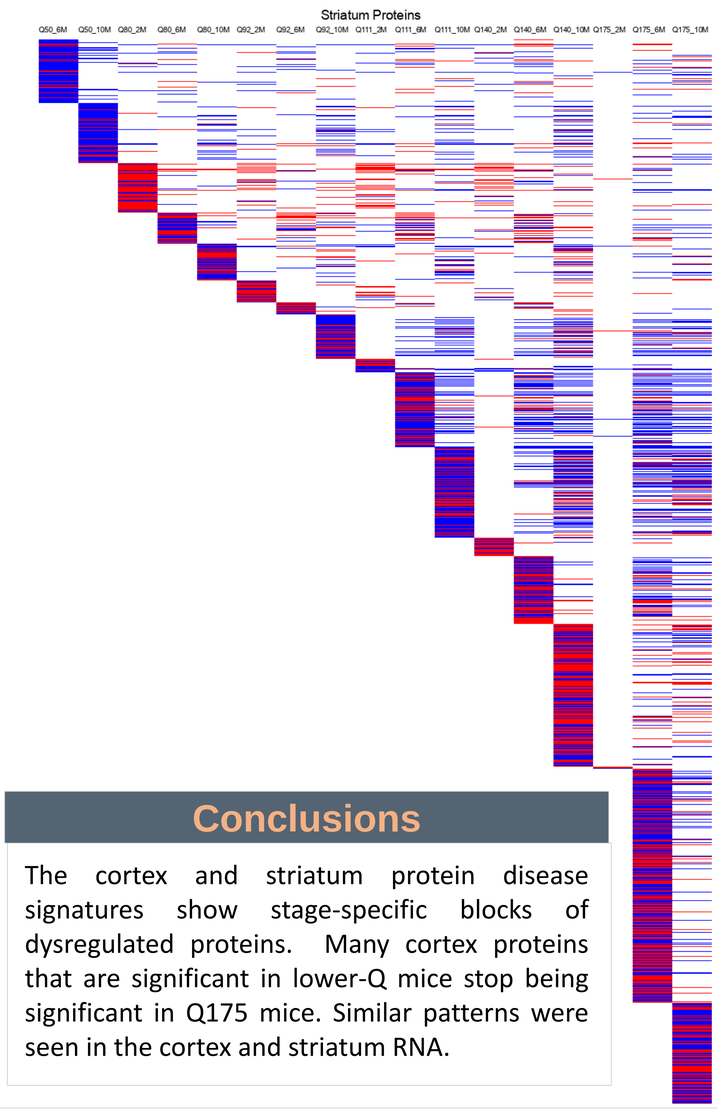
Q Length	Significant Genes
Q50	276 (48%)
Q80	170 (30%)
Q92	181 (32%)
Q111	170 (30%)
Q140	174 (31%)
Q175	58 (10%)

Results

Gene set enrichment was performed for the blocks of cortex proteins, and different GO terms were enriched in each stage. For example, regulation of neuron migration occurs early in disease progression (Q50 6M; Nsmf, Cdh1, Nrg3, and Adam17), while actin-mediated cell contraction occurs much later (Q175 6M; Pde4b and Scn4b). Red gene symbols increase in the disease state while blue gene symbols decrease.

Stage	GO BP	Genes
Q50 6M	Cell-cell junction assembly	Pkp3/Heg1/Cdh1/Prkch/Epb41l3/Mpp7/Actn4
	Regulation of membrane potential	Grik1/Usp53/Htt/Slmap/Cnih2/Zmynd8/Nrcam/Mapt/Mtmr2/Cacna1d/Ank2
	Regulation of protein dephosphorylation	Vcan/2810408A11Rik/Nsmf/Htt/Mprip
	Regulation of neuron migration	Nsmf/Cdh1/Nrg3/Adam17
	Protein localization to cell periphery	F11r/Pkp3/Cdh1/Slmap/Kif1b/Epb41l3/Ank2/Ephb2
Q80 2M	Response to drug	Slc1a2/Atp4a/Abcb1b/Gstm1/Slco1a6
Q80 10M	Neural plate development	Epb41l5/Htt
Q92 2M	Microtubule-based movement	Map4/Ttc21b/Ttc30b
	Regulation of mitochondrial membrane potential	Ngfr/Bco2
Q92 6M	Amino acid transport	Slc6a6/Trh/Grik1/Slc38a7/Slc17a8
	Regulation of gamma-aminobutyric acid secretion	Trh/Grik1
Q111 6M	Positive regulation of MAP kinase activity	Pde5a/Magi3/Tnik
	Positive regulation of mRNA metabolic process	Paf1/Nanos1
	Vasoconstriction	Pde5a/Htr1a
	Microtubule polymerization or depolymerization	Htr1a/Cryab
	Blood circulation	Pde5a/Htr1a/Pomc
Q140 10M	'De novo' protein folding	Cct4/Dnajb1
	Positive regulation of ATPase activity	Tpm1/Dnajb1
Q175 2M	Positive regulation of interleukin-2 biosynthetic process	Prkcq
	Positive regulation of NF-kappaB import into nucleus	Prkcq
	Positive regulation of telomere maintenance via telomere lengthening	Prkcq
Q175 6M	cAMP catabolic process	Pde10a/Pde4b
	Actin-mediated cell contraction	Pde4b/Scn4b
	Negative regulation of kinase activity	Ppp1r1b/Apoe

The striatum proteins similarly showed stage-specific expression (plot below). Similar patterns were seen in the RNA for both the cortex and the striatum (data not shown).



Conclusions

The cortex and striatum protein disease signatures show stage-specific blocks of dysregulated proteins. Many cortex proteins that are significant in lower-Q mice stop being significant in Q175 mice. Similar patterns were seen in the cortex and striatum RNA.

References

1. Alexandrov V et al., *Nat Biotechnol* 2016, 34:838-844.
2. Langfelder P et al., *Nat Biotechnol* 2016, 19:623-633.

17th Annual HD Therapeutics Conference, 2022