Association Between Osteopontin and Human Abdominal Aortic Aneurysm


Objectives—In vitro and animal studies have implicated osteopontin (OPN) in the pathogenesis of aortic aneurysm. The relationship between serum concentration of OPN and variants of the OPN gene with human abdominal aortic aneurysm (AAA) was investigated.

Methods and Results—OPN genotypes were examined in 4227 subjects in which aortic diameter and clinical risk factors were measured. Serum OPN was measured by ELISA in two cohorts of 665 subjects. The concentration of serum OPN was independently associated with the presence of AAA. Odds ratios (and 95% confidence intervals) for upper compared with lower OPN tertiles in predicting presence of AAA were 2.23 (1.29 to 3.85, P=0.004) for the population cohort and 4.08 (1.67 to 10.00, P=0.002) for the referral cohort after adjusting for other risk factors. In 198 patients with complete follow-up of aortic diameter at 3 years, initial serum OPN predicted AAA growth after adjustment for other risk factors (standardized coefficient 0.24, P=0.001). The concentration of OPN in the aortic wall was greater in patients with small AAAs (30 to 50 mm) than those with aortic occlusive disease alone. There was no association between five single nucleotide polymorphisms or haplotypes of the OPN gene and aortic diameter or AAA expansion.

Conclusions—Serum and tissue concentrations of OPN are associated with human AAA. We found no relationship between variation of the OPN gene and AAA. OPN may be a useful biomarker for AAA presence and growth.


Key Words: abdominal aortic aneurysm • genetic polymorphisms • osteopontin

Abdominal aortic aneurysm (AAA; OMIM#100070) is associated with considerable morbidity and mortality in the Western world.1 The identification, monitoring, and selection of patients for treatment presently rely on aortic imaging and clinical judgment.2 The cause of AAAs remains unclear. Identified risk factors for AAA include increasing age, male gender, smoking, coronary heart disease (CHD), hypertension, dyslipidemia, and a positive family history.3–5 A number of studies have examined the association between circulating and genetic factors with the presence or expansion of AAAs; however, there have been no replicated findings to date.6–9 Because accumulation of macrophages and lymphocytes is a consistent observation in biopsies of human AAA, circulating markers of inflammation may play a role in the identification of patients with AAA or those more likely to expand to a size requiring treatment.9,10

See page 439

Osteopontin (OPN, SPP1, OMIM*166490) is a phosphor-ylated acidic glycoprotein that has been implicated in a large number of physiological and pathological processes including bone remodelling, vascular calcification, and tumor metastasis.11,12 Experimental studies support a role for OPN in promoting inflammation, proteolysis, and atherosclerosis, which are all integral processes in AAA.12–18 In fact, OPN has been specifically linked with development of AAA in an animal model.19 In humans, common variants of the OPN gene have been associated with immunologic diseases and also related to the serum concentrations of the protein.20–23 The aim of the present study was to assess the relationship between serum OPN concentrations, polymorphisms of the OPN gene, and AAA presence and growth in humans.

Methods

Study Design
We tested the following hypotheses in this study: (1) Serum (a) and tissue (b) OPN concentrations are associated with the presence of AAA; (2) Serum OPN concentration is associated with the growth rate of AAA; (3) Selected single nucleotide polymorphisms (SNPs) in the OPN gene are associated with the presence of AAA; (4) Selected single nucleotide polymorphisms (SNPs) in the OPN gene...
are associated with the growth rate of AAA. To examine these hypotheses we assessed patients from two different populations. Firstly, we used patients from the Western Australia Screening Study. DNA was available from 4227 subjects of which 689 had an AAA (maximum aortic diameter ≥30 mm) and 3538 did not (maximum aortic diameter <30 mm). This sample was used to examine hypothesis three. Serum was available from 233 patients with AAA from this population. We selected a similar number of controls (n=233) from the screened population who did not have an AAA. Selection of these subjects was based on aortic diameter (19 to 22 mm) and a similar age range. A previous study suggested normal aortic diameter in this population and age group for men is 19 to 22 mm. This group of screen detected patients with AAA, and matched controls were used to assess hypothesis 1a. To further examine the value of OPN as a biomarker for AAA we examined a second population. This cohort was a consecutive series of patients referred to a tertiary vascular referral center for aortic assessment. 132 of this referral population were found to have an AAA, and 67 did not have an aortic aneurysm. This population of referral detected AAAs (n=132), and controls (n=67) were also used to assess hypothesis 1a. From this referral population 24 patients underwent open aortic surgery and were used to assess hypothesis 1b. Hypothesis 4 was examined in a subset of patients with AAA (n=644) from the screened population in which 6 monthly ultrasound surveillance was carried out for a minimum of 12 months (median 54 months). Hypothesis 2 was examined in a subset of patients in which serum was available, who had been identified from the screened (n=146) or referral (n=52) populations in which ultrasound follow-up was available for a minimum of three years. The study was approved by the relevant ethics committees, and informed consent was obtained from participating patients.

### Clinical Data and Definitions

To adjust for other risk factors for AAA development and progression we collected a variety of clinical characteristics including: age, gender, hypertension, diabetes, dyslipidemia, coronary heart disease (CHD), smoking history, peripheral arterial disease (PAD), family history of AAA, and blood pressure (BP) (see supplementary methods for definitions available online at http://atvb.ahajournals.org).

### Assessment of Aortic Diameter

The maximum transverse and anteroposterior diameter of the infrarenal abdominal aorta was measured by an experienced vascular sonographer at each site utilizing a 3.75 MHz probe and their own validated ultrasound equipment (Toshiba Capasee, Philips HDI 5000; GE Logic 9). The reproducibility of aortic measurements is regularly assessed in these vascular laboratories. For example, the reproducibility of aortic readings was recently assessed in one laboratory by two repeat readings on 10 patients with AAA. The coefficient of repeatability was 1.2 mm, and the limit of agreement −0.8 to 1.6 mm.

### Analysis of Serum and Aortic OPN

Blood was collected from patients after an overnight fast. Serum was stored at −80°C until later batch assessment of OPN concentrations using ELISA according to manufacturer’s instructions and expressed as ng/mL (R&D Systems). This assay was selected because a previous study demonstrated excellent recovery and intra- and interassay reproducibility in our laboratory. In a subgroup of patients undergoing open aortic surgery for occlusive disease (n=6) or AAA (n=18), biopsies taken from the anterior wall of the aorta opposite the inferior mesenteric artery were available for measurement of tissue OPN concentration. The cohort included 6 patients with small AAAs (<50 mm) in which the decision to carry out aortic intervention was also influenced by concurrent symptoms of occlusive iliac atheroma. Aortic biopsies were ground under liquid nitrogen (LN2), and proteins were extracted and quantified as previously described. The concentration of OPN within aortic biopsies was measured using ELISA (R&D Systems) and expressed as pmol/L per mg of protein. Western analysis was used to confirm the results of ELISA as previously described. The localization of OPN in aortic biopsies was investigated with immunohistochemistry as previously reported.

### C-Reactive Protein Assay

C-reactive protein (CRP) was measured by a high-sensitivity assay, with the use of the particle-enhanced immunonephelometry system on the BNII analyzer (Dade Behring).

### Genotyping

Fifty-five single nucleotide polymorphisms (SNPs) have been identified in the human OPN gene (SPP1, http://www.ncbi.nlm.nih.gov/projects/SNP). Many of these are within noncoding regions of the gene and have not been related to any functional effects. We selected 5 SNPs (rs4754, location exon 6, rs9138, location 3′-UTR, rs1226616, location exon 7, rs1126772, location 3′-UTR, and rs11730582, location promoter) based on previous association with autoimmune diseases and ability to determine haplotypes predicting serum OPN concentrations. Genotyping was performed using real-time quantitative polymerase chain reaction (PCR) utilizing custom Taqman probes (see supplemental methods, available online at http://atvb.ahajournals.org, and Table 1). Error rate was assessed by duplicate analysis of 10% of the samples and found to be <0.05%.

### Statistical Analyses

The characteristics of patients with AAA and controls were compared with Pearson Chi-Squared for nominal variables and Mann-Whitney U test for continuous variables (Tables 1 and 2). We used binary logistic regression analysis to examine the relationship between serum OPN and CRP with AAA allowing for other known determinants of aneurysm (Table 3). The ability of serum OPN and CRP to distinguish the patients with AAA was investigated using receiver operating characteristic (ROC) curves and area under the curve (AUC). The association of aortic OPN concentration with aortic diameter group was examined with Kruskal Wallis and Dunnnett T3 posthoc tests (Figure 1). Trend analysis revealed that change in aortic diameter over time approximated linear, therefore the association of serum OPN with aortic expansion was assessed.

### Table 1. Comparison of Patients With and Without AAA Undergoing Serum OPN Assessment in the Population Cohort (n=466)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>AAA</th>
<th>No AAA</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aortic diameter, mm</td>
<td>35.0±5.0</td>
<td>21.0±1.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No.</td>
<td>233</td>
<td>233</td>
<td></td>
</tr>
<tr>
<td>Age, years</td>
<td>76.0±5.9</td>
<td>74.6±7.4</td>
<td>0.29</td>
</tr>
<tr>
<td>Male</td>
<td>233</td>
<td>233</td>
<td>1.0</td>
</tr>
<tr>
<td>Hypertension</td>
<td>113</td>
<td>81</td>
<td>0.005</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>18</td>
<td>12</td>
<td>0.28</td>
</tr>
<tr>
<td>Dyslipidemia</td>
<td>149</td>
<td>122</td>
<td>0.009</td>
</tr>
<tr>
<td>Current smoker</td>
<td>32</td>
<td>14</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>156</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>Non smoker</td>
<td>45</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>CHD</td>
<td>86</td>
<td>31</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CAD</td>
<td>17</td>
<td>14</td>
<td>0.63</td>
</tr>
<tr>
<td>PAD</td>
<td>60</td>
<td>30</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CRP, mg/dl</td>
<td>4.92±8.73</td>
<td>3.41±8.09</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>OPN, ng/ml</td>
<td>77.19±31.86</td>
<td>63.33±29.14</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Nominal variables are presented as numbers and continuous variables as mean±SD. CHD indicates coronary heart disease; CAD, cerebral artery disease; PAD, peripheral artery disease; CRP, C-reactive protein; OPN, osteopontin.
with linear multiple regression analysis to allow for other known determinants of AAA progression (initial aortic diameter, diabetes, and smoking history). SNPs were tested for deviations from Hardy-Weinberg equilibrium using an exact Markov-Chain Monte Carlo (MCMC) test.26 The association between OPN genotype and haplotype frequencies were estimated using the Bayesian-MCMC framework implemented in Phase v2.1.1 and statistical significance was assessed by means of permutation testing.29

### Results

**Serum OPN and AAA**

We examined the association between serum OPN and AAA in 2 cohorts. Firstly, we compared serum OPN in 233 patients with AAA (220 patients with 30 to 49 mm diameter AAAs and 3 patients with 50 to 54 mm diameter AAAs) and 233 controls with normal aortic diameter (19 to 22 mm) identified from a population screening study (Table 1). Secondly, we examined serum from a cohort of 199 patients referred for aortic assessment at a tertiary vascular center (31 patients with ≥50 mm diameter AAAs, 101 patients with 30 to 49 mm diameter AAAs, and 67 patients with aortic diameter <30 mm; Table 2). Serum OPN was significantly higher in patients with AAA (Tables 1 and 2). OPN was independently associated with the presence of AAA after adjusting for other risk factors in both cohorts. Subjects with serum OPN in the highest tertile were more than twice as likely to have an AAA compared with those with serum OPN in the lowest tertile after allowing for other risk factors in the population cohort (Table 3). Independent odds ratio for upper and middle OPN tertiles were 4.08 (95% CI 1.67 to 10.00, *P* = 0.002) and 3.06 (95% CI 1.30 to 7.23, *P* = 0.01) in the referral cohort. Serum CRP was also associated with the presence of AAA in the population cohort only (Table 1). The association of serum OPN with AAA was independent of CRP level (Table 3). Mean serum OPN was 77.4±43.5 ng/mL in patients with 30 to 49 mm AAAs compared with 77.3±44.5 ng/mL in those with 50 to 80 mm AAAs in the referral cohort, *P* = 0.99. The AUC from ROC analyses of serum OPN in determining the presence of AAA was 0.65 (95% CI 0.60 to 0.70, *P* <0.0001) compared with 0.61 (95% CI 0.56 to 0.66, *P* <0.0001) for CRP in the population cohort. Combining serum OPN and CRP did not improve accuracy in detecting AAA (AUC 0.66, 95% CI 0.61 to 0.70). An OPN concentration of 57.2 ng/mL gave the most accurate diagnosis of AAA but only had a sensitivity of 73% and a specificity of 52%. For screening purposes we require a test with high sensitivity. An OPN

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>AAA (n=132)</th>
<th>No AAA (n=67)</th>
<th><em>P</em> Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serum OPN (ng/mL)</td>
<td>73.6±43.60</td>
<td>58.6±43.83</td>
<td>0.001</td>
</tr>
<tr>
<td>CRP (mg/dl)</td>
<td>8.65±21.57</td>
<td>9.09±29.80</td>
<td>0.88</td>
</tr>
</tbody>
</table>

The reference comparisons are to subjects without the risk factor or with values in the lowest tertile. CHD indicates coronary heart disease; OPN, osteopontin; CRP, C-reactive protein.
Aortic Wall Concentrations of OPN
The concentration of OPN in the aortic wall was measured in 24 patients undergoing open aortic surgery for occlusive disease or AAA. The patients were divided into 3 groups based on aortic diameter as aortic occlusive disease only (diameter <30 mm, n=6, mean diameter 20.8±4.2 mm), small AAA and occlusive disease (diameter 30 to 50 mm, n=6, mean diameter 44.4±6.8 mm), and large AAA (diameter >50 mm, n=12, mean diameter 65.5±16.4 mm). Aortic OPN concentration was associated with diameter group (P=0.01), being higher in patients with small AAAs compared with those with aortic occlusive disease (P=0.04) or large AAA (P=0.008; Figure 1). The higher concentration in biopsies from small AAAs was confirmed by Western blotting (data not shown). Immunohistochemistry demonstrated that OPN was most markedly expressed within the intimal atheroma of aortic biopsies (Figure 2). In biopsies taken from small AAAs OPN was also demonstrated within the media and adventitia (Figure 2b).

Serum OPN and AAA Growth
We assessed the association of serum OPN concentration with aneurysm growth (mean increase in maximum aortic diameter over three years 3.5±3.3 mm, range −5 to 14.2 mm) in 198 patients (mean age 76.7±5.0 years) with small AAAs (mean aortic diameter 34.0±3.8 mm) followed by ultrasound surveillance for 3 years (complete follow-up). Serum OPN was correlated with aortic diameter change (r=0.24, P=0.001). After adjusting for other known risk factors for aortic expansion (initial aortic diameter, diabetes mellitus, and smoking history) serum OPN predicted AAA growth (F=5.42, standardized coefficient=0.24, P=0.001).

OPN Genotype and AAA Presence and Growth
The characteristics of the 4227 patients genotyped for OPN genetic polymorphisms are shown in Table 4. All loci were in Hardy-Weinberg equilibrium in control samples and were therefore tested for association with AAA (Table 5). No significant associations were found between any of the genotyped loci and AAA or aortic diameter (data not shown).
example, mean serum OPN concentration for subjects with polymorphisms at rs11226616 (previously associated with OPN concentration) were 68.08±28.62, 72.40±37.88 and 74.05±37.88 for those who were common homozygotes (n=245), heterozygotes (n=189) or rare homozygotes (n=32), P=0.47.

Discussion
In this study we provide evidence to support an association between OPN and human AAA. Serum concentrations of OPN were significantly higher in subjects with AAA, and the association between OPN level and AAA was independent of other risk factors for AAA. We compared OPN with the inflammatory marker CRP in terms of predicting AAA presence (Tables I through 3). We found that unlike CRP, OPN predicted AAA presence in both cohorts examined. Serum OPN was also associated with faster growth of small AAAs, which we have previously demonstrated is not the case for CRP.9 Using ROC curves we found serum OPN concentration to be better at identifying individuals with AAA than CRP but the sensitivity and specificity achieved were too poor to support its value as a screening test. The concentration of OPN within biopsies of small AAAs was greater than demonstrated in occlusive aortic disease or large AAAs. However, no association between variants of the OPN gene, including haplotypes, and AAA was found.

OPN has been shown to promote inflammation by a number of mechanisms, including supporting macrophage and T cell chemotaxis and adhesion, prolonging lymphocyte survival, and enhancing cell mediated immunity.13,14 OPN is also involved in the activation of proteolytic pathways such as paracrine augmentation of pro-MMP9 activity.18 Deficiency of and therapy to reduce OPN protects against aneurysm formation in an animal model.19,30 These findings coupled with the serum studies outlined in this project suggest a role for this cytokine in AAA development and progression. It is possible that the high serum concentrations of OPN in patients with AAA simply reflect a response to AAA formation or are related to the environmental risk factors for the disease. The lack of association between genetic polymorphism for OPN and AAA could be interpreted as supporting this view. However, given the complex pathophysiology of AAA, it is likely that any single genetic variant will only have a weak effect on the development or progression of the condition. The demonstration of high concentrations of OPN in biopsies of small but not large AAAs by comparison to patients with aortic occlusive disease alone also brings in to question the importance of OPN in aneurysm progression. It should be noted that the majority of the patients in this study had small rather than large AAAs. Thus the association of serum OPN with AAA refers to this population. We postulate that OPN plays a role in the early development and progression of AAA. High concentrations of OPN in the aorta of some patients with aortic occlusive disease could encourage the excessive positive remodelling and medial destruction typical of AAA. A larger sample of patients with aortic biopsies and serum will be required to investigate this in more detail.

In conclusion our findings support animal data suggesting a role for OPN in the initial development of AAA. Further studies are required to assess whether serum OPN combined with other biomarkers may be useful in identifying patients with AAA or predicting disease progression. We found no evidence to support the importance of genetic polymorphism in OPN in the pathogenesis of AAA.

Acknowledgments
Thanks to the participants and staff involved in the Western Australian AAA Screening Study and Health in Men Study.

Sources of Funding
This project is supported by The Townsville Hospital Private Practice Fund and grant number RO1 HL080010–01 from the National Institutes of Health and NHMRC Practitioner Fellowships (431503 and 458505 to J.G. and P.E.N.).

Disclosures
None.

References


Association Between Osteopontin and Human Abdominal Aortic Aneurysm

Arterioscler Thromb Vasc Biol. 2007;27:655-660; originally published online December 14, 2006;

DOI: 10.1161/01.ATV.0000255560.49503.4e

Arteriosclerosis, Thrombosis, and Vascular Biology is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2006 American Heart Association, Inc. All rights reserved.
Print ISSN: 1079-5642. Online ISSN: 1524-4636

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://atvb.ahajournals.org/content/27/3/655

Data Supplement (unedited) at:
http://atvb.ahajournals.org/content/suppl/2006/12/15/01.ATV.0000255560.49503.4e.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Arteriosclerosis, Thrombosis, and Vascular Biology can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Arteriosclerosis, Thrombosis, and Vascular Biology is online at:
http://atvb.ahajournals.org//subscriptions/
Supplemental Data

Methods

Definitions: Hypertension was defined by a history of high blood pressure, receiving treatment to reduce blood pressure or a systolic blood pressure $\geq 140/80$ [1]. Diabetes was defined by a fasting blood glucose $\geq 7.0$ mmol/L, or history of, or treatment for hyperglycaemia [2]. Dyslipidaemia was defined by a history of or treatment for high cholesterol, or a total cholesterol $\geq 4.5$ mM, triglycerides $\geq 2$ mM, low density lipoprotein $\geq 2.5$ mM or high density lipoprotein $\leq 1.0$ mM [3]. CHD was defined by a history of angina, myocardial infarction or previous treatment by coronary medications or intervention. Smoking status was classified into current smokers (smoked within the last month), ex-smokers (given up for more than one month) and never smokers based on history. Peripheral artery disease (PAD) was recorded if the patient reported a history of intermittent claudication, ankle brachial pressure index was $<0.9$ or the patient had been surgically treated with lower limb angioplasty, bypass or amputation. Cerebral artery disease (CAD) was defined by a history of stroke, transient ischemic attack or previous cerebral revascularization. A positive family history was recorded if a patient reported that a mother, father, brother or sister had or died from an AAA.

OPN ELISA: The R&D quantikine assay is provided with detailed manufacturer’s specifications including assessment of cross reactivity, precision, linearity and sensitivity. No significant cross reactivity was found to a large number of other proteins including mouse and bovine OPN, enterokinase, thrombin, MMP-3 and -7.
**Genotyping of OPN SNPS using Real-Time Quantitative PCR:** This technique involved the use of custom TaqMan probes (Applied Biosystems, USA), labelled with fluorogenic reporter dyes. The probe consisted of either a FAM or VIC dye with a quencher dye attached. While the probe is intact, the quencher dye significantly reduces the fluorescence emitted by the reporter dye (FAM or VIC) by Förster resonance energy transfer (FRET) through space. TaqMan probes incorporate minor groove technology (MGB) on the 3’ end. The MGB molecule binds to the minor groove of the DNA helix, improving hybridisation-based assays by stabilising the MGB-probe/template complex. Detection was achieved with 5’ nuclease chemistry by means of exonuclease cleavage of a 5’ allele-specific dye label, which generates the permanent assay signal. A non-fluorescent quencher is attached to the MGB probe which virtually eliminates any background fluorescence normally associated with traditional quencher [4]. A number of optimisation experiments demonstrated that 10ng of DNA per reaction gave the best genotyping results. Master Mix containing probes and primers was added to each well of the DNA pre-printed 384 well MJ Research Black, Hard-Shell, Thin-Wall micro plates (Supplemental Table I). The 384 well plates were then amplified in an MJ research PTC – 200 Peltier Thermocycler with a 384 well block, under the following conditions: 1) 50˚C for 2 min, 2) 95˚C for 10 min, 3) 92˚C for 15 sec and 4) 60˚C for 1 min. Steps 3 & 4 were repeated 39 times. The plates were then spun briefly at 1500rpm using a Quantum Scientific 2-5 Sigma Plate Spinner, and read using a Perkin Elmer Life Science Wallac VICTOR\(^2\) 1420 Multi-label Counter. If satisfactory separation of the fluorescent markers across the graph was achieved after 40 cycles then the results were kept and used, else the plate was re-lidded and run for a further 5 cycles in the thermocycler and once again re-checked. This process was repeated until satisfactory separation was achieved. An
example of a fluorescence graph with satisfactory spread is shown in Supplemental Figure 1.
References


### Supplemental Table I: Probes for OPN genotyping.

<table>
<thead>
<tr>
<th>SNP rs Number</th>
<th>ABI Assay ID</th>
<th>Probe</th>
</tr>
</thead>
<tbody>
<tr>
<td>4754</td>
<td>C---1840817-10</td>
<td>ATGATATGGATGATGAAAGATGATGA[C/T]GACCATGTGGA CAGCCAGGACTCCA</td>
</tr>
<tr>
<td>1126616</td>
<td>C---1840818-10</td>
<td>AGTCCAGATTATATAAGCGAAAGC[C/T]AATGATGAGA GCAATGAGCATTCCG</td>
</tr>
<tr>
<td>1126772</td>
<td>C---8826999-10</td>
<td>TATCTATTTTGGTCGGAATAACT[A/G]ATGTGGTTGAT AATTAGTTCGGTT</td>
</tr>
<tr>
<td>9138</td>
<td>C---8826997-10</td>
<td>TCTCATGAAATAGAAATTATGTGAA[A/C]GCAAAAAAAT ACTTTACCCCNCTT</td>
</tr>
<tr>
<td>11730582</td>
<td>C---1840808-10</td>
<td>GAGTAGTAAAGGACAGGCAAGTT[C/T]TCTGAACCTCCT TGCAGGCTTGACA</td>
</tr>
</tbody>
</table>

Yellow= labelled with VIC.

Blue= labelled with FAM.

The probe for the first named nucleotide in the SNP (eg G if it is G/A) was labelled VIC. The probe for the second named nucleotide in the SNP (in this case A) was labelled FAM. If the probe was reverse strand, the G allele would still be VIC (even though the probe would contain a C at that point) and the A allele would be FAM (even though the probe would contain a T at that point).
Supplemental Table II: Pairwise LD measures.

<table>
<thead>
<tr>
<th></th>
<th>rs1126772</th>
<th>rs9138</th>
<th>rs4754</th>
<th>rs1126616</th>
<th>rs11730582</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1126772</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>rs9138</td>
<td>0.74</td>
<td>0.99</td>
<td>0.99</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>rs4754</td>
<td>0.74</td>
<td>0.98</td>
<td>0.99</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>rs1126616</td>
<td>0.74</td>
<td>0.98</td>
<td>0.98</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>rs11730582</td>
<td>0.25</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
<td></td>
</tr>
</tbody>
</table>

Lower triangle is D'; upper triangle is $r^2$. 


Supplemental Table III: Estimated haplotype frequencies and their standard errors.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>No AAA (SE)</th>
<th>AAA (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AACCC</td>
<td>0.465 (0.001)</td>
<td>0.472 (0.003)</td>
</tr>
<tr>
<td>AACCT</td>
<td>0.272 (0.001)</td>
<td>0.275 (0.003)</td>
</tr>
<tr>
<td>CGTTT</td>
<td>0.204 (0.0007)</td>
<td>0.197 (0.002)</td>
</tr>
<tr>
<td>CATTC</td>
<td>0.047 (0.0006)</td>
<td>0.047 (0.0015)</td>
</tr>
<tr>
<td>Other</td>
<td>0.012 (NA)</td>
<td>0.009 (NA)</td>
</tr>
</tbody>
</table>

Rare haplotypes with a frequency < 0.01 are grouped into the other category.

Permutation (n = 1000) for differences in haplotype frequencies p=0.345. NA= Not applicable
The majority of samples have clustered into distinct groups across the graph which makes assigning genotypes a simpler process.