

# UC Davis

## UC Davis Previously Published Works

### Title

Proteomic analysis of hair shafts from monozygotic twins: Expression profiles and genetically variant peptides

### Permalink

<https://escholarship.org/uc/item/8gm056bn>

### Journal

Proteomics, 17(13-14)

### ISSN

1615-9853

### Authors

Wu, Pei-Wen  
Mason, Katelyn E  
Durbin-Johnson, Blythe P  
[et al.](#)

### Publication Date

2017-07-01

### DOI

10.1002/pmic.201600462

### Supplemental Material

<https://escholarship.org/uc/item/8gm056bn#supplemental>

Peer reviewed



# Proteomic analysis of hair shafts from monozygotic twins: Expression profiles and genetically variant peptides

Pei-Wen Wu<sup>1</sup>, Katelyn E. Mason<sup>2</sup>, Blythe P. Durbin-Johnson<sup>3</sup>, Michelle Salemi<sup>4</sup>, Brett S. Phinney<sup>1</sup>, David M. Rocke<sup>3</sup>, Glendon J. Parker<sup>2,5</sup>, Robert H. Rice<sup>1</sup>

<sup>1</sup> Forensic Science Graduate Program and Department of Environmental Toxicology, University of California, Davis

<sup>2</sup> Forensic Science Center, Lawrence Livermore National Laboratory

<sup>3</sup> Division of Biostatistics, Department of Public Health Sciences, Clinical and Translational Science Center Biostatistics Core, University of California, Davis

<sup>4</sup> Proteomics Core Facility, University of California, Davis

<sup>5</sup> Protein-Based Identification Technologies

## Table of Contents

Table S1. Summary of subject data from questionnaires.

Table S2. Publicly available GPM accessions.

Table S3. Detection of genetically variant peptides in subjects.

Table S4. Proximity matrix of co-occurrence between twin sample weighted GVP profiles calculated weighting values by minor allele frequency.

Table S5. Proximity matrix of co-occurrence between twin sample weighted GVP profiles calculated without weighting values by minor allele frequency.

Figure S1. Agglomerative hierarchical clustering dendrogram of twin sample genetically variant peptide profiles.

**Table S1.** Summary of subject data from questionnaires. Included are age, sex, hair length sampled, whether the twins shared their placenta and numbers of protein differences within MZ pairs. Also included are lifestyle characteristics that were distinct within the MZ pair.

Twin pair	Age	Sex	Hair Length (in)	Lifestyles different in ...	Shared Placenta	Protein Differences
A <sub>1</sub>	73	Male	3	Geography Medications Types of exercise Hair care product	Yes	3
A <sub>2</sub>			3			
B <sub>1</sub>	66	Male	1.5	Geography Home Medications Medical conditions Alcohol Stress	Yes	0
B <sub>2</sub>			1			
C <sub>1</sub>	28	Female	10		Yes	2
C <sub>2</sub>			10			
D <sub>1</sub>	84	Female	10	Geography Medications Brush hair times	unknown	0
D <sub>2</sub>			7			
E <sub>1</sub>	76	Female	6	Geography Home (Pet) Sleep Diet	unknown	1
E <sub>2</sub>			5-7			
F <sub>1</sub>	19	Female	20	Diet Brush hair times Types of exercise	unknown	1
F <sub>2</sub>			22			
G <sub>1</sub>	20	Male	3	Hair care product Exercise times	Yes	0
G <sub>2</sub>			5			
H <sub>1</sub>	19	Female	12		unknown	6
H <sub>2</sub>			8.5			
I <sub>1</sub>	25	Male	2-3		Yes	3
I <sub>2</sub>			2-3			
J <sub>1</sub>	59	Male	6	Geography Diet Medications Work Brush hair times	separate	0
J <sub>2</sub>			3			

**Table S2.** Publicly available GPM accessions.

sample name	gpm#	spectra#	prot#	pept#	unique #	FPR
F2-4	GPM32100014557	33399	678	131823	1720	0.77
F2-1	GPM32100014558	38476	724	16284	1847	0.77
F1-4	GPM32100014573	38741	703	16280	1824	0.81
F1-2	GPM32100014563	35620	669	13941	1761	0.81
A1-2	GPM32100015970	35996	723	14845	1768	0.79
A1-3	GPM32100015980	38432	693	15224	1681	0.8
A2-1	GPM32100015978	32267	616	12451	1472	0.82
A2-4	GPM32100015985	39041	740	16544	1764	0.74
C1-4	GPM32100014655	31440	612	11273	1270	0.83
C1-2	GPM32100014656	32918	633	11717	1268	0.86
C2-4	GPM32100014715	32191	606	11802	1220	0.83
C2-2	GPM32100014763	37365	690	14381	1401	0.81
D1-2	GPM32100015800	31084	574	10773	1204	0.83
D1-1	GPM32100015806	38330	724	15793	1599	0.8
D2-2	GPM32100015808	37395	707	14388	1558	0.84
D2-1	GPM32100015990	36064	633	13813	1430	0.85
G1-1	GPM32100015809	37055	698	16772	1856	0.68
G1-2	GPM32100016016	33468	678	13186	1725	0.83
G2-1	GPM32100014935	35241	732	13891	1789	0.85
G2-2	GPM32100016021	39537	751	17036	1948	0.79
B1-2	GPM32100016031	32617	626	12494	1495	0.79
B1-3	GPM32100016128	38190	664	15014	1594	0.79
B2-1	GPM32100016129	29736	554	11162	1271	0.82
B2-2	GPM32100015823	39132	701	15672	1679	0.8
H1-3	GPM32100016414	31967	614	11319	1328	0.89
H1-1	GPM32100016415	35608	636	13997	1409	0.84
H2-1	GPM32100015968	35484	651	15042	1557	0.76
H2-2	GPM32100016473	37118	647	14939	1512	0.85
E1-2	GPM32100016545	37866	726	14934	1725	0.78
E1-3	GPM32100016547	31664	785	16094	1802	0.81
E2-3	GPM32100016560	39447	821	16213	1894	0.76
E2-1	GPM32100016566	36267	726	14156	1690	0.83
I1-1	GPM32100060073	35204	616	14035	1444	0.81
I1-3	GPM32100060098	30970	542	11943	1209	0.79
I2-3	GPM32100060075	35468	716	13192	1559	0.82
I2-2	GPM32100060101	33333	660	11562	1459	0.88
J2-1	GPM32100060074	31922	540	12568	1354	0.72
J2-2	GPM32100060078	36415	680	13933	1574	0.82
J1-1	GPM32100060100	36415	691	12747	1603	0.84
J1-4	GPM32100060095	27096	510	10215	1089	0.72

[http://human.thegpm.org/tandem/thegpm\\_tandem.html](http://human.thegpm.org/tandem/thegpm_tandem.html)

**Table S3.** Detection of genetically variant peptides in subjects. Proteomic datasets from two biological duplicates were processed for each subject and detected genetically variant peptides tabulated. The number of datasets containing an observation is indicated. The genetically variant peptides are listed in alphabetical order according to gene name. The genetically variant peptides are indicated as the corresponding dbSNP accession number and nucleotide allele (rs#\_nuc; [www.ncbi.nlm.nih.gov/SNP/](http://www.ncbi.nlm.nih.gov/SNP/)). The genotype frequency of each allele in the European and African population is also indicated (EUR\_GF, AFR\_GF; [www.internationalgenome.org](http://www.internationalgenome.org)). These frequencies were used for the weighting of detected peptides in the analysis. The presence of a peptide in only one dataset was considered sufficient to impute the status of the corresponding SNP allele. Individual subjects in each twin set are paired alphanumerically and color-coded.







**Table S4.** Proximity matrix of co-occurrence between twin sample weighted GVP profiles calculated weighting values by minor allele frequency. Color scale (red to green) shows the degree of co-occurrence. Red indicates higher magnitudes of co-occurrence (increased similarity) and green indicates lower magnitudes of co-occurrence (increased dissimilarity) between twin sets.

	F2	F1	A1	A2	C1	C2	D1	D2	G1	G2	B1	B2	H1	H2	E1	E2	i1	i2	J2	J1
F2	97.000	93.000	86.000	86.000	80.000	80.000	85.000	81.000	81.000	80.000	89.000	87.000	82.000	83.000	81.000	81.000	79.000	82.000	79.000	80.000
F1	93.000	97.000	88.000	88.000	82.000	82.000	87.000	83.000	83.000	84.000	89.000	87.000	84.000	83.000	83.000	83.000	81.000	84.000	81.000	82.000
A1	86.000	88.000	97.000	91.000	81.000	81.000	84.000	84.000	84.000	83.000	86.000	84.000	85.000	84.000	84.000	86.000	84.000	87.000	80.000	85.000
A2	86.000	88.000	91.000	97.000	85.000	85.000	86.000	86.000	84.000	81.000	86.000	86.000	87.000	86.000	86.000	86.000	88.000	89.000	86.000	89.000
C1	80.000	82.000	81.000	85.000	97.000	93.000	88.000	90.000	78.000	79.000	84.000	84.000	93.000	92.000	90.000	90.000	88.000	85.000	86.000	85.000
C2	80.000	82.000	81.000	85.000	93.000	97.000	88.000	90.000	80.000	77.000	86.000	88.000	91.000	90.000	90.000	88.000	88.000	87.000	86.000	85.000
D1	85.000	87.000	84.000	86.000	88.000	88.000	97.000	89.000	81.000	80.000	87.000	87.000	90.000	89.000	87.000	87.000	85.000	88.000	85.000	84.000
D2	81.000	83.000	84.000	86.000	90.000	90.000	89.000	97.000	81.000	80.000	85.000	85.000	90.000	89.000	89.000	89.000	91.000	88.000	87.000	88.000
G1	81.000	83.000	84.000	84.000	78.000	80.000	81.000	81.000	97.000	92.000	85.000	83.000	80.000	79.000	83.000	83.000	83.000	88.000	83.000	84.000
G2	80.000	84.000	83.000	81.000	79.000	77.000	80.000	80.000	92.000	97.000	82.000	80.000	81.000	80.000	82.000	84.000	80.000	83.000	80.000	81.000
B1	89.000	89.000	86.000	86.000	84.000	86.000	87.000	85.000	85.000	82.000	97.000	95.000	88.000	89.000	87.000	85.000	81.000	84.000	83.000	84.000
B2	87.000	87.000	84.000	86.000	84.000	88.000	87.000	85.000	83.000	80.000	95.000	97.000	88.000	89.000	87.000	85.000	81.000	84.000	83.000	84.000
H1	82.000	84.000	85.000	87.000	93.000	91.000	90.000	90.000	80.000	81.000	88.000	88.000	97.000	96.000	94.000	94.000	86.000	85.000	86.000	85.000
H2	83.000	83.000	84.000	86.000	92.000	90.000	89.000	89.000	79.000	80.000	89.000	89.000	96.000	97.000	93.000	93.000	85.000	84.000	85.000	84.000
E1	81.000	83.000	84.000	86.000	90.000	90.000	87.000	89.000	83.000	82.000	87.000	87.000	94.000	93.000	97.000	95.000	87.000	86.000	89.000	86.000
E2	81.000	83.000	86.000	86.000	90.000	88.000	87.000	89.000	83.000	84.000	85.000	85.000	94.000	93.000	95.000	97.000	87.000	86.000	87.000	86.000
i1	79.000	81.000	84.000	88.000	88.000	88.000	85.000	91.000	83.000	80.000	81.000	81.000	86.000	85.000	87.000	87.000	97.000	92.000	87.000	90.000
i2	82.000	84.000	87.000	89.000	85.000	87.000	88.000	88.000	88.000	83.000	84.000	84.000	85.000	84.000	86.000	86.000	92.000	97.000	88.000	91.000
J2	79.000	81.000	80.000	86.000	86.000	86.000	85.000	87.000	83.000	80.000	83.000	83.000	86.000	85.000	89.000	87.000	87.000	88.000	97.000	90.000
J1	80.000	82.000	85.000	89.000	85.000	85.000	84.000	88.000	84.000	81.000	84.000	84.000	85.000	84.000	86.000	86.000	90.000	91.000	90.000	97.000

**Table S5.** Proximity matrix of co-occurrence between twin sample weighted GVP profiles calculated without weighting values by minor allele frequency. Color scale (red to green) shows the degree of co-occurrence. Red indicates higher magnitudes of co-occurrence (increased similarity) and green indicates lower magnitudes of co-occurrence (increased dissimilarity) between twin sets.

	F2	F1	A1	A2	C1	C2	D1	D2	G1	G2	B1	B2	H1	H2	E1	E2	i1	i2	J2	J1
F2	167.000	163.000	156.000	156.000	149.000	149.000	154.000	150.000	151.000	150.000	158.000	156.000	151.000	152.000	151.000	151.000	148.000	152.000	148.000	150.000
F1	163.000	167.000	158.000	158.000	151.000	151.000	156.000	152.000	153.000	154.000	158.000	156.000	153.000	152.000	153.000	153.000	150.000	154.000	150.000	152.000
A1	156.000	158.000	167.000	161.000	150.000	150.000	153.000	153.000	154.000	153.000	155.000	153.000	154.000	153.000	154.000	156.000	153.000	157.000	149.000	155.000
A2	156.000	158.000	161.000	167.000	154.000	154.000	155.000	155.000	154.000	151.000	155.000	155.000	156.000	155.000	156.000	156.000	157.000	159.000	155.000	159.000
C1	149.000	151.000	150.000	154.000	167.000	163.000	158.000	160.000	147.000	148.000	154.000	154.000	163.000	162.000	159.000	159.000	158.000	154.000	154.000	154.000
C2	149.000	151.000	150.000	154.000	163.000	167.000	158.000	160.000	149.000	146.000	156.000	158.000	161.000	160.000	159.000	157.000	158.000	156.000	154.000	154.000
D1	154.000	156.000	153.000	155.000	158.000	158.000	167.000	159.000	150.000	149.000	157.000	157.000	160.000	159.000	156.000	156.000	155.000	157.000	153.000	153.000
D2	150.000	152.000	153.000	155.000	160.000	160.000	159.000	167.000	150.000	149.000	155.000	155.000	160.000	159.000	158.000	158.000	161.000	157.000	155.000	157.000
G1	151.000	153.000	154.000	154.000	147.000	149.000	150.000	150.000	167.000	162.000	154.000	152.000	149.000	148.000	153.000	153.000	152.000	158.000	152.000	154.000
G2	150.000	154.000	153.000	151.000	148.000	146.000	149.000	149.000	162.000	167.000	151.000	149.000	150.000	149.000	152.000	154.000	149.000	153.000	149.000	151.000
B1	158.000	158.000	155.000	155.000	154.000	156.000	157.000	155.000	154.000	151.000	167.000	165.000	158.000	159.000	156.000	154.000	151.000	153.000	151.000	153.000
B2	156.000	156.000	153.000	155.000	154.000	158.000	157.000	155.000	152.000	149.000	165.000	167.000	158.000	159.000	156.000	154.000	151.000	153.000	151.000	153.000
H1	151.000	153.000	154.000	156.000	163.000	161.000	160.000	160.000	149.000	150.000	158.000	158.000	167.000	166.000	163.000	163.000	156.000	154.000	154.000	154.000
H2	152.000	152.000	153.000	155.000	162.000	160.000	159.000	159.000	148.000	149.000	159.000	159.000	166.000	167.000	162.000	162.000	155.000	153.000	153.000	153.000
E1	151.000	153.000	154.000	156.000	159.000	159.000	156.000	158.000	153.000	152.000	156.000	156.000	163.000	162.000	167.000	165.000	156.000	156.000	158.000	156.000
E2	151.000	153.000	156.000	156.000	159.000	157.000	156.000	158.000	153.000	154.000	154.000	154.000	163.000	162.000	165.000	167.000	156.000	156.000	156.000	156.000
i1	148.000	150.000	153.000	157.000	158.000	158.000	155.000	161.000	152.000	149.000	151.000	151.000	156.000	155.000	156.000	156.000	167.000	161.000	155.000	159.000
i2	152.000	154.000	157.000	159.000	154.000	156.000	157.000	157.000	158.000	153.000	153.000	153.000	154.000	153.000	156.000	156.000	161.000	167.000	157.000	161.000
J2	148.000	150.000	149.000	155.000	154.000	154.000	153.000	155.000	152.000	149.000	151.000	151.000	154.000	153.000	158.000	156.000	155.000	157.000	167.000	159.000
J1	150.000	152.000	155.000	159.000	154.000	154.000	153.000	157.000	154.000	151.000	153.000	153.000	154.000	153.000	156.000	156.000	159.000	161.000	159.000	167.000

**Figure S1.** Agglomerative hierarchical clustering dendrogram of twin sample genetically variant peptide profiles. Clustering as in Figure 4, but detected genetically variant peptides were not weighted according to the minor allelic frequency.

