Glaucoma Biomarker Discovery

Kwok Peng Ng*, John S Crabb*, Xiaorong Gu*, Elisa Bala*§, Ellen Simpson*, Xiuzhen Yue*, Edward Rockwood*, Scott Smith*, Scott Anthony§, Stacia S Yaniglos§, Neal Peachey*§, Ramaswami Krishnadas¶ and John W Crabb*

*Cole Eye Institute and Lerner Research Institute, Cleveland Clinic; §Louis Stokes VA Medical Center, Cleveland, OH; ¶Aravind Eye Hospital, Madurai, India

Cleveland Clinic

COLE EYE INSTITUTE

ABSTRACT

Purpose: To develop a blood test for predicting susceptibility to primary open angle glaucoma (POAG) and for monitoring the efficacy of POAG therapeutics.

Methods: Blood was collected from clinically documented POAG and age-matched normal control (CTRL) donors at the Cole Eye Institute, Cleveland Clinic, Louis Stokes Cleveland VA Medical Center and at the Aravind Eye Hospital, Madurai, India. Plasma were fractionated on C18 reverse phase magnetic beads and analyzed by MALDI TOF/TOF mass spectrometry. Bioinformatic cluster analysis for peptidomic patterns and cross validation was performed with GeneSpring software.

Results: Plasma from 429 donors was analyzed in this preliminary study. For USA donors (n = 378), MALDI TOF mass spectrometric analyses of 249 POAG and 138 CTRL plasma has allowed correct prediction of 96% as either POAG or normal based on peptidomic profiles. For Indian donors (n = 42), mass spectrometric analyses of 22 POAG and 20 CTRL provided 90% correct prediction as either POAG or normal. The peptide patterns from Indian and US plasma exhibit distinct differences. Relative to normal controls, the mean intensity of peptide m/z 1896 was elevated about 95% in the Indian POAG plasma and 160% in US POAG plasma. Peptide m/z 1896 has been sequenced and determined to be from complement C4B.

Conclusions: These preliminary results provide proof of principle that plasma peptidomic patterns offer potential utility in predicting POAG susceptibility and monitoring therapeutic efficacy. Peptidomic pattern differences between USA and Indian plasma may reflect ethnic and environmental differences between the populations. The significance of the elevated complement C4B peptide in POAG plasma remains to be determined. Such findings may provide insights into the mechanism of POAG pathogenesis.

METHODS

Case-Control Study Design. Participants were enrolled in a control and POAG group. Disease progression were evaluated primarily based on Humphrey visual fields tests. All subjects were aged 30 years or older. Subjects were recruited at the Cole Eye Institute in the clinics of glaucoma clinicians Drs E Rockwood and S Smith, at the Louis Stokes Cleveland VA Medical Center under the direction of the Chief-of-Staff in the VA Eye Clinic, Dr S Yaniglos and at the Aravind Eye Hospital, Maduria, India in the clinic of Dr R Krishnadas. All subjects received a comprehensive ophthalmological examination by one of the clinicians, including (a) best-corrected visual acuity, (b) confrontation visual fields, (c) Amsler Grid performance, (d) pupil reaction, and (e) intraocular pressure. After the pupils are dilated, the fundus were examined with biomicroscopy and indirect ophthalmoscopy. Donor health data and family histories were collected using standardized questionnaires. Included in patient information were ethnicity, gender, age, and possible confounding factors such as diabetes mellitus, systemic hypertension and myopia. Special emphasis were placed on identifying siblings and/or parents with POAG in the family. After obtaining informed consent, a research nurse recorded health and family histories, drew blood (~30 ml), and plasma and DNA pellets were prepared according to our protocols. Donors were assigned a code and all clinical information maintained in a confidential database with the biomarker analysis results. DNA samples were archived for screening of glaucoma susceptibility genes based on results from the proposed plasma biomarker studies. Genetic screening will be performed in the laboratory of Dr S Hagstrom, Cole Eye Institute and will be initiated in and supported by subsequent grant applications.

Indian plasma donors have been included in the proposed research because of (i) the significant numbers of POAG donors that can be readily obtained in India, (ii) the opportunity to detect ethnic differences in POAG plasma biomarkers relative to the USA population, (iii) the prevalence of consanguineous pedigrees in India that can provide more information for mapping genes responsible for glaucoma than families without consanguinity, and (iv) the enhanced opportunities for mechanistic insights into glaucoma pathology from the proposed proteomic analyses.

Plasma Fractionation. Reverse-phase-derivatized magnetic beads were mixed with plasma, the magnetic particles pulled to the side with a magnet, the supernatant aspirated, the beads washed, bound components eluted with acetonitrile and mixed with MALDI matrix and deposited on the MALDI target plates.

Mass Spectrometry and Data Analysis. Mass spectra of peptides recovered from plasma samples was determined in a-cyano-4-hydroxy-cinnaminic acid matrix under optimized instrument settings on the Applied Biosystems model 4800 MALDI TOF/TOF MS instruments over the mass range ~0.7-4 kDa. Spectra was acquired from 1000 laser shots per sample as in preliminary studies. External calibration was used, with the calibrant located centrally to eight plasma samples on the sample target plate. A m/z peak was generated for each sample and duplicates within a tolerance of 0.15% was discarded. Binning of the m/z peaks was done with GeneSpring software and mass values that do not vary significantly across POAG and control samples filtered out. A one way ANOVA nonparametric test (Mann-Whitney U test) at $p < 1e-6$ was used to identify m/z signals that represent statistically significant differences between POAG and control samples. The GeneSpring clustering tool assigns similarity by Pearson correlation and is commonly used in DNA microarray analysis. We used this software to perform cross validation by the "leave-one-out" method and independent test set validation to test the accuracy of the cluster assignments.

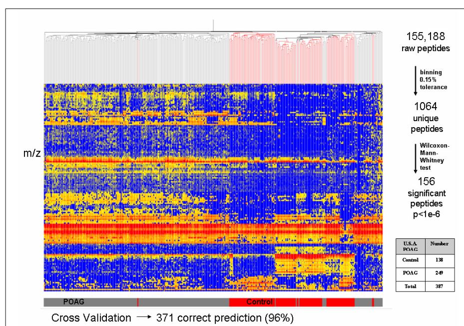


Figure 1. Peptide Mass Signatures, POAG Plasma from the U.S.A. American POAG (n = 249) and normal (n = 138) donor plasma were analyzed by MALDI TOF MS following reverse phase fractionation. A total of 155,188 raw peptides were obtained from the 387 samples. After binning at 0.15% tolerance, 1064 unique peptides remained. Peptide totaling 156 were determined to exhibit statistically significant differences (p=1e-6) between POAG and normal plasma. Hierarchical clustering using the Pearson correlation function in GeneSpring software generated the above display where columns represent donor plasma and rows represent peptide mass-to-charge (m/z) values. The color bars indicate peptide mass and intensity and blue indicates peptide absence; red indicates normal and gray denotes POAG plasma. Cross validation correctly predicted 371 of the 387 plasma samples as either POAG or normal (96% accuracy).

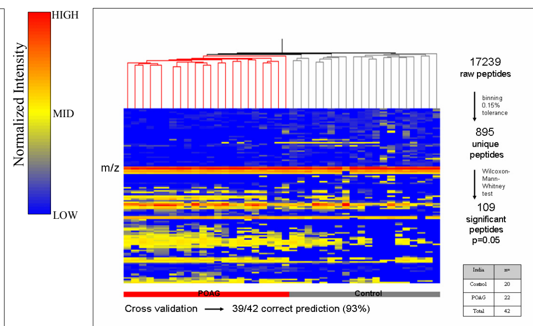


Figure 2. Peptide Mass Signatures, POAG Plasma from India. Indian POAG (n = 22) and normal (n = 20) donor plasma were analyzed by MALDI TOF MS following reverse phase fractionation. A total of 17,239 raw peptides were obtained from the 42 samples. After binning at 0.15% tolerance, 895 unique peptides remained. Peptide masses totaling 109 were determined to exhibit statistically significant (p=0.05) between POAG and normal plasma. Hierarchical clustering using the GeneSpring software generated the above display where columns represent samples and rows represent peptide mass values. The color bars indicate peptide mass and intensity and blue indicates peptide absence; red indicates POAG and yellow denotes normal control plasma. Cross validation correctly predicted 38 of the 42 plasma samples as either POAG or normal (93% accuracy).

Table 1. Plasma Sample Summary

	Number and Gender	Average Age	Age Range	Ethnicity
Control USA-Cleveland	138 10F/128M	69	39-88	65% Caucasian (90) 30% African American 4% Others (3 Hispanic; 1 Lebanese; 1 Egyptian; 1 N.A.)
POAG USA-Cleveland	249 142F/107M	71	34-94	58% Caucasian (144) 38% African American (94) 4% Others (5 Asian; 4 Hispanic; 2 Native American)
Control India-Madurai	20 8F/12M	53	40-70	Indian
POAG India-Madurai	22 4F/18M	61	39-75	Indian

Table 2. Cross-validation of Training Set

Condition	True Value	Prediction	P value ratio	glaucoma votes	glaucoma P value	control votes	control P value
3527	glaucoma	glaucoma	0.000399	10	0.000399	0	1
3533	glaucoma	glaucoma	0.000399	10	0.000399	0	1
3534	glaucoma	glaucoma	0.00609	9	0.00609	1	1
3535	glaucoma	glaucoma	0.000399	10	0.000399	0	1
3536	glaucoma	glaucoma	0.000399	10	0.000399	0	1
3537	glaucoma	control	0.156	3	0.96	7	0.149
3538	glaucoma	glaucoma	0.0401	8	0.0399	2	0.994
544	control	control	0.000399	0	1	10	0.000399
554	control	control	0.000399	0	1	10	0.000399
564	control	control	0.000399	0	1	10	0.000399
573	control	control	0.000399	0	1	10	0.000399
654	control	control	0.0401	2	0.994	8	0.0399
11 correct predictions, 1 incorrect predictions, 0 not predicted					92%		

Test set (n=12): 7 POAG and 5 normal controls. Test set data is not part of the original n=60 training set.
Training set (n=60): 30 POAG and 30 normal controls
The identity of the test set was predicted using the training set and 136 significant peaks (p=0.001).

Model testing and validation using K-nearest-neighbor. We tested our prediction model using the Class Prediction Tool in *GeneSpring*. This table shows results from a randomly selected group of patients (n=12) that was predicted against one of our training sets (n=60). Testing of random samples not part of the original training set serves to evaluate the accuracy of the peptidomic profiling methods used to identify POAG plasma.

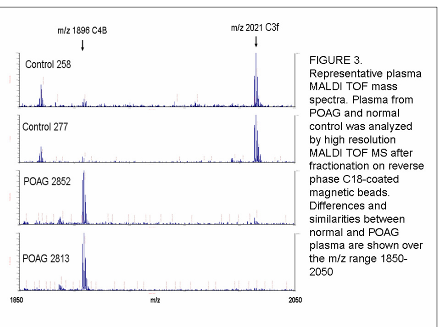
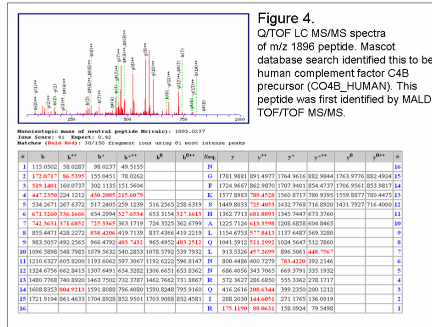


FIGURE 3. Representative plasma MALDI TOF mass spectra. Plasma from POAG and normal control was analyzed by high resolution MALDI TOF MS after fractionation on reverse phase C18-coated magnetic beads. Differences and similarities between normal and POAG plasma are shown over the m/z range 1850-2050



Monoisotopic mass of neutral peptide Mr(calc): 1895.0237
Ions Score: 41 Expect: 0.42
Matches (Bold Red): 30/150 fragment ions using 81 most intense peaks

#	b	b++	b*	b*++	b0	b0++	Seq.	y	y++	y*	y*++	y0	y0++	#
1	115.0502	58.0287	98.0237	49.5155			N							16
2	172.0717	86.5395	155.0451	78.0262			G	1781.9881	891.4977	1764.9616	882.9844	1763.9776	882.4924	15
3	319.1401	160.0737	302.1135	151.5604			F	1724.9667	862.9870	1707.9401	854.4737	1706.9561	853.9817	14
4	447.2350	224.1212	430.2085	215.6079			K	1577.8983	789.4528	1560.8717	780.9395	1559.8877	780.4475	13
5	534.2671	267.6372	517.2405	259.1239	516.2565	258.6319	S	1449.8033	725.4053	1432.7768	716.8920	1431.7927	716.4000	12
6	671.3260	336.1666	654.2994	327.6534	653.3154	327.1613	H	1362.7713	681.8893	1345.7447	673.3760			11
7	742.3631	371.6852	725.3365	363.1719	724.3525	362.6799	A	1225.7124	613.3598	1208.6858	604.8465			10
8	855.4471	428.2272	838.4206	419.7139	837.4366	419.2219	L	1154.6753	577.8413	1137.6487	569.3280			9
9	983.5057	492.2565	966.4792	483.7432	965.4952	483.2512	Q	1041.5912	521.2992	1024.5647	512.7860			8
10	1096.5898	548.7985	1079.5632	540.2853	1078.5792	539.7932	L	913.5326	457.2699	896.5061	448.7567			7
11	1210.6327	605.8200	1193.6062	597.3067	1192.6222	596.8147	N	800.4486	400.7279	783.4220	392.2146			6
12	1324.6756	662.8415	1307.6491	654.3282	1306.6651	653.8362	N	686.4056	343.7065	669.3791	335.1932			5
13	1480.7768	740.8920	1463.7502	732.3787	1462.7662	731.8867	R	572.3627	286.6850	555.3362	278.1717			4
14	1608.8353	804.9213	1591.8088	796.4080	1590.8248	795.9160	Q	416.2616	208.6344	399.2350	200.1212			3
15	1721.9194	861.4633	1704.8928	852.9501	1703.9088	852.4581	I	288.2030	144.6051	271.1765	136.0919			2
16							R	175.1190	88.0631	158.0924	79.5498			1

Figure 4. Q/TOF LC MS/MS spectra of m/z 1896 peptide. Mascot database search identified this to be human complement factor C4B precursor (CO4B_HUMAN). This peptide was first identified by MALDI TOF/TOF MS/MS.

m/z 1896 — control n=138 mean 31767.55 ± SE 2248.14; POAG n=249 mean 75264.99 ± SE 4109.4; * p<0.0001

m/z 2377 — control n=138 mean 524.15 ± SE 30.95; POAG n=249 mean 1558.54 ± SE 50.79; * p<0.0001

m/z 2021 — control n=138 mean 131029.98 ± SE 10921.65; POAG n=249 mean 16648.82 ± SE 2597.58; * p<0.0001

FIGURE 5. Statistical analysis of complement factor peptides in plasma of normal control vs POAG patients. Normalized intensities for the following peptides (m/z 1896, 2021 and 2377) from each donor plasma was analyzed using one-way analysis of variance test. The peptides were sequenced and identified as peptide fragments from complement factors C4B and C3. C4B peptides with m/z values of 1896 and 2377 were 1.4x and 2x higher in POAG plasma respectively. C3 peptide (m/z 2021) was ~7x lower in POAG compared to normal control plasma.

m/z 2377 C4B — DDPDAPLQPVTPLQLFEGRRN
m/z 2021 C3f — SSKITHRIHWESASLLR
m/z 1896 C4B — NGFKSHALQLNNRQIR

POAG CONTROL

Crossvalidation → 338/387 (87%)
***Sensitivity 0.86**
Specificity 0.91

Figure 6. Analysis of plasma peptide profiling based on three complement factor peptides. Columns represent sample; rows are m/z peaks with mono-isotopic mass on the left and peptide sequence on the right. Class prediction by *GeneSpring* based on the 3 peptides correctly predicted 338 out of 387 (87.3%) samples with test accuracy of 0.86 sensitivity and 0.91 specificity. An ELISA based test developed against these and other peptides that could potentially be used as a prognostic tool for POAG. The identity of peptides such as these may provide mechanistic clues into the etiology of this disease.

***Sensitivity** refers to the proportion of people with disease who have a positive test result
Specificity refers to the proportion of people without disease who have a negative test result.

relative density (au): control 494, control 515, control 517, control 519, POAG 3047, POAG 3408, POAG 3409, POAG 3415, POAG 3452
0.144 0.082 0.069 0.048 0.1 0.255 2.63 0.496 0.277

Control n=4; POAG n=5; *p<0.05

Figure 7. Corroboration of plasma peptide profiling data. Plasma from normal control (n=4) and POAG (n=5) was albumin depleted by acetone/TCA precipitation, 30ug separated by SDS-PAGE and transferred onto Immobilon-P membrane and probed with an anti-human C4 antibody (CompTech, Inc.) An immunoreactive band (~75 kDa), possibly C4b-B fragment of C4 was detected. Densitometric analysis of the anti-C4 immunoreactive bands showed 3-fold higher levels in POAG compared to normal control plasma.

FUTURE DIRECTIONS

The long term goal is the development of routine diagnostic technology for use in clinical medicine to predict POAG susceptibility, essentially as serum cholesterol measurements are used in risk management for cardiovascular disease. To accomplish this goal, validation of the predictive value of POAG plasma biomarkers will be sought in subsequent years of studies. Future studies will be required to compare biomarker composition across broader population ethnic groups with larger numbers of samples to establish biomarkers for population subgroups and to extend the study to other forms of the disease like primary angle closure glaucoma (PACG). In future studies, polyclonal antibodies will be generated to constituent peptides identified in POAG plasma patterns and tested by ELISA for diagnostic utility. MALDI-TOF mass spectrometry and ELISA methods will then be used in a prospective case-control study to validate the prognostic utility of POAG peptidomic patterns and POAG-specific antibodies in predicting POAG susceptibility.