BioRDF: Microarray Use Case

Kei Cheung, Ph.D.
Yale Center for Medical Informatics

An Example Neuoscience Microarray Experiment

- Dunckley et al. Gene expression correlates of neurofibrillary tangles in Alzheimer's disease. Neurobiology of Aging 27(10): 1359-1371

Abstract and Keywords

- **Abstract:**

  Neurofibrillary tangles (NFT) constitute one of the cardinal histopathological features of Alzheimer's disease (AD). To explore in vivo molecular processes involved in the development of NFTs, we compared gene expression profiles of NFT-bearing entorhinal cortex neurons from 19 AD patients, adjacent non-NFT-bearing entorhinal cortex neurons from the same patients, and non-NFT-bearing entorhinal cortex neurons from 14 non-demented, histopathologically normal controls (ND). Of the differentially expressed genes, 225 showed progressively increased expression (AD NFT neurons > AD non-NFT neurons > ND non-NFT neurons) or progressively decreased expression (AD NFT neurons < AD non-NFT neurons < ND non-NFT neurons), raising the possibility that they may be related to the early stages of NFT formation. Immunohistochemical studies confirmed that many of the implicated proteins are dysregulated and preferentially localized to NFTs, including apolipoprotein J, interleukin-1 receptor-associated kinase 1, tissue inhibitor of metalloproteinase 3, and casein kinase 2, beta. Functional validation studies are underway to determine which candidate genes may be causally related to NFT neuropathology, thus providing therapeutic targets for the treatment of AD.

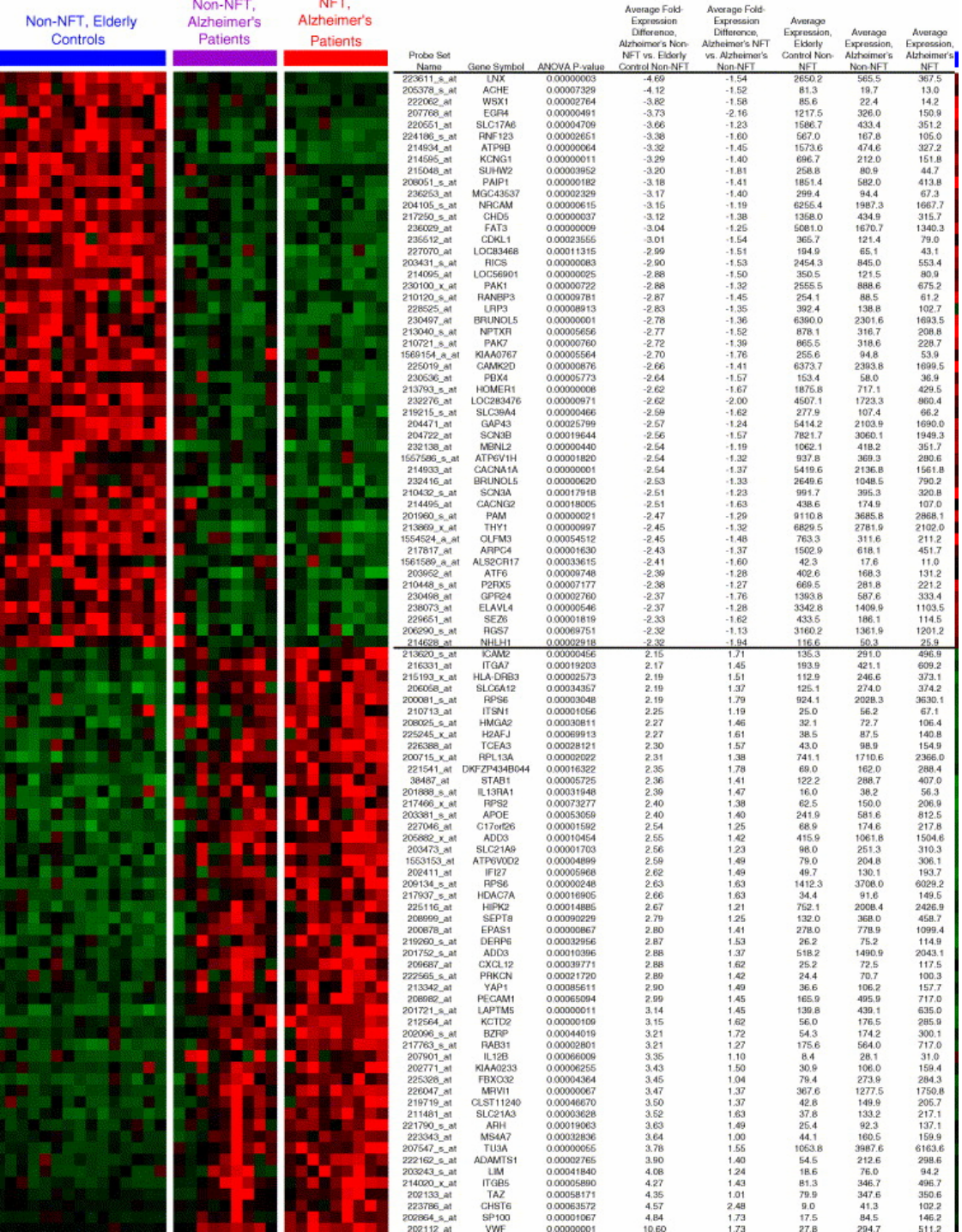
- **Keywords:** Alzheimer's disease; Neurofibrillary tangles; Microarray; Gene expression; Dementia; Neurodegeneration; NFT; Laser capture microdissection

Structured Digital Abstract

- A translation table or “cast of characters” listing all named biological entities like genes, proteins, metabolites and other objects in the article, and relates their human-readable names to precise database identifiers.
- A list of the main results described in simple ontologies using a controlled vocabulary — for example, interactions (“protein A binds to protein B”), phenotypes (“mutation C suppresses deletion D”), and protein modifications (“protein E is phosphorylated at residue F by protein kinase G”).
- Standard evidence codes for how the results were obtained — for example, “affinity purification' or 'mass spectrometry”

Application of SDA

- Pilot application of SDA in FEBS Letter (with links to MINT)
- Other applications by other journals?
  - e.g., gene lists with link to databases
  - How to represent gene lists in RDF with link to existing ontologies (genes, experimental context, provenance, samples, neuroscience domain knowledge, etc)



Probe Set Name	Gene Symbol	ANOVA P-value	Average Fold-Expression Difference, Alzheimer's Non-NFT vs. Elderly Control Non-NFT	Average Fold-Expression Difference, Alzheimer's NFT vs. Alzheimer's Non-NFT	Average Expression, Elderly Control Non-NFT	Average Expression, Alzheimer's Non-NFT	Average Expression, Alzheimer's NFT
223611_s_at	LNX	0.00000003	-4.69	-1.54	2650.2	565.5	367.5
205378_s_at	ACHE	0.00007329	-4.12	-1.52	81.3	19.7	13.0
222062_at	WSX1	0.00002764	-3.82	-1.58	85.6	22.4	14.2
207768_at	EGR4	0.00000491	-3.73	-2.16	1217.5	326.0	150.9
220551_at	SLC17A6	0.00004709	-3.66	-1.23	1586.7	433.4	351.2
224186_s_at	RNF123	0.00002651	-3.38	-1.60	567.0	167.8	105.0
214934_at	ATP9B	0.00000064	-3.32	-1.45	1573.6	474.6	327.2
214595_at	KCNG1	0.00000011	-3.29	-1.40	696.7	212.0	151.8
215048_at	SUHW2	0.00003952	-3.20	-1.81	258.8	80.9	44.7
208051_s_at	PAIP1	0.00000182	-3.18	-1.41	1851.4	582.0	413.8
236253_at	MGC43537	0.00002329	-3.17	-1.40	299.4	94.4	67.3
204105_s_at	NRCAM	0.00000615	-3.15	-1.19	6255.4	1987.3	1667.7
217250_s_at	CHD5	0.00000037	-3.12	-1.38	1358.0	434.9	315.7
236029_at	FAT3	0.00000009	-3.04	-1.25	5081.0	1670.7	1340.3
235512_at	CDKL1	0.00023555	-3.01	-1.54	365.7	121.4	79.0
227070_at	LOC83468	0.00011315	-2.99	-1.51	194.9	65.1	43.1
203431_s_at	RICS	0.00000083	-2.90	-1.53	2454.3	845.0	553.4
214095_at	LOC56901	0.00000025	-2.88	-1.50	350.5	121.5	80.9
230100_x_at	PAK1	0.00000722	-2.88	-1.32	2555.5	888.6	675.2
210120_s_at	RANBP3	0.00009781	-2.87	-1.45	254.1	88.5	61.2
228525_at	LRP3	0.00008913	-2.83	-1.35	392.4	138.8	102.7
230497_at	BRUNOL5	0.00000001	-2.78	-1.36	6390.0	2301.6	1693.5
213040_s_at	NPTXR	0.00005656	-2.77	-1.52	878.1	316.7	208.8
210721_s_at	PAK7	0.00000760	-2.72	-1.39	865.5	318.6	228.7
1569154_a_at	KIAA0767	0.00005564	-2.70	-1.76	255.6	94.8	53.9
225019_at	CAMK2D	0.00000876	-2.66	-1.41	6373.7	2393.8	1699.5
230636_at	PBX4	0.00005773	-2.64	-1.57	153.4	58.0	36.9
213793_s_at	HOMER1	0.00000008	-2.62	-1.67	1875.8	717.1	429.5
232276_at	LOC283476	0.00000971	-2.62	-2.00	4507.1	1723.3	860.4
219215_s_at	SLC39A4	0.00000466	-2.59	-1.62	277.9	107.4	66.2
204471_at	GAP43	0.00025799	-2.57	-1.24	5414.2	2103.9	1690.0
204722_at	SCN3B	0.00019644	-2.56	-1.57	7821.7	3060.1	1949.3
232138_at	MBNL2	0.00000440	-2.54	-1.19	1062.1	418.2	351.7
1557586_s_at	ATP6V1H	0.00001820	-2.54	-1.32	937.8	369.3	280.6
214933_at	CACNA1A	0.00000001	-2.54	-1.37	5419.6	2136.8	1561.8
232416_at	BRUNOL5	0.00000620	-2.53	-1.33	2649.6	1048.5	790.2
210432_s_at	SCN3A	0.00017918	-2.51	-1.23	991.7	395.3	320.8
214495_at	CACNG2	0.00018005	-2.51	-1.63	438.5	174.9	107.0
201960_s_at	PAM	0.00000021	-2.47	-1.29	9110.8	3685.8	2868.1
213869_x_at	THY1	0.00000997	-2.45	-1.32	6829.5	2781.9	2102.0
1554524_a_at	OLFM3	0.00054512	-2.45	-1.48	763.3	311.6	211.2
217817_at	ARPC4	0.00001630	-2.43	-1.37	1502.9	618.1	451.7
1561589_a_at	ALS2CR17	0.00033615	-2.41	-1.60	42.3	17.6	11.0
203952_at	ATF6	0.00009748	-2.39	-1.28	402.6	168.3	131.2
210448_s_at	P2RX5	0.00007177	-2.38	-1.27	669.5	281.8	221.2
230498_at	GPR24	0.00002760	-2.37	-1.76	1393.8	587.6	333.4
238073_at	ELAVL4	0.00000546	-2.37	-1.28	3342.8	1409.9	1103.5
229651_at	SEZ6	0.00001819	-2.33	-1.62	433.5	186.1	114.5
206290_s_at	RGS7	0.00069751	-2.32	-1.13	3160.2	1361.9	1201.2
214628_at	NHLH1	0.00002918	-2.32	-1.94	116.6	50.3	25.9
213620_s_at	ICAM2	0.00000456	2.15	1.71	135.3	291.0	496.9
216331_at	ITGA7	0.00019203	2.17	1.45	193.9	421.1	609.2
215193_x_at	HLA-DRB3	0.00002573	2.19	1.51	112.9	246.6	373.1
206068_at	SLC6A12	0.00034357	2.19	1.37	125.1	274.0	374.2
200081_s_at	RPS6	0.00003048	2.19	1.79	924.1	2028.3	3630.1
210713_at	ITSN1	0.00001056	2.25	1.19	25.0	56.2	67.1
208025_s_at	HMGA2	0.00030811	2.27	1.46	32.1	72.7	106.4
225245_x_at	H2AFJ	0.00069913	2.27	1.61	38.5	87.5	140.8
226388_at	TCEA3	0.00028121	2.30	1.57	43.0	98.9	154.9
200715_x_at	RPL13A	0.00002022	2.31	1.38	741.1	1710.6	2366.0
221541_at	DKFZP434B044	0.00016322	2.35	1.78	69.0	162.0	288.4
38487_at	STAB1	0.00005725	2.36	1.41	122.2	288.7	407.0
201888_s_at	IL13RA1	0.00031948	2.39	1.47	16.0	38.2	56.3
217466_x_at	RPS2	0.00073277	2.40	1.38	62.5	150.0	206.9
203381_s_at	APOE	0.00053059	2.40	1.40	241.9	581.6	812.5
227046_at	C17orf26	0.00001592	2.54	1.25	68.9	174.6	217.8
205882_x_at	ADD3	0.00010454	2.55	1.42	415.9	1061.8	1504.6
203473_at	SLC21A9	0.00001703	2.56	1.23	98.0	251.3	310.3
1553153_at	ATP6V0D2	0.00004899	2.59	1.49	79.0	204.8	306.1
202411_at	IFI27	0.00005968	2.62	1.49	49.7	130.1	193.7
209134_s_at	RPS6	0.00000248	2.63	1.63	1412.3	3708.0	6029.2
217937_s_at	HDAC7A	0.00016905	2.66	1.63	34.4	91.6	149.5
225116_at	HIPK2	0.00014885	2.67	1.21	752.1	2008.4	2426.9
208999_at	SEPT8	0.00090229	2.79	1.25	132.0	368.0	458.7
200878_at	EPAS1	0.00000867	2.80	1.41	278.0	778.9	1099.4
219260_s_at	DERP6	0.00032956	2.87	1.53	26.2	75.2	114.9
201752_s_at	ADD3	0.00010396	2.88	1.37	518.2	1490.9	2043.1
209687_at	CXCL12	0.00039771	2.88	1.62	25.2	72.5	117.5
222565_s_at	PRKCN	0.00021720	2.89	1.42	24.4	70.7	100.3
213342_at	YAP1	0.00085611	2.90	1.49	36.6	106.2	157.7
208982_at	PECAM1	0.00065094	2.99	1.45	165.9	495.9	717.0
201721_s_at	LAPTM5	0.00000011	3.14	1.45	139.8	439.1	635.0
212564_at	KCTD2	0.00000109	3.15	1.62	56.0	176.5	285.9
202096_s_at	BZRP	0.00044019	3.21	1.72	54.3	174.2	300.1
217763_s_at	RAB31	0.00002801	3.21	1.27	175.6	564.0	717.0
207901_at	IL12B	0.00006009	3.35	1.10	8.4	28.1	31.0
202771_at	KIAA0233	0.00006255	3.43	1.50	30.9	106.0	159.4
225328_at	FBXO32	0.00004364	3.45	1.04	79.4	273.9	284.3
226047_at	MRVI1	0.00000067	3.47	1.37	367.6	1277.5	1750.8
219719_at	CLST11240	0.00046670	3.50	1.37	42.8	149.9	205.7
211481_at	SLC21A3	0.00003628	3.52	1.63	37.8	133.2	217.1
221790_s_at	ARH	0.00019063	3.63	1.49	25.4	92.3	137.1
223343_at	MS4A7	0.00032836	3.64	1.00	44.1	160.5	159.9
207547_s_at	TU3A	0.00000055	3.78	1.55	1053.8	3987.6	6163.6
222162_s_at	ADAMTS1	0.00002765	3.90	1.40	54.5	212.6	298.6
203243_s_at	LIM	0.00041840	4.08	1.24	18.6	76.0	94.2
214020_x_at	ITGB5	0.00005890	4.27	1.43	81.3	346.7	496.7
202133_at	TAZ	0.00058171	4.35	1.01	79.9	347.6	350.6
223786_at	CHST6	0.00063572	4.57	2.48	9.0	41.3	102.2
202864_s_at	SP100	0.00001067	4.84	1.73	17.5	84.5	146.2
202112_at	VWF	0.00000001	10.60	1.73	27.8	294.7	511.2

Non-NFT normal vs. Non-NFT-AD vs NFT-AD

Significant neuronal gene expression changes occur in Alzheimer's disease prior to NFT formation. We performed a one-way ANOVA to identify statistically significant genes that show consistently increasing or decreasing expression across the data sets of ND non-NFT neurons, AD non-NFT neurons, and AD NFT neurons. Shown are the top 100 genes from this analysis. Cluster representation was generated by GeneCluster software. Prior to figure generation, data was mean centered for each gene by subtracting the mean gene expression from the expression values so that the mean expression for each gene is 0. Data was subsequently normalized on a gene-by-gene basis by multiplying each value by a scale factor S, such that the sum of the squares of the values for each gene is 1.0. Data shown include the probe ID, gene symbol, ANOVA *p*-value, the average fold change for the comparison of AD non-NFT neurons vs. control non-NFT neurons, the average fold change for AD NFT neurons vs. AD non-NFT neurons, and the average expression values for each gene across all three data sets.