Research Article

iMedPub Journals
www.imedpub.com

Journal of Clinical Epigenetics
ISSN 2472-1158

2019
Vol.5 No.1:1

DOI: 10.21767/2472-1158.100086

Addiction and Immune Enriched Genomic Regions may be Influenced by Selection for Hepatitis B Resistance

Latifa F Jackson[1,2,3*], Maksim Shestov[1,4], Kim Burrows[1,2] and Fatimah LC Jackson[3,5]

[1]School of Biomedical Engineering, Sciences and Health Systems, Drexel University, Philadelphia, PA

[2]Pediatrics and Child Health, College of Medicine, Howard University, Washington, DC

[3]W. Montague Cobb Research Laboratory, College of Arts and Sciences, Howard University Washington, DC

[4]Genomics and Computational Biology Program, University of Pennsylvania, Philadelphia, PA

[5]Biology, College of Arts and Sciences, Howard University, Washington, DC

***Corresponding author:** Latifa F Jackson

Latifa.f.jackson@gmail.com

2201 Georgia Avenue NW, Room 421G, Washington DC 20059, USA

Tel: +202-806-6954

Citation: Jackson LF, Shestov M, Burrows K, Jackson FLC (2019) Addiction and Immune Enriched Genomic Regions may be Influenced by Selection for Hepatitis B Resistance. J Clin Epigenet Vol.5 No.1:1

Abstract

Finding the genetic markers that influence infectious and chronic disease phenotypes has been an area of significant biological study. Understanding complex disease related traits like addiction has been hampered by the lack of functional insights in to the human genome. We hypothesized that environmental factors such as geographical location, relative pathogen burden and infection rates will identify allele frequency differences in for immune and addiction gene hotspots between populations that are consistent with natural selection within human populations living predominantly in tropical environments.

To test whether there are correlative relationships between ecoregions, disease and population allele frequencies, we use genes contained in addiction and immunity curated by NCBI. Immune associated genes were identified from NCBI gene lists using immune related search terms. These terms were added to 587 genes previously identified as being involved in opiate, dopamine, and GABA reception addiction. These genes were then projected onto the genome to identify cluster regions of genetic importance for immunity and addiction. Clusters were defined as regions of the genome with more than 15 genes within a 1.5Mb linear genomic window. When addiction and immunity gene lists were combined, we found that they created three hotspots located on chromosomes 11, 17, and 19. Human polymorphism data was surveyed from the 1054 individuals comprising 51 populations of the Human Genome Diversity Panel, 1148 individuals comprising the 11 sample populations of the HapMap Project and the 1092 individuals representing the 1000 Genomes dataset. Our analyses demonstrate that when human populations are grouped into tropical versus non-tropical living groups, significant differences in allele frequencies at the hotspot located on chromosome 11 for 5 polymorphisms were found.

Keywords: Addiction; Immunity; Bioinformatics; Genomics; Geography

Received: March 01, 2019; **Accepted:** March 27, 2019; **Published:** April 03, 2019

Introduction

Human immunological interactions with their environment have long been considered the substrate for natural selection [1]. While this has been the acknowledged paradigm in evolutionary medicine, few connections have been made between concerning the evolutionary relationships between the complex chronic diseases underlying substance addiction and selection for immunity [2,3]. For example, addiction phenotypes in chronically substance abusing individuals bear striking similarities to immunological compromised individuals. This is thought to be because of an as yet under-characterized interplay between addiction phenotypes and immunity phenotypes. While the relationship has been documented between chronic substance addiction and degenerative immunity, the converse reaction is less well characterized. For example, opioids are known to effect host defenses [4] with heroin addicts presenting higher prevalence of infectious disease than those non-heroin abusers.

Our previous observations of immune genes sharing genome

locality with addiction genes at addiction hotspots motivated our interest in further understanding whether addiction alleles might arise in ethnic populations as a result of natural selection against infectious disease at immune loci. Specifically, we hypothesize that alleles associated with addiction that lie within hotspots adjacent to immune functioning genes will be correlated to infectious disease response and by extension will have allele frequencies that are tied to global distributions of infectious disease prevalence, where the associations are of a sufficient duration to be the substrate for natural selection. To date this has been a challenging endeavor due to the lack of global population sets within which to test these hypotheses.

One approach to studying natural selection in humans has been to examine single genes in a population to directly assess selection caused by some environmental effect (i.e. HBB and malaria, SLC24A5 and UV exposure) [5,6]. An analysis conducted on a global population set has demonstrated success in identifying variation in allele frequencies between populations taking into account diet, subsistence strategy and ecoregions [7]. Indeed this approach can be seen as a major innovation in multivariate analysis for factors affecting gene frequencies in human populations. No portrait of environmental selection pressures would be complete without the inclusion of disease status and pathogen density. The interplay of nutrition, location, and exposure to pathogens are compelling external forces that impact individual survival.

By examining correlations between environmental and disease conditions with allele frequencies, we will be able to search for allele frequencies consistent with signatures of selection on multi-locus traits [7]. This method is useful in identifying adaptations (whether tolerance or resistance focused) for complex infectious diseases. Asian populations provide an opportunity to examine how environmental effects affect complex traits because they exhibit a variety of subsistence strategies, population histories and exposures to infectious disease. They also live in a variety of ecological regions, nutritional contexts and latitudes. These various conditions make them a living laboratory for studies in natural selection. We specifically ask whether selection against infectious disease burden plays a role as an evolutionary driver of addiction genetics.

I hypothesize that environmental factors such as geographical location, relative pathogen burden and infection rates will identify allele frequency differences in ethnic population for immune and addiction gene hotspots. We further hypothesize that these hotspots are consistent with natural selection within human populations living predominantly in tropical environments. This finding would establish a critical link between the agent of natural selection, the genetic process and a complex disease that lies within the close proximity to an allele under selection.

Methods

Identification of immune and addiction gene clusters

To test whether there are correlative relationships between diet, ecoregions, disease and population allele frequencies, we use genes contained in addiction and immune gene step curated by NCBI. Immune associated genes were identified from NCBI gene lists using search terms, 'adaptive immunity,' 'innate immunity,' 'autoimmunity,' and 'Th1/Th2.' These terms were added to 587 genes previously identified as being involved in opiate, dopamine, and GABA reception addiction. These genes were then projected onto the genome to identify cluster regions of genetic importance for immunity and addiction. Clusters were defined as regions of the genome with more than 25 genes within a 2.5Mb linear genomic window.

Annotation of genes in the addiction and immunity windows

To determine the functional role that these NCBI Genes for addiction and immunity as well as candidate genes hotspots play, we performed functional annotation for these three hotspot regions. Genes located within hotspots were considered in two ways in statistical enrichments: all genes in the hotspot window, and only those previously linked to addiction. All genes in the hotspots were annotated using DAVID's Bioinformatics Resources Tool software [8,9] for biological process, molecular function, and KEGG pathways [10,11]. Functional enrichments were quantified using $p<0.05$ and Benjamini score analysis cutoffs of 0.01 [12].

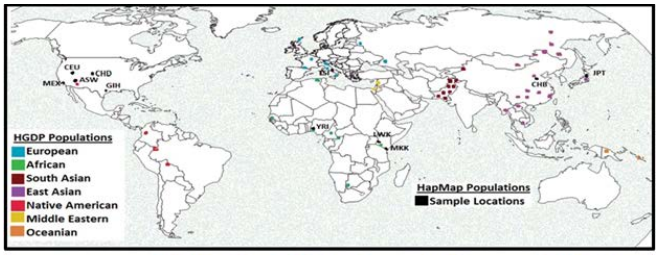
Populations studied and determination of environmental context

In order to understand how populations vary at these key addiction and immunity clusters, we surveyed human polymorphism data from the 1054 individuals comprising 51 populations of the Human Genome Diversity Panel (HGDP), and the 1078 individuals comprising the 11 sample populations of the HapMap project dataset. The set of polymorphism shared amongst these populations was identified and cross population comparisons were made on allele frequencies. **Figure 1** shows the sampled populations.

To determine how populations fit into ecological regions, a map with population sampling locations was cross-referenced to global Bailey's ecoregions maps [13]. This map characterizes the environmental conditions under which study populations live. **Figure 1** also makes a comparison between those populations living in tropical environments as identified by the Bailey's ecoregion map for a global population set. We used sampling locations as the determinant for the locality of the population. For this analysis, the Human Genome Diversity Panel populations were more appropriate to differentiate between those populations Population sample locations were cross-referenced with the World Health Organizations' public health indicators (available online at: http://apps.who.int/gho/data/?theme=main). Disease prevalence and pathogen load data for various infectious diseases including: malaria (*Plasmodium vivax and P. falciparum*), Cholera (*Vibrio cholera*), Polio (*Poliovirus* spp), Schistosomiasis (*Schistosoma japonicum* and *S. mansoni*) and Yellow Fever (*Yellow fever virus*) were obtained from World Health Organization datasets to characterize the epidemiological environments within which population samples live along with additional health indicators which might determine the role that these pathogens might play in determining the in allele frequency variation.

Geographical mapping of correlated variants

To further assess the effect of Hepatitis viral disease on the



	Populations	SNPs typed	N. of Individuals	# of Tropical populations
HapMap Populations	11	~10 million through Phase III	1078	1
Human Genome Diversity Panel	51	650,000	1054	15

Figure 1 Ethnic populations surveyed for immunity and addiction crosstalk from the HapMap and Human Genome Diversity Project. The HapMap populations (populations described in **Table 1**) and the HGDP populations: HGDP Africans (green)- Moazibite, Mandenka, Yoruba, Biaka, Mbuti, San, NE Bantu, and SAf Bantu. HGDP Asians (purple); N. Asia: Oroquen, Daur, N. Han, Hezhen, Japanese, Uygur, Xibo); C. China: Han, Yi, She, Tu; S.E. Asia: Naxi, Lahu, Dai, Miao, Cambodian), S. Asian (burgundy): HGDP Europeans (blue): Adygei, Basque, French, North Italian, Orcadian, Russian, Sardinian, and Tuscan; and HGDP Oceanians (orange): Melanesia, Papua.

Acute Inflammatory Response Pathway, we subsampled the 10 genes most closely associated with this process (AHSG, APCS, C2, CEBPB, CFHR1, KRT1, LBP, MBL2, ORM1, and SIGIRR). These were all part of our immune function NCBI Gene lists. This subset of genes was identified using the Molecular Signatures database housed in the Gene Sets Enrichment Analysis database (available online at: http://www.broadinstitute.org/gsea/index.jsp). Acute Inflammatory Response functions in short lived antigenic challenge as is demonstrated from infectious diseases such as Hepatitis B and Hepatitis C. All single nucleotide polymorphisms (SNPs) typed in the 10 genes were assessed for their allele frequencies following the Hancock methodology [7]. This refers specifically to SNPs typed in the HGDP populations using the Illumina 650Y platform. SNPs were filtered to exclude those that had minor allele frequencies that fell above 0.90 and below 0.10. These sites were then sorted by geographical region of sample origin into major sub continental regions (Sub-Saharan Africa, Central Asia, Northern Asia, Central China, Southern Asia, and Oceania).

To test whether there is a correlation between Hepatitis B, acute inflammatory response and geography, a Pearson correlation analysis was performed between the variables of mean acute inflammatory response frequency and disease [14]. Candidate and neutral SNP sets were combined and a Spearman Rank Correlation was performed to determine whether candidate SNPs were found to be statistical outliers.

Results

Addiction and immunity hotspots

Our analyses identified three hotspots containing genes from both addiction and immunity NCBI Gene lists. We note that these three genomic hotspots were identified despite the lack of intersection of gene sets between the addiction and immunity. Addiction and immunity hotspots were located on chromosomes 11, 17 and 19. **Tables 1 and 2** identifies the genes in this study with NCBI identified genes bolded and those gene names were color coded to represent the category of gene list with which these genes were initially identified. Hotspots all shared genes from immunity and addiction gene sets. The chromosome 11 hotspot contained 10 genes from autoimmune, cocaine, alcohol and innate immunity lists along with seven genes previously unidentified as participating in addiction and immunity phenotypes. The chromosome 17 hotspot contains 11 genes from innate immunity, alcohol, morphine, and Th1 along with 44 previously unassociated genes. And finally the chromosome 19 hotspot locus contains 18 genes from alcohol, GHB, autoimmune, innate immunity and Th2.

Addiction and immunity hotspot annotation

We used David's Functional Annotation web tool to annotate the genes that were identified in each hotspot and to determine

Table 1: Genes identified at the intersection of Addiction and Immunity. Three hotspot regions were identified located on chromosome 17, 11, and 19. Each hotspot had genes identified through multiple addiction and immunity gene lists.

Chr	Location	Genes in Window
11	46,406,640-47,606,115	CREB3L1, DGKZ, MDK, CHRM4, AMBRA1, HARBI1, ATG13, ARHGAP1, ZNF408, SNORD67, F2 (autoimmune and innate), LRP4, C11orf49, CKAPS, ARFGAP2, ACP2 (Autoimmune and alcohol), NR1H3
17	39,393,369-41,277,468	KRT16, KRT42P, EIF1, GAST, HAP1, JUP, LEPREL4, KLHL10, FKBP10, ACLY, TTC25, CNP, DNAJC7, NKIRAS2, ZNF385C, DHX58, KAT2A, HSPB9, RAB5C, HCRT, GHDC, STAT5B, STAT5A (TH2 and Autoimmune), STAT3, PTRF, ATP6V0A1, NAGLU, HSD17B1, MLX, COASY, PSMC3IP, FAM134C, TUBG1, TUBG2, CCR10, PLEKHH3, CNTNAP1, EZH1, RAMP2, VPS25, WNK4, CNTDN1, BECN1, PSME3, AOC3, AOC2, G6PC, AARSD1, PTGES3L, RPL27, IF135, RUNDC1, VAT1, RND2, BRCA1
19	47,870,466-49,085,208	SULT2A1, BSPH1, ELSPBP1, CABP5, PLA2G4C, LIG1, CARD8, ZNF114, CCDC114, TMEM143, EMP3, SYGR4, KDELR1, GRIN2D (alcohol and cocaine), GRWD1, KCNJ14, CYTH2, SULT2B, FAM83E, SPACA4, RPL18, SPHK2, DBP, CA11, NTN5, FUT2, MAMSTR, RASIP1, IZUMO1, FGF21 PLEKHA4, PPP1R15A, TULP2, BCAT2, HSD17B14, DHDH, BAX, FTL, GYS1, RUVBL2, LHB, CGB, CGB2, CGB1, CGB5, CGB8, CGB7, NTF4, KCNA7, SNRNP70, LIN7B , PPFIA3, HRC, TRPM4, SLC6A16, CD37, TEAD2, DKKL1, CCDC155, ALD16A1, SLC17A7, PIH1D1, FLT3LG (Th1 and Adaptive), RPL13A, SNORD32A, RPS11, FCGRT, RCN3, NOSIP, PRRG2, RRAS, PRR12, SCAF1, IRF3

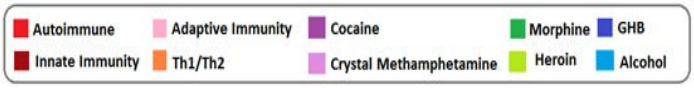


Table 2: Functional Annotation of Addiction and Immunity Hotspot Crosstalk Regions. Genes identified in each hotspot were annotated from biological process, molecular function and pathway participation. Annotation was undertaken using DAVID functional annotation with a p-value cutoff of 0.05.

	Biological Processes		Molecular Function		KEGG Pathways
CHR		p		p	
17	JAK-STAT Cascade	5.4^{-5}	Phosphate binding	0.01	Acute myeloid leukemia
	Growth hormone receptor signaling pathway	2.5^{-4}	GTPase activity	3.6^{-3}	Adipocytokine signaling pathway
	Response to growth hormone	6.6^{-3}	Steroid hormone receptor binding	5.8^{-3}	-
	Homeostatic process	0.01	Ion binding	0.04	-
	Eating behavior	0.02			-
11	Catalytic activity	0.01	Nucleotide regulator activity	0.01	-Regulation of actin cytoskeleton -Neuroactive ligand-receptor interaction -Cholinergic synapse -Hepatitis B -Complement and coagulation cascades -Peroxisome -Hypertrophic cardiomyopathy -Nucleotide excision repair -Dilated cardiomyopathy -Alcoholism
	Protein import	0.02	Enzyme activator activity	3.0^{-3}	
	Regulation of cellular protein metabolism	0.04	-	-	
	Negative regulation of endocytosis	0.02	-	-	
	Regulation of phosphorylation	0.04	-	-	
19	Fertilization	3.4^{-3}	Hormone activity	6.6^{-4}	-Ribosome
	Neurotransmitter transport	4.0^{-3}	-	-	-
	Cell-cell signaling	7.7^{-3}			
	Single fertilization	0.02			

the role that each hotspot played in addiction and immunity disorders. The chromosome 19 hotspot is located between 47.8 and 49 Mb of the Hg18 build of the human genome.

Annotation of tropical identified SNPs

We further annotated the top five polymorphisms that showed allele frequency differences between populations that lived in temperate environments and those living in tropical environments as illustrated in **Figure 1.** These five SNPs were identified in the HGDP populations: rs11818969, rs17790342, rs12417519, rs752849, and rs901746. Annotation of these five SNPS varying showed that they exhibit addiction, mental health and immune function. These SNPs are further characterized in **Table 3**. Additionally we typed these five polymorphisms in the 11 HapMap populations to understand whether these populations showed similar patterns to the HGDP population datasets. Our analyses confirm that they HapMap and HGDP populations show allelic congruence at these five sights, suggesting this pattern is a true representation of what is happening in human populations.

The rs11818969 polymorphism found in the autophagy/beclin-1 regulator 1 gene (AMBRA1) is intimately involved in the development of the nervous system [15]. It has been shown to be a component of a complex network between autophagy, cell growth and cell death of crucial importance to neural development [16]. In particular, the rs11818969 polymorphisms have been identified as Schizophrenia associated polymorphism [17]. A follow up study of this polymorphism has extended this work to find that it specifically alters impulsivity-related behavioral and neural traits [18]. We note that the AMBRA2 gene is more than four times more prevalent in African populations than in Europeans and more than 15 times more prevalent than in East Asian populations of the HapMap. In an African American HapMap sample collected in the south-western United States (ASW) the frequency was found to be 0.592.

There were three polymorphisms, rs17790342, rs12417519, and rs752849, which are located in the Chromosome 11 open reading

Table 3: Top five SNPs have shared patterns of allele frequency differences in HapMap populations as in HGDP populations. SNPs are identified with their location type and gene. We then typed these genes in the HapMap populations and found congruence in the tropical populations of the HapMap and those of the HGDP.

					Africans				Europeans		East Asian		
SNP s	Chr	Location	Gene	Type	LWK	MKK	YRI	ASW	CEU	TSI	CHB	CHD	JPT
rs11819869	11	46500680	AMBRA1	Intron	0.644	0.451	0.566	0.592	0.159	0.153	0.049	0.041	0.047
rs17790342	11	47067706	C11orf49	Intron	0.089	0.091	0.062	0.173	0.088	0.091	0.037	0.047	0.041
rs12417519	11	47069397	C11orf49	Intron	0.828	0.818	0.845	0.837	0.781	0.812	0.354	0.333	0.314
rs752849	11	47115327	C11orf49	Intron	0.533	0.563	0.442	0.49	0.775	0.727	0.927	0.935	0.93
rs901746	11	47200319	DDB2	Intron	0.483	0.486	0.606	0.571	0.301	0.222	0.744	0.753	0.738

frame 49 (C11orf49). While open reading frame regions in the genome are not well characterized, these particular SNPs have been identified as participating in the liver interactome [19]. The liver is the site of metabolism of xenobiotics so these SNPs may have metabolic function.

Finally the rs901746 polymorphism is located in the Damage-Specific DNA Binding Protein 2 (DDB2). DDB2 is involved in the repair of UV damage to DNA. The DDB2 gene participates in a complex that mediates the ubiquitylation of histones H3 and H4, which facilitates the cellular response to DNA damage. Additionally, the DDB2 genes have been implicated in lung cancer susceptibility [20] and most importantly in the destabilized the Hepatitis B viral protein X [21]. This destabilization of the viral protein X is thought to be a key component in the prevention of viral particle proliferation.

We used GWAS3D software to look for local and long range interactions identified from genome wide association studies, regulatory variation. In particular it is good for characterizing noncoding phenotypically associated variants that underlie the molecular mechanisms of complex traits. Our analysis identified three long range interactions with AMBRA1 **(Figure 2).** The first occurs between the AMBRA1 gene, and 2 genes located in a chromosome 8 interaction window: ZNF705D is a zinc finger protein that is thought to be involved in transcriptional regulation, while FAM66D, is a human specific gene that is known to interact with Tetrachlorodibenzodioxin. According to the Comparative Toxicogenomics Database, Tetrachlorodibenzodioxin has more than 1000 documented interactions with AHR (N=3255 Interactions) and CYP1A19 (N=1795 Interactions) and is also used as a treatment for drug-induced liver damage.

The second region with AMBRA1 interactions is a 10Kb chr3 interaction region (chr3:136520001-136530000) that lies just upstream of the SLC35G2 gene. While this gene is not well annotated, there are 13 ENCODE-identified different transcription factors found bundled together (FOS, KAP1, JUN, MEF2, NFIC, BATF, ATF2, USF1, USF2, CTCF, GATA3, and RUNX3). The third regional interaction is a local one located between the AMBRA1 gene and a local chr11 region (chr11:46180001-46190000). There is one polymorphism identified as rs7128538 which has been associated with Systemic Sclerosis [22].

Two regions were shown to have strong enhancer signatures: chr11:46046791-46539727 and chr11:46446962-46516078.

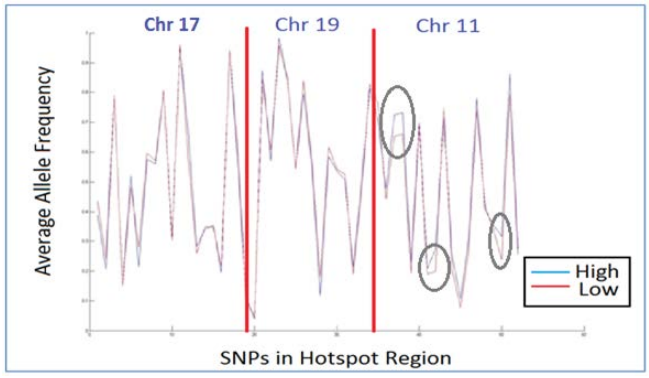


Figure 2 Five SNPs underlying the Chr 11 hotspot vary between high tropical versus low tropical living HGDP populations. Allele frequencies were plotted in the tropical living climates (blue) versus those living in temperate climates (Red). The inset shows the hotspot locations.

These signatures are generated from the ENCODE analysis of H3K4me1, H3K27ac, P300, and DHS. These enhancers represent epigenetic enhancers of on genomic sequences. Finally there was one identified Conservation Region of GERP++ Elements located at chr11:46499803-46500030. GERP++ Elements are constrained elements in multiple alignments and are estimates potential functional constraint. These are summarized in **Table 4**.

Acute inflammatory response and hepatitis B

Following up on this observation that rs901746 is a polymorphism involved in Hepatitis B prevention, and our hypothesis that infectious disease might be the driving force shaping allele frequencies at adjacent addiction and mental health sites, we wanted to identify whether we could identify if there was a correlation between Hepatitis B infection prevalence in a global set of human populations and these populations. When we studied the 10 genes most closely associated with acute inflammation in HGDP Africans (Africa- Mandenka, Yoruba, Biaka, Mbuti, San, NE Bantu, SAf Bantu), Asians (N. Asia: Oroquen, Daur, N Han, Hezhen, Japanese, Uygur, Xibo); C. China: Han, Yi, She, Tu; S. Asia: Naxi, Lahu, Dai, Miao, Cambodian) and Oceanians (Oceania: Melanesia, Papua),we found that there was a trend towards significance (p=0.08) for alleles correlation to hepatitis B prevalence distributions described from sentinel surveillance conducted in 2004 [23].

Table 4: GWAS3D analysis of the tropical segregating polymorphisms showed interacting SNPs, transcriptions factors and ACP2, a gene involved in de-phosphorylation. We characterized the set of SNPs that were frequency divergent in tropical and non-tropical living populations. The set had interacting SNPs that were locally located, often in the same gene with the exception of rs901746 which interacted with rs2167079 at ACP2, an adjacent gene. Three of the interacting SNPs sat within transcription factors, with the ACP2 gene SNP (rs2167079) sat at the intersection of 62 different transcription factors.

dbSNP ID	Gene	SNP Functional Annotation	Interacting dbSNP ID	CHR: Location	Gene	Location	Transcription Factors?
rs11819869	AMBRA1	Schizophrenia associated	rs7130141	11:46499874	AMBRA1	Intronic	Yes- EBF1
rs17790342	C11orf49	Liver interactome	rs12576831	11:47082255	C11orf49	Intronic	No
rs12417519	C11orf49	Liver interactome	rs11601798	11:47158392	C11orf49	Intronic	No
rs752849	C11orf49	Liver Interactome	rs7940473	11:47182353	C11orf49	Intronic	Yes- CTCF
rs901746	DDB2	Inhibition of Hepatitis B Protein X	rs2167079	11:47270255	ACP2	Coding	Yes- 62 TFs

Mean allele frequencies were calculated within an ethnic group sample and then across a geographical region **(Figures 3 and 4).** These candidate regions were compared to neutral SNPs identified using HOMINID coordinates representing 71 regions of the human genome that are far from genes/motifs and are thought to be consistent with neutral evolutionary processes. On the x axis-the mean minor allele frequency (MAF) for African populations was 0.3418, while the mean pan Asian MAF mean was 0.3781. This was significantly different (p=0.00436). When Asian populations were grouped regionally, Central Chinese sample locations (Han, Yi, She, and Tu) had a MAF mean almost identical to African populations (0.3428), and the central Chinese population differed significantly from N. Asia and S. Asia regional means (p=0.0045 and p=0.0022, respectively) using the Student's T-test statistic to determine significance.

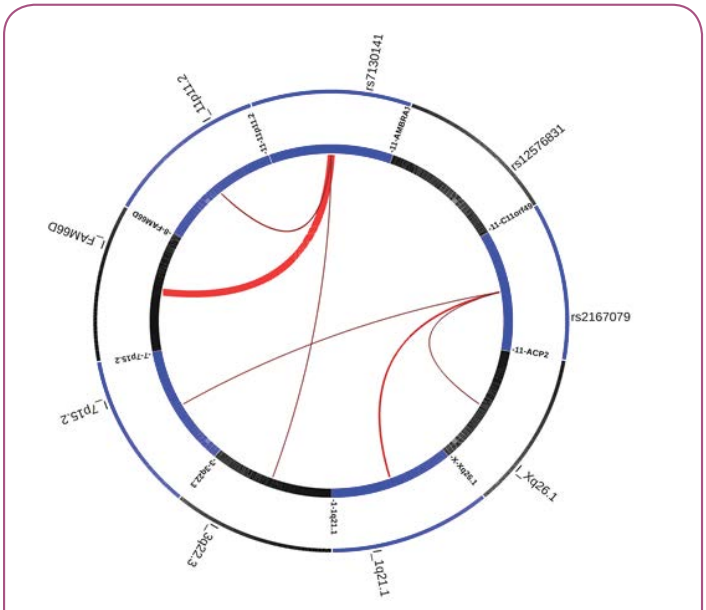


Figure 3 GWAS3D analysis of the five single nucleotide polymorphisms found to differ in tropical versus non-tropical living populations show that SNPs have local and long range interactions. AMBRA1 showed 2 long range interactions on chromosomes 3 and 8 along with a local interaction on chromosome 11. The ACP2 gene (rs2167079), an interacting partner of rs901746, had three interactions: at Xq26.1, 1q21.1, and 7p15.2.

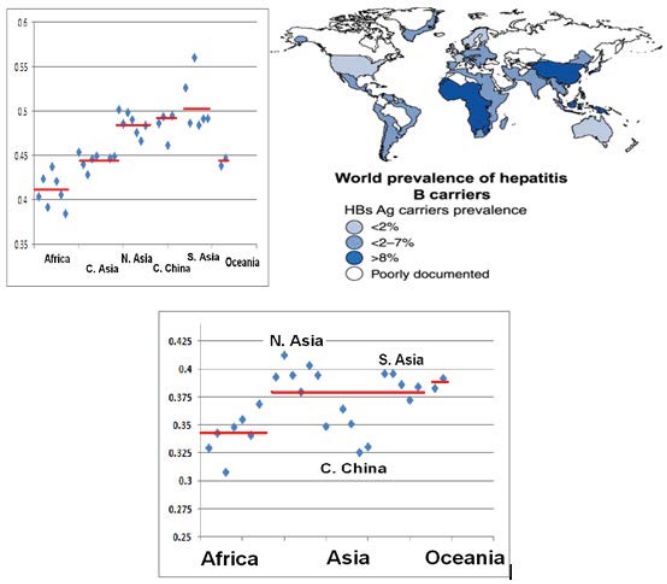


Figure 4 Neutral SNPs in global human populations and the Acute Inflammatory response gene set surveyed in a diverse set of African and East Asian populations. Neutral SNPs only show significant differences between Africans and non-African populations. A Pearson correlation of Hepatitis B prevalence and mean acute inflammatory response allele frequency shows a trend toward correlation (p=0.08).

To test whether there is a correlation between Hepatitis B, acute inflammatory response and geography, we performed a Pearson correlation analysis between the variables of mean acute inflammatory response frequency and disease [14]. This result is consistent with the central Chinese region being a nexus of Hepatitis B transmission. The pattern observed in genes of the acute inflammatory response pathway is consistent with the observation that acute inflammation is an integral part of the Hepatitis B infection process [19]. When HGDP sampling locations are overlaid with Hepatitis B prevalence maps, there is trend towards congruence between high Hepatitis B prevalence and low MAF frequencies [23].

Discussion

Our findings show that genes involved in opiate, dopamine and GABA addiction do form three genomic hotspot regions

in conjunction with immunity regulating genes within the human genome. These positions are located on three separate chromosomes: chr 11, chr 17, and chr19. Functional annotation of these joint addiction and immunity hotspot regions confirmed that the genes previously identified in either addiction or immune surveys share broad functional classifications with those candidate genes that are their hotspot neighbours. Finally when we survey genetic polymorphisms that underlie these three genomic regions in a global distribution of human populations we find that when populations are grouped by their locality in tropical ecological zones, we identified polymorphisms that significantly differ at the chromosome 11 hotspot region. Further annotation of these climate and frequency divergent polymorphisms showed that they have been identified as playing key roles in liver interactome function, schizophrenia and Hepatitis B infection. A follow up analysis of the relationship between Hepatitis B and the acute inflammation process points to a correlation between Hepatitis B and acute inflammation pathways in African and Asian populations living in those areas of the continent that have high hepatitis burden. Taken together, our results point to a functional trifecta existing between immunity, addiction and mental health related SNPs at the chromosome 11 locus. We can infer that hepatitis B is at least a contributing factor in the addiction and mental health allele frequencies that are seen in tropical living populations.

When these addiction and immunity genomic hotspots were functionally annotated using the David's functional annotation web tool [8,9,24,25], we found that while hotspots were not well annotated with common computational tools, the genes underlying these regions represent candidates for both mental health and infectious disease genes. Interestingly, alcoholism was one of the KEGG pathways identified as participating in the metabolic dynamics of these genomic regions. This gives additional support to our conjecture that the SNPs for C11orf49 may indeed play a role in the liver interactome in a manner that potentially creates susceptibility to alcoholic substances.

We considered the relationship between tropical-living populations, as a climatological surrogate for multiple pathogen loads. These analyses determined that those populations living in highly tropical environments showed the highest polymorphism frequency differences to their temperate climate-living population comparisons. This was true when populations were considered without respect to ethnic origin, confirming that human population locality and more specifically proximity to pathogens was sufficient to explain the observed allele frequency differences. This finding further supports the idea that local climate and specifically potential disease burden appear to play a significant role in shaping both immunological but also co-located non-immunological phenotypes.

Previous studies examining the role of climate-either as a surrogate for pathogen load [26] or as a function of climate change [27] have shown that local climate affects the availability of the pathogen substrates that are thought to drive natural selection in human populations. It is therefore not surprising that selection for immune fitness can lead to hitchhiking of addiction associated SNPs. The proximity of these polymorphisms indicates that their evolutionary histories are intertwined. This concept now gives us a fertile ground on which to build further analyses between the addiction alleles and immune-regulatory elements under infectious disease based selection. Furthermore, there is evidence that opioid peptides may have first arisen as modulators of cellular immune function- where morphine down regulates immune processes in addiction, an action/function that it appears to normally perform [2]. This strengthens our argument that these putatively adaptive features that increased human fitness in disease rich environments may now be causing secondary effects when the immune pressure is removed, or when xenobiotics such as the addictive substances we study here are given in non-homeostatic doses.

Conclusion

Our assessments of GWAS3D interacting polymorphisms showed that the five polymorphisms identified have both local and long distance interactions. These interactions include rs2167079, a coding SNP that sits in at least 62 transcription factors. These analyses demonstrate that addiction genes do indeed form genomic hotspots with immune genes, which these immune hotspots share functional cohesiveness, and that when SNPs at these addiction and immunity hotspots were compared between HGDP populations living in tropical versus temperate climates, they identified SNPs associated with hepatitis B, the liver interactome, and schizophrenia. A follow on analysis of a subset of immune genes involved in the acute inflammatory response show congruence with reported hepatitis B endemicity geographic distributions in central Asia and sub Saharan Africa. This analysis begs follow up in understanding how pervasive this phenomenon is in complex disorders and their associated immune responses.

References

1. Haldane JB (1956) Natural selection in man. Acta Genet Stat Med 6: 321-332.
2. Stefano GB, Fricchione G, Goumon Y, Esch T (2005) Pain, immunity, opiate and opioid compounds and health. Med Sci Monit 11: MS47-MS53.
3. Zhang R, Miao Q, Wang C, Zhao R, Li W, et al. (2013) Genome-wide DNA methylation analysis in alcohol dependence. Addict Biol 18: 392-403.
4. Alonzo NC, Bayer BM (2002) Opioids, immunology, and host defenses of intravenous drug abusers. Infect Dis Clin North Am 16: 553-569.
5. Lamason RL, Mohideen MA, Mest JR, Wong AC, Norton HL, et al. (2005) SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. Science 310: 1782-1786.
6. Wood ET, Stover DA, Slatkin M, Nachman MW, Hammer MF (2005) The beta -globin recombinational hotspot reduces the effects of strong selection around HbC, a recently arisen mutation providing resistance to malaria. Am J Hum Genet 77: 637-642.
7. Hancock AM, Witonsky DB, Gordon AS, Eshel G, Pritchard JK,

et al. (2008) Adaptations to climate in candidate genes for common metabolic disorders. PLoS genetics 4: e32.

8. Huang DW, Sherman BT, Zheng X, Yang J, Imamichi T, et al. (2009) Extracting biological meaning from large gene lists with DAVID. Curr Protocols Bioinform 27: 1-13.
9. Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4: 44-57.
10. Kanehisa M (2002) The KEGG database. Novartis foundation symposium 247: 91-101.
11. Kanehisa M, Goto S, Kawashima S, Nakaya A (2002) The KEGG databases at GenomeNet. Nucleic acids Res 30: 42-46.
12. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I (2001) Controlling the false discovery rate in behavior genetics research. Behav Brain Res 125: 279-284.
13. Bailey RG (2004) Identifying ecoregion boundaries. Environ Manage 34 Suppl 1: S14-S26.
14. Babor F, Grund S, Siepermann M, Oommen PT, Kuhlen M, et al. (2012) Epidemiology and clinical characteristics of pandemic (H1N1) 2009 influenza infection in pediatric hemato-oncology and hematopoietic stem cell transplantation patients. Transpl Infect Dis 14: 589-594.
15. Fimia GM, Stoykova A, Romagnoli A, Giunta L, Di Bartolomeo S, et al. (2007) Ambra1 regulates autophagy and development of the nervous system. Nature 447: 1121-1125.
16. Behrends C, Sowa ME, Gygi SP, Harper JW (2010) Network organization of the human autophagy system. Nature 466: 68-76.
17. Rietschel M, Treutlein J (2013) The genetics of alcohol dependence. Ann N Y Acad Sci 1282: 39-70.
18. Heinrich A, Nees F, Lourdusamy A, Tzschoppe J, Meier S, et al. (2013) From gene to brain to behavior: schizophrenia-associated variation in AMBRA1 alters impulsivity-related traits. Eur J Neurol 38: 2941-2945.
19. Wang J, Huo K, Ma L, Tang L, Li D, et al. (2011) Toward an understanding of the protein interaction network of the human liver. Mol sys biol 7: 536.
20. Hu Z, Zhang Z, Kim JW, Huang Y, Liang TJ (2006) Altered proteolysis and global gene expression in hepatitis B virus X transgenic mouse liver. Br J Virol 80: 1405-1413.
21. Bergametti F, Sitterlin D, Transy C (2002) Turnover of hepatitis B virus X protein is regulated by damaged DNA-binding complex. Br J Virol 76: 6495-6501.
22. Allanore Y, Saad M, Dieude P, Avouac J, Distler JH, et al. (2011) Genome-wide scan identifies TNIP1, PSORS1C1, and RHOB as novel risk loci for systemic sclerosis. PLoS genetics 7: e1002091.
23. Lavanchy D (2004) Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. J Viral Hepat 11: 97-107.
24. Huang da W, Sherman BT, Tan Q, Collins JR, Alvord WG, et al. (2007) The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol 8: R183.
25. Huang DW, Sherman BT, Tan Q, Kir J, Liu D, et al. (2007) DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nuc Acid Res 35: W169-W175.
26. Graczyk TK, Lucy FE, Tamang L, Mashinski Y, Broaders MA, et al. (2009) Propagation of human enteropathogens in constructed horizontal wetlands used for tertiary wastewater treatment. Applied and environmental microbiology 75: 4531-4538.
27. Mills JN, Gage KL, Khan AS (2010) Potential influence of climate change on vector-borne and zoonotic diseases: a review and proposed research plan. Environmental health perspectives 118: 1507-1514.