UNIVERSITY OF CALIFORNIA, SAN DIEGO

In Silico, In Vitro, In Vivo and In Populo:

Regulatory Genetics of Single Nucleotide Polymorphisms

in the

Phenylethanolamine N-Methyltransferase Promoter

A Dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy

in

Bioinformatics

by

Juan Lorenzo Rodriguez Flores

Committee in charge:

Professor Daniel T. O'Connor, Chair
Professor Vineet Bafna
Professor Shankar Subramanian
Professor Glenn Tesler
Professor Wei Wang
Professor Michael Ziegler

2009

The Dissertation of Juan Lorenzo Rodriguez Flores is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2009

DEDICATION

I dedicate this dissertation to my family. The Rodriguez, Flores, Quesada and Forastieri of Puerto Rico, a tribe whose history, accomplishments and adventures constantly remind me of who I am, where I came from and what I am capable of.

I also dedicate this dissertation to Tatiana, who first danced with me the day I advanced to candidacy and who shortly after became my soulmate and the love of my life. She provided infinite support, motivation and nourishment for my graduation.

EPIGRAPH

We have been told we cannot do this by a chorus of cynics who will only grow louder and more dissonant in the weeks to come. We've been asked to pause for a reality check. We've been warned against offering ... false hope.

But in the unlikely story that is America, there has never been anything false about hope....

Yes we can.

Barak H. Obama

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

SNP: Single Nucleotide Polymorphisms

PNMT: Phenylethanolamine N-Methyltransferase

EMSA: Electrophoretic Mobility Shift Assay

LD: Linkage Disequilibrium

GWAS: Genome Wide Association Scan

KO: Knockout

PWM: Position Weight Matrix

PFM: Position Frequency Matrix

SBP: Systolic Blood Pressure

DBP: Diastolic Blood Pressure

BMI: Body Mass Index

FW: Forward Primer

RV: Reverse Primer

KU: Kaiser Unrelated

SDT: San Diego Twin Study

MAF: Minor Allele Frequency

ANOVA: Analysis of Variance

mRNA: messenger Ribonucleic Acid

DNA: Deoxyribonucleic Acid

ATP: Adenine Tri-Phosphate

AMP: Adenine Mono-Phosphate

cAMP: cyclic AMP

PKA: Phosphokinase A

HSL: Hormone Sensitive Lipase

A: Adenine

T: Thymine

G: Guanine

C: Cytosine

A.A: A haplotype

G-161A: SNP at position -161 with G allele ancestral (chimp) and A mutant.

-161A: A allele of -161 SNP

G-367: G allele of -367 SNP.

LIST OF ILLUSTRATIONS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Bettie Graham and the National Human Genome Research Institute for funding my research with a Minority Predoctoral Fellowship.

Chapters 1, 2, 3, 4 of this dissetation in part are currently being prepared for submission for publication. Chapter 1 was co-authored by Gen Wen, Ryan Friese, Manjula Mahata, Maple Fung, Rany Salem and Bruce Hamilton. Chapter 2 was co-authored by Chih Min Tang and Kuixing Zhang. Chapter 3 was co-authored by Sushil K Mahata, Stephane Chiron, and Maite Courel. Chapter 4 was co-authored by Fujun Liu and Michael Ziegler. All chapters were co-authored by Daniel O'Connor. The dissertation author was the primary investigator and author of this material.

VITA

1994-1998	Massachusetts Institute of Technology, Cambridge, M.A. Bachelor of Science, Biology
2001 - 2001	Neurobehavioral Research Institute, Corte Madera, C.A. Software Engineer
2002 - 2002	Jacobson Lab, Viral Immunology Section, Neuroimmunology Branch, NINDS, NIH, Bethesda, MD Biologist
2003 - 2006	University of California, San Diego Master of Science, Bioinformatics
2004 -2005	Science Applications International Corp, San Diego, C.A. Java Software Engineering Intern
2007-2009	University of California, San Diego Doctor of Philosophy, Bioinformatics

AWARDS & FELLOWSHIPS

2002 - 2003	Minority Access to Science Engineering and Math Program, Graduate Fellowship, National Science Foundation.
2004 - 2005	San Diego Export Center Fellowship.
2005 - 2006	San Diego Fellowship, University of California.
2006 – 2009	Ruth L. Kirschstein National Research Service Award, National Human Genome Research Institute.

TEACHING EXPERIENCE

2007	Biological Data Analysis, Teaching Assistant
2008	Biological Sequence and Structure: Algorithms and Analysis, Teaching Assistant.

PUBLICATIONS

Rao, F., Zhang, L., Wessel, J., Zhang, K., Wen, G., Kennedy, B.P., Rana, B.K., Das, M., **Rodriguez-Flores, J.L.**, Smith, D.W., et al. (2007). Tyrosine hydroxylase, the rate-limiting enzyme in catecholamine biosynthesis: discovery of common human genetic variants governing transcription, autonomic activity, and blood pressure in vivo. Circulation 116, 993-1006.

Chen, Y., Rao, F., **Rodriguez-Flores, J.L.**, Mahapatra, N.R., Mahata, M., Wen, G., Salem, R.M., Shih, P.A., Das, M., Schork, N.J., et al. (2008). Common genetic variants in the chromogranin A promoter alter autonomic activity and blood pressure. Kidney Int 74, 115-125.

Chen, Y., Rao, F., **Rodriguez-Flores, J.L.**, Mahata, M., Fung, M.M., Stridsberg, M., Vaingankar, S.M., Wen, G., Salem, R.M., Das, M., et al. (2008). Naturally occurring human genetic variation in the 3'-untranslated region of the secretory protein chromogranin A is associated with autonomic blood pressure regulation and hypertension in a sex-dependent fashion. J Am Coll Cardiol 52, 1468-1481.

Rao, F., Zhang, L., Wessel, J., Zhang, K., Wen, G., Kennedy, B.P., Rana, B.K., Das, M., **Rodriguez-Flores, J.L.**, Smith, D.W., et al. (2008). Adrenergic polymorphism and the human stress response. Ann N Y Acad Sci 1148, 282-296.

Biswas, N., **Rodriguez-Flores, J.L.**, Courel, M., Gayen, J.R., Vaingankar, S.M., Mahata, M., Torpey, J.W., Taupenot, L., O'Connor, D.T., and Mahata, S.K. (2009). Cathepsin L Co-Localizes with Chromogranin a in Chromaffin Vesicles to Generate Active Peptides. Endocrinology.

Scott A. Jackson, I.P., J Eugene leClerc, Thomas A. Cebula, Mark K. Mammel, Michael L. Kotewicz, Amit Mukherjee, Sushil K. Mahata, **Juan L. Rodriguez Flores**, Manjula Mahata, Daniel T. O'Connor, and Jannavi R. Srinivasan (2009). Principles of Functional Genomic Analysis. In Genomics: Fundamentals and Applications, D.B.C. Supratim Choudhuri, ed. (Informa Health Care).

FIELDS OF STUDY

Major Field: Bioinformatics

Studies in Disease Genetics
Professor Daniel T O'Connor

Studies in Bioinformatics
Professor Vinet Bafna
Professor Wei Wang

Studies in Probability and Statistics
Professor Glenn Tesler

Studies in Phenylethanolamine N-Methyltransferase Biology
Professor Michael Ziegler

Studies in Systems Biology
Professor Shankar Subramaniam

ABSTRACT OF THE DISSERTATION

In Silico, In Vitro, In Vivo and In Populo:
Regulatory Genetics of Single Nucleotide Polymorphisms in the
Phenylethanolamine N-Methyltransferase Promoter

by

Juan Lorenzo Rodriguez Flores
Doctor of Philosophy in Bioinformatics
University of California, San Diego, 2009
Professor Daniel T. O'Connor, Chair

In this dissertation we address the following questions. (1) Are common PNMT promoter SNPs associated with obesity *in populo*? (2) Do the associated and LD-tagged SNPs disrupt mammalian-conserved regulatory motifs predicted *in silico* and verified *in vitro*? (3) Can we replicate *in vitro* haplotype-specific and transcription-factor-specific regulatory effects

consistent with the association? (4) Can we build an *in silico* model of chromaffin cell changes in response to PNMT deficiency in humans based on *in vivo* mRNA expression profies from PNMT KO vs WT mice?

The most significant findings of this dissertation are as follows. Two common SNPs, one which destroys a Sp1 activation motif (rs3764351) and one which creates a Sox17 inhibitory motif (rs876493), together generate a low-activity PNMT promoter which is associated with obesity in European-American females. We identify the SNPs by resequencing, predict function by phylogenetic footprinting, then validate the bioinformatic prediction in EMSA and Luciferase promoter reporter assays. In an GO clustering analysis of adrenal-specific transcriptome changes in PNMT KO versus WT mice we identify an enrichment for obesity-associated and lipid metabolism genes among those differentially expressed. Together, these results present a genetic, molecular and cellular mechanism whereby the obesity-associated SNPs alter chromaffin cell function to produce the obesity phenotype. The experimental results can be used for pharmacogenomics research, while the genetic evidence can be used to help identify women at risk for obesity, a lifespan-reducing disease with prevalence > 20% in the USA.

INTRODUCTION

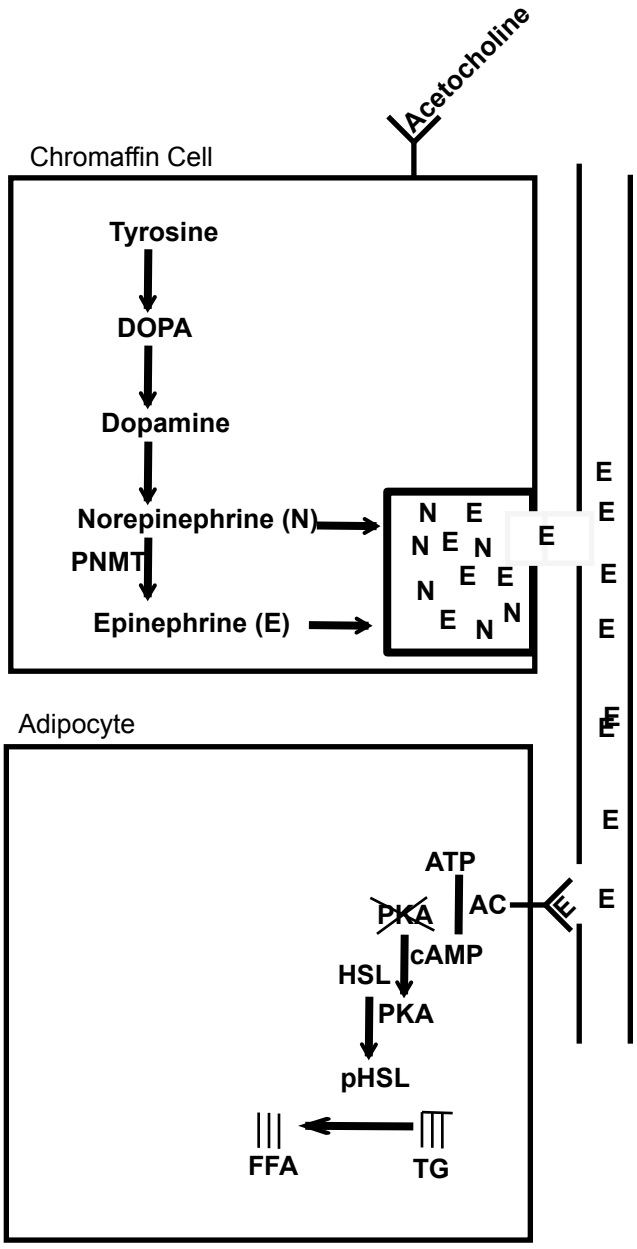

Illustration 1. The hypothetical PNMT-BMI connection. Cartoon illustration of how epinephrine synthesized in the chromaffinc cell acts on white adipoctyes to stimulate lipolysis. Epinephrine synthesized in the chromaffin cell cytoplasm is stored in chromaffin granules, whose contents are relased into the blood upone Acetocholine stimulation of nicotinic receptors. E travels to the adipocyte, where it activates beta-2-adrenergic receptors. This g-protein-coupled receptor activates adenylate cyclase (AC), which converts ATP to cAMP, leading to PKA (phosphokinase a) activation. PKA phosphorylates Hormone Sensitive Lipase (HSE aka LIPE), which breakes triglicerides into Free Fatty Acids (FFA).

The hypothetical PNMT-BMI connection

In chromaffin cells of the adrenal medulla, the catecholamine hormone epinephrine (E) is biochemically synthesized by addition of a methyl group to norepinephrine (N), a reaction catalyzed by the enzyme phenylethanolamine N-methyltransferase (PNMT). During stress E is released into the bloodstream, acting on adipocyte beta-adrenergic receptors to accelerate lipolysis (Qi et al., 2009). Assuming that PNMT promoter activity is directly proportional to cytoplasmic PNMT enzyme quantity and hence the quantity of epinephrine available for secretion into the bloodstream during stress, we hypothesize that common low-activity PNMT promoter haplotypes increase risk for obesity, a complex common disease (defined as BMI > 30) with prevalence > 20% in 49 US states (CDC survey).

The genetic history of PNMT regulatory SNPs

Two common G/A SNPs in the 500bp human PNMT promoter were discovered by Wu and Comings in 1999 and later resequenced in European-American, African-American and Japanese populations (Ji et al., 2005; Saito et al., 2001; Wu and Comings, 1999). In NCBI dbSNP these SNPs are known as rs876493 and rs3764351, located -367 and -161 bp 5' of the 23 bp 5' UTR (Kaneda et al., 1988), bases -390 and -184 with respect to the ATG codon. Reported linkage disequilibrium in resequenced individuals is high (D' > 0.9 in HapMap CEU population) (Thorisson et al., 2005).

Numerous genetic associations for these common promoter SNPs and for a rare SNP in Intron 1 exist in the literature on PNMT. This includes promoter SNP associations to hypertension in African-Americans (Cui et al., 2003), early-onset Alzheimer disease (Mann et al., 2001), multiple sclerosis (Mann et al., 2002), weight loss with sibutramine in women (Peters et al., 2003), and reward dependence in Japanese women (Yamano et al., 2008). Although two independent studies failed to find association with hypertension in European (Kepp et al., 2007) and European-American (Cui et al., 2003) subjects, a third study which maximized statistical power by sampling the extreme high/low blood pressure individuals from a population of 50,000 European-Americans did observe a significant association to diastolic blood pressure (Rana et al., 2007). The only GWAS study with statistically significant hits in the genetic neighborhood of PNMT was an association to childhood asthma on a 112 kb interval on chromosome 17q21 (Moffatt et al., 2007).

Three recent publications point to rare SNPs in Intron 1 with potential for regulatory function. Kepp et al. observe sequence conservation and selective pressure on PNMT intron 1 (Kepp et al., 2007), suggesting the presence of a glucocorticoid response element (GRE) and an insulin response element (IRE). The GRE predicted by Kepp et al. overlaps with an intronic SNP (Intron 1, position 280) which Ji et al. associated with circulating epinephrine levels during exercise (Ji et al., 2008). Another rare intronic SNP

rs200173 was also associated at genome-wide significance with impulsive behavior in ADHD patients (Oades et al., 2008).

The genetic story of PNMT is peculiar and unconventional, showing evidence of gene by environment, sex-specific effects and molecular heterosis (Comings and MacMurray, 2000). Intuitively it makes sense that a genetic association test of resting blood pressure will not detect effects of stress-response genes. Recent work in both humans and knockout mice support the stress environment hypothesis. A significant difference in blood pressure between knockout and wild-type mice was evident only during periods of exercise stress (Bao et al., 2007). Likewise a change in circulating epinephrine levels during exercise was associated with the above-mentioned intronic SNP (Ji et al., 2005; Ji et al., 2008). Furthermore, molecular heterosis and sex-specific effects were observed by three indepenent studies of distinct phenotypes; DBP (Rana et al., 2007), BMI (Peters et al., 2003) and MS (Mann et al., 2002). Thus the application of an additive genetic model to a heterosis effect may be the reason why hypertension associations were not replicated.

The only experimental evidence for functional regulatory SNPs include luciferase promoter reporter assays comparing (G-367A.G-161; rs3764351.rs876493) PNMT promoter haplotypes, EMSA and luciferase experiments testing for allele-specific binding and promoter activity changes for the rare Intron 1 SNP (I1G280A, rs200173). A significant decrease in activity for the double-mutant A.A haplotype was observed (versus wild-type

G.G), and allele-specific binding and increased transcriptional activity was observed for the A-allele at I1G280A.

None of the prior genetic studies identified allele-specifc transcription factors binding, nor do they attempt to explore epistatic interactions with other mutations. Extensive work in the PC12 rat pheochromocytoma cell line suggests orthologous regulatory motifs for Sp1, Egr1, Ap2, Maz and GRE in the 1 kb proximal promoter, as reviewed (Wong and Tank, 2007). Kepp et al. and Ji et al. suggest GRE and IRE elements in intron 1, while Kaneda et al. predicted Sp1 and GRE elements in the 1.5 kb promoter. Upstream and downstream genes in the epinephrine synthesis and signalling pathway such as tyrosine hydroxylase and the beta-2 adrenergic receptor show independent effects but were never tested for epistasis.

Summary of findings

This work identifies a BMI-increasing effect of a common promoter SNP (rs876493) and demonstrates this effect is cause by a decrease in PNMT promoter activity with the mutant allele. In chapter 1 we resequence a diversity panel to identify 2 regulatory (G-367A and G-161A) SNPs in the 0.5 kb proximal promoter common in 4 ethnicities (European, African, Asian, Mexican), and establish the A.A and G.G as the two most common haplotypes (> 87% in all populations). We demonstrate a BMI-increasing effect of the A-allele at -161 in unrelated European-Americans, replicated in twins by a

linkage effect. In chapter 2 we identify mammalian-conserved Sp1 and Sox17 binding motifs with preferential binding for the G-367 and -161A alleles respectively, confirmed both computationally and experimentally by phylogenetic footprinting and EMSA, supershift assays. In chapter 3 we then demonstrate allele-specific transcription activation by Sp1 on G-367 and repression by Sox17 on -161A, dependent on co-transfection of Sp1 and Sox17. This hapolotype-specific result in PC12 cells (a rat cancer cell model of the chromaffin cell), combined with the observed haplotype frequencies (A.A 43%, G.G 46%) allows us to conclude the BMI-increasing effect associated with the -161A allele is caused by a decrease in promoter activity for the A.A haplotype versus the G.G haplotype conserved in all primates. Finally in chapter 4 we propose a connection between PNMT promoter activity decrease and BMI increase by demonstrating an enrichment of lipid metabolism and obesity-associated genes differentially expressed in chromaffin cells of a PNMT KO mouse when compared to a WT mouse.

Significance

This work advances the forefront of human PNMT regulatory SNP functional annotation by discovering a novel transcription inhibition motif with allele-specific effect on PNMT. This Sox17 motif is conserved only in primates and depends on Sp1 co-transfection. Furthermore it is the first analysis of adrenal-specific expression changes in a PNMT KO mouse vs wild-type.

CHAPTER 1.
ARE COMMON PNMT PROMOTER SNPS ASSOCIATED WITH OBESITY *IN POPULO*?

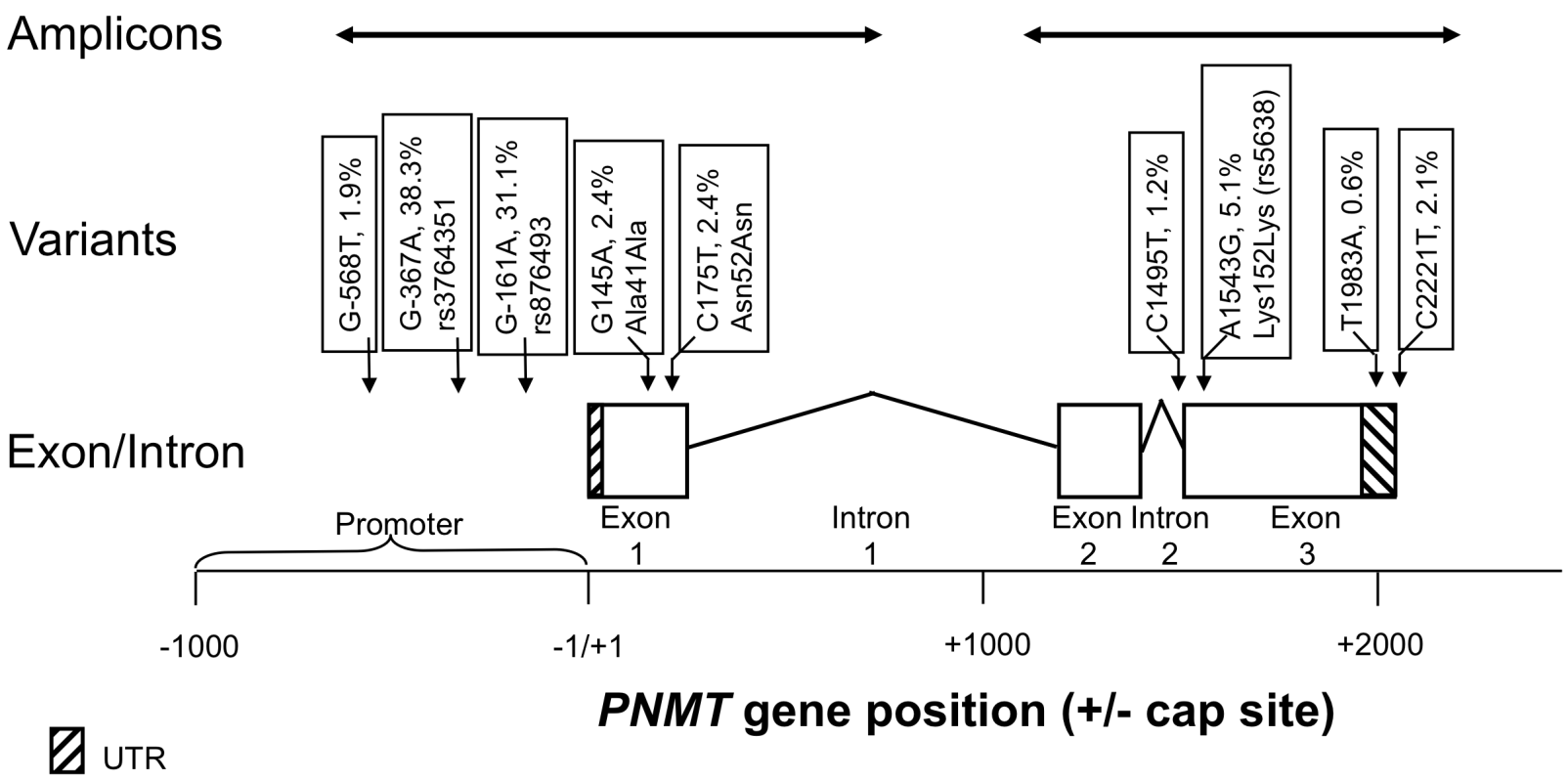

Figure 1.1. Human Phenylethanolamine N-Methyltransferase: Systematic polymorphisms discovery in n=160 chromosomes. Shown is the PNMT gene structure based on genomic DNA clone X52730. Variant frequencies combined from 80 individuals in 4 ethnic populations (Americans with Mexican, African, Asian and European ancestry).

Table 1.1. Common SNP frequency, minor allele and Hardy-Weinberg Equilibrium Test in resequenced populations.

			G-367A			
Ethnicity	Obs HET	Pred HET	HW p-val	% Geno	MAF	M.A.
All	0.506	0.495	1	100	0.448	A
Asian	0.625	0.5	0.7078	100	0.5	A
African	0.48	0.435	1	100	0.32	A
Mexican	0.5	0.492	1	100	0.438	A
European	0.45	0.489	0.9875	100	0.425	G

			G-161A			
Ethnicity	ObsHET	PredHET	HWpval	%Geno	MAF	M.A.
All	0.473	0.452	0.9355	96.1	0.345	A
Asian	0.625	0.469	0.5018	100	0.375	A
African	0.36	0.385	1	100	0.26	A
Mexican	0.357	0.436	0.8128	87.5	0.321	A
European	0.579	0.494	0.8709	95	0.447	A

Table 1.2. Haplotype frequency and linkage disequilibrium in resequenced ethnic populations.

		Haplotype Frequency -367.-161				Linkage Disequilibrium	
Ethnicity	Chrom.	A.A	G.G	A.G	G.A	D'	r^2
European	2n=46	0.43	0.46	0.11	0	1	0.6
African	2n=50	0.68	0.26	0.06	0	1	0.8
Mexican	2n=32	0.57	0.31	0.13	0	1	0.6
Asian	2n-32	0.5	0.38	0.13	0	1	0.6

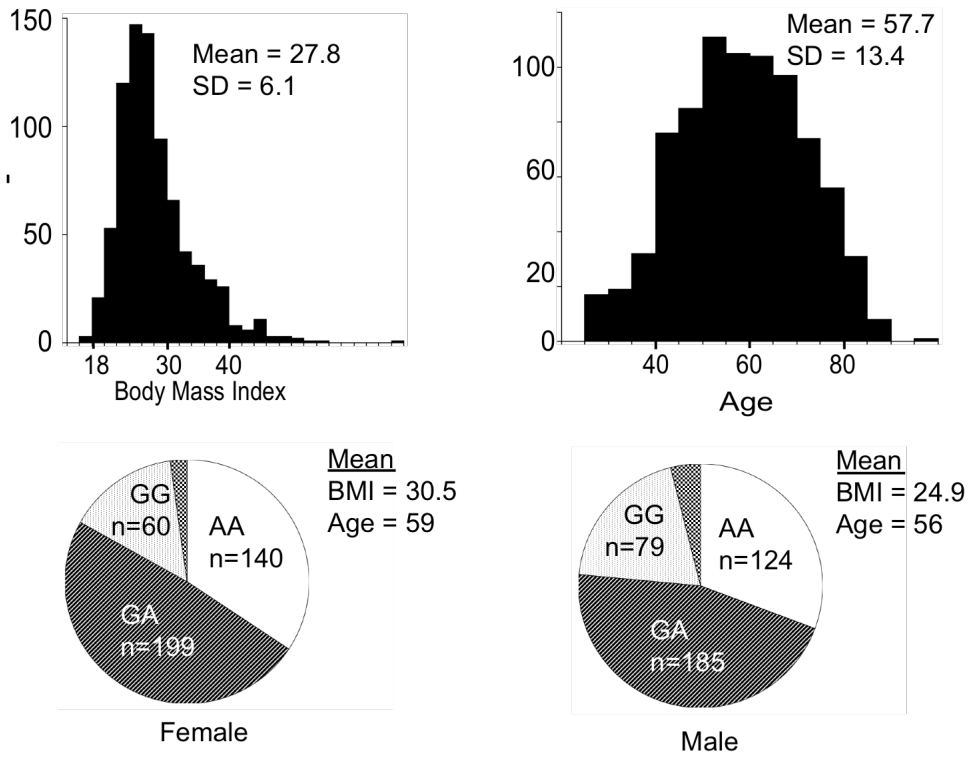

Figure 1.2. Descriptive statistics of unrelated European-American population sample. Frequency histograms of BMI and Age for 817 European-American individuals. Pie charts at bottom show genotype frequencies, males and females separated. Mean and standard deviation for the whole group shown above each histogram; sex-specific means for each phenotype shown to the left of each pie chart.

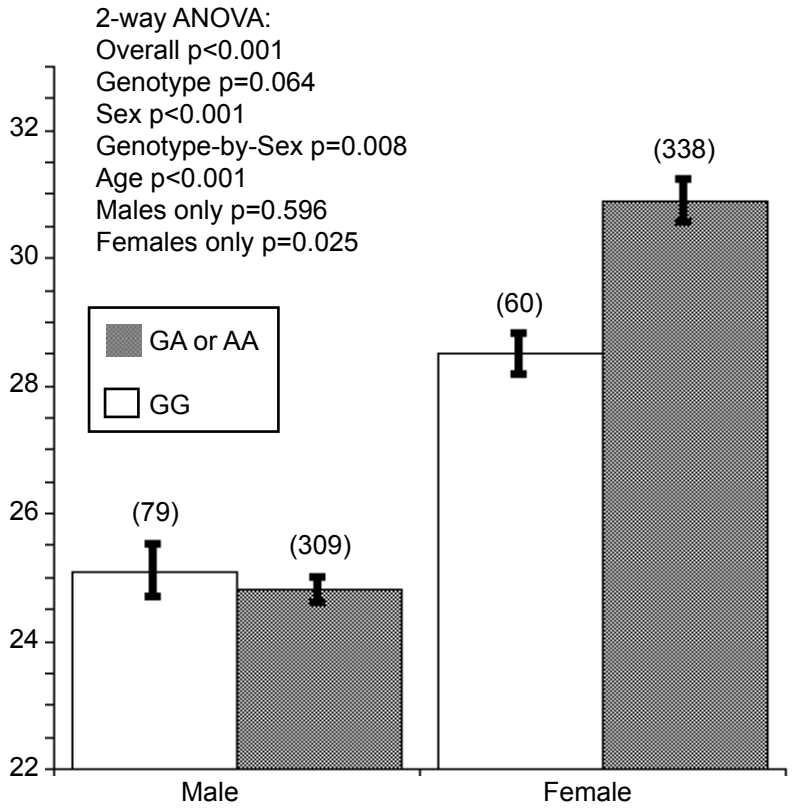

Figure 1.3. PNMT Promoter G-161A: Association to BMI in unrelated European-Americans. Bar graph of mean BMI +/- 1 SE for genotype by sex groups (male on the left, female on the right; GG at G-161A in white, GA or AA in grey). Number of individuals shown in parenthesis above each bar. Above graph are ANOVA results for test of significant difference in BMI between genotype and genotype by sex groups.

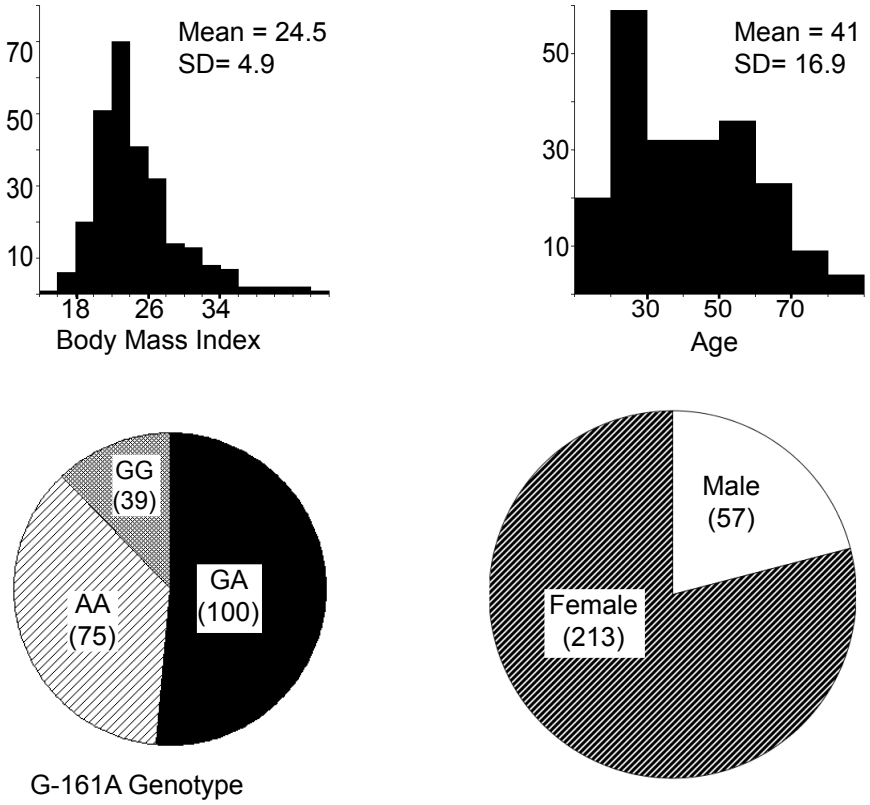

Figure 1.4. Descriptive Statistics of European-American twin population sample. Frequency histograms of BMI and Age for 214 European-American female twins. Mean and standard deviation for the whole group shown above each histogram. Pie charts at bottom show genotype and gender frequencies.

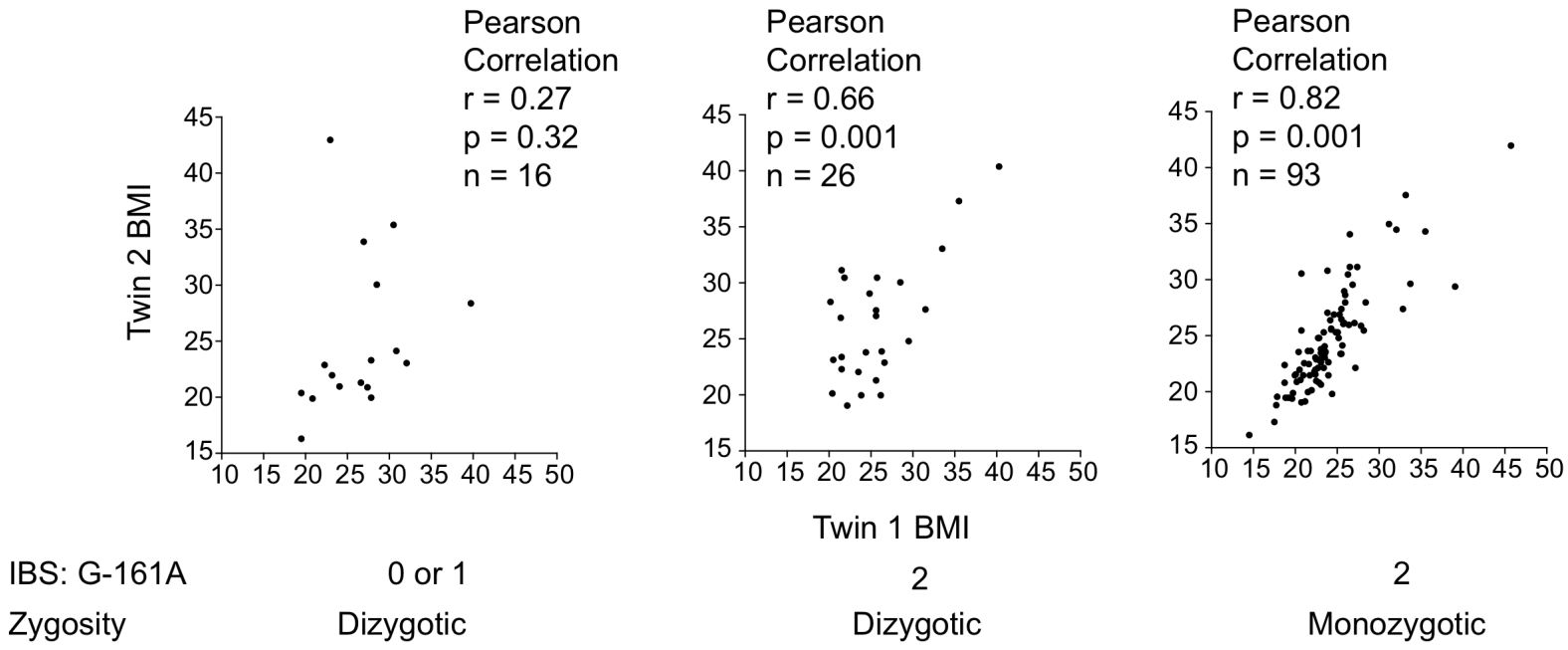

Figure 1.5. PNMT G-161A and zygosity effect on BMI correltation between twins. Based on the observed correlations we can calucalte the heritability and variance componemnt for BMI as in a classical twin study.

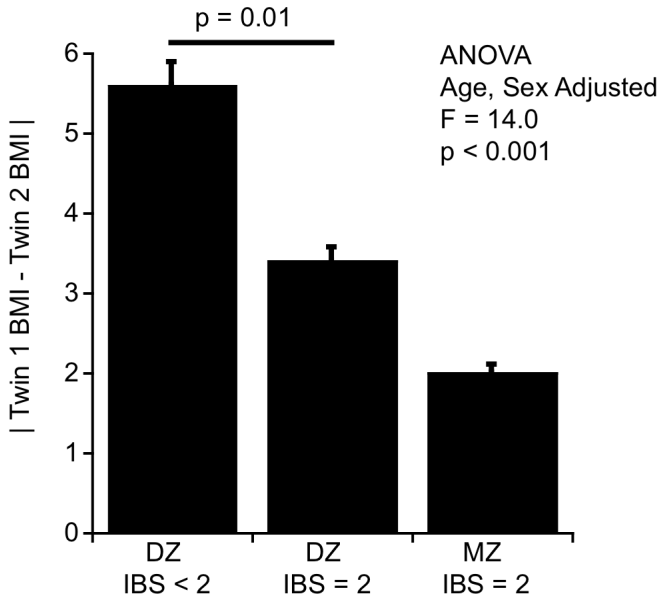

Figure 1.6. PNMT G-161A effect on BMI: Mean absolute value of BMI difference between siblings in twin pairs grouped by IBS status ans zygosity. I

INTRODUCTION

The work described in chapter 1 began in 2003 as part of a candidate gene polymorphism discovery resequencing project led by Dan O'Connor. At the time online resources such as the HapMap and dbSNP were in their infancy, GWAS was still a fantasy, and the standard method to answer the question, "What mutations exist in PNMT and what is their frequency?" was Sanger sequencing on an ABI Prism apparatus.

Although sequening technology and online SNP databases are rapidly advancing, once the genotypes are available in tab-delimited format the basic statistical analysis remains the same. For a quantitative trait we can compare the phenotype means and variances between genotype groups using an ANOVA test. However the heated debate revolves around how to minimize false positives by controlling for unaccounted or unknown effects on the group means, answering the question: Are there environmental or genetic factors which differ between the genotype groups which are capable of influencing the phenotype means?

Are common PNMT promoter SNPs associated with obesity *in populo*? Another approach which evolved since the commencement of this dissertation project is preference for studying rare versus common SNPs. In this dissertation we focus on 2 SNPs common in all populations resequenced. A basic tenet of population genetics and evolutionary theory is: the mutations with the lowest frequency are selected against due to deleterious phenotype,

while common mutations are expected to be neutral. An exception to this rule would be an evolutionarily advantageous mutation which becomes common over time thanks to increased fitness of individuals carrying the mutation. This theory is well established for nonsynonymous coding SNPs (Ng and Henikoff, 2006) and was (dissapointingly) verified in plethora of GWAS studies focusing on common SNPs (Bogardus, 2009; Maher, 2008).

In this chapter we identify 2 common SNPs in the PNMT promoter, genotye these in both unrelated individuals and twin-pairs, and then demonstrate both an association and a linkage effect of one SNP on body mass index (BMI). After estimating haplotype frequences for all 4 possible combinations of the 2 SNPs, we determine that the genetic association can be attributed to either SNP since they are in close proximity on both the genetic (LD) and genomic (bp) scales.

RESULTS

SNP Discovery

Only 2 out of 9 SNPs identified in resequencing were common in all 4 ethnicities (G-367A and G-161A, see Figure 1.1). After aligning the flanking sequences with the reference genome (BLAT webserver, http://genome.ucsc.edu), we confirmed these as rs3764351 and rs876493, first discovered by Wu and Comings (Wu and Comings, 1999). Combined-population SNP frequencies, positions with respect to the 5' cap (Refseq X2730, Nagatsu 1989), and dbSNP id's (if known) are reported in Figure 1.1. Ethnicity-specific minor allele frequencies and Hardy-Weinberg test p-values are shown in Table 1.1. Both regulatory SNPs show no significant deviation from HWE. A is the minor allele for both SNPs in all populations, with the exception G-367A in Europeans, where A-allele frequency is 46% and G-allele frequency is 43%.

Haplotype Phasing and Linkage Disequilibrium

Linkage disequilibrium was high between these SNPs in all populations (D' = 1.0, $r^2 > 0.6$, Table 1.2). Haplotype inference using PHASE reported no individuals with a G.A (-367.-161) haplotype, and > 87% of individuals with either G.G (chimpanzee) or A.A (double-mutant) haplotypes. The A.A haplotype was major in all non-European populations, although the difference is < 3%.

Descriptive Statistics

Frequency histograms plus mean and standard deviations of SBP, DBP, BMI and age for the three groups are reported in Figure 1.2 (KU) and Figure 1.4 (SDT). The KU group had the highest means for all 4 phenotypes (mean DBP = 79, SBP = 132, BMI = 28, age = 58) versus SDT (mean DBP = 71, SBP = 131, BMI = 25, age = 41). The KU blood pressure histograms show a bi-modal distribution (by design, individuals were selected from BP extremes of 50,000 individuals).

Selection of SNPs for Genetic Tests

We observed high genotype error rates at G-367A both in pyrosequencing and maldi-tof sequencing assays. The > 87% combined frequencies of the G.G and A.A (G-367A.G-161A) haplotypes, as well as the 1.0 LD betwen the SNPs. Based on these results we only tested G-161A for association in the KU and SDT groups.

Tests of Association and Linkage

We observed a significant difference in mean BMI between genotypes in the KU group females only. The ANOVA p-value for the 2-way gene by sex was < 0.001 (Figure 1.3). The p-value for a t-test comparing the two female genotype groups (dominance model, GG versus GA or AA) was 0.01. We

sought to replicate this result in female twins, and observed a significant linkage effect ($p < 1.0 e -8$) in a variance components model which simultaneously models an association effect but tests the difference in log-likelihood between VC models with and without a dominant genetic effect.

In Figure 1.5 we illustrate the correlation in twin BMI, for the purpose of calculating the variance components of BMI, based on the classical twin study ACE model, where heritability $h^2 = 2(r(MZ) - r(DZ)) = 2(0.82-0.47) = 0.8$ is the A component, the shared environment $c^2 = r(DZ) - 0.5\ h^2 = 0.47 - 0.4 = 0.07$; and the unique environment component is $e^2 = 1 - r(MZ) = 1 - 0.82 = 0.18$.

METHODS

PNMT Promoter Resequencing

DNA samples from 80 individuals of 4 different self-identified ethnicities (by ancestry, European, African, Mexican or Asian) were sequenced on an ABI 3100 Genetic Analyzer (Applied Biosystems) using a forward primer in the 5' upstream region and a reverse primer in intron 1:

FW 5'-CACAGCTCACCTCCCCTATC-3'

RV 5'-CACAGCTCACCTCCCCTATC-3'

with the amplicon mapping to X52730:910-1990. A separate pair of primers were used to resequence from Exon 2, Intron 2, Exon 3 and the 3'UTR:

FW 5'-CACCAGGACCCTCTTCCTCT-3'

RV 5'-GAGCCACACACAGCACTGAC-3'

with the amplicon mapping to X52730:2185-3718.

Chromatogram files were converted to fasta format and assembled into amplicons using PHRED (Ewing et al., 1998) and PHRAP (de la Bastide and McCombie, 2007). Polymorphisms were identified using POLYPHRED (Nickerson et al., 1997), excluding polymorphism calls of quality > 1. Flanking sequences for identified polymorphisms were aligned to the reference human genome using BLAT (Kent, 2002). Flanking sequences and alleles for all known SNPs in the PNMT 1 kb promoter and Exon 1 were compared to the resequenced amplicons to distinguish known from discovered SNPs.

Haplotypes were inferred for each ethnicity using PHASE (Stephens et

al., 2001). Linkage disequilibrium was calculated and genotype/allele frequency accuracy was verified using the Hardy-Weinberg equilibrium test using HAPLOVIEW (Barrett et al., 2005).

Human Subjects

All individuals studied were residents of the USA whose self-reported ancestry is European. We selected for genetic analysis 817 unrelated European-American individuals from the San Diego Kaiser health network (KU), 214 (107 pairs) female twins from the San Diego Twin Registry (SDT). Details of individual recruitment and genetic associations not reported here for the recruited individuals were previously reported elsewhere (Zhang et al., 2007).

Genotyping by Pyrosequencing

Based on SNP frequency and potential for functional study invitro we selected for genotyping by pyrosequencing the two common regulatory SNPs G-367A and G-161A in the KU and SDT population samples. The complete genotyping protocol is described in detail elsewhere (Jackson et al., 2009). Based on the SNP flanking sequences we selected PCR amlification primers using Biotage Assay Design Software. Given the proximity of the SNPs one pair of PCR primers were used for both SNPs

FW 5'-TCCACCTCGAATCAGGAAAG-3'

RV Biotin-5'-ACCCATTCATCCATCTCCCT-3'

After amplification, amplified and biotinylated DNA in 96-well plates was combined with streptavidin beads (5 min shake) then lifted from plates by vacum suction, washed in 70% EtOH, 1.0 M NaOH, sterile H_20 then combined with sequencing primers

G-367A : 5'-AGGTGAATGGCTGCG-3'

G-161A : 5'-CGCCCCACTCACCTC-3'

for 1 min at 96^{o} C. Plates were then genotyped in a Biotage HS96A. SNP genotypes were determined from pyrophosphate-dependent luminescence peaks comparing observed and expected using manufacturer software per specifcations (Biotage). Plates with > 30% failed genotypes were repeated, and primers with > 50% failed genotypes were redesigned and verified by gradient PCR.

Statistical Test of Association and Linkage

All statistical analysis was performed in SPSS (Spss Inc.), Plink (Purcell et al., 2007) and KaleidaGraph (Synergy Software) for cross-validation. BMI was calculated as kg/m^2. Single blood pressure measurements were used for the descriptive statistics in KU and SDT groups. Genotype BMI means for the KU were compared by both ANOVA (SPSS) and linear regression (Plink). SDT test of linkage while modeling association using the orthogonal model and a

dominance genetic effect QTDT (Abecasis et al., 2000). Input IBD matrices were calculated using MERLIN (Abecasis et al., 2002).

CONCLUSION

We identified 2 SNPs in the PNMT promoter common in 4 major North American ethnic populations (European, African, Asian, Mexican). One of these SNPs, G-367A, turned out to be very difficult to genotype accurately. Based on linkage disequilibrium and haplotype frequency, we proceeded with extensive genotyping and genetic tests using G-161A. We identified a significant BMI-increasing effect of the -161 A allele, which by linkage diseqilibrium corresponds to the AA haplotype in European-Americans (GA was a rare, unobserved haplotype). Based on the extensive rodent model literature, we hypotesize a decrease in PNMT promoter activity due for this haplotype and mutation. This concludes the *in populo* portion of the dissertation. In the next chapter we use *in silico* and *in vitro* methods to determine which of the two SNPs (or haplotype combination) is functional in the genetic effect observed in Chapter 1.

Acknowledgements

The work in this chapter in part is currently being prepared for submission for publication. It was co-authored by Gen Wen, Ryan Friese, Manjula Mahata, Bruce Hamilton, Maple Fung, Rany Salem and Daniel O'Connor. The dissertation author was the primary investigator and author of this material.

CHAPTER 2.
DO THE ASSOCIATED AND LD-TAGGED SNPS DISRUPT MAMMALIAN-CONSERVED REGULATORY MOTIFS PREDICTED *IN SILICO* AND VERIFIED *IN VITRO*?

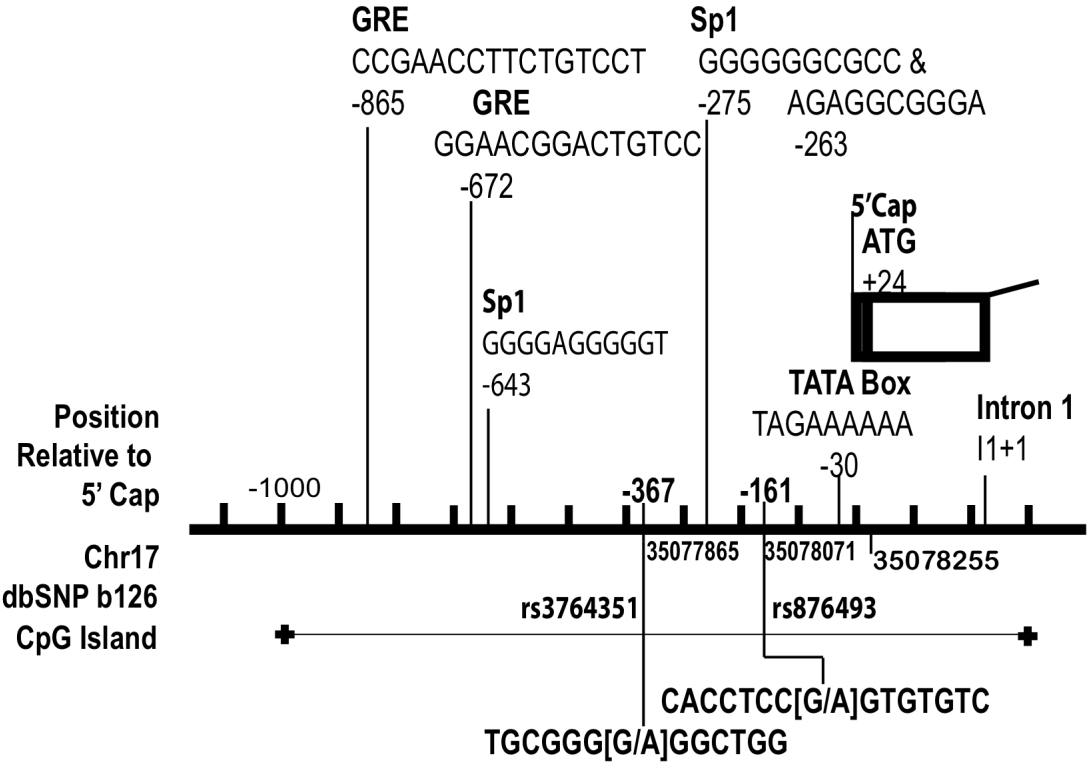

Figure 2.1. Regulatory Motifs and Genetic Polymorphisms in Human Phenylethanolamine N-Methyltransferase. Along the top are putative GRE, Sp1, Egr1 motifs in the human PNMT promoter (Nagatsu1989), as well as the TATA box and UTR/exon/intron boundaries (RefSeq X52730; Nagatsu 1989). For the 2 SNPs studied, henceforth referred to as G-367A and G-161A, relative positions are shown (from top to bottom) with respect to the 5' UTR cap site, and chromosome 17 in NCBP build 36 of the human genome.
At the bottom are dbSNP rsIDs. Crosses mark the boundaries of a CpG island spanning the 1 bb promoter and Exon 1.

Figure 2.2. Mammalian PNMT Promoter Alignment: SNPs and Flanking Sequence. On the left are manually constructed phylogenetic trees for the 10 mammalian species included in the alignment. On the left are sections of Clustal-W alignments of the orthologous PNMT promoters. The promoter sequences for chimp, bonobo, and rat were obtained by resequencing. All other sequences were obtained from Multi-Z exon-anchored promoter alignments availabe at genome.ucsc.edu. Both alleles of G-367A and G-161A are shown in lowercase. The orthologous alllele in aligned mammalian motifs also in lowercase.

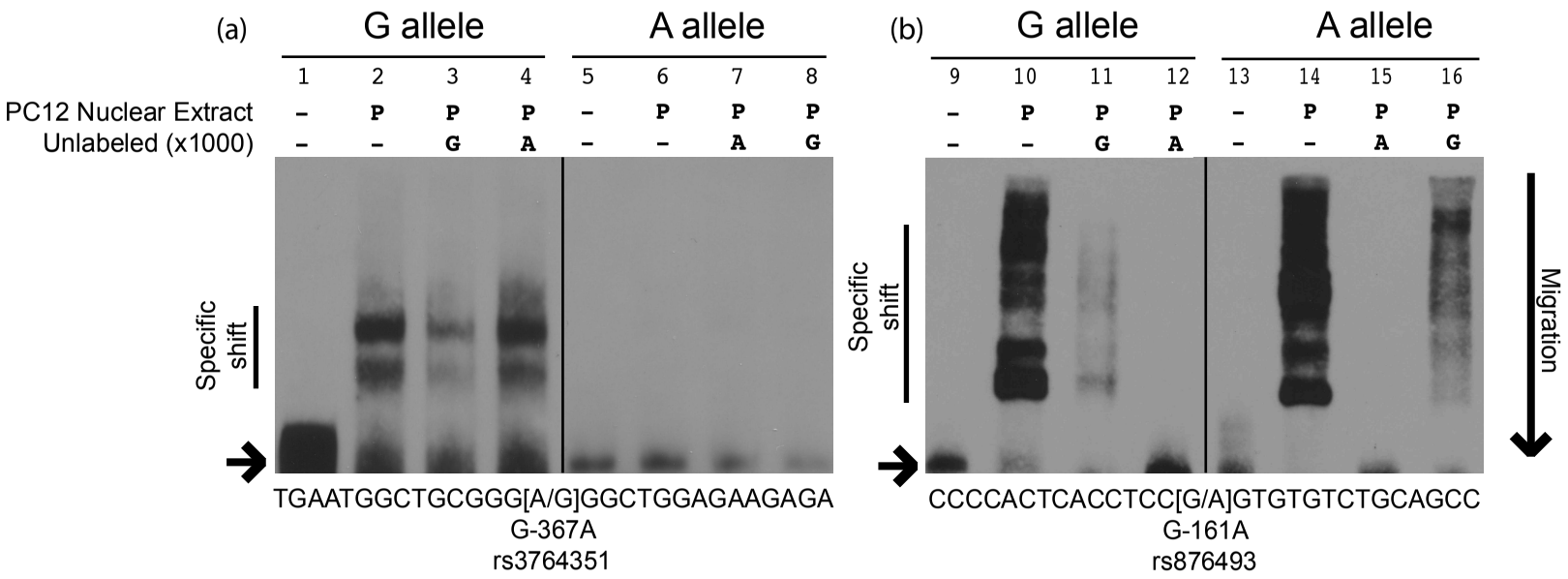

Figure 2.3. Allele-specific EMSA of G-367A and G-161A with PC12 nuclear extract. The forward strand of the dsOligo tested for allele-specific shift shown along the bottom of the gel (identified by promoter position and rsID), with brackets around the polymorphic site. Above the gel is a legend of binding reaction contents (from top to bottom): labelled oligonucleotide allele (G or A), PC12 nuclear extract (P or - for absence), unlabeled competitor oligonucleotide allele (G, A or - for none). Lanes 1-4 for G-367, lanes 5-8 for -367A, lanes 9-12 for G-161, and lanes 13-16 for -161A. Horizontal bar on leftof gel highlights allele-specific shifted bands and arrow marks unshifted labelled probe. Top-to-bottom migration indicated by arrow on right side.

Table 2.1. Phylogenetic footprinting scores for G-367 on alignment with rat.

Matrix	% ID	Dir	% Score	Factor	Motif
M00196	92	+	80	Sp1	**********GGGgGGCTGGAGA*****
MA0055	92	+	75	Myf	**********GGGgGGCTGGAG******
M00189	92	-	73	AP-2	******CTCCAGCCcCCC**********
MA0079	90	+	79	SP1	**********GgGGCTGGAG******
M00008	90	+	78	Sp1	**********GgGGCTGGAG******
MA0092	90	+	78	Thing1-E47	**********gGGCTGGAGA*****
MA0028	90	+	73	Elk-1	**********gGGCTGGAGA*****
M00075	90	+	73	GATA-1	**********GgGGCTGGAG******
M00253	88	-	91	cap	********CCAGCCcC**********
MA0100	88	+	74	c-MYB_1	**********GGgGGCTG*********
M00255	86	+	75	GC_box	**********GGGgGGCTGGAGAA****
M00257	86	-	73	RREB-1	******CTCCAGCCcCCCGC*********
M00255	86	+	71	GC_box	**********GGgGGCTGGAGAAG***
M00085	85	-	73	ZID	***CTTCTCCAGCCcC**********
M00159	85	+	71	C/EBP	**********gGGCTGGAGAAGA**
M00084	85	+	71	MZF1	*********GCGGGgGGCTGGA*******
M00189	83	-	90	AP-2	**********AGCCcCCCGCAG*******
M00378	83	-	87	Pax-4	***CTTCTCCAGCCc**********
M00222	81	+	73	Hand1/E47	*********GCGGGgGGCTGGAGAA****
MA0079	80	+	84	SP1	*********GCGGGgGGCT**********
M00008	80	+	83	Sp1	*********GCGGGgGGCT**********
MA0003	78	-	93	AP2alpha	**********GCCcCCCGC*********
M00141	78	+	75	Lyf-1	********TGCGGGgGG**********
MA0057	70	+	74	MZF_5-13	*****GGCTGCGGGg**********
M00141	67	+	73	Lyf-1	******GCTGCGGGg**********

Table 2.2. Phylogenetic footprinting scores for -161A on alignment with rat.

Matrix	% Id	Dir	% Score	Factor	Motif
MA0039	70	-	76	Gklf	**********tGGAGGTGAG*****
M00255	64	-	74	GC_box	**********tGGAGGTGAGTGGG*
M00189	58	+	72	AP-2	***CACTCACCTCCa**********

Table 2.3. Known regulatory motifs and positions in rat PNMT promoter.

Sp1	Egr-1	GR	Ap2	MAZ	GCM
-165	-168	-539	-103	-45	-414
-45	-48	-759	-587		
		-773	-674		

Table 2.4. Top scoring allele-specific motifs in JASPAR database.

Model ID	Factor	% Score	Dir	Allele Motif	Allele Score
				G-367	
MA0133	BRAC1	83	-1	CCAGCCc	0.11
MA0118	Macho-1	93	1	CGGGgGGCT	0.1
MA0079	SP1	84	1	GCGGGgGGCT	0.08
MA0123	ABI4	82	-1	CAGCCcCCCG	0.08
MA0056	MZF1_1-4	85	1	CGGGgG	0.04
MA0003	TFAP2A	93	-1	GCCCcCCGC	0.02
				-161A	
MA0118	Macho-1	83	-1	CtGGAGGTG	0.06
MA0103	Sox17	82	1	TCCaGTGTG	0.15
MA0078	sna	86	1	CCaGTG	0.03
MA0006	Arnt-Ahr	80	1	aGTGTG	0.02

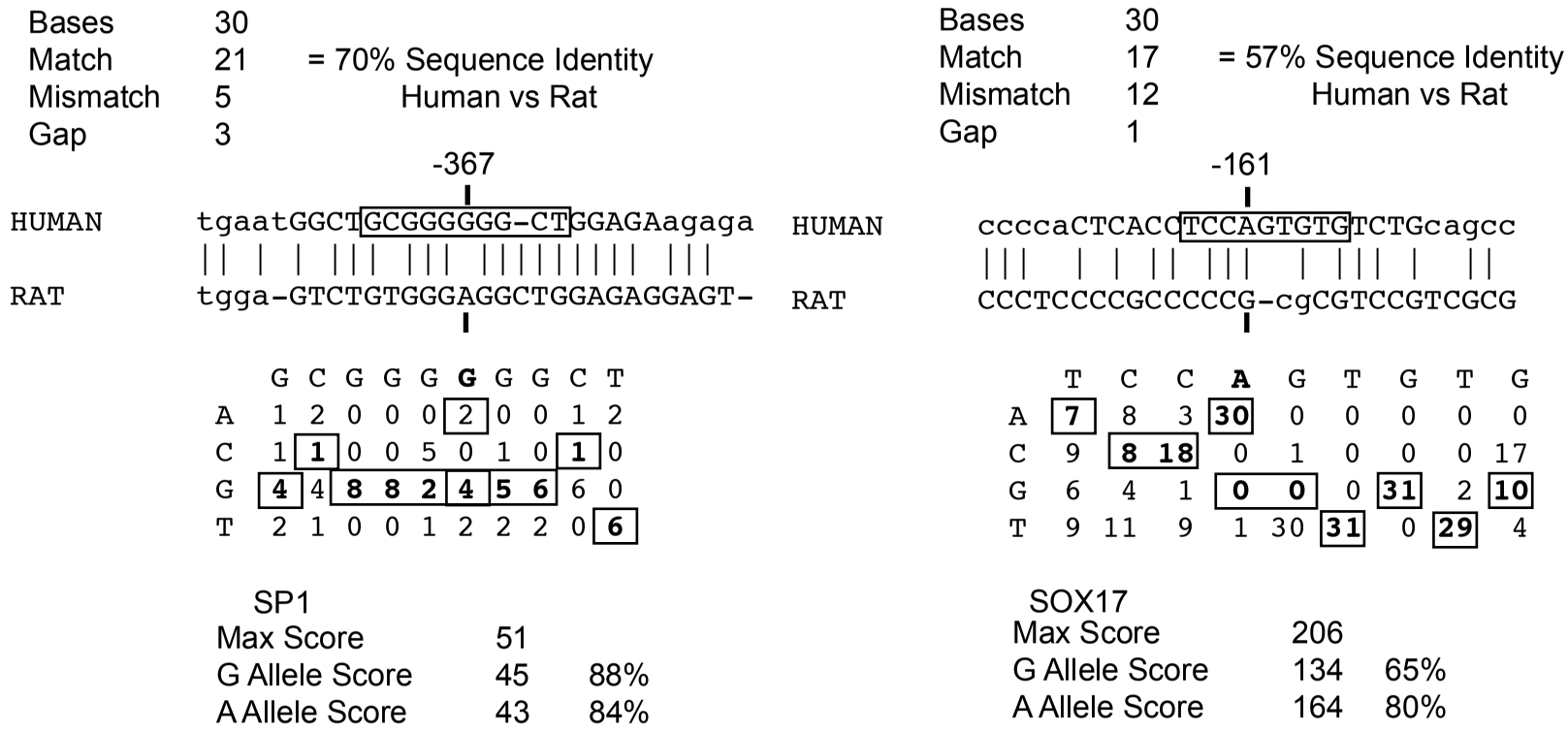

Figure 2.4. G-367A and G-161A flanking motifs compared to high-scoring PFMs from the JASPAR database. Along the top are human-rhesus alignments of G-367A (left) and G-161A (right), limited to sequences used in EMSA, and alleles showing specific shift (G at -367 and A at -161). Capitalized bases are segments of >75% sequence identity over a 5 bp window. Below the alignments are PFMs for Sp1 and Sox17. Sequence above the PFM is the putative PNMT promoter motif, with SNP allele in bold. Boxes show the scores for each nucleotide in the putative motif. Two boxes at the SNP site show the difference in score between alleles.

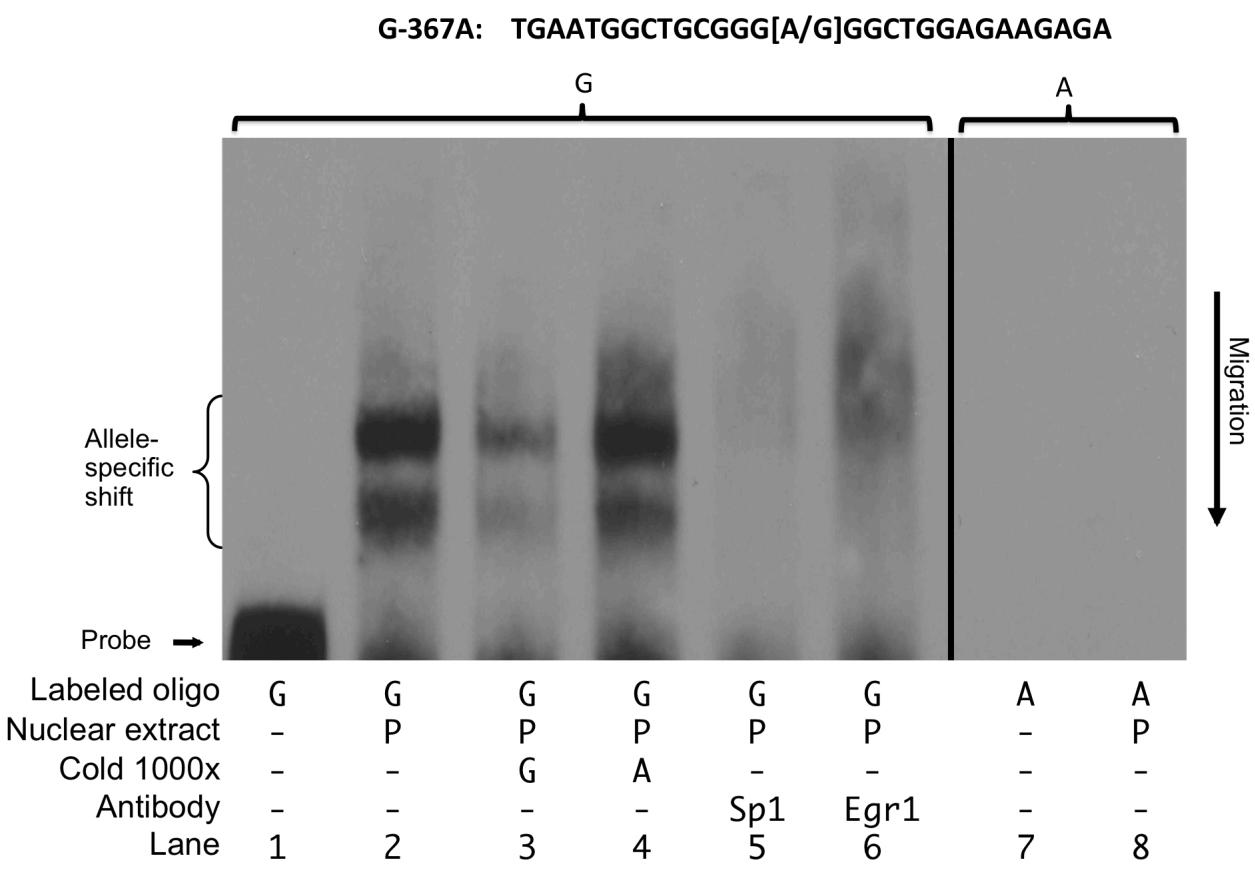

Figure 2.5. PNMT promoter variant G-367A EMSA with PC12 nuclear extract, Sp1 & Egr1 Supershift. Lanes 1-4 are the basic EMSA identical to Figure 2.3. Lanes 5 and 6 show antibody supershift assays for Sp1 and Egr1. No supershift is shown for -367A due to the lack of shifted bands in the basic EMSA.

INTRODUCTION

At the end of chapter 1 we have an association and linkage effect of PNMT promoter polymorphism G-161A on BMI. Based on linkage disequilibrium we are unable to determine with certainty which SNP, be it G-367A or G-161A (or both) is functionally responsibe for the effect, as G.G and A.A (G-367A.G-161A) haplotypes represent >90% of European-American individuals. In this chapter we combine *in silico* prediction and biochemical *in vitro* validation methods to determine the most likely transcription factors with allele-specific binding to G-367A and G-161A motifs, in preparation for the haplotype-based experiments of chapter 3.

RESULTS

Mammalian Promoter Alignment

Alignment sections covering the EMSA oligonucleotides are shown in Figure 2.2. All primate sequences have a G allele at both G-367A and G-161A. All mammalian sequences contain a 5-7 G motif flanking G-367A. A mammalian-conserved motif flanking G-161A is not apparent, however the primate motif is nearly-identical between human, chimp and rhesus.

Based on the primate sequences, G appears to be the ancestral allele in humans and A the mutant for both SNPs, as all primates posess a G-allele at the SNP sites (Figure 2.2, lowercase).

Allele-Specific Electrophoretic Mobilitiy Shift (EMSA)

EMSA experiments were performed to test potential transcription factor binding on the promoter regions surrounding rs876493 and rs3764351. Our results demonstrate A-allele specific shift for rs876493 (G-161A) and G-allele specific shift for rs3764351 (G-367A).

As shown in Figure 2.3, in lane 2, we observed 2 distinct shifted bands for the G-allele of G-367A not present for the A-allele (lanes 6-8). Specific binding was determined by reduction in band intensity for lane 3 in the presence of 1000x molar excess G-allele specific comptetitor oligonucleotide, but not with A-allele competitors (Fig. 2.3, lane 4). Therefore, the -367A allele

alters potential transcription factor binding, abolishing the 2 specific bands found in Figure 2.3 lane 2.

For the G-161A EMSA, numerous shifted bands were observed for G- and A-allele oligonucleotides (Figure 2.3 lanes 10,14). Those for G-allele are non-specific due to complete elimination by A-allele competitors (Fig. 2.3 lane 12). The shifted bands for A-allele (Figure 2.3 lane 14) were removed by A-allele competitor oligonucleotide (lane 15) but showed weak to no competition with excess G-allele oligonucleotide (Fig. 2.3 lane 16). Thus, the A-161 SNP appears to introduce new binding site(s) for unknown transcription factors.

Transcription Factor Binding Prediction

We used phylogenetic footprinting to predict the transcription factors with allele-specific binding shifts observed in the EMSA assays. Highest-scoring predictions for the G-allele of G-367A and the A-allele of G-161A are shown in Tables 2.1 and 2.2 respectively. Phylogenetic footprinting was conducted using the TFBS Perl library (Lenhard and Wasserman, 2002) through the Conreal webserver. Results in the tables are ordered by decreasing % sequence identity between human and rat motifs. The highest scoring motif for G-367 allele was Sp1, while the the highest-scoring motif for --161A was Gklf.

Comparing these results to what is known about the rat PNMT promoter, both Sp1 (predicted binding to -367) and Ap2 (predicted binding to

both SNPs) high-scoring candidates with known roles in PNMT regulation (Table 2.3) sites and motifs specific to the G-allele of G-367A and Phylogenetic footprinting using the RAVEN webserver also predicted conserved Sox17 and Ahr/Arnt motifs at rs876493. The presence of known rat PNMT promoter motifs at at -165 and -103 respectively established Sp1 and Ap2 as the most likely candidates. GEO expression profiles for PC12 cells (GDS1039, http://ncbi.nlm.nih.gov/geo) confirm expression of both factors: Sp1 at > 75%, Sox17 at > 25%. GEO expression levels for other putative factors include: Egr1 (aka Nab1, NGFI-A) > 90%, Ap2 (Tcfap2a) > 40%.

METHODS

Mammalian Promoter Sequencing & Alignment

We obtained DNA samples from the Coriell repository for 3 primates (P. troglodytes NG06939, P. troglodytes versus NA03448, M. mulatta NG07109), and resequenced these using the human primers from Chapter 1 above. Rat PNMT promoter was sequenced by R Fries. Base calling, amplicon assembly and polymorphism detection as described above. Mouse, dog, cat, horse, cow, promoter sequences (1 kb upstream of the ATG codon) were obtained from genome.ucsc.edu (Karolchik et al., 2008). Promoter orientation for each sequence was determined by Clustal-W alignment of Exon 1 coding sequence. Promoter sequences were aligned using Clustal-W with a high gap penalty. Phylogenetic tree of the mammals was constructed manually using taxonomy information from NCBI.

Electrophoretic Mobilitiy Shift Assay

Biotinylated double-stranded 29-mer oligonucleotides (ValueGene; Biotin 3' End DNA Labeling Kit, Pierce, Rockford, IL, USA) with SNP at position 15 were incubated 45 minutes with nuclear extracts of PC12 pheochromocytoma for 20 minutes in the presence of 0.1 μg/μl salmon sperm DNA to minimize nonspecific binding. Nuclear extracts were prepared from PC12 cells grown in 10 cm plates by Qproteome Nuclear Protein Kit following manufacturer specifications (Qiagen (Avery-Kiejda et al., 2008)). In separate

reactions, either specific competitor (i.e. unlabled same-allele oligonucleotides) or non-specific competitor (i.e. unlabeled opposite-allele oligonucleotides) was added at 1000x concentration to assay allele-specific binding. DNA-protein complexes were electrophoresed in 5% acrylamide gels for 1 hour at 100 V in 0.5x TBE then transferred to nylon membranes at 380 mA in 0.5x TBE and cross-lined with 1200 nm UV light for 1 minute. Complexes were then visualized on X-ray film using horseradish-peroxidase-conjugated streptavidin per manufacturer protocol (Chemiluminescent Nucleic Acid Detection Module, Pierce, Rockford, IL, USA (Matata and Galinanes, 2002; Oltra et al., 2003)).

Prediction Allele-Specific Conserved Regulatory Motifs

We predicted transcription factors with allele-specific binding to motifs containing G-367 or G-161A using phylogenetic footprinting (Wasserman and Sandelin, 2004). Human and rat PNMT promoter sequences were pairwise aligned using the LAGAN (Brudno et al., 2003), as implemented by the CONREAL web server (Berezikov et al., 2005; Berezikov et al., 2004). Motifs with > 50% human-rat homology (length > 10 bp) were scored on publicly available position-weight matrices (PWMs) from Jaspar (Vlieghe et al., 2006) and Transfac v.8.2 (Wingender et al., 2000) databases. Motifs scoring > 80% of maximum in one allele with a difference in score between alleles were considered positive hits (Summarized in Figure 2.4 using Sox17 and Sp1 as

examples). Results were cross-verified using the RAVEN web server (Andersen et al., 2008), which uses 7-vertebrate PhastCons alignments for identification of conserved motifs. Candidate hits were checked for adrenal, chromaffin cell or PC12 expression activity in GEO expression array repository (Andersen et al., 2008) and prioritized for experimental validation if known to regulate the well-characterized rat PNMT promoter (Wong and Tank, 2007).

CONCLUSION

This chapter puts forth potential regulatory function at both SNPs, but with a more convincing story at G-367A, where the EMSA binding is completely abolished in the A-allele. This result is consistent with a GGGGGG motif conserved in all mammals sequenced, from chimp to horse. Phylogenetic footprinting suggests the most likely candidate for the G-367 specific shift is Sp1, which is expressed in PC12 cells and has a known regulatory role in rat PNMT promoter activity.

The G-161A story is less clear, at all three levels (alignment, EMSA and phylofoot). The motif is conserved in primates but not all mammals, and only 2 putative binding motifs scoring above 70% are identified in human-rat alignments with the -161A motif. We observe the same set of shifted bands for both alleles in the EMSA, while competitor assays point to A-allele specific binding.

Combined with the results of Chapter 1, for G-367 we can hypothesize that a evolutionarily-conserved Sp1 motif is disrupted by the -367A allele, significantly weakening PNMT promoter activity, which results in decreased availability of epinephrine in the chromaffin cell, and this deficiency causes the increased obesity in A.A haplotype females. However no sex-specific pattern is clear from these results, with the exception of Sox17, predicted to bind with -161A. Sox17 is in the class of SrY-related testis-differentiation transcription factors. In chapter 3 we evaluate the top-candidate predicted factors from

chapter 2 at the level of promoter haplotypes. In addition to observations of allele-specific effects, at the haplotype level it is possible that motifs at both SNPs interact in regulatory function.

Acknowledgements

The contents of this chapter in part is currently being prepared for submission for publication. It was co-authored by Chih Min Tang, Kuixing Zhang and Daniel O'Connor. The dissertation author was the primary investigator and author of this material.

CHAPTER 3.

CAN WE REPLICATE *IN VITRO* HAPLOTYPE-SPECIFIC AND TRANSCRIPTION-FACTOR-SPECIFIC REGULATORY EFFECTS CONSISTENT WITH THE ASSOCIATED DISEASE MODEL?

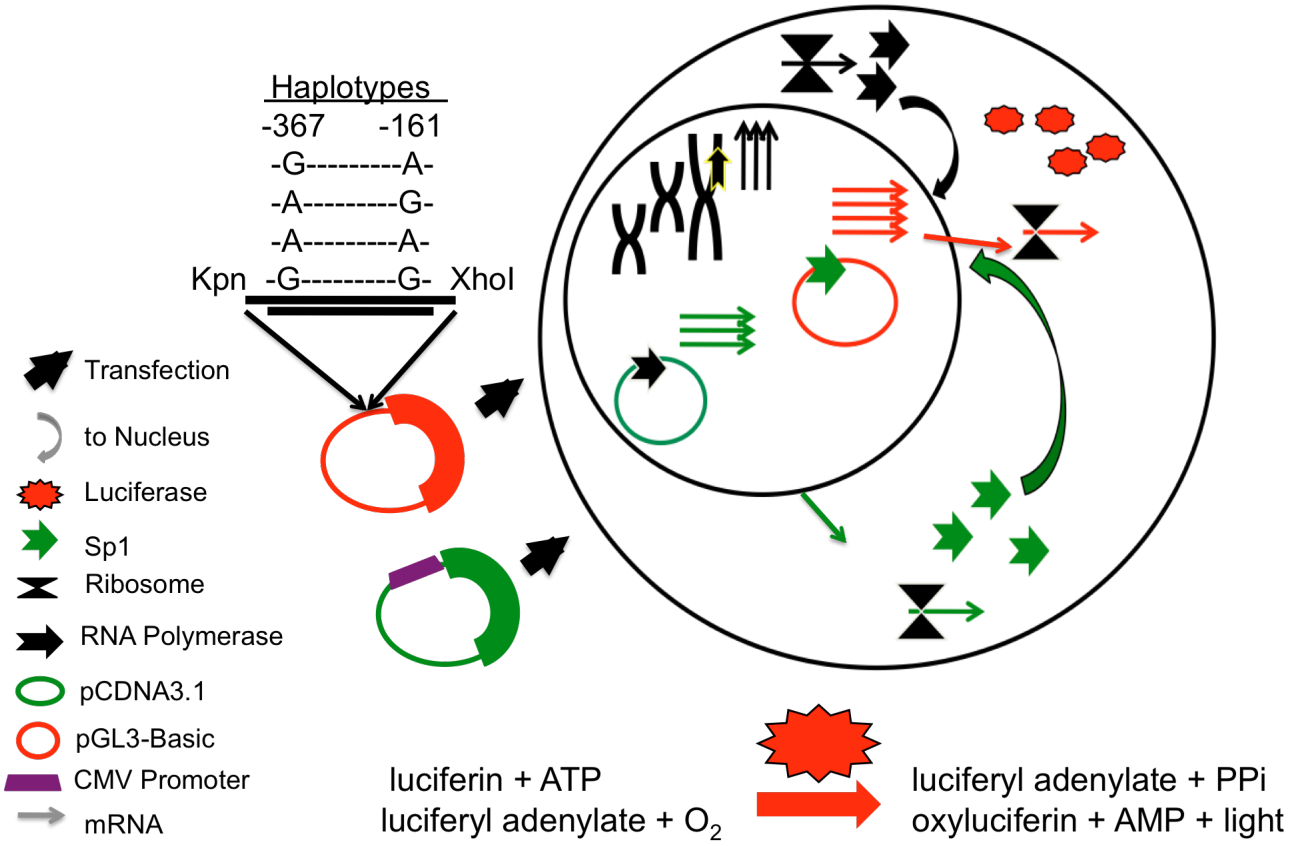

Figure 3.1. Cartoon overview of the promoter reporter assay. Two plasmids are transfered into the PC12 cell nucleus. The CMV-driven Sp1-expressing plasmid is overexpressed, producing mRNA which returns to the nucleus as protein; a transcription factor which binds to the PNMT promoter to drive Luciferase mRNA production. When the cells are lysed, the quantity of Luciferase protein is compared between 4 promoter haplotyes.

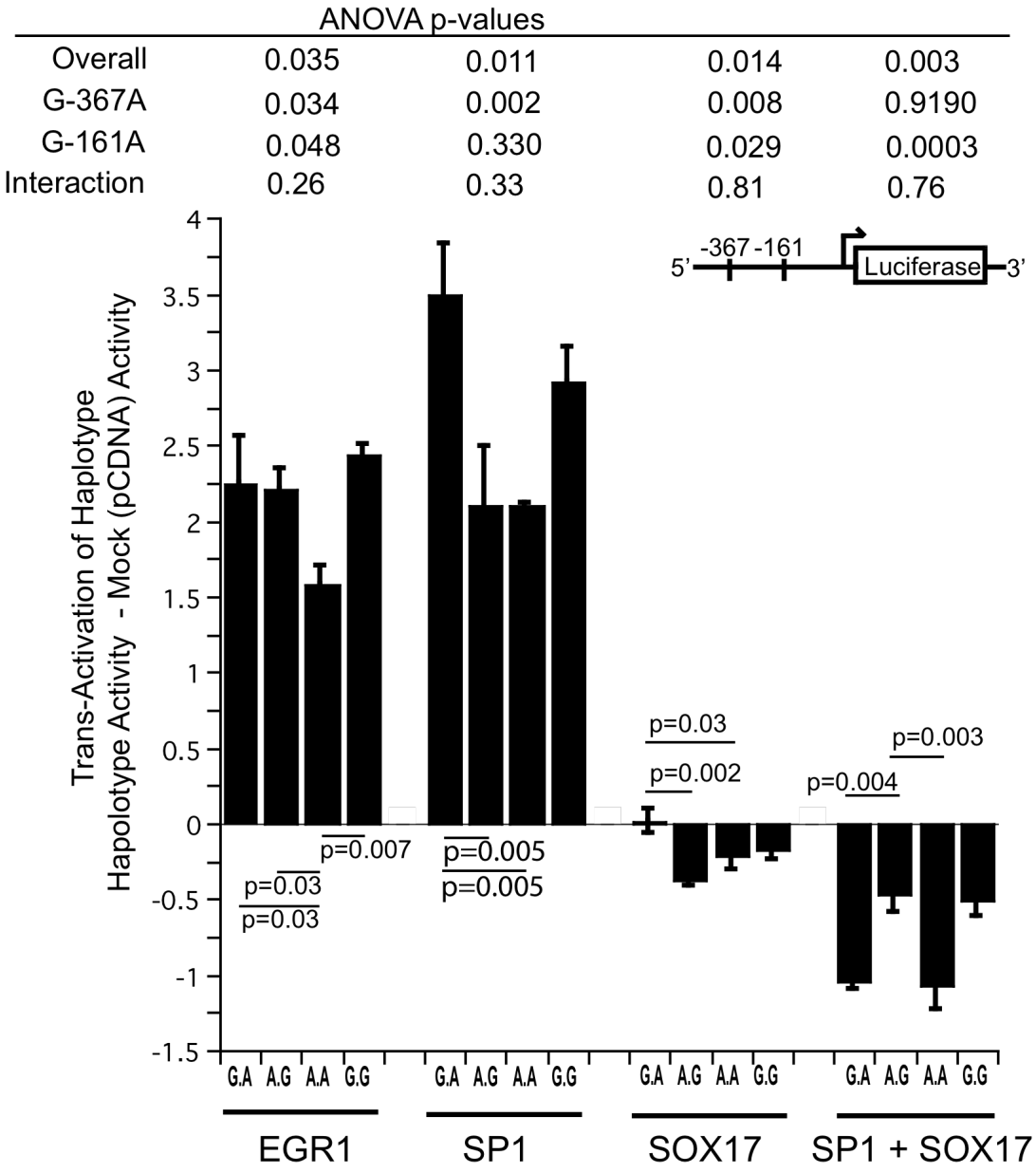

Figure 3.2. Allele-specific trans-activation of PNMT promoter haplotypes in PC12 cells. Y axis shows promoter trans-activation expressed as the increase in luciferase activity of transfected promoter haplotypes with co-transfection of cmv-driven transcription factors. Activity of mock (pCDNA3.1 plasmid) co-transfection is subtracted from transcription factor activity. Units on Y axis represent x fold over-expression normalized by empty vector (pGL3-Basic) baseline activity. To account for cell number, lysate volume and transfection efficiency variations, raw RLU luciferase signal is normalized by total protein quantified by the Bradford assay. Mean of four wells +/- 1 standard error.
Along X axis are listed the transfected promoter haplotypes (G-367A . G-161A) and transcription factors for each well (Sp1, Egr1, Sox17 and Sp1 + Sox17). Along top are ANOVA p-values for haplotype effects, allele effects, and pairwise haplotype comparison.

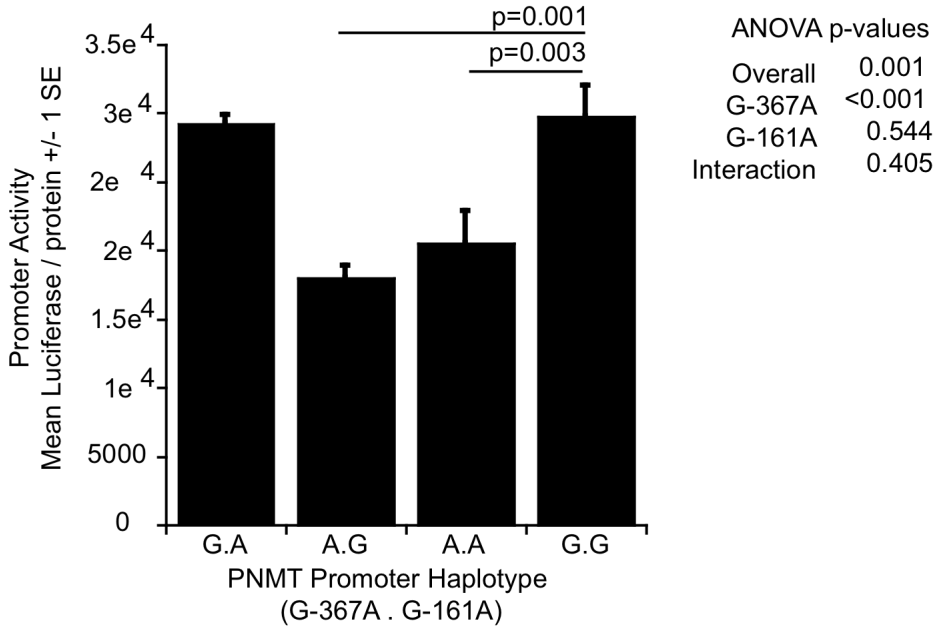

Figure 3.3 PNMT prmoter haplotype activity in PC12 cells with endogenous transcription factors. Along X axis are the four PNMT promoter haplotypes under comparison via luciferase promoter reporter assay in PC12 cells. Y axis shows luciferase luminescence activity normalized by total protein. Endogenouse transcription factors refers to the set of transcription factors present in the PC12 cell which control transcription activity.

INTRODUCTION

In this chapter we seek to determine experimentally if the allele-specific binding observed in chapter 2 for the G allele at -367 and the A allele at -161 results in increased or decrased PNMT promoter activity. We test this by constructing a promoter reporter plasmid which contains all 4 possible PNMT 428 bp promoter haplotypes (G-367A.G-161A) driving expression of Firefly luciferase, and comparing the activity in response to co-transfection of the transcription factors predicted in chapter 2.

RESULTS

Promoter Haplotype Activity In PC12 Cells

To test the combined effects of rs876493 (G-161A) and rs3764351 (G-367A) on PNMT promoter function, luciferase reporter constructs for different G-367A.G-161A haplotypes (G.A, A.G, A.A, G.G) were generated and transiently transfected into PC12 cells. Promoter activity was not significantly different from pGl3-Basic background for both -367A haplotypes (A.A, A.G), whereas both G-367 haplotypes did produce activity significantly above baselne ($p < 0.001$, Figure 3.1. far left).

We observed significant 4x expression for the G-367 haplotypes with Sp1 co-transfection ($p < 0.001$). We did not observe a difference in fold-overexpression when normalizing by pcDNA3.1 versus pGL3-Basic. With Egr1 co-transfection all but the A.A haplotype were significantly above baseline at $p < 0.01$ significance threshold (Figure 3.1. third set from left-to-right). No promoter activity above background was observed for Ap2, SrY, and Ahr/Arnt co-transfection (data not shown).

Luciferase activity for all haplotypes and pGl3-Basic negative control with or without co-transfection of indicated transcription factor plasmids or mock pcDNA3.1 promoter-only vector was normalized using total protein quantification (Bradford assay).

These results suggest a non-polymorphic Egr1 active motif within the 428 bp human PNMT promoter sequence and Sp1 active motif with preferential stimulation of the G allele at G-367A.

Sox17 and Sp1 Opposite Directional Effects

Figure 3.1 shows significant activity below baseline for all 4 Sox17 co-transfected haplotypes (far right), with AG haplotype lowest. This suggested consideration of an inhibitory effect, either by direct binding to the promoter or an epigeneitc effect. The PNMT promter does have a CpG island spanning the haplotype we are studying above, thus I suspect epigenetic promoter inactivation by CpG island hypermethylation. In order to differentiate the two possibilities we co-transfected Sp1 and Sox17, with pcDNA and Ap2 as respective negative dual-transfection controls. If in fact Sox17 has a direct regulatory effect on G-161A, then we would expect promoter activity lower for Sp1+Sox17 versus Sp1 alone. If the effect of Sox17 is to upregulate PNMT promoter inactivation by CpG island methylation, then we would expect activity similar to Sox17-only transfection for the dual transfection.

METHODS

Luciferase Promoter Reporter Plasmids

Human PNMT promoter reporter plasmids were constructed as previously described (Chen et al., 2008). GG and AG haplotype fragments corresponding to PNMT -428 to +10 bp (with respect to cap site of 19 bp 5' UTR, X52730 : 1327 - 1793, chr 17 : 35077803 - 35078242) were amplified from genomic DNA of known subject who is GG homozygous or AG heterozygous. The haplotype fragments were subcloned by sticky-end ligation into the upstream promoter site of the pGL3-Basic vector (Promega Inc., Madison, WI, USA) between XhoI and KpnI restriction digest sites (New England Biolabs, Ipswich, MA, USA). GA and AA haplotypes were made by site-directed mutagenesis per manufacturer specifications (QuickChange II; Stratagene, La Jolla, CA, USA) with rs876493 A-allele mutagenic primers using the aforementioned GG and AG haplotype plasmids as template, respectively. Successful ligations were verified by electrophoresis on 1% agarose of maxi-prepped plasmids digested with KpnI and XhoI; correct sequence and orientation were verified by sequencing with RV3 reverse primer (Promega Inc., Madison, WI, USA).

Plasmids were amplified in DH5alpha cells (Invitrogen, Carlsbad, CA, USA), and purified by plasmid maxiprep (PureLink HiPure Plasmid Maxiprep Kit, Invitrogen, Carlsbad, CA, USA). Plasmid concentrations were quantified by SmartSpec (BioRad, Hercules, CA, USA) and adjusted to 1.0 µg/µl before

transfection. We also prepared a negative control pGL3-Basic promotorless vector and a mock co-transfection negative control pCDNA3.1 (CMV promoter without transcription factor sequence).

Transient Transfection And Reporter Assay

PC12 rat adrenal pheochromocytoma cells were grown in high-glucose Dulbecco's modified Eagle media supplemented with penicillin G (100 U/ml), streptomycin sulfate (100 mg/ml), 5% fetal bovine serum and 10% horse serum. PC12 cells were splitted to 12-well plates and reached 50 - 60% confluence overnight before transfection. Plasmid DNA was incubated in 100 µl per well of serum-free media with TransfectIN reagent (BioRad) for 15 minutes, then added to 400 µl PC12 media per well. Cells were transiently transfected for 4 hours with 1µg/well haplotype plasmid or pGl3-basic vector then washed with serum-free media followed by 1 ml PC12 media. Firefly luciferase activity was measured after 48 hours incubation at 37 °C, 6% CO_2. Luciferase activity was normalized by total protein per well measured by Bradford assay (Protein Assay, BioRad, Hercules, CA, USA) in a Wallac spectrophotometer. Linear range of protein concentration was verified using serial dilutions of BSA as a standard.

In addition to testing promoter activity with endogenous PC12 transcription factors, co-transfection of transcription factors, including CMV-driven Ahr/Arnt, Egr1, Ap2, Sox17 or Sp1 plasmids (Open Biosystems,

Huntsville, AL, USA), was tested at 0.1ug/well. Dual co-transfection of Sp1 with Sox17 or Sp1 with pcDNA was also tested.

Calculation of Promoter Activity

The two readings obtained in each experiment are (1) quantity of light produced after adding luciferin and ATP to the 50 µl lysate from one well in a 12-well plate where PC12 cells were transiently transfected with one of 5 possible pGL3-Basic plasmids (empty promoter or one of 4 PNMT promoters driving luciferase activity) and a CMV-driven trans-activating plasmid (mock without transcription factor gene or a transcription factor such as Sp1, Egr1, Sox17); and (2) the quantity of protein in 10 µl of lysate measured by A595 absorbance in the Bradford assay. We refer to these two readings as RLU and A595. To calculate the activity (PA) of the G.G haplotype we subtract background (buffer) reading then divide by total volume. This haplotype-activity number is then normalized to total protein (also background subtracted and divided by volume). The activity shown in Figure 3.3 is the mean ACTIVITY for 4 wells +/- 1 standard error. In order to compare experiments, we normalize all values to the empty vector pGL3-basic activity. Finally, in order to control for effects of plasmid co-transfection on trans-activation, we subtract the activity of the mock co-transfection of the pcDNA plasmid, which contains the promoter used to drive transcription factor expression.

CONCLUSION

Luciferase promoter reporter assays clearly showed a significant increase in promoter activity for G-367 haplotypes (G.G and G.A) versus -367A haplotypes. We also observed G-allele-specific stimulation of promoter activity with Sp1 co-transfection. Egr1 was the only other factor which stimulated PNMT promoter activity above baseline for any haplotype (we also tested Ap2alpha, Ap2beta, Ahr/Arnt, Sry). Hence this result confirms the expected Egr1/Sp1 regulation of the human PNMT promoter predicted by the rat model, and suggests Sp1 allele-specific stimulation at G-367A.

Furthermore we identified an interesting Sp1-dependent inhibitory effect of Sox17 on the -161A allele. When only Sox17 is co-transfected, the A.G haplotype shows the greatest inhibitory effect, while the other 3 haplotypes are at or significantly indistinguishable from baseline promotorless activity. However, in the presence of Sp1 we observe repression of the G.A and A.A haplotypes 2x that of the A.G and A.A haplotypes.

Acknowledgements

The contents of this chapter in part is currently being prepared for submission for publication. It was was co-authored by Sushil K Mahata, Stephane Chiron, Maite Courel and Daniel O'Connor. The dissertation author was the primary investigator and author of this material.

CHAPTER 4.

HOW CAN WE CONNECT THE *IN POPULO* GENETIC ASSOCIATION WITH THE *IN VITRO* MOLECULAR AND CELLULAR RESULTS?

Table 4.1. Mouse mRNA transcripts: significant decrease of expression in chromaffin cells of male PNMT KO mice, VAMPIRE analysis with Bonferroni correction.

Name	Unigene ID	Name	Unigene ID
ACTA1	MM.214950	PNMT	MM.57030
AGT	MM.301626	PPP1R1A	MM.143788
ARSK	MM.196399	PRL8A2	MM.291450
CD300LG	MM.266885	PTPRN	MM.2902
DIRAS2	MM.29362	RETN	MM.1181
FRZB	MM.427436	SERPINB1A	MM.20144
FRZB	MM.427436	SKIV2L2	MM.291029
GALC	MM.5120	SLC6A4	MM.300318
GNAS	MM.125770	SMARCE1	MM.379086
ISL1	MM.42242	SOX4	MM.240627
MAPT	MM.1287	SPTLC1	MM.240336
MSI2	MM.400451	SRD5A1	MM.422833
NPY	MM.154796	SRD5A1	MM.422833
NRXN1	MM.312068	TH	MM.1292
OLFM3	MM.54183	TLX2	MM.37
PDE10A	MM.87161	TMEM130	MM.126793
		VGF	MM.389697

Table 4.2. Mouse mRNA transcripts: significant increase of expression in chromaffin cells of male PNMT KO mice, VAMPIRE analysis with Bonferroni correction.

Name	Unigene ID	Name	Unigene ID
ABTB2	MM.35850	HMGCR	MM.316652
ACSM3	MM.334199	HSPA1A	MM.6388
ACVR1	MM.689	HSPA1B	MM.372314
ALB	MM.16773	LSM12	MM.259079
ALDH1A1	MM.457973	LYZ	MM.177539
APOA1	MM.26743	MALAT1	MM.298256
ARRDC3	MM.423137	MUP1	MM.458383
AYTL1	MM.21463	NEK5	MM.113940
AYTL1	MM.21463	NLGN2	MM.151293
CIDEA	MM.449	PDK4	MM.235547
COX8B	MM.3841	PLXNB3	MM.275600
CPT1B	MM.358582	PPARA	MM.212789
CR1	MM.423001	S100A8	MM.21567
CUEDC1	MM.374846	SERPINA1A	MM.439693
CYP3A11	MM.332844	SERPINA1B	MM.439693
DSP	MM.355327	SERPINA1C	MM.439694
EGR3	MM.103737	SLC15A2	MM.281804
ERDR1	MM.391385	SLC25A3	MM.298
FABP1	MM.22126	SLC27A2	MM.290044
FABP3	MM.388886	SSR1	MM.409778
FADS1	MM.30158	TTR	MM.2108

Table 4.3. Obesity-associated genes differentially expressed in PNMT KO vs WT mouse chromaffin cells.

mRNA	Symbol	Name
DOWN	AGT	Angiotensinogen (serpin peptidase inhibitor, clade A, member 8)
DOWN	NPY	Neuropeptide Y (Npy)
DOWN	ISL1	ISL1 transcription factor
UP	CIDEA	Cell death-inducing DNA fragmentation factor, alpha subunit-like effector A (Cidea)
UP	HMGCR	3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase
UP	CYP3A11	Cytochrome P450, family 3, subfamily a, polypeptide 11)
UP	APOA1	Apolipoprotein A-I
UP	PPARA	Peroxisome proliferator activated receptor alpha
UP	SERPINA1B	Serine (or cysteine) preptidase inhibitor, clade A, member 1b
DOWN	GNAS	GNAS (guanine nucleotide binding protein alpha stimulating) complex locus
UP	TTR	Transthyretin
UP	ALDH1A1	Aldehyde dehydrogenase family 1, subfamily A1
UP	HSPA1B	Heat shock protein (hsp68) mRNA
UP	PDK4	Pyruvate dehydrogenase kinase, isoenzyme 4
UP	FABP1	Fatty acid binding protein 1, liver
UP	MUP1	Major urinary protein 4
DOWN	VGF	VGF nerve growth factor inducible
UP	COX8B	Cytochrome c oxidase, subunit VIIIb
DOWN	RETN	Adipose-specific cysteine-rich secreted A12-alpha precursor
DOWN	FRZB	Frizzled-related protein
DOWN	PDE10A	CAMP/cGMP phosphodiesterase (Pde10a)
UP	ARRDC3	Arrestin domain containing 3
DOWN	PNMT	Phenylethanolamine-N-methyltransferase (Pnmt)
DOWN	TH	Tyrosine hydroxylase (Th)

Table 4.4. Clusters of GO terms for PNMT KO vs. WT transcriptome.

Group name	p-value	q-value
Endothelial cell proliferation	1.56E-07	1.85E-07
Phenylethanolamine N-methyltransferase activity	1.96E-05	3.00E-04
Hormone activity	2.20E-09	8.82E-07
Extracellular region	1.60E-09	6.10E-09
Lipid metabolic process	5.00E-10	2.19E-08
Membrane lipid metabolic process	7.33E-06	2.71E-01
Phospholipid metabolic process	1.84E-05	5.63E-01
Steroid biosynthetic process	2.09E-07	1.25E-02
Cholesterol biosynthetic process	3.55E-07	5.29E-03
Response to stress	1.39E-05	2.89E-02
Digestion	4.86E-05	3.99E-04
Lipoprotein binding	2.04E-06	3.33E-06
High-density lipoprotein binding	2.92E-08	1.37E-02
Steroid metabolic process	1.47E-05	2.08E-01
Cholesterol metabolic process	3.86E-05	6.47E-01
Glucocorticoid metabolic process	8.11E-07	7.49E-03
Lipid biosynthetic process	3.10E-09	5.28E-04
Response to temperature stimulus	3.88E-07	3.89E-04
Response to hormone stimulus	4.16E-05	8.61E-04
Sterol transporter activity	2.40E-07	7.29E-04
Sterol transport	3.19E-06	6.13E-03
Sterol biosynthetic process	9.45E-07	2.04E-02
Sequestering of lipid	6.50E-09	2.46E-05
Digestive process	1.81E-06	1.57E-05
Cholesterol absorption	8.60E-09	2.37E-02
Regulation of cholesterol absorption	2.90E-09	2.69E-07
Cholesterol transport	2.87E-06	8.97E-01
Adrenal gland development	6.96E-07	1.52E-04
Thyroid hormone metabolic process	8.11E-07	4.47E-02
Catecholamine biosynthetic process	4.13E-05	3.42E-02
Hormone metabolic process	0.00E+00	0.00E+00
Negative regulation of enzyme activity	2.64E-05	4.88E-04
Response to peptide hormone stimulus	1.24E-06	2.73E-02
Blood vessel endothelial cell migration	7.38E-08	8.89E-02
Endothelial cell migration	8.11E-07	2.65E-05
Lipid digestion	8.60E-09	1.55E-04
Cellular lipid metabolic process	5.00E-10	2.57E-08
Intestinal absorption	3.54E-07	1.91E-01

Table 4.5. Transfac cluster of differentially expressed transcription factors in PNMT WT vs. KO transcriptome, GOBY analysis.

Transfac ID	Name	Count	p-value
Transfac:T00029	AP-1	4	1.236E-07
Transfac:T00108	C/EBPalpha	5	3.3E-09
Transfac:T00370	HNF-3	5	0
Transfac:T00371	HNF-3alpha	5	0
Transfac:T00372	HNF-4alpha1	4	2E-10
Transfac:T00384	HSF1	2	0.000673648
Transfac:T00459	C/EBPbeta	5	0
Transfac:T00972	HSF2	2	0.000673648
Transfac:T01049	HNF-3beta	4	2.9E-09
Transfac:T01050	HNF-3gamma	4	0
Transfac:T01051	FOXA4a	4	0
Transfac:T01104	HNF-3	4	0
Transfac:T01830	FOXA4b	4	0
Transfac:T02294	FOXI1a	4	0
Transfac:T02422	HNF-4alpha2	4	2E-10
Transfac:T02429	HNF-4alpha1	4	2E-10
Transfac:T02445	FOXC2	4	0
Transfac:T02493	CWH-1	4	0
Transfac:T02494	CWH-2	4	0
Transfac:T03256	HNF-3beta	4	0
Transfac:T03257	HNF-6alpha	7	0
Transfac:T03258	HNF-6beta	7	0
Transfac:T04166	FOXD3	4	0

Table 4.6. GOBY and DAVID clusters of GO annotations with lowest p-values.

	GOBY: LIPID METABOLIC PROCESS p = 5e-10	GOBY: HORMONE METABOLIC PROCESS p = 2.2e-9	DAVID: DEVELOPMENTAL PROCESS p = 6.2e-4	OBESITY GENE MAP p = 0.0008
DOWN	GALC	NPY	AGT	
DOWN	SRD5A1	PRL8A2	FRZB	AGT
DOWN		RETN	GNAS	FRZB
DOWN		VGF	ISL1	GNAS
DOWN			TH	ISL1
DOWN			VGF	NPY
DOWN				PNMT
DOWN				RETN
DOWN				TH
DOWN				VGF
UP	ACSM3	ACSM3	ALDH1A1	ALDH1A1
UP	APOA1	APOA1	APOA1	APOA1
UP	CIDEA	CIDEA	CIDEA	CIDEA
UP	CPT1B	CPT1B	HMGCR	COX8B
UP	FABP3	FABP3	HSPA1B	CYP3A11
UP	HMGCR	HMGCR	PPARA	FABP1
UP	LPCAT2	LPCAT2		HMGCR
UP	SLC27A2	SLC27A2		HSPA1A
UP				MUP1
UP				PDK4
UP				PLXNB3
UP				PPARA
UP				SERPINA1A
UP				SERPINA1B
UP				TTR

INTRODUCTION

In chapter 1 we observed an association between G-161A and BMI in European-American women, with a BMI-increasing effect of the -161A allele. This corresponds to the A.A haplotype, since we did not observe the G.A haplotype in European-Americans. In chapter 2 we demonstrated G-367 and -161A allele-specific binding of transcription factors extracted from PC12 cells. In chapter 3 our luciferase promoter reporter assays identified the A.A haplotype as the least-active promoter, when in the presence of Egr1, Sp1, and Sox17+Sp1 transcription factors (as well as endogenous factors). In this chapter we seek to answer the question: "how does the decrease in PNMT promoter activity for the A.A haplotype affect chromaffin cell function?". This is the first step in answering the question: "how does decreased PNMT promoter activity lead to obesity in women?".

The traditional method for answering this question is a literature review. As explained in the introduction, one clear model involves the interaction between the chromaffin cell and the adipocyte. Chromaffin cells synthesize and store epinephrine and norepinephrine in chromaffin granules, catecholamine hormones interconverted by PNMT. Upon sympathetic nervous stimulation of chromaffin cells via nicotinic receptors, chromaffin cells release the contents of chromaffin granules into the bloodstream. The epinephrine travels in the blood to the adipocytes, where it stimulates beta2-adrenergic receptors. This stimulation on the adipocyte cell membrane results in

conversion of AMP to cAMP, in the standard GPCR response. The cAMP activates PKA, which then activates hormone sensitive lipase (HSL). This lipase catalyzes the breakdown of triglicerides in the adipocyte (Illustration 1). Hence we expect a decrease in PNMT promoter activity to cause a decrease in triglyceride catabolism and therefore triglycerides accumulate in the adipocyte and obesity develops.

An alternative approach is to look at the transcriptomic changes which occur in the chromaffin cell in a PNMT KO mouse. Complete elimination of PNMT expression is an extreme of decreased expression, but it can provide clues for a system-wide association test of interacting genes. We hypothesize that the set of differentially expressed genes in PNMT KO are more likely to be associated with obesity or hypertension than a random collection of genes. To accomplish this in chapter 4 we first identify differerentially expressed genes and then use Fisher's exact test to test for enrichment of GO annotation clusters.

RESULTS

Identification and Clustering of Differentially Expressed Transcripts

All chips passed quality criteria test. Raw intensity data was uploaded to the VAMPIRE microarray analysis suite website for analysis. All transcripts were tested for significant difference in expression between KO and WT, with a Bonferroni correction of multiple testing. Significantly differentially expressed genes listed in Table 4.1 and 4.2, of note is downregulation of TH, a gene upstream of PNMT in the catecholamine synthesis pathway (Figure 4.1). Clustering of GO annotations using the MGI webserver show the DNA binding proteins Sox4 (down), Tlx2 (down), Atb2 (up), Egr3 (up), Ppara (up), Mgmt (aka Agt, down), Isl1 (down), Smarce1 are all significantly differentially expressed. We observe significant enrichment of obesity associated, lipid metabolic process, hormone metabolic process, and developmental process (Table 4.6).

METHODS

Identification of Differentially Expressed Transcripts

Raw intensity data in .cel format (Hubner & Schulz, Max Delbruk Center, Berlin, Germany) from 6 wild type and 6 PNMT KO mice (all male) was uploaded to the VAMPIRE (Hsiao et al., 2005; Hsiao et al., 2004) microarray analysis suite website for automated identification of differentially expressed genes and GO (Ashburner et al., 2000)/Transfac (Wingender et al., 1996) annotation clusters using GOBY (part of VAMPIRE). We also utilized DAVID (Huang da et al., 2009), which is identical to GOBY and uses a Fisher's exact test to calculate p-values.

Test for BMI QTL's in PNMT KO Gene Clusters

We tested the set of KO mouse differentially expressed genes (KOTome) for association to obesity in humans in order to establish a mechanism for catecholamine malfunction in obesity. For comparison, we also tested the genes of the Obesity Gene Map (Rankinen et al., 2006) (OGM) and replicated the most recent set of obesity GWAS hits. We compared the distribution of p values versus the entire 500K SNPs, as well as a random selection of SNPs of the same size. Tag-SNPs for the OGM and the KOTome. All statistical tests were performed in Plink, using genotype and phenotype data obtained from NCBI dbGAP. Mouse-human ortholog names were obtained from the Mouse Genome Informatics website, then positions for the

human genes were downloaded from the UCSC genome browser. SNPs within 30 kb of gene boundaries were included in the analysis.

CONCLUSION

This analysis shows the results of complete absence of PNMT expression, rather than the decreased expression observed for the A.A promoter haplotype. The list of differentially expressed genes includes 6 transcription factors and significant GO annotation enrichment for lipid metabolism and developmental processes. Furthermore, there is a significant enrichment for obesity-associated genes in candidate gene studies.

Acknowledgements

The contents of chapter in part is currently being prepared for submission for publication. It was co-authored by Fujun Liu, Michael Ziegler, Daniel O'Connor. The dissertation author was the primary investigator and author of this material.

DISCUSSION:

RECOMMENDATIONS

FOR

FUTURE INVESTIGATION

GWAS and Resequencing Study Design

The genetic story of PNMT is complicated by gene*sex and gene by environment effects. In this study we observe an association effect in females, while Ji et al observed an effect only during exercise. Likewise the PNMT KO mouse blood pressure is significantly different from that of the wild type mouse only during exercise. This result makes sense, because epinephrine is a stress hormone, and we'd expect a stress situation is necessary to show the difference between genotypes. Hence it is not surprising that the recent GWAS studies for hypertension and obesity did not observe a significant effect of PNMT nor other catecholamine synthesis and signalling pathway genes, despite extensive research in rodent models.

Hence a GWAS research needs to include diet and exercise as part of the study design, either as a factor tested or controlled for in association tests. Furthermore, important lifestyle factors such as smoking and personal stress sholud be taken into account. It is a waste of time to study stress genetics when the phenotype is measured while subjects are at rest.

Furthermore, in terms of SNP selection for resequencing and genotyping projects must be aware of sequence-specific variation in genotype accuracy. We encountered insurmountable difficulties with the G-367A SNP as have others. This is in part due to the presence of a CpG island in the PNMT promoter, a sequence feature frequently encountered in mammalian

regulatory regions. If a repeat is present in the genotyped region, then there will be a lack of primer specificity and erroneous genotypes will be obtained.

Phylogenetic Footprinting

The prevailing transcription factor binding motif databases, Jaspar and Transfac, are limited to human-rodent conserved motifs. In this study we observe that although G-161A is in a motif not conserved in human-rodent alingments, it is conserved in human-rhesus alignments. Thus the putative Sox17 regulatory motif discovered in this dissertation is missed in phylogenetic footprinting analysis. As more primate genome sequences become available, it is valuable to search for primate-specific transcription factors which could demonstrate why we are so different from mice.

EMSA and Antibody Tests of Transcription-Factor Binding

The EMSA is a conventional method which can only query one sequence versus many transcription factors at the same time. This is the inverse of a ChIP-ChIP method, where one transcription factor is queried against all possible sequence motifs. This approach is powerfull for exploring cell-specific transcriptional environments, for example the set of transcription factors expressed in the chromaffin cell versus an adipocyte. This information is valuable for stem cell research, where the "regulome" determines the state of a cell and its differentiation or de-differentiation process. EMSA technology

needs an upgrade to a method where the shift-causing proteins can be easily purified and sequenced by MS. Antibodies are a much less reliable tool, as their specificity and interference with oligonucleotide binding is unknown.

Luciferase Promoter Reporter Assays

Prior to this dissertation, the primary use of luciferase promoter reporter assays was to detect an increase in transcription activity for a promoter inserted into the pGl3-Basic vector. In this dissertation we demonstrate allele-specific transcription repression by Sox17. This observation was possible thanks to a novel method for calculating trancriptional activity, where the baseline activity of the pGl3-Basic empty vector (without a promoter inserted) is used for normalization. Eukaryote transcriptional regulation includes enhancers, inhibitors and insulators, all of which have the potential for allele-specific changes in binding and activity. This advance adds to the toolkit of regulatory polymorphism phenotype discovery.

Transcriptomic Analyis

The KO vs. WT PNMT transcriptome analysis does not fit well with the rest of the dissertation, since it is limited to male mice and the association is in female individuals. However it is interesting to note the list of transcription factors differentially expressed. A potential avenue for bioinformatics research

would be to search the differentially expressed genes for motifs of the differentially expressed transcription factors.

Cis-Regulation versus Trans-Regulation

Although Chromatin Immunoprecipitation is an obvious candidate for further study to verify the Sp1-dependent repression via Sox17, a negative result would not provide definitive proof. It is possible that Sp1/Sox17 regulate an intermediate transcription factor which actually has the effect on the SNP.

Epigenetic SNPs

CpG islands such as the one present in the PNMT and Sox17(Zhang et al., 2008) promoters are targets for epigenetic silencing of transcription activity via methylation of Cytidine. The methylation status of the PNMT promoter can be measured by bisulfite treatment before sequencing (Haines et al., 2001), can be inhibited *in cella* using the demethylating agent 5-Aza-2 deoxycytidine (Parilla et al., 2006), or can be created in vitro prior to transfection using SssI methylase.

REFERENCES

Abecasis, G.R., Cardon, L.R., and Cookson, W.O. (2000). A general test of association for quantitative traits in nuclear families. Am J Hum Genet *66*, 279-292.

Abecasis, G.R., Cherny, S.S., Cookson, W.O., and Cardon, L.R. (2002). Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet *30*, 97-101.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T.*, et al.* (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet *25*, 25-29.

Bao, X., Lu, C.M., Liu, F., Gu, Y., Dalton, N.D., Zhu, B.Q., Foster, E., Chen, J., Karliner, J.S., Ross, J., Jr.*, et al.* (2007). Epinephrine is required for normal cardiovascular responses to stress in the phenylethanolamine N-methyltransferase knockout mouse. Circulation *116*, 1024-1031.

Bogardus, C. (2009). Missing heritability and GWAS utility. Obesity (Silver Spring) *17*, 209-210.

Chen, Y., Rao, F., Rodriguez-Flores, J.L., Mahapatra, N.R., Mahata, M., Wen, G., Salem, R.M., Shih, P.A., Das, M., Schork, N.J.*, et al.* (2008). Common genetic variants in the chromogranin A promoter alter autonomic activity and blood pressure. Kidney Int *74*, 115-125.

Comings, D.E., and MacMurray, J.P. (2000). Molecular heterosis: a review. Mol Genet Metab *71*, 19-31.

Cui, J., Zhou, X., Chazaro, I., DeStefano, A.L., Manolis, A.J., Baldwin, C.T., and Gavras, H. (2003). Association of polymorphisms in the promoter region of the PNMT gene with essential hypertension in African Americans but not in whites. Am J Hypertens *16*, 859-863.

Haines, T.R., Rodenhiser, D.I., and Ainsworth, P.J. (2001). Allele-specific non-CpG methylation of the Nf1 gene during early mouse development. Dev Biol *240*, 585-598.

Hsiao, A., Ideker, T., Olefsky, J.M., and Subramaniam, S. (2005). VAMPIRE microarray suite: a web-based platform for the interpretation of gene expression data. Nucleic Acids Res *33*, W627-632.

Hsiao, A., Worrall, D.S., Olefsky, J.M., and Subramaniam, S. (2004). Variance-modeled posterior inference of microarray data: detecting gene-expression changes in 3T3-L1 adipocytes. Bioinformatics *20*, 3108-3127.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Ji, Y., Salavaggione, O.E., Wang, L., Adjei, A.A., Eckloff, B., Wieben, E.D., and Weinshilboum, R.M. (2005). Human phenylethanolamine N-methyltransferase pharmacogenomics: gene re-sequencing and functional genomics. J Neurochem *95*, 1766-1776.

Ji, Y., Snyder, E.M., Fridley, B.L., Salavaggione, O.E., Moon, I., Batzler, A., Yee, V.C., Schaid, D.J., Joyner, M.J., Johnson, B.D.*, et al.* (2008). Human phenylethanolamine N-methyltransferase genetic polymorphisms and exercise-induced epinephrine release. Physiol Genomics *33*, 323-332.

Kaneda, N., Ichinose, H., Kobayashi, K., Oka, K., Kishi, F., Nakazawa, A., Kurosawa, Y., Fujita, K., and Nagatsu, T. (1988). Molecular cloning of cDNA and chromosomal assignment of the gene for human phenylethanolamine N-methyltransferase, the enzyme for epinephrine biosynthesis. J Biol Chem *263*, 7672-7677.

Kepp, K., Juhanson, P., Kozich, V., Ots, M., Viigimaa, M., and Laan, M. (2007). Resequencing PNMT in European hypertensive and normotensive individuals: no common susceptibilily variants for hypertension and purifying selection on intron 1. BMC Med Genet *8*, 47.

Lenhard, B., and Wasserman, W.W. (2002). TFBS: Computational framework for transcription factor binding site analysis. Bioinformatics *18*, 1135-1136.

Maher, B. (2008). Personal genomes: The case of the missing heritability. Nature *456*, 18-21.

Mann, M.B., Wu, S., Rostamkhani, M., Tourtellotte, W., MacMurray, J., and Comings, D.E. (2001). Phenylethanolamine N-methyltransferase (PNMT) gene and early-onset Alzheimer disease. Am J Med Genet *105*, 312-316.

Mann, M.B., Wu, S., Rostamkhani, M., Tourtellotte, W., MacMurray, J.P., and Comings, D.E. (2002). Association between the phenylethanolamine N-methyltransferase gene and multiple sclerosis. J Neuroimmunol *124*, 101-105.

Moffatt, M.F., Kabesch, M., Liang, L., Dixon, A.L., Strachan, D., Heath, S., Depner, M., von Berg, A., Bufe, A., Rietschel, E.*, et al.* (2007). Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. Nature *448*, 470-473.

Ng, P.C., and Henikoff, S. (2006). Predicting the effects of amino acid substitutions on protein function. Annu Rev Genomics Hum Genet *7*, 61-80.

Oades, R.D., Lasky-Su, J., Christiansen, H., Faraone, S.V., Sonuga-Barke, E.J., Banaschewski, T., Chen, W., Anney, R.J., Buitelaar, J.K., Ebstein, R.P.*, et al.* (2008). The influence of serotonin- and other genes on impulsive behavioral aggression and cognitive impulsivity in children with attention-deficit/hyperactivity disorder (ADHD): Findings from a family-based association test (FBAT) analysis. Behav Brain Funct *4*, 48.

Parilla, N.W., Hughes, V.S., Lierl, K.M., Wong, H.R., and Page, K. (2006). CpG DNA modulates interleukin 1beta-induced interleukin-8 expression in human bronchial epithelial (16HBE14o-) cells. Respir Res *7*, 84.

Peters, W.R., MacMurry, J.P., Walker, J., Giese, R.J., Jr., and Comings, D.E. (2003). Phenylethanolamine N-methyltransferase G-148A genetic variant and weight loss in obese women. Obes Res *11*, 415-419.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J.*, et al.* (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet *81*, 559-575.

Qi, L., Saberi, M., Zmuda, E., Wang, Y., Altarejos, J., Zhang, X., Dentin, R., Hedrick, S., Bandyopadhyay, G., Hai, T.*, et al.* (2009). Adipocyte CREB Promotes Insulin Resistance in Obesity. Cell Metabolism *9*, 277-286.

Rana, B.K., Insel, P.A., Payne, S.H., Abel, K., Beutler, E., Ziegler, M.G., Schork, N.J., and O'Connor, D.T. (2007). Population-Based Sample Reveals Gene-Gender Interactions in Blood Pressure in White Americans. Hypertension *49*, 96-106.

Rankinen, T., Zuberi, A., Chagnon, Y.C., Weisnagel, S.J., Argyropoulos, G., Walts, B., Perusse, L., and Bouchard, C. (2006). The human obesity gene map: the 2005 update. Obesity (Silver Spring) *14*, 529-644.

Saito, S., Iida, A., Sekine, A., Miura, Y., Sakamoto, T., Ogawa, C., Kawauchi, S., Higuchi, S., and Nakamura, Y. (2001). Identification of 197 genetic variations in six human methyltranferase genes in the Japanese population. J Hum Genet *46*, 529-537.

Scott A. Jackson, I.P., J Eugene leClerc, Thomas A. Cebula, Mark K. Mammel, Michael L. Kotewicz, Amit Mukherjee, Sushil K. Mahata, Juan L. Rodriguez Flores, Manjula Mahata, Daniel T. O'Connor, and Jannavi R. Srinivasan (2009). Principles of Functional Genomic Analysis. In Genomics: Fundamentals and Applications, D.B.C. Supratim Choudhuri, ed. (Informa Health Care).

Thorisson, G.A., Smith, A.V., Krishnan, L., and Stein, L.D. (2005). The International HapMap Project Web site. Genome Res *15*, 1592-1593.

Wilk, J.B., Chen, T.H., Gottlieb, D.J., Walter, R.E., Nagle, M.W., Brandler, B.J., Myers, R.H., Borecki, I.B., Silverman, E.K., Weiss, S.T.*, et al.* (2009). A genome-wide association study of pulmonary function measures in the Framingham Heart Study. PLoS Genet *5*, e1000429.

Wingender, E., Dietze, P., Karas, H., and Knuppel, R. (1996). TRANSFAC: a database on transcription factors and their DNA binding sites. Nucleic Acids Res *24*, 238-241.

Wong, D.L., and Tank, A.W. (2007). Stress-induced catecholaminergic function: transcriptional and post-transcriptional control. Stress *10*, 121-130.

Wu, S., and Comings, D.E. (1999). Two single nucleotide polymorphisms in the promoter region of the human phenylethanolamine N-methyltransferase PNMT gene. Psychiatr Genet *9*, 187-188.

Yamano, E., Isowa, T., Nakano, Y., Matsuda, F., Hashimoto-Tamaoki, T., Ohira, H., and Kosugi, S. (2008). Association study between reward dependence temperament and a polymorphism in the phenylethanolamine N-methyltransferase gene in a Japanese female population. Compr Psychiatry *49*, 503-507.

Zhang, L., Rao, F., Zhang, K., Khandrika, S., Das, M., Vaingankar, S.M., Bao, X., Rana, B.K., Smith, D.W., Wessel, J.*, et al.* (2007). Discovery of common human genetic variants of GTP cyclohydrolase 1 (GCH1) governing nitric oxide, autonomic activity, and cardiovascular risk. J Clin Invest *117*, 2658-2671.

Zhang, W., Glockner, S.C., Guo, M., Machida, E.O., Wang, D.H., Easwaran, H., Van Neste, L., Herman, J.G., Schuebel, K.E., Watkins, D.N.*, et al.* (2008). Epigenetic inactivation of the canonical Wnt antagonist SRY-box containing gene 17 in colorectal cancer. Cancer Res *68*, 2764-2772.