UNIVERSITY OF CALIFORNIA, SAN DIEGO

Systems Analysis of Model Organisms in the Study of Human Disease Phenotypes

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Bioinformatics and Systems Biology

by

Merril Joy Gersten

Committee in charge:

Professor Shankar Subramaniam, Chair
Professor Gabriel G. Haddad, Co-Chair
Professor Christopher K. Glass
Professor Trey Ideker
Professor Wei Wang

2011

The Dissertation of Merril Joy Gersten is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____
Co-Chair

_____
Chair

University of California, San Diego

2011

DEDICATION

To my parents

Who inspired me

To my children

Whom I try to inspire

And to my husband

Who makes all things possible

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SUPPLEMENTAL FIGURES

LIST OF SUPPLEMENTAL TABLES

ACKNOWLEDGEMENTS

I would first like to thank my advisor, Professor Shankar Subramaniam, who encouraged me to embark on a second career in science, and who has been a source of guidance, inspiration, insight and support throughout my Ph.D. work. I would also like to thank Professor Gabriel Haddad who graciously agreed to become my co-advisor, welcoming me into his laboratory to work on the Hypoxia Tolerance project. I am indebted to Professor Trey Ideker, in whose laboratory I completed the SIVE project, and who provided me with a unique environment in which to learn Systems Biology. The other members of my committee, Professors Christopher Glass and Wei Wang, have been generous with their time, insights and advice, for which I am deeply appreciative. I would also like to thank Professor Howard Fox, in whose previous laboratory at The Scripps Research Institute I conducted the experimental work for the SIVE project.

I am grateful for help and scientific discussions provided by friends and colleagues I have worked with, especially Dr. Timothy Ravasi, Dr. Silpa Suthram, Dr. Han-Yu Chuang, Dr. Sourav Bandyopadhyay, and Samad Lotia (Ideker Lab); Dr. Mehrdad Alirezaei, Dr. Maria Cecilia Garibaldi Marcondes, and Claudia Flynn (Fox Lab); Dr. Dan Zhou, Dr. Priti Azad, Orit Gavrialov, and Mary Hsiao (Haddad Lab); Dr. Ashok Dinasarapu, Dr. Brian Saunders, Dr. Shakti Gupta, Dr. Pang Ko, and Eugene Ke (Subramaniam Lab). I would also like to acknowledge fellow students in the Bioinformatics and Bioengineering programs, especially Dr. Mary Pacold,

Tejaswini Narayanan, Stan Luban and Vipul Bhargava, and my friends, particularly Sally Cooley and Patricia Kadosch.

Finally, I would like to thank my husband and children for their unwavering patience and support, which have been invaluable in helping me reach this goal.

Chapter 2, in full, is a re-editing of the materials published in Gersten, M., Alirezaei, M., Marcondes, M.C., Flynn, C., Ravasi, T., Ideker, T. & Fox, H.S. An integrated systems analysis implicates EGR1 downregulation in simian immunodeficiency virus encephalitis-induced neural dysfunction. J Neurosci 29, 12467-76 (2009). The dissertation author was the primary investigator and author of this paper.

Chapter 3, in part, is a re-editing of materials published in Zhou, D., Udpa, N., Gersten, M., Visk, D.W., Bashir, A., Xue, J., Frazer, K.A., Posakony, J.W., Subramaniam, S., Bafna, V. & Haddad, G.G. Experimental selection of hypoxia-tolerant *Drosophila melanogaster*. Proc Natl Acad Sci U S A 108, 2349-54 (2011). The dissertation author was a co-second author responsible for analyzing data and contributing to the writing of the paper. Chapter 3 is also, in part, a re-editing of materials currently being prepared for submission for publication: Gersten, M., Zhou, D., Haddad, G.G., & Subramaniam, S. Wnt pathway activation increases hypoxia tolerance in *Drosophila melanogaster*, **in preparation**. The dissertation author is the primary investigator and author of this paper.

Chapter 4 is, in part, a re-editing of materials currently being prepared for submission for publication: Gersten, M., Zhou, D., Haddad, G.G., & Subramaniam, S. Wnt pathway activation increases hypoxia tolerance in *Drosophila melanogaster*, **in preparation**. The dissertation author is the primary investigator and author of this paper.

Chapter 5 is, in part, a re-editing of materials currently being prepared for submission for publication: Gersten, M., Zhou, D., Haddad, G.G., & Subramaniam, S. Wnt pathway activation increases hypoxia tolerance in *Drosophila melanogaster*, **in preparation**. The dissertation author is the primary investigator and author of this paper.

VITA

1972	Bachelor of Arts	Barnard College
1976	Doctor of Medicine	Cornell University Medical College
2011	Doctor of Philosophy	University of California, San Diego

RECENT PUBLICATIONS

Gersten, M., Zhou, D., Haddad, G.G., & Subramaniam, S. Wnt pathway activation increases hypoxia tolerance in *Drosophila melanogaster*, **in preparation**.

Narayanan, T., **Gersten, M**., Subramaniam, S. & Ananth Grama. Modularity Detection in Protein-Protein Interaction Networks, **submitted**.

Zhou, D., Udpa, N., **Gersten, M**., Visk, D.W., Bashir, A., Xue, J., Frazer, K.A., Posakony, J.W., Subramaniam, S., Bafna, V. & Haddad, G.G. Experimental selection of hypoxia-tolerant *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* **108**, 2349-54 (2011).

Bandyopadhyay, S., Chiang, C.Y., Srivastava, J., **Gersten, M**., White, S., Bell, R., Kurschner, C., Martin, C.H., Smoot, M., Sahasrabudhe, S., Barber, D.L., Chanda, S.K. & Ideker, T. A human MAP kinase interactome. *Nat Methods* **7**, 801-5 (2010).

Gersten, M., Alirezaei, M., Marcondes, M.C., Flynn, C., Ravasi, T., Ideker, T. & Fox, H.S. An integrated systems analysis implicates EGR1 downregulation in simian immunodeficiency virus encephalitis-induced neural dysfunction. *J Neurosci* **29**, 12467-76 (2009).

ABSTRACT OF THE DISSERTATION

Systems Analysis of Model Organisms in the Study of

Human Disease Phenotypes

by

Merril Joy Gersten

Doctor of Philosophy in Bioinformatics and Systems Biology

University of California, San Diego, 2011

Professor Shankar Subramaniam, Chair
Professor Gabriel G. Haddad, Co-Chair

Complex phenotypes are distinguished by the interplay of multiple interacting molecules and pathways. Effective study of these phenotypes requires a comprehensive approach utilizing unbiased, genome-scale datasets, including

transcriptomics, proteomics and genome sequencing. These datasets can be computationally analyzed to identify known pathways and processes likely to contribute to the phenotype and can be integrated to produce models and make testable predictions. The latter analyses are aided by incorporation of proteome-wide protein-protein interaction data, which can be represented as a network and which may permit the identification of novel functional modules that contribute to the pathogenesis or physiology of the complex phenotype.

I used a systems approach to study two complex phenotypes in animal models of the human conditions. In Chapter 2, I looked at rhesus macaques suffering from SIV encephalopathy (SIVE), a model for human HIV-Associated Dementia (HAD). Previous work using a human microarray had identified significant upregulation of inflammatory molecules in monkeys suffering from SIVE but little significant gene downregulation. I integrated gene expression data obtained using a newly available rhesus-specific microarray with a large human protein-protein interaction network I constructed from multiple sources, and then applied a module-finding algorithm to identify modules that discriminated between control and SIVE monkeys. I identified EGR1, which plays a role in memory and learning, as a candidate gene and further work led to a model linking infection-associated downregulation of EGR1 to the cognition deficits seen in HAD.

In Chapters 3, 4, and 5, I investigated hypoxia tolerance in *Drosophila melanogaster* that had been adapted to 4% O_2 over generations of selection at progressively lower oxygen tension. Hypoxia contributes to the morbidity and

mortality of several important human diseases, including myocardial infarction and stroke, and plays a role in the chemo- and radio-resistance of solid tumors. Understanding mechanisms of hypoxia tolerance may help design new therapeutic approaches to these diseases. In Chapter 3, I analyzed the genome sequences of control and hypoxia-tolerant flies, and in Chapter 4, I analyzed gene expression data. Polymorphisms and gene expression changes identified in the hypoxia-tolerant flies both pointed to involvement of the Wnt signaling pathways in acquisition of hypoxia tolerance. In Chapter 5, I confirmed Wnt signaling involvement through experimental studies of Wnt pathway gene overexpression and knockdown.

1. INTRODUCTION

Complex traits or phenotypes, including many common diseases, do not follow simple Mendelian, or monogenic, inheritance patterns. Rather, they are complex, or multifactorial in origin, reflecting a combination of environmental and genetic influences on phenotype expression. Mendelian diseases are most commonly characterized by a single non-synonymous polymorphism in a conserved coding region[1], with the resultant residue change exerting a strong deleterious effect. In contrast, complex traits and diseases are typically associated with multiple additive genetic changes, no single one of which is associated with a strong phenotype[2]. The set of genetic changes may impact multiple genes within a signaling, metabolic or regulatory pathway, as well as multiple interacting pathways; and different individuals, particularly if unrelated, may harbor different subsets of polymorphisms leading to the phenotype[3]. Mendelian diseases have been amenable to study using the reductionist approach which dominated biology during the twentieth century, and which focuses on intensive study of individual biological components. However, complex phenotypes require a more comprehensive approach, capable of performing an integrated analysis of interacting molecules and pathways to understand how perturbations of the *system* lead to trait expression or disease pathogenesis. Such a paradigm has become feasible with the development of 1) high-throughput genome-scale technologies that interrogate, in an unbiased fashion, how DNA sequences and abundances of transcripts, proteins and other molecules are altered in the disease or phenotype; and 2) powerful computational methodologies for analyzing and modeling

the plethora of data produced, and making predictions that can be tested experimentally[4,5].

Systems Biology analysis of complex phenotypes starts with a comprehensive quantitative or qualitative measurement of genes, transcripts, proteins or other molecules in control and phenotypic populations and determines which specific biomolecules differ between the two groups. In my dissertation I analyzed two types of genome-scale data – gene expression and genomic sequencing data.

Gene expression was analyzed using species-specific oligonucleotide microarrays. In this methodology, sequence information is used to synthesize hundreds of thousands of small transcript-complementary single-stranded probes which are bound to specific sites on a solid substrate[6]. RNA from control and disease/experimental groups is purified and each sample is hybridized to a microarray; the strength of the signal at a particular site reflects the abundance of RNA in the sample able to bind to the probe. After background subtraction, array normalization and expression value estimation, statistical methods can be applied to determine which genes are over- or under-expressed in the disease/experimental group relative to control[7]. More recently, a new technique for transcriptomic profiling has been developed, RNA-Seq, which uses deep sequencing to determine transcript abundance and offers the advantages of increased dynamic range and the ability to reveal transcriptional sequence variation[8].

Up until the last several years, the Sanger di-deoxy method of DNA sequencing, first developed in 1977, dominated the field of genome analysis and was

responsible for many sequencing accomplishments, including completion of the human genome [9]. More recently, "next-generation" sequencing technologies have been introduced which produce massive numbers of sequence reads per experiment[10]. Data generated using Illumina's Genome Analyzer was analyzed in this dissertation. In this methodology fragmented, adapter-ligated, single-stranded genomic DNA is bound to the surface of a flow cell channel; millions of individual DNA molecules are then amplified to form dense clusters in each channel of the 8-channel flow cell. Clusters are sequenced in parallel using a "sequencing by synthesis" technology; in each sequencing cycle, A, C, G and T labeled reversible terminator deoxynucleoside triphosphates (dNTP) are added and the complementary dNTP (most often) binds to each nascent nucleic acid chain. The labeling dye is imaged to identify the incorporated base and then cleaved to allow the next round of synthesis and base identification. Each raw read base has an assigned quality score. (http://www.illumina.com/documents/products/techspotlights/techspotlight_sequencing.pdf).

Other kinds of data can also be used to probe differences between control and phenotypic populations. These include proteome and phospho-proteome analyses performed using mass spectrometry[11,12] and genome-wide transcription factor binding using ChIP-Chip or ChIP-Seq[13].

Output from genome-wide comparisons of control and phenotypic populations can be analyzed to identify known pathways[14-16] and processes[17] which are statistically enriched for molecules that distinguish the two groups. This type of first-order

analysis can provide important clues to mechanisms involved in pathogenesis or adaptation when relatively little is known about the phenotype. However, the insights from this type of analysis are limited by gaps in biomolecule annotation and are necessarily constrained to pathways and processes that have already been described. These issues are particularly limiting for the study of phenotypes where much is already understood and identification of more subtle mechanistic effects is sought. One approach to this problem has been to construct gene co-expression networks[18]. Another approach, which, in addition to incorporating prior knowledge may also better accommodate small numbers of replicate samples, involves integration of an orthogonal dataset.

Proteins are the workhorses of the cell, responsible for both regulating and executing biological functions. To accomplish these roles proteins generally work together in pathways or as parts of a molecular complex. Cataloguing the protein-protein interactions in an organism and assembling them into a protein-protein interaction network (PPIN) is another means of addressing the limitations imposed by incomplete pathway/process information. High-throughput techniques have been developed to identify proteome-wide interactions, enabling PPIN construction for several model eukaryotic organisms[19-22], as well as for human[23-25]. In these networks, proteins are represented by nodes and interactions between proteins are denoted by edges, and they can be visualized by a variety of software programs, including Cytoscape[26], which was used in the current work.

The main high-throughput techniques used to identify protein interactions are yeast-two-hybrid[27] (Y2H) and tandem affinity purification combined with mass spectrometry[28] (TAP-MS). Y2H makes use of the fact that the DNA-binding (BD) and transcription activating (AD) domains of a transcription factor (TF) can be expressed separately, but need to be physically close to each other in order to drive transcription. In this technique, bait proteins are fused to the BD and prey proteins are fused to the AD of a transcription factor; binding of the bait and prey protein functionally reconstitutes the TF which then drives expression of a reporter gene. In TAP-MS, a bait protein, fused with a TAP tag, is expressed in a host cell, allowing it to interact with its protein partners. The TAP-tagged protein is then retrieved by affinity purification, in the process co-immunoprecipitating the proteins it has bound, which are then identified using mass spectrometry. Protein interactions gleaned from low-throughput experiments or orthology, as well as other types of interactions (as they become available) can be added to constructed networks; the latter include interactions between a protein kinase and its protein substrates and those between a TF and the genes to whose promoter it binds.

Analysis of PPIN by themselves can provide insights into the cell's overall functional organization and can be used to predict functions of uncharacterized proteins[29,30]. When they are integrated with functional genomic datasets, such as gene expression, they can be analyzed to find connected modules or partitions of the network[31-34] whose expression level distinguishes phenotype from control. These

modules are hypothesized to comprise a functional pathway or complex whose dysregulation may contribute to the phenotype.

In my dissertation, I explore two complex phenotypes using systems analyses of relevant animal models: Human Immunodeficiency Virus (HIV) -Associated Dementia (HAD) in a rhesus macaque model, and hypoxia tolerance in *Drosophila melanogaster*. A recent publication summarized many of the advantages of using model organisms in the study of human diseases, despite advances in human genome sequencing and disease modeling[35]. These include the ability to eliminate genetic heterogeneity, control environmental variables and implement *in vivo* experimental manipulations and validations precluded in humans due to ethical and/or technical considerations.

HAD is the most severe form among a spectrum of HIV-associated neurocognitive disorders (HAND), with patients displaying cognitive, motor and behavioral deficits. Prior to the advent of highly active antiretroviral therapy (HAART), 15-25% of infected individuals manifested severe symptoms. With the availability of effective therapy, the incidence of HAD has decreased, but as a result of increased survival rates, the prevalence of less severe neurologic impairment has increased[36,37]. Simian Immunodeficiency Virus (SIV) infection of rhesus macaque is a well-established model for HIV, and SIV Encephalitis (SIVE) has been particularly useful for the study of HAD. A prior gene expression study of macaque cortex using human microarrays had identified upregulation of interferon-induced genes in the monkeys with SIVE[38,39], which coincides with the major role inflammatory mediators,

released by virus-activated microglia, play in the pathogenesis of HAD[36]. However, imprecise correlation of neurologic impairment with numbers of infected cells in brain[40] points to complexity in the pathogenesis of this chronic infectious disorder. In Chapter 2, I examined gene expression in macaque hippocampus using the (then) recently available Affymetrix rhesus microarray. I was interested in identifying novel, and in particular downregulated molecules, that might play a role in pathogenesis. I integrated multiple sources to compile a large human protein-protein interaction network, integrated the rhesus gene expression results (after mapping Affymetrix rhesus probesets to human genes), and applied a module-finding algorithm[32] designed to identify modules whose gene expression significantly differentiated SIVE from control monkeys. Through this process, I identified EGR1, an early response gene involved in memory and learning[41-43] as a possible candidate gene. Further work demonstrated downregulation of EGR1 in SIV-infected hippocampus *in situ*, and in differentiated human neuroblastoma cells treated with CCL8, a product of activated microglia, leading to the proposal that downregulation of EGR1 in SIVE may occur as a consequence of the host response to infection, leading to deficits in cognition.

Hypoxia refers to oxygen deprivation to a tissue or organism, and plays a key pathogenic role in the outcome of a variety of pathologic conditions, including airway obstruction, stroke, and myocardial infarction and in the increased radio- and chemoresistance of solid tumors. Although all metazoans require oxygen, there is significant variability among species with respect to the degree and duration of

hypoxia that can be tolerated, though low oxygen selection may allow the evolution of individuals with increased tolerance to reduced oxygen supply. An understanding of the mechanisms that facilitate survival at low oxygen tension may provide clues to improved approaches for the treatment of disorders in which hypoxia plays a major or contributory role in morbidity and mortality.

There are several features that distinguish *Drosophila* as a highly useful model organism. Approximately 75% of human disease-related genes have fly orthologs[44,45]. In addition, it has a short generation time, a sequenced genome and many available genetic tools. It has been a key system for the study of developmental processes, vision and signal transduction[44]. It offers particular advantages in studying hypoxia due to the significant tolerance of adult flies to extreme hypoxia and even anoxia[46], and ventilation of its tissues with the ambient atmospheric O_2 tension so that the actual O_2 level in tissue is known[47].

To facilitate study of hypoxia tolerance in flies, the Haddad lab used *D. melanogaster* in a long-term selection strategy to generate a population of flies able to reproduce and thrive at 4% O_2, a level lethal to the original parental lines[48-50]. In Chapter 3, I used a fine-grained approach to analyze genomic sequences from pools of hypoxia-adapted flies and a control pool maintained in parallel. In a first analysis of the data I identified several polymorphisms in Notch signaling pathway-related genes, which supported an earlier microarray study[50] suggesting a contribution of Notch pathway activation to the tolerant phenotype. Network evidence of Notch crosstalk with other polymorphism-containing pathways, and consideration of the reproductive

bottleneck that preceded achievement of tolerance to 4% O_2, suggested that additional adaptive mechanisms were likely needed to achieve this level of tolerance. I reanalyzed the sequencing data using more stringent criteria and identified the Wnt signaling pathway to contain many genes with one or more polymorphisms.

In Chapter 4, I asked whether there was physiological evidence for Wnt pathway involvement in the sequenced flies, and performed a gene expression analysis of flies adapted to 4% O_2. In order to address other related questions, I looked at flies early post-eclosion, as well as 3rd instar larvae and adults, and included populations of adapted flies that were maintained at 4% O_2 or transferred to a normoxic environment. I found that gene expression was primarily altered in flies early post-eclosion that were maintained in hypoxia. In these samples I identified numerous Wnt pathway genes that were differentially expressed, adding support to the gene sequencing results.

In Chapter 5, I sought experimental validation of a role for Wnt signaling in hypoxia tolerance. Using the GAL4-UAS system, I studied overexpression of Wnt pathway activators and knockdown of Wnt pathway inhibitors, to evaluate the impact of Wnt signaling activation on fly eclosion rates in 5% O_2. Results from studying multiple genes in two different cell types, neurons and hemocytes, provided evidence that Wnt activation improves fly tolerance to hypoxia.

2. AN INTEGRATED SYSTEMS ANALYSIS IMPLICATES EGR1 DOWNREGULATION IN SIVE- INDUCED NEURAL DYSFUNCTION

Abstract

HIV-associated dementia (HAD) is a syndrome occurring in HIV-infected patients with advanced disease that likely develops as a result of macrophage and microglial activation as well as other immune events triggered by virus in the central nervous system. The most relevant experimental model of HAD, rhesus macaques exhibiting SIV encephalitis (SIVE), closely reproduces the human disease and has been successfully used to advance our understanding of mechanisms underlying HAD. In this study we integrate gene expression data from uninfected and SIV-infected hippocampus with a human protein interaction network and discover modules of genes whose expression patterns distinguish these two states, to facilitate identification of neuronal genes that may contribute to SIVE/HIV cognitive deficits. Using this approach we identify several downregulated candidate genes and select one, EGR1, a key molecule in hippocampus-related learning and memory, for further study. We show that EGR1 is downregulated in SIV-infected hippocampus and that it can be downregulated in differentiated human neuroblastoma cells by treatment with CCL8, a product of activated microglia. Integration of expression data with protein interaction data to discover discriminatory modules of interacting proteins can be usefully employed to prioritize differentially expressed genes for further study. Investigation of

EGR1, selected in this manner, indicates that its downregulation in SIVE may occur as a consequence of the host response to infection, leading to deficits in cognition.

Introduction

HIV-associated dementia (HAD) is characterized by motor disorders, behavioral changes and cognitive dysfunction, including deficits in learning and memory[51]. Since HIV does not infect neurons, possible pathogenic mechanisms include direct toxicity of HIV proteins and indirect toxicity mediated by HIV-infected or -activated cells. Several HIV proteins have been associated with neurotoxicity[52], however, clinical signs of HAD correlate better with microglial activation than with quantitative measures of brain infection[40]. These observations have suggested that HAD likely develops as a secondary neuronal response to macrophage/microglial activation precipitated by HIV infection[40,51].

SIV infection of rhesus macaques is a well-established model for HIV infection, particularly useful for studying HAD[39,53]. Infected monkeys may develop the same clinical disturbances as people with HAD, and they display similar neuropathologic changes, termed SIV encephalitis (SIVE)[39]. Furthermore, transient depletion of CD8+ cells soon after infection ensures that SIVE develops rapidly in a large proportion of the infected animals[39], thus facilitating investigation.

Our previous microarray studies of infected human and rhesus frontal lobe using the Affymetrix HG_U95Av2 (human) chip detected both upregulated and downregulated human genes[54], and upregulated rhesus genes, particularly interferon-

induced genes[38,39]. Microarray studies, while useful, are subject to the limitations imposed by incomplete knowledge of pathway membership and insensitivity to pathway members not regulated via transcriptional control. One means to address this limitation is to integrate microarrays with other genomic-level data, such as protein interaction networks and to extract modules of physically-interacting genes with coherent expression patterns[31-34] considered likely to reflect known or novel pathways or complexes. In protein interaction networks nodes represent proteins and are connected to their interacting partners by edges, which represent interactions identified through either small-scale or high-throughput experiments.

We used the Affymetrix Rhesus Genome chip to compare hippocampal gene expression in uninfected rhesus macaques versus those with histopathologic evidence of SIVE. We mapped Affymetrix probesets to human gene identifiers and integrated the expression data with a human protein interaction network constructed from publicly available datasets. We were particularly interested to determine whether downregulated genes might be detectable under these conditions. We anticipated that greater efficiency of rhesus gene detection using the rhesus chip (efficiency ~ 70% using the human chip[55]), coupled with use of hippocampal tissue (relatively enriched for neurons) and integration of protein interaction data, would facilitate detection of neuronal genes whose downregulation contributes to HAD. We found that EGR1, an early response gene involved in memory and learning[41-43] was significantly downregulated and present in a differentially expressed network module, and we further investigated how EGR1 may become downregulated in SIVE.

Results

Expression analysis of SIV-infected macaques

Hippocampal RNA from nine uninfected and nine SIV-infected rhesus macaques was hybridized to Affymetrix Rhesus Macaque Genome chips and expression values calculated using the Robust Multichip Average[56] (see Methods). A summary of the treatment history of the monkeys studied is provided in Supplemental Table S2.1. Of 18,313 macaque genes that could be unequivocally mapped to human Entrez Gene IDs by best reciprocal Blast (see Methods), 720 were considered to be significantly upregulated and 106 significantly downregulated (CyberT[57] $p \leq 0.001$, PPDE ~95%; see Methods, Supplemental Table S2.2). Analysis of the differentially expressed genes for known functions and pathway membership using DAVID[58] showed significant enrichment of upregulated genes for KEGG pathways and/or GO Biological Processes relating to immune, inflammatory and stress responses, apoptosis, cell proliferation, and signaling cascades (JAK-STAT, NFκB), but little functional enrichment of downregulated genes.

Identification of transcriptionally active network modules

To facilitate identification of genes whose downregulation may contribute to neuronal dysfunction in HAD, macaque expression data were integrated with a human protein interaction network (10,906 proteins, 81,211 edges) assembled from several publicly available datasets (see Methods). The integrated network was then searched

to identify “discriminatory network modules” using PinnacleZ[32]. “Discriminatory network modules” are connected subnetworks in the protein interaction map whose average mRNA expression level significantly discriminates between two experimental classes; significance is assessed using permutation testing (see Methods). Figure 2.1 provides an overview of the experimental and bioinformatic methodology.

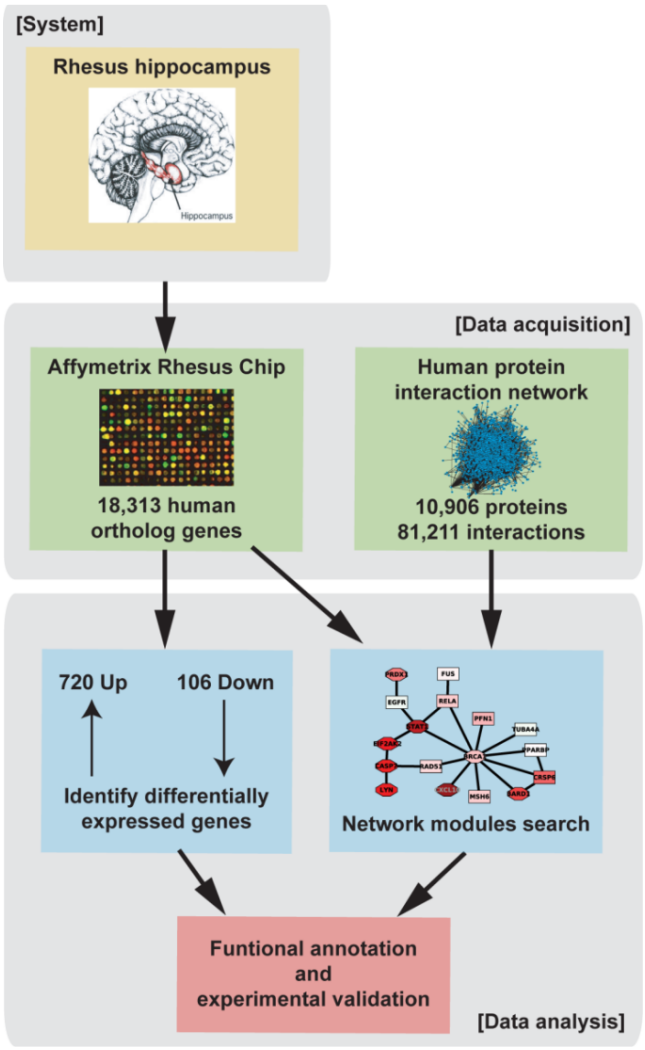


Figure 2.1 Overview of analysis process.
Expression data from uninfected and SIV-infected rhesus hippocampus was integrated with a human protein-protein interaction network assembled from public sources to facilitate identification of novel SIVE-associated gene modules using a greedy search algorithm.

We recovered 447 modules containing a range of 2 to 20 proteins with discriminatory scores significantly better than those of randomly generated modules. Module expression was largely coherent, as modules contained predominantly upregulated (N=429) or downregulated (N=18) proteins in addition to protein products of non-differentially expressed genes. As shown in Supplemental Figure S2.1, the effect of averaging expression values of genes in significant modules is to reduce intensity and increase homogeneity within the experimental group (SIVE or non-SIVE) relative to individual gene expression values. We noted that 31 proteins appeared in 5% or more of the significant modules and 30 of these proteins form a connected module within the protein interaction network (Figure 2.2A). This group of high-frequency proteins is significantly enriched (Benjamini[59] corrected p-value < 0.05; DAVID[58]) for intracellular signaling cascade (N=14), response to stress (N=13), cell cycle (N=10), and apoptosis (N=9), key processes in SIVE. Most of these proteins are either highly connected in the network (>50 interactions; N=22) or correspond to highly differentially expressed genes (p-value <10-5; N=13), or both (N=9), providing a rationale for their frequent selection in significant modules. To reduce redundancy among significant modules containing these high-frequency proteins while maximizing recovery of modules containing significantly downregulated genes, we separated the modules into those containing ≥ 1 downregulated genes and those containing only ≥ 1 upregulated genes, then in each case removed modules showing substantial overlap with higher-scoring modules (Jaccard Index > 0.15). Our final set of 110 modules represents the union of these two filtering procedures and is available

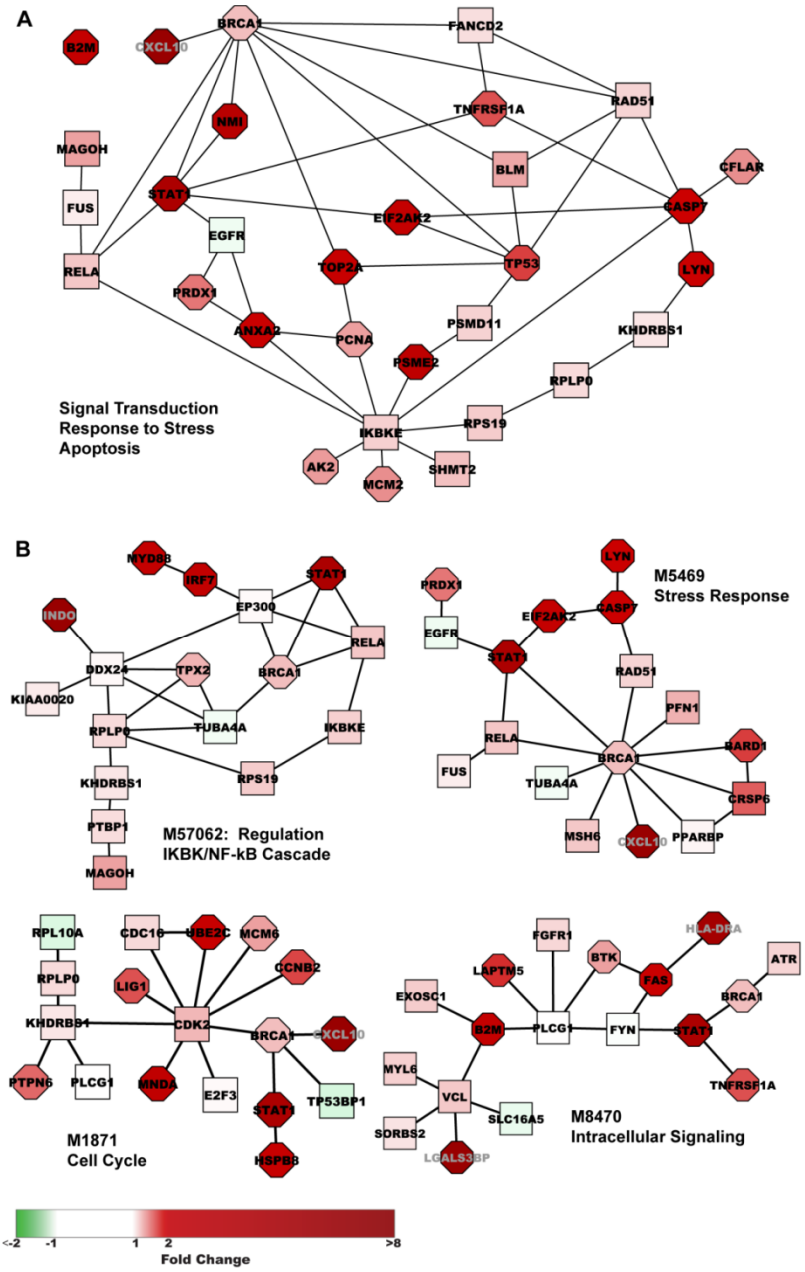


Figure 2.2 Highest scoring significant modules.
(A) 30 of 31 proteins present in ≥ 22 (5%) of the significant modules form a connected module within the PPIN. (B) The four highest scoring filtered modules, M57062, M5469, M1871 and M8470, are up-regulated modules with top GO annotations (BiNGO plugin to Cytoscape) relating to regulation of I-kappaB kinase/NF-kappaB cascade (MYD88, RELA, IKBKE), stress response (CXCL10, PRDX1, EGFR, EIF2AK2, RELA, BRCA1, BARD1, RAD51, MSH6), cell cycle (UBE2C, E2F3, KHDRBS1, CDK2, CDC16, MCM6, CCNB2, LIG1), and intracellular signaling (FAS, ATR, FGFR1, BRCA1, STAT1, FYN, BTK, PLCG1), respectively. Fold-change in gene expression relative to uninfected monkeys is depicted by intensity of node color: red (upregulated) or green (downregulated). Significance of differential expression as determined by Cyber T p-value is represented by node shape: rectangular ($p > 10^{-3}$), octagonal ($\leq 10^{-3}$).

in the CellCircuits Database[60] at http://www.cellcircuits.org/Gersten 2009/. Modules are identified using the Entrez Gene ID of the seed protein for the module (see Methods).

Functional enrichment of active network modules

Figure 2.2B presents the four top-scoring modules, all upregulated, which share STAT1 and BRCA1 but include other proteins yielding top GO annotations relating to regulation of I-kappaB kinase/NF-kappa B cascade (M57062), stress response (M5469), cell cycle (M1871), and intracellular signaling (M8470), respectively (Benjamini[59] corrected $p<0.05$; BiNGO plugin[61] to Cytoscape[26]). Figure 2.3A presents the top-scoring downregulated module, M9644, which is significantly enriched for genes participating in signal transduction. Notably, this module includes three neuronal signaling proteins (CNTNAP1[62], CDK5[63], ITGAV[64]) and two ion channels (KCNB1, KCNMA1). Proteins in this module are connected via the tyrosine kinase, SRC; several of these proteins, or related proteins, have been predicted (CBL, CDK5, PIK3R2; KCNB2, ITGB1,3,4; http://networkin.info/search.php)[65] or demonstrated (SH3PXD2A)[66] to be SRC substrates. Most of the genes are only minimally, if at all, downregulated; nevertheless, expression of the module protein members, taken as a unit, significantly distinguishes between uninfected animals and those with SIVE. The full set of 110 modules contains 749 proteins, including 189 differentially expressed genes (175 upregulated and 14 downregulated), which represents 33% of the 570 differentially expressed genes present in the protein

interaction network. The large number of non-differentially expressed genes in the modules provide connectivity among differentially expressed genes and in some cases may reflect masking of neuronal differential expression as a result of expression in non-neuronal hippocampal cells and/or functionality not regulated via transcription. For example, the activity of SRC (Figure 2.3A) is modulated by phosphorylation[67].

Network module analysis implicates EGR1 in SIVE-induced dementia

Past SIV/HIV microarray studies have focused on upregulated genes which are more striking in their differential expression and hence easier to confirm using small-scale experiments. However, these upregulated genes have predominantly reflected inflammatory responses, whereas many neuron-related functional annotations in a prior human study were associated with downregulated genes[54]. To identify novel HAD-related candidate genes we focused on downregulated genes present in significant modules. Of the 14 genes with significant p-values ≤ 0.001, six had a fold-reduction ≥ 1.5 (~ 33% reduction). We selected four of these genes, which have been

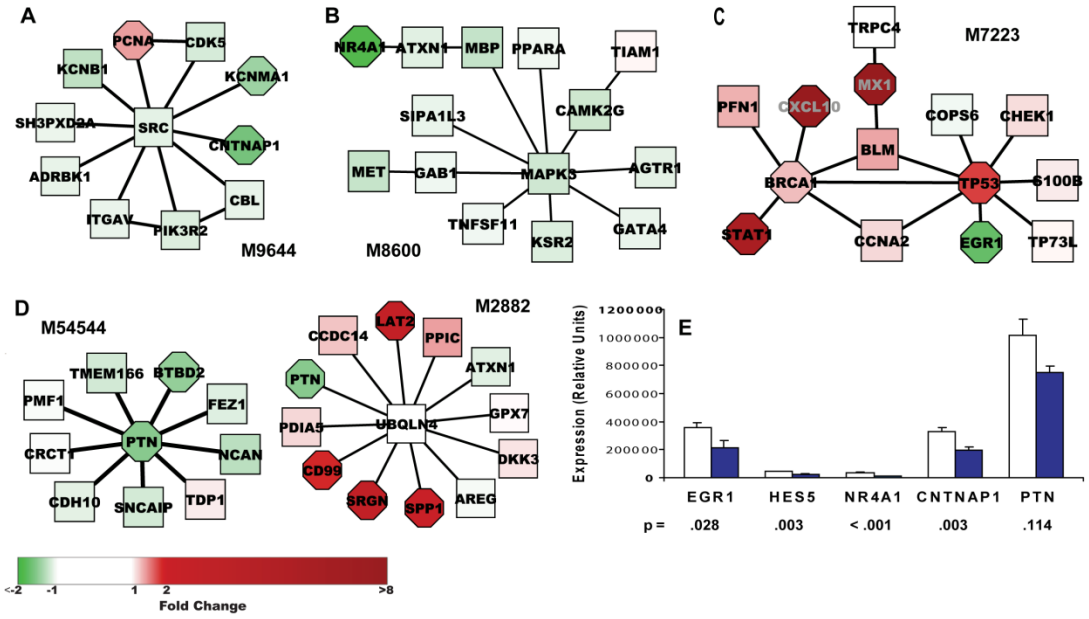


Figure 2.3 Significant modules containing selected downregulated genes and RT-PCR confirmation.
(A-D) Significant modules in which downregulated genes appear. Module M9644 (Fig 2.3A) was the highest scoring downregulated module and is significantly enriched for genes participating in signal transduction (PCNA, ITGAV, ADRBK1, CDK5, CBL, CNTNAP1, PIK3R2, SRC). Fold-change in gene expression relative to uninfected monkeys is depicted by intensity of node color: red (upregulated) or green (downregulated). Significance of differential expression as determined by Cyber T p-value is represented by node shape: rectangular (p >10-3), octagonal (≤10-3). (E) Hippocampal tissue from each of the nine uninfected and nine SIVE monkeys was tested by RT-PCR for four downregulated genes present in significant modules and HES5. Normalization was done using housekeeping genes 18S, TBP and GAPDH. Figure shows mean expression ± SEM in terms of relative units; t-test was performed on log-transformed data. Unfilled bars show control data; solid blue bars show SIVE data.

associated with brain physiology, as well as HES5, a significantly downregulated gene with >2.0 fold reduction not included in a significant module, and performed RT-PCR to confirm the downregulation seen in the microarray data. Figure 2.3A-D presents the modules in which these gene products appear. CNTNAP1 (M9644) is a transmembrane protein which may act as a signaling subunit of contactin[62] and is required, together with contactin, for normal organization of paranodal junctions between glia and axons in myelinated fibers[68]. NR4A1 (M8600) is a nuclear receptor

and an early response gene that is specifically upregulated during memory consolidation[69]. It is connected to the module via ATXN1, whose absence in Sca1 null mice is associated with deficits in learning[70]. EGR1 (see below) is an early response transcription factor critical for certain types of memory and learning[41-43]. PTN (M54544, M2882) is an extracellular matrix-associated protein that promotes neurite outgrowth[71] and that has been implicated in both reduced LTP and faster learning in mice[72]. Among its neighbors in M54544, PTN is connected to SNCAIP, which associates with alpha-synuclein and has been identified as a component of Lewy bodies and other neurodegenerative-associated protein inclusions[73]. As shown in Figure 2.3E, RT-PCR confirmed significant downregulation in SIVE hippocampus for 4 of the 5 genes tested (EGR1, NR4A1, CNTNAP1, and HES5, but not PTN).

We were particularly intrigued by the inclusion of EGR1 (aka Zif268, KROX-24, NGFI-A), a zinc finger transcription factor, among the group of downregulated genes in the discriminatory modules. In addition to being downregulated in the current SIVE hippocampus dataset, it has been found to be downregulated in gene expression studies of experimental models of other neurodegenerative diseases, including Huntington's[74], age-related cognitive impairment[75], and scrapie[76]. As shown in Figure 2.3C, EGR1 appears in a module (M7223) which contains mostly upregulated proteins involved in regulation of transcription, response to stress, programmed death, and cell cycle. EGR1 is linked to the module via TP53, a transcription factor which contributes to HAD via induction of apoptotic pathways and perhaps other more subtle effects[77]. The seed protein for this module, TRPC4, is a member of a family of non-selective

cation channels; it was found to be highly expressed in several brain regions of the adult rodent, including frontal cortex, hippocampus, and dentate gyrus, and may be involved in learning and memory[78]. The association of downregulated EGR1 with other proteins that may impact memory is of interest, given the well-described role of EGR1 in synapse plasticity, memory and learning[41-43], and led us to select EGR1 for further study.

EGR1 expression in neurons

After confirming by RT-PCR downregulation of EGR1 in SIVE hippocampus (Figure 2.3E), we asked if the observed downregulation of EGR1 was occurring in neurons. Tissue from the hippocampus from four of the animals was available for examination of EGR1 protein expression by immunohistochemistry. Examination of the hippocampus revealed a distinctly reduced expression of EGR1 within hippocampal neurons (Figure 2.4). Quantification revealed a significant reduction of EGR1 expression by 42% in SIVE relative to control (average ± SD expression in the CA1 region in control brains 78.24 ± 17.33, in SIVE brains 45.69 ± 7.62, $p=0.014$). Therefore the reduction in EGR1 mRNA in the hippocampus is indeed reflective of reduced protein expression in neurons.

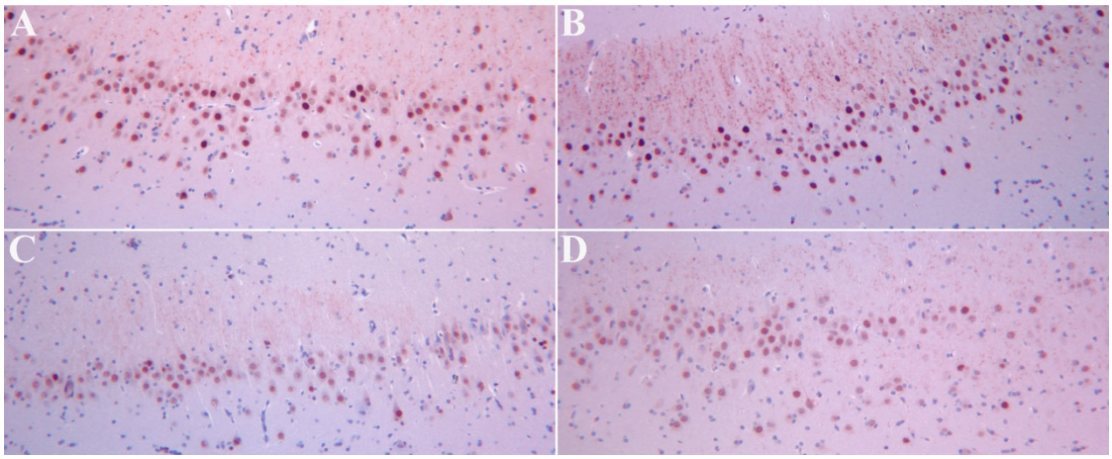


Figure 2.4 EGR1 expression in hippocampal neurons.
Deparaffinized hippocampal sections from control monkeys (A, B) and monkeys with SIVE (C,D) were immunohistochemically stained for EGR1 content by immunohistochemistry. Compared with control sections, staining of sections from SIVE monkeys is less intense and includes fewer heavily stained cells. Quantification (see Methods) revealed a 42% reduction of EGR1 in SIVE relative to control, p=0.014.

CCL8 downregulates EGR-1 in differentiated SK-N-MC cells

We next questioned whether neuronal downregulation of EGR1 might be a result of microglial activation, thereby contributing to impaired cognition. To facilitate experimental manipulation we used SK-N-MC, a human neuroblastoma cell line which can be differentiated toward a neuronal phenotype by exposure to retinoic acid[79]. We first evaluated the effect of supernatant prepared from cultured SIV-infected macrophages (see Methods). As shown in Figure 2.5A, retinoic acid-differentiated SK-N-MC cells showed a decrease in both constitutive and NGF-stimulated EGR1 expression after 2-day pretreatment with a 1:50 dilution of

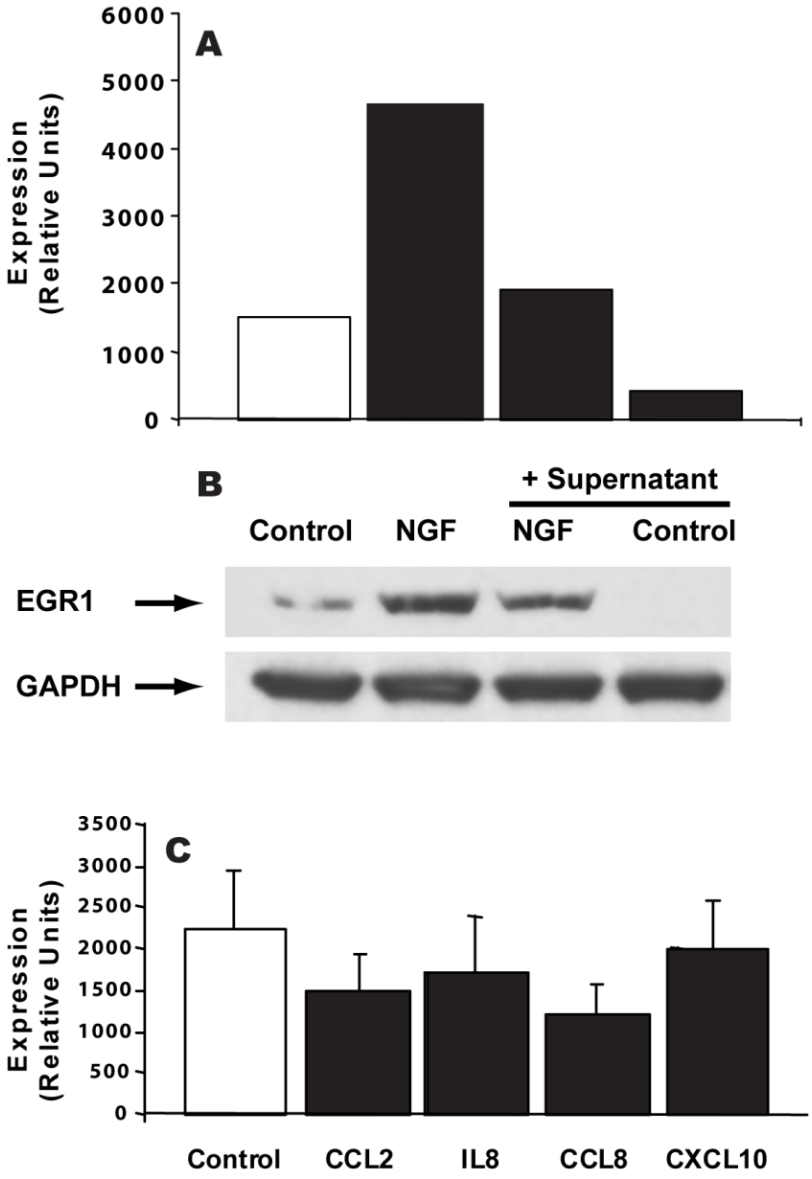


Figure 2.5 CCL8 downregulated EGR1 in differentiated SK-N-MC cells.
(A) Culture supernatant from SIV-infected macrophages inhibited both constitutive and NGF-induced (30 minute treatment) EGR1 expression in 4-day differentiated SK-N-MC cells after 2-day pretreatment at 1:50 supernatant dilution. Normalization was done using housekeeping genes 18S and GAPDH; figure shows expression in terms of relative units. (B) In cells cultured at the same time and under the same conditions as for (A), constitutive and NGF-induced (2 hour treatment) EGR1 protein was also reduced by 2-day pretreatment with 1:50 dilution of supernatant from infected macrophages. (C) Mean EGR1 expression (normalized to 18S and GAPDH; relative units) ± SEM of five experiments in which 5-day differentiated SK-N-MC cells were treated for 2.5-3 hours with one of the indicated chemokines at 100 ng/ml is presented. Inhibition of EGR1 by CCL8 was significant at $p < 0.01$ as assessed by paired sample one-way ANOVA (Prism: GraphPad, La Jolla, CA) performed on log-transformed data followed by Dunnett's post-test.

supernatant; this decrease in expression was associated with a decrease in EGR1 protein, as shown in Figure 2.5B.

Since chemotactic cytokines have been linked to both macrophage infiltration into the brain and effects on neurons in HAD[80,81], we examined chemokine expression in our dataset to identify microglia products that might be present in the supernatant and responsible for EGR1 downregulation. CCL2, CCL8, IL8 (CXCL8) and CXCL10 were the most significantly upregulated chemokines in monkey SIVE-hippocampus (23.6-, 16.8-, 4.8- and 36.3-fold, respectively) and are known to bind to receptors CCR2, CCR2/CCR3, CXCR1/CXCR2, and CXCR3, respectively. We confirmed that retinoic acid-treated SK-N-MC express these chemokine receptors (Supplemental Figure S2.2). To eliminate potential secondary effects, we evaluated the effect of these chemokines on EGR1 expression after a brief (2.5-3 hour) exposure, using a standard chemokine concentration (100 ng/ml, ~ 12 nM). In these experiments, cells were exposed to chemokines shortly (2-3 hours) after removal from serum-containing medium. As shown in Figure 2.5C, CCL2, CCL8 and IL8 each appear to reduce EGR1 expression by ~30%, similar in magnitude to the 40% reduction seen in SIV-infected hippocampus, whereas no reduction is seen after treatment with CXCL10. Paired sample one-way ANOVA using Prism (GraphPad, La Jolla, CA) was performed on log-transformed data and confirmed significant differences among treatment groups (p-value = 0.0065); however, Dunnett's post-test revealed only CCL8 as significantly different from control (p-value < 0.01).

Discussion

Integration of hippocampal gene expression data from nine SIV-infected and nine uninfected monkeys with human protein interaction data recovered significant upregulated modules enriched for well-described SIVE/HIVE processes, as well as several downregulated modules which include candidate neuronal genes whose downregulation may contribute to the neurologic dysfunction observed in monkeys and humans with SIV/HIV infection. Notably, many of the genes included in these modules are themselves not significantly differentially expressed, but taken together, each module significantly distinguishes between control and infected animals. This is considered to reflect perturbation of the same molecular pathways in different SIVE monkeys, even as individual animals may differ in perturbed expression of specific genes within those pathways. Dysregulation of molecular networks are responsible for functional pathology, leading to the change in state from normality to disease.

We selected several of the downregulated genes for quantitative RT-PCR confirmation and were able to confirm significant downregulation of EGR1, NR4A1 and CNTNAP1 in infected monkey hippocampus, despite the modest (< 2-fold) downregulation of these of the genes. In addition to verifying that these genes are indeed downregulated in vivo, this confirmation provides support for our method of mapping Affymetrix rhesus probesets to the human protein network.

One of these genes, EGR1, has been well-described as a participant in hippocampal functions of learning and memory[41-43] and it seems plausible, given the impairment of these functions in SIVE/HAD, that the observed downregulation occurs

in neurons, rather than in other hippocampus cell types. Indeed, examination of hippocampal tissue sections revealed reduced expression in neurons in animals with SIVE.

Supporting this finding, others have reported that the HIV-Tat protein reduces NGF-stimulated EGR1 expression in SK-N-MC cells, the neuroblastoma cell line used in our studies, acting via inhibition of the MAPK / Erk1/2 pathway[82]. Since EGR1 is reported as downregulated in models of several other neurodegenerative diseases[74-76], we asked whether EGR1 downregulation might be a consequence, at least in part, of macrophage/microglia activation present in many of these diseases, rather than the result of SIV infection per se. We evaluated several chemokines that were highly upregulated in SIV-infected hippocampus and demonstrated that acute treatment with CCL8 reduces EGR1 expression in differentiated SK-N-MC cells by ~30%. Although this appears to be a modest reduction, it is similar to the ~50% reduction seen in heterozygous mouse EGR1 knockout mutants which display the same signs of cognitive impairment (defects of late LTP and spatial learning) as those seen in homozygous mutants[41,42].

Chemokines bind to G-protein coupled receptors and activate different signaling pathways, including MAPK, PKB/ AKT, and JAK/STAT pathways, depending on receptor(s) activated and cell type[83,84]. Although EGR1 is activated via the MAPK1/2 pathway[82] and would therefore be expected to be induced by chemokine stimulation, inhibition of this pathway could occur via “RAF1-AKT crosstalk”, in which RAF1 kinase activity (and downstream MAPK/ERK activation) is suppressed

by PKB/AKT phosphorylation of Ser259 in a ligand-, concentration-, and time-course-dependent manner[85,86] (Figure 2.6B). This inhibitory mechanism could be particularly effective after CCL8 stimulation, as it has been reported that binding of CCL8 to CCR2, in contrast to CCL2 binding, fails to induce nuclear translocation of activated ERK1 and downstream gene induction[87]. AKT activity has been associated with neuroprotection in models of corticosterone-induced[88] and ischemia-induced[89] cell death. AKT activity is also required for CXCR2-mediated survival of cerebellar granule cells[90], and is induced after CCL2 stimulation of spinal cord neurons[91]. The latter two observations underscore the ability of chemokines to mediate neuroprotective, as well as neuroinflammatory functions. Downregulation of EGR1 might then occur as part of a neuroprotective response, with the associated untoward effect in hippocampus of reduced neuroplasticity.

Although only CCL8 produced statistically significant EGR1 downregulation in our study, CCL2 and IL8 also appeared to inhibit EGR1 expression, and failure to observe significance for CCL2 and IL8 could reflect the dose-dependency of chemokine effects. The possibility of EGR1 downregulation by IL8 is particularly interesting, as IL8 has been reported both to mediate neuroprotection[90] and to reduce rat hippocampal long-term potentiation (LTP)[92]; suppression of EGR1 might then work in tandem with the mechanism proposed by the authors, i.e., reduction of Ca2+ currents by IL8[92].

We selected EGR1 for further study based on its inclusion in significant protein-interaction modules; however, as a transcription factor, the effects of EGR1

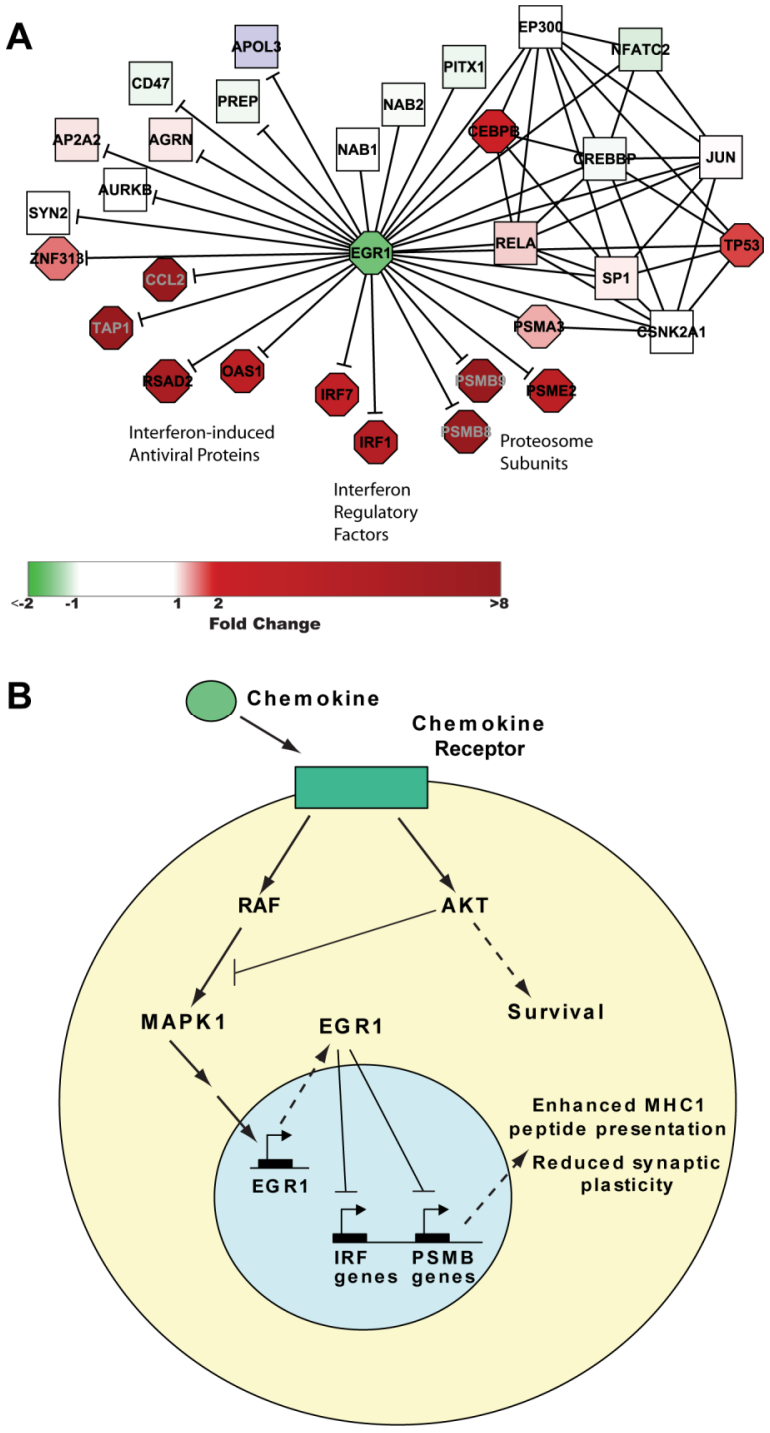


Figure 2.6 Proposed model of EGR1 downregulation in SIVE.
(A) EGR1 protein binding partners in the protein interaction network (—) and high-confidence gene targets (—|) taken from Table 1 in James et al[93], in which Irf7 appears twice, once with p=0.0025 and once with p=0.05. Rat gene names were associated with human Entrez Gene IDs using Entrez Gene Symbols and aliases. (B) Schematic depiction of proposed mechanism of EGR1 downregulation by chemokines in SIV-infected hippocampus and consequent reduced neuroplasticity.

downregulation may be better understood in terms of its target genes, amplifying the effect of its altered expression through its control of expression of many other genes. Interestingly, a recent microarray study of rat PC12 neuronal cells transfected with an EGR1 expression vector reported putative targets of EGR1 regulation[93], some of which were later confirmed experimentally[94]. All but three of the genes were downregulated by EGR1 overexpression. We analyzed the putative targets reported to have at least one EGR response element in the promoter region, (excluding MHC genes, which could not be reliably mapped to a human Entrez ID) for their expression in our SIVE dataset. Of the 100 genes mapped to a human Entrez ID (see Supplemental Table S2.3) 94 had expression data, and 25 were differentially expressed. Of the 24 upregulated genes, all were reported as downregulated in the EGR1-overexpressing in PC12 cells; the single downregulated gene was EGR1, reported to upregulate itself. In this SIVE setting of downregulated EGR1, the set of putative EGR1-downregulated target genes is highly enriched for upregulated genes (hypergeometric p-value = 2.5×10^{-12}) with respect to hippocampal SIVE gene expression in the complete set of mapped genes, providing support for their identification by James et al[93] as targets of EGR1 repression in neurons.

We then focused on a high-confidence subset of putative EGR1 targets whose differential expression was reported as significant at p-value <0.05, all 17 of which were downregulated in the EGR1-overexpressing in PC12 cells, and added them to the human network subset of EGR1-protein interactions as directed interactions. As shown in Figure 2.6A, expression of these genes in the SIVE dataset, in the context of

downregulated EGR1, was either significantly upregulated (N=10), unchanged (N=6), or not determined; in no case was expression significantly downregulated (significant enrichment for upregulated genes: hypergeometric p-value = 7.3×10^{-11}). This high-confidence set of upregulated putative targets of EGR1 regulation includes three clusters, interferon-induced antiviral proteins, OAS1 and RSAD2, interferon regulatory factors, IRF1 and IRF7, and three interferon gamma-induced proteasome subunits, PSMB8, PSMB9, and PSME2, as well as TAP1, essential for MHC Class I antigen processing. IRF comprise a family of transcription factors that not only induce Type I interferon transcription but also regulate transcription of interferon-induced genes, as well as play diverse roles in a variety of cell types[95,96]. In particular, IRF1 regulates interferon gamma-stimulated induction of proteosome immunosubunits PSMB8, PSMB9, and PSMB10 whose replacement in the proteosome facilitates MHC Class I antigen processing and antiviral immunity[97]. Proteasome activity has also been implicated in neuronal plasticity, associated with both a facilitating[98] and suppressive[99] effect, with the overall outcome likely related to the precise balance of active subunits influencing specific protein degradation[94].

These results, taken together, suggest a model in which downregulation of EGR1 in SIVE is a consequence of chemokine expression, brought about in the host for immune stimulation and protection of the CNS. In turn, downregulation of EGR1 promotes changes, both directly and via interferon regulatory factors, in proteosomal subunit selection that are detrimental to neuronal plasticity and cognition (Figure 2.6B). In any case, by linking activated/infected macrophages to cognitive deficit in

SIVE via downregulation of EGR1, an essential transcription factor in hippocampal neuronal function, our studies have identified a major factor in the indirect toxicity of microglia/macrophages on neurons.

Materials and Methods

Ethics Statement

Experiments were carried out under the approval and guidelines of The Scripps Research Institute's Animal Care and Use Committee. Anesthetic and analgesic agents were used to alleviate any pain or discomfort during procedures.

Human Protein Interaction Network

The human protein interaction network was compiled using the following sources: Human Protein Reference Database (HPRD, 01JAN07)[100], Biomolecular Interaction Network Database (BIND, 25MAY06)[101], REACTOME (28FEB07)[102], Database of Interacting Proteins (DIP, 19FEB07)[103], published human yeast two-hybrid[24,25] and mass spectrometry[23] datasets, and orthology-predicted PPI[104]. All interactions for which both proteins could be mapped to Entrez Gene IDs were included. Complexes identified in HPRD and BIND were treated using the matrix paradigm[105] to provide consistency with REACTOME's treatment of complexes. After removing redundancies, the final network contains 10,906 proteins (nodes) and 81,211 interactions (edges). Although high-throughput protein-interaction data is known to be

noisy[106], we elected to avoid pre-filtering the network using probabilistic edge weights, opting instead to filter by means of data integration, as described below.

Rhesus Macaques

We used 18 male rhesus macaques free of SIV, type D simian retrovirus and Herpes B virus obtained from Covance (Alice, TX), LABS of Virginia (Yemassee, SC) and Charles River Laboratories (Key Lois, FL). Experiments were carried out under the approval and guidelines of The Scripps Research Institute's Animal Care and Use Committee. Nine monkeys were intravenously inoculated with a cell-free stock of a derivative of SIVmac251 that had been subjected to in vivo passage[107,108]; four of these monkeys additionally underwent a CD8+ cell depleting regimen in the acute period after viral inoculation[39]. Two of the control monkeys were similarly treated with anti-CD8 antibody. The SIV-infected monkeys were sacrificed after development of neurological signs of simian AIDS, after a range of 56-132 (mean, 92; median, 93) days post-inoculation. Necropsy was performed after terminal anesthesia with ketamine, xylazine and pentobarbital, followed by intracardiac perfusion with sterile PBS containing 1 U/ml heparin to clear the brain of blood-borne cells. All SIV-infected monkeys had neuropathological evidence of SIVE on post-mortem histopathological examination. A portion of hippocampus was preserved frozen at necropsy for later RNA extraction. A summary of the monkeys used in this study is presented in Supplemental Table S2.1.

Microarray

Total RNA was purified from hippocampal samples using TRIzol Reagent (Invitrogen, Carlsbad, CA), centrifuged to remove cellular debris, then purified using the RNeasy mini kit (Qiagen, Valencia, CA). RNA quantity was assessed by 260-nm UV absorption and quality verified by Agilent BioAnalyzer analysis of ribosomal RNA bands. RNA was reverse-transcribed and hybridized to the Affymetrix Rhesus Genome Array. This platform contains 52,024 rhesus probesets representing ~ 47,000 transcripts. Each sample was tested in duplicate and results were processed using Affymetrix software (MAS 5.0) to produce CEL files, which were submitted to R/Bioconductor Robust Multi-Chip Analysis (RMA)[56] to yield normalized log2 expression (intensity) values for each probeset in the 36 arrays. Fold-change and minimum expression (intensity) thresholds were not applied in order to facilitate detection of genes that might be significantly regulated in only a subset of cells in the mixed cell population present in the hippocampus samples. Differential gene expression using mean subject values was determined using Cyber T[57] software, which calculates significance using a Bayesian approach to regularize t-tests and also calculates a Posterior Probability of Differential Expression (PPDE), the probability that a gene at a given p-value is differentially expressed (http://cybert.microarray.ics.uci.edu/help/index.html, Allison *et al*[109]). Differential expression p-values were calculated by Cyber T using log-transformed data, and a p-value of 0.001 corresponded to a PPDE of approximately 95%. The microarray data have been deposited in NCBI's Gene Expression Omnibus and are accessible through

GEO Series accession number GSE13824 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc= GSE13824).

To facilitate integration of the microarray results with the human PPI network, it was necessary to associate Affymetrix probeset identifiers with human Entrez gene identifiers. Affymetrix probesets were mapped to human Entrez Gene IDs by reciprocally blasting probeset target sequences (downloaded from the Affymetrix website) and the human genome (sequences downloaded from Refseq ftp site) keyed to Entrez ID and selecting only the best reciprocal hits that had bit scores ≥ 50. This approach was taken, rather than using Inparanoid[110] software, to avoid multiple rounds of mapping between the Ensembl ID used by Inparanoid, and Entrez Gene ID. Using this method, it was possible to map 18,313 Affymetrix probesets to human Entrez Gene IDs, including 10,209 of the 10,906 proteins present in the human PPI network. Seventeen mappings had bit scores <60 with e-values ranging between 3e-05 and 5e-08 and %-identity 80.43-100%; all other mappings had e-values ≤ 4e-08. Applying a Cyber T p-value cutoff of 0.001 identified 826 genes (Affymetrix probesets mapped to Entrez Gene IDs) as significantly differentially expressed, of which 106 were downregulated and 720 were upregulated.

Immunohistochemistry

Formalin fixed, paraffin embedded tissue blocks containing the hippocampus from four monkeys in each group were sectioned at 5 µm thickness and picked up on glass slides. Following deparaffinization, antigen retrieval was carried out by heating

to 95° C in 0.01 M citrate buffer, pH 6.39, for 40 minutes, then left for 20 minutes to steep. Sections were blocked in 0.5% casein, followed by the addition of a 1/50 dilution of rabbit monoclonal anti-EGR1 antibody (clone 15F7, Cell Signaling Technology, Beverly, MA) overnight at 4° C. Following washes, signal was detected using the SuperPicture broad spectrum secondary antibody-horseradish peroxidase polymer reagent (Invitrogen, Carsbad, CA) and developed with the NovaRed chromogen (Vector Laboratories, Burlingame, CA), followed by a haematoxylin counterstain (Sigma-Aldrich). Controls included omission of the primary antibody and use of irrelevant primary antibodies. Image capture was performed with a Spot RT Color CCD camera with Spot RT software (Spot Diagnostic Instruments, Sterling Heights, MI) using a Leica Diaplan microscope (Leica, Deerfield, IL). Adobe Photoshop CS3 (Adobe Systems, San Jose, CA) was used for quantification. Photomicrographs were converted to black and white using the green filter, inverted, and for each monkey the signal in 20 consecutive neuronal nuclei from the CA1 region was quantified using the mean histogram density measurement. Background was determined for each section, and was subtracted from the average nuclear signal, yielding the average staining signal from each monkey. The two groups were compared using a unpaired two-tailed t-test.

Data Integration

To integrate differential gene expression with protein interaction data we used a recently described algorithm[32] that identifies network modules whose expression

significantly discriminates between two experimental conditions. The algorithm overlays gene expression values on corresponding network proteins and, starting from every protein (seed) in the protein network, greedily adds interactions to identify the module starting from each seed whose mean expression calculated for each sample best discriminates between the two sample types. This process yields as many candidate modules as there are proteins in the network; non-significant subnetworks are then filtered out using three types of permutation testing. This method was shown to improve classification of metastatic versus non-metastatic breast cancer and to capture well-established prognostic markers not regulated at the level of transcription[32]. Parameters were chosen as described in Chuang et al.[32], except that t-test rather than mutual information was used to determine significance. We ran the algorithm five times and selected modules that scored as significant in at least three of the runs. The algorithm is available as PinnacleZ, a plugin to Cytoscape[26].

Cell Culture

SK-N-MC human neuroblastoma cells were purchased from ATCC and maintained in culture as recommended. For each experiment, cells were plated a minimum of 9 hours or overnight at 2.5 x 10^5 cells/ml in 6-well plates (chemokine experiments; 2 ml/well) or 100 cm dish (supernatant experiment; 14 ml/dish), treated for 4 or 5 days with 10 μM trans-retinoic acid (EMD/Calbiochem, San Diego, CA) then switched to Neurobasal (NB, Invitrogen/Gibco) medium (without glutamate) supplemented with B-27 (Invitrogen/Gibco). Cells were kept in supplemented NB

medium for 2-3 hours or overnight prior to reagent addition. Macrophage supernantant was prepared as described below; recombinant human β-NGF (Millipore/Chemicon, Billerica, MA), and recombinant human CCL2, CCL8, CXCL8 (IL8) and CXCL10 (Peprotech, Rocky Hill, NJ) were purchased.

Flow Cytometry

SK-N-MC cells were plated overnight at $2x10^5$/ml (2 ml/6-well plate well), treated for 4 days with 10 μM trans-retinoic acid (Calbiochem) then switched to Neurobasal (NB) medium (without glutamate) supplemented with B-27 (Gibco) for overnight culture. Cells were harvested in Versene, centrifuged and resuspended at $1x10^7$/ml in Flow Buffer (PBS pH 7.2 containing 2% Fetal Bovine Serum and 0.2% Sodium Azide). Cells were then incubated for 30 minutes at 4^oC in the presence of fluorescently-labeled antibodies (1 ug/10^6 cells). The antibodies used in the assays were: Biotin-mouse anti-human CCR-2 (R&D Systems Cat. No. FAB151B); PE-mouse anti-human CD193 (CCR3) (BD Biosciences Cat. No. 558165); FITC-mouse anti-human CD181 (CXCR1) (BD Biosciences Cat. No. 555939); APC-mouse anti-human CD182 (CXCR2) (BD Biosciences Cat. No. 551127); PE-mouse anti-human CD183 (CXCXR3) (BD Biosciences Cat. No. 557185); PE-mouse anti-human TrkA (High-afinity NGF receptor) (RD Systems Cat No. FAB1751P); Biotin-mouse anti-human CD271 (Low-affinity NGF receptor) (BD Biosciences Cat. No. 557195). Following a wash step wells containing biotinilated antibodies were further incubated for 20 minutes with streptavidin-APC or streptavidin-PE (BD Pharmingen). Propidium

Iodide was used to differentiate viable cells and acquisition was immediately performed in a FACScalibur (BD Biosciences). Cells were analyzed in FlowJo software (Tree Star, Inc., OR).

Macrophage supernatant

Macrophages/microglia were isolated by Percoll gradient from disrupted, collegenase-digested brains of rhesus macaques with SIVE and maintained in RPMI 1640 supplemented with MCSF at 50 ng/ml (Peprotech) up to 30 days. After day 7, supernatant was harvested daily, and samples with p27 > 1 ng/ml were filtered (MW cutoff 30 kDa) and filtrates stored at -80^{o}C.

Western Blot

Following treatment, cells were washed twice in ice-cold PBS, scraped, centrifuged and then resuspended in ice-cold RIPA lysis buffer [50 mM HEPES buffer, 1 mM sodium orthovandate, 10 mM sodium pyrophosphate, 10 mM NaF, 1% NP-40, 30 mM p-nitrophenyl phosphate (MP Biomedicals, Solon, OH), 1 mM PMSF, 1× complete protease inhibitor cocktail (Roche Diagnostics, Indianapolis, IN)]. Samples were incubated on ice for 30 min with periodic vortexing, then centrifuged at high speed for 25 min at 4^{o}C, after which supernatants containing soluble proteins were collected. Total protein concentration was determined with the bicinchoninic acid assay, using bovine serum albumin standards. Proteins were separated on a 4-12% Bis-Tris gradient gel (Invitrogen) and transferred electrophoretically to 0.45 μM

pore PVDF membrane (Invitrogen). Nonspecific antibody binding was blocked by 5% nonfat dried milk for 1 hr at room temperature. Immunoblotting was carried out with antibody against EGR1 (1:1000, Cell Signaling mAb 15F7; Danvers, MA), followed by secondary antibody (1:10,000 HRP conjugated anti rabbit IgG; GE Healthcare, Little Chalfont, UK). The blot was developed with 1:1 solution of Super Signal West Pico Chemiluminescent Substrate and Luminol/Enhancer (Thermo Fisher Scientific, Rockford, IL). The blot was then stripped using ReStore® Western Blot stripping buffer (Thermo Fisher Scientific) and re-probed for GAPDH (loading control; Millipore/Chemicon, MAB374).

RT-PCR

Probes (labeled with FAM and TAMRA) and primers were designed to detect human and/or rhesus genes, as appropriate and were obtained from Eurogentec (San Diego, CA). For the housekeeping genes (18S, GAPDH and TBP), the sequences used are the same as given in Wikoff et al.[111]. For the other genes, the sequences (given as forward – for, reverse – rev, and probe – pro) were EGR1: for ACGCCGAACACTGACATTT, rev GCTGTGGAAACAGGTAGTCG, pro TCCAGTACCCGCCTCCTGCC; PTN: for AATGGCAGTGGAGTGTGTGT, rev CTTCATGGTTTGCTTGCACT, pro CGGCTCCAGTCCGAGTGCCT; NR4A1: for TTCCTGGAGCTCTTCATCCT, rev GAAGATGAGCTTGCCCTCTC, pro CGCCTGGCATACAGGTCTAAGCC; HES5: for AAGCCAGTGGTGGAGAAGAT, rev TAGCTGACAGCCATCTCCAG, pro

CGACCGCATCAACAGCAGCA; CNTNAP1: for CTTCAAGACCGAGGAGAAGG, rev CAGCTCGATCGTCACGTAGT, pro ACGGGCTTCTGCTGCACGC. RT-PCR was performed using the Stratagene MXPro 3000 instrument. Figures show gene expression normalized by housekeeping genes in terms of relative units; statistical testing was performed on log-transformed data.

Supplemental Figures

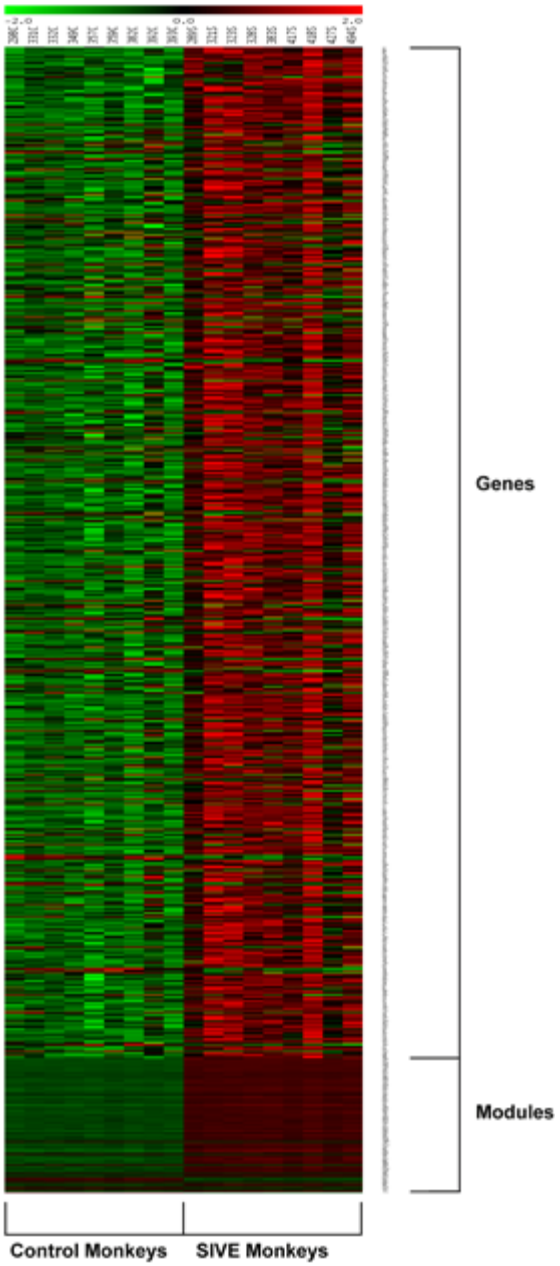


Figure S2.1 Heat map comparing expression of individual genes and significant modules. Expression values were normalized by gene and mean normalized expression values for each experimental animal (Control or SIVE) were calculated for the set of filtered significant modules. Compared with individual gene expression, average module expression yields a more homogeneous, though less intense, differentiation between uninfected and SIVE animals. Figure was generated by TMeV software [http://www.tm4.org/mev.html].

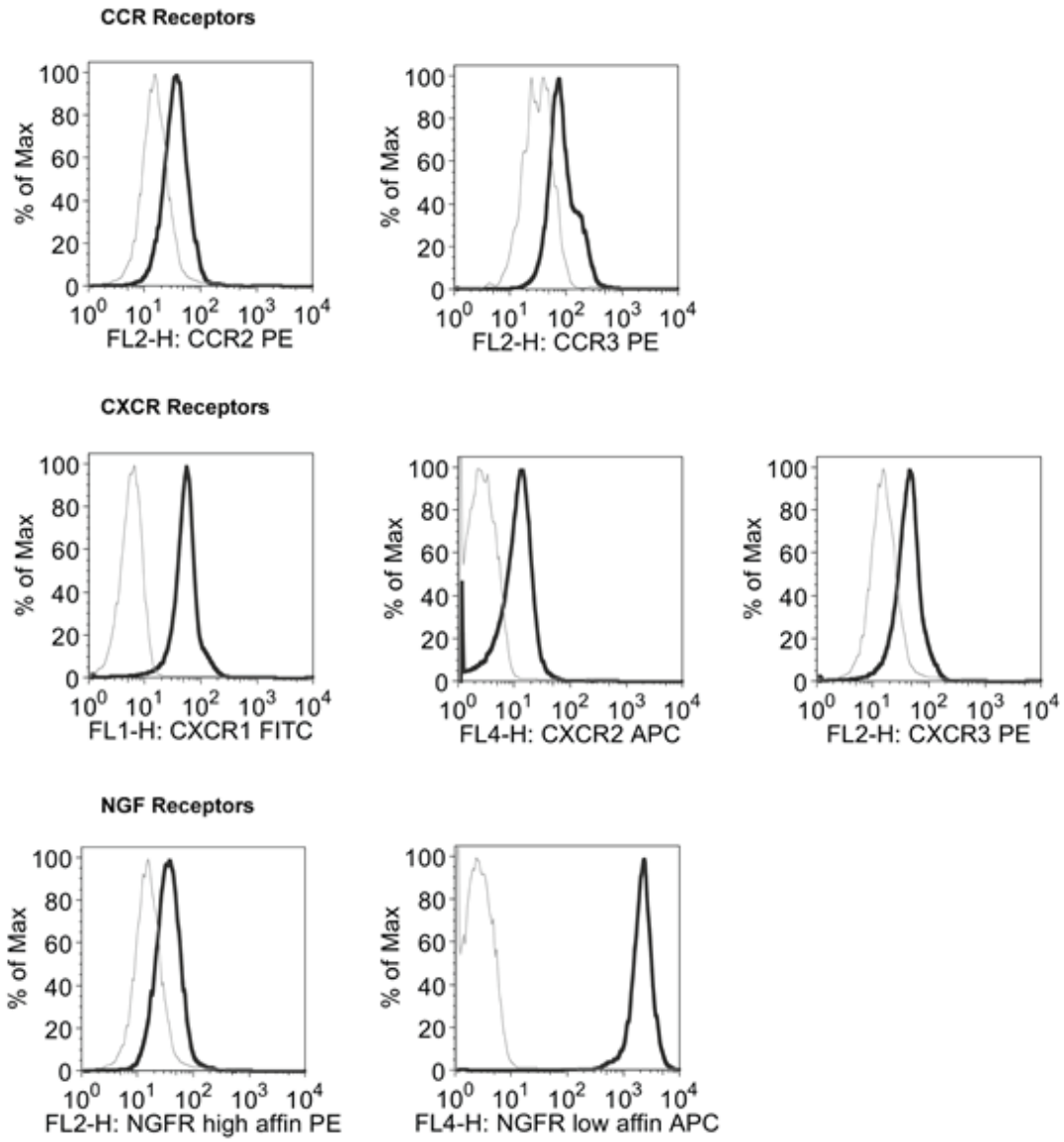


Figure S2.2 Multiple chemokine receptors are expressed on differentiated SK-N-MC cells. This flow cytometry analysis was performed by C. Marcondes[112]. Cells were differentiated for 4 days with retinoic acid prior to FACS analysis for chemokine receptors CCR2 (ligands: CCL2, CCL8), CCR3 (CCL8), CXCR1 (IL8), CXCR2 (IL8), CXCR3 (CXCL10), and high- and low-affinity NGF receptors.

Supplemental Tables

Table S2.1 Treatment History of Rhesus Monkeys
The age of the monkey and time after viral inoculation that the infected animals were sacrificed, and hippocampus harvested for RNA, are given. Animals that received intravenous antibody treatment are are also indicated. The brain viral load was determined on a sample of frontal lobe, containing both grey and white matter, and is given in log(10) viral genome equivalents per microgram RNA. Quantification was performed by the bDNA method by Siemens Reference Lab, Emeryville, CA. NA = not applicable.

CONTROL					
Monkey	**Age (years)**	**Sex**	**anti-CD8 treatment**	**Days infected**	**Brain Viral Load**
298	5	Male	No	NA	NA
331	6	Male	No	NA	NA
332	5	Male	No	NA	NA
349	3	Male	No	NA	NA
357	4	Male	No	NA	NA
359	8	Male	No	NA	NA
382	5	Male	No	NA	NA
392	4	Male	Yes	NA	NA
393	3	Male	Yes	NA	NA
SIVE					
289	6	Male	Yes	72	5.41
321	6	Male	Yes	108	5.77
323	6	Male	Yes	93	6.68
328	6	Male	Yes	79	6.20
383	5	Male	No	132	6.06
417	2	Male	No	56	5.89
418	5	Male	No	82	6.76
427	4	Male	No	104	6.63
494	5	Male	No	105	5.96

Table S2.2 Differential Gene Expression in Monkeys with SIVE.
Selection based on CyberT p ≤0.001. Affymetrix monkey probeset IDs are given together with the human genes to which they were mapped.

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
Downregulated					
MmuSTS.91.1.S1_at	285313	IGSF10	1.49E-06	-2.205	1.00E+00
MmugDNA.30177.1.S1_at	388585	HES5	2.07E-07	-2.102	1.00E+00
MmugDNA.25181.1.S1_at	8630	HSD17B6	1.33E-07	-2.051	1.00E+00
MmuSTS.4085.1.S1_at	1843	DUSP1	2.37E-05	-1.948	9.95E-01
MmugDNA.18685.1.S1_s_at	3164	NR4A1	1.26E-05	-1.862	9.99E-01
MmugDNA.20291.1.S1_at	3547	IGSF1	1.80E-06	-1.832	9.99E-01
MmugDNA.42663.1.S1_at	7138	TNNT1	3.15E-04	-1.823	9.92E-01
MmuSTS.1096.1.S1_at	79957	PAQR6	9.42E-06	-1.739	9.98E-01
MmuSTS.4112.1.S1_at	1958	EGR1	7.86E-07	-1.717	1.00E+00
MmugDNA.3399.1.S1_at	649504		3.45E-05	-1.704	9.97E-01
MmuSTS.4649.1.S1_at	23620	NTSR2	2.00E-05	-1.699	9.94E-01
MmunewRS.661.1.S1_at	8364		1.16E-05	-1.697	9.92E-01
MmugDNA.36630.1.S1_at	8490	RGS5	3.08E-06	-1.677	9.99E-01
MmugDNA.30510.1.S1_at	80864	EGFL8	5.21E-04	-1.656	9.82E-01
MmugDNA.43242.1.S1_at	8506	CNTNAP1	1.07E-06	-1.618	9.99E-01
MmuSTS.754.1.S1_at	10815	CPLX1	2.21E-03	-1.616	9.53E-01
MmugDNA.43402.1.S1_at	122622	ADSSL1	2.37E-05	-1.612	9.97E-01
MmugDNA.13626.1.S1_at	5740	PTGIS	4.62E-04	-1.608	9.75E-01
MmugDNA.15888.1.S1_at	94234	FOXQ1	1.83E-05	-1.598	9.97E-01
MmugDNA.13861.1.S1_at	79152	FA2H	7.81E-04	-1.594	9.77E-01
MmugDNA.42660.1.S1_at	112703		1.77E-05	-1.592	9.98E-01
MmugDNA.17369.1.S1_at	93377	TMEM10	9.13E-04	-1.582	9.47E-01
MmuSTS.4753.1.S1_at	11197	WIF1	4.27E-04	-1.577	9.65E-01
MmugDNA.9634.1.S1_at	887	CCKBR	8.76E-04	-1.565	9.80E-01
Mmu.2136.1.S1_at	1675	DF	3.36E-03	-1.559	9.48E-01
MmugDNA.20999.1.S1_at	398	ARHGDIG	7.43E-05	-1.554	9.89E-01
MmugDNA.2357.1.S1_at	338645		1.28E-03	-1.541	9.47E-01
MmuSTS.2842.1.S1_at	23237		4.79E-04	-1.538	9.75E-01
MmugDNA.40384.1.S1_at	4056	LTC4S	9.07E-04	-1.521	9.68E-01
MmugDNA.3340.1.S1_at	5764	PTN	6.32E-05	-1.521	9.96E-01
MmugDNA.12381.1.S1_s_at	284723	SLC25A34	9.26E-04	-1.511	9.78E-01
MmugDNA.23555.1.S1_at	94032		7.36E-04	-1.499	9.71E-01
MmuSTS.2035.1.S1_at	5156	PDGFRA	2.49E-05	-1.498	9.94E-01
MmuSTS.2696.1.S1_at	1734	DIO2	2.84E-04	-1.497	9.72E-01
MmugDNA.8781.1.S1_at	1297	COL9A1	1.69E-03	-1.497	9.51E-01
MmugDNA.27557.1.S1_at	166647	GPR125	4.35E-05	-1.496	9.94E-01
MmugDNA.40342.1.S1_at	387066		8.88E-05	-1.492	9.92E-01

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmugDNA.26480.1.S1_at	349236		7.98E-04	-1.488	9.71E-01
MmugDNA.12222.1.S1_at	1179	CLCA1	6.50E-05	-1.486	9.94E-01
MmuSTS.3752.1.S1_at	8325	FZD8	3.86E-04	-1.485	9.75E-01
MmugDNA.20650.1.S1_at	285613		3.67E-04	-1.468	9.79E-01
MmugDNA.22295.1.S1_at	6664	SOX11	1.79E-04	-1.467	9.78E-01
MmugDNA.29391.1.S1_at	55643	BTBD2	7.54E-04	-1.464	9.61E-01
MmugDNA.11974.1.S1_at	54205	CYCS	1.61E-05	-1.463	9.95E-01
MmugDNA.16349.1.S1_at	2913	GRM3	7.86E-05	-1.461	9.86E-01
MmugDNA.1162.1.S1_at	65108	MARCKSL1	1.98E-05	-1.455	9.95E-01
MmugDNA.19653.1.S1_at	6623	SNCG	2.96E-04	-1.454	9.68E-01
MmuSTS.3252.1.S1_at	6404	SELPLG	9.73E-04	-1.450	9.50E-01
MmugDNA.31779.1.S1_at	56301	SLC7A10	7.71E-04	-1.449	9.68E-01
MmuSTS.2076.1.S1_at	5454	POU3F2	5.27E-04	-1.449	9.86E-01
MmugDNA.21469.1.S1_at	2947	GSTM3	1.61E-05	-1.449	9.94E-01
MmugDNA.40858.1.S1_at	53919	SLCO1C1	2.49E-04	-1.446	9.83E-01
MmuSTS.12.1.S1_at	55502	HES6	6.49E-05	-1.445	9.94E-01
MmuSTS.4093.1.S1_at	1908	EDN3	3.32E-04	-1.438	9.89E-01
MmugDNA.12551.1.S1_s_at	732056		1.10E-04	-1.438	9.84E-01
MmugDNA.12650.1.S1_at	143282		1.12E-04	-1.429	9.89E-01
MmuSTS.170.1.S1_at	27297	RCP9	3.36E-04	-1.424	9.67E-01
MmuSTS.4083.1.S1_at	9143	SYNGR3	5.46E-05	-1.423	9.88E-01
MmugDNA.8277.1.S1_s_at	3778	KCNMA1	1.25E-03	-1.419	9.53E-01
MmugDNA.5031.1.S1_at	2840	GPR17	1.73E-04	-1.409	9.86E-01
MmuSTS.1598.1.S1_at	266743	NXF	3.56E-04	-1.409	9.88E-01
MmugDNA.33484.1.S1_at	85358	SHANK3	1.31E-04	-1.400	9.80E-01
MmugDNA.30027.1.S1_at	23231	KIAA0746	2.74E-04	-1.397	9.72E-01
MmuSTS.1894.1.S1_at	9547	CXCL14	6.68E-05	-1.395	9.87E-01
MmuSTS.2072.1.S1_s_at	5354	PLP1	4.07E-04	-1.392	9.49E-01
MmugDNA.38889.1.S1_at	51181	DCXR	2.03E-04	-1.391	9.74E-01
MmugDNA.31713.1.S1_at	200407		6.09E-04	-1.382	9.48E-01
MmugDNA.29861.1.S1_at	770	CA11	3.18E-04	-1.375	9.61E-01
MmugDNA.39371.1.S1_at	10591	C6orf108	7.31E-04	-1.372	9.48E-01
MmuSTS.2106.1.S1_at	5507	PPP1R3C	2.75E-04	-1.358	9.68E-01
MmugDNA.6593.1.S1_at	6305	SBF1	4.64E-04	-1.354	9.76E-01
MmuSTS.1387.1.S1_at	27133	KCNH5	9.78E-04	-1.354	9.54E-01
MmugDNA.1734.1.S1_at	6676	SPAG4	8.38E-04	-1.337	9.80E-01
MmugDNA.2197.1.S1_at	11261	CHP	1.35E-03	-1.328	9.53E-01
MmugDNA.17036.1.S1_at	56955	MEPE	1.29E-03	-1.321	9.60E-01
MmugDNA.28744.1.S1_at	441005		9.83E-04	-1.321	9.59E-01
MmugDNA.726.1.S1_at	4715	NDUFB9	1.10E-04	-1.320	9.89E-01
MmuSTS.272.1.S1_at	3381	IBSP	6.93E-04	-1.317	9.79E-01
MmugDNA.40575.1.S1_at	266722	HS6ST3	1.81E-04	-1.316	9.91E-01
MmugDNA.17905.1.S1_at	146439		5.55E-04	-1.313	9.69E-01

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmuSTS.2464.1.S1_at	2827	GPR3	8.58E-04	-1.304	9.69E-01
MmugDNA.36148.1.S1_at	1592	CYP26A1	1.29E-03	-1.304	9.60E-01
MmugDNA.19182.1.S1_at	6002	RGS12	1.72E-03	-1.300	9.55E-01
MmuSTS.78.1.S1_at	10930	APOBEC2	2.11E-04	-1.299	9.89E-01
MmugDNA.12351.1.S1_at	388323		2.26E-03	-1.299	9.47E-01
MmugDNA.39175.1.S1_at	6655	SOS2	2.55E-04	-1.297	9.83E-01
MmugDNA.15647.1.S1_at	146330	FBXL16	8.52E-05	-1.289	9.82E-01
MmuSTS.3016.1.S1_at	190	NR0B1	7.02E-05	-1.285	9.94E-01
MmugDNA.20299.1.S1_at	368	ABCC6	1.14E-03	-1.280	9.55E-01
MmugDNA.23494.1.S1_at	10166	SLC25A15	1.13E-04	-1.267	9.94E-01
MmugDNA.1488.1.S1_at	339976	FLJ36180	1.20E-03	-1.266	9.52E-01
MmugDNA.21566.1.S1_at	5455	POU3F3	1.10E-03	-1.265	9.54E-01
MmugDNA.40761.1.S1_at	124783		6.87E-04	-1.263	9.77E-01
MmugDNA.3541.1.S1_at	9923	ZBTB40	1.63E-03	-1.260	9.63E-01
MmugDNA.20029.1.S1_at	11123	DSCR1L2	1.79E-03	-1.243	9.52E-01
MmuSTS.2328.1.S1_at	122970	PTE2B	8.14E-04	-1.242	9.65E-01
MmugDNA.19469.1.S1_s_at	151254	ALS2CR11	1.95E-04	-1.235	9.89E-01
MmugDNA.2170.1.S1_at	11238	CA5B	1.03E-03	-1.217	9.56E-01
MmugDNA.15046.1.S1_at	83872	HMCN1	8.19E-04	-1.215	9.73E-01
MmugDNA.28903.1.S1_at	3228	HOXC12	9.42E-04	-1.215	9.61E-01
MmugDNA.43496.1.S1_at	283551		1.85E-03	-1.203	9.61E-01
MmugDNA.5237.1.S1_at	727961		1.86E-03	-1.186	9.54E-01
MmugDNA.13171.1.S1_at	283461		7.13E-04	-1.182	9.70E-01
MmugDNA.41659.1.S1_at	126638	RPTN	1.50E-03	-1.174	9.48E-01
MmunewRS.488.1.S1_at	389012		1.37E-03	-1.166	9.48E-01
MmunewRS.324.1.S1_at	5105	PCK1	7.74E-04	-1.148	9.69E-01
Upregulated					
MmugDNA.21995.1.S1_at	4534	MTM1	4.55E-04	1.189	9.74E-01
MmuSTS.686.1.S1_at	3574	IL7	1.80E-03	1.195	9.47E-01
Mmu.11859.1.S1_at	284293	LOC284293	4.39E-04	1.197	9.83E-01
MmugDNA.18745.1.S1_at	200316	APOBEC3F	7.69E-04	1.205	9.60E-01
MmugDNA.3959.1.S1_at	220042	FLJ25416	1.76E-03	1.218	9.50E-01
MmuSTS.1653.1.S1_at	7272	TTK	7.01E-05	1.218	9.94E-01
MmugDNA.27361.1.S1_at	6372	CXCL6	7.93E-06	1.220	9.98E-01
MmugDNA.24387.1.S1_at	3676	ITGA4	1.33E-04	1.222	9.85E-01
MmunewRS.48.1.S1_at	159371	TMEM20	9.97E-04	1.226	9.47E-01
MmugDNA.37505.1.S1_at	137075	CLDN23	1.06E-04	1.229	9.87E-01
MmuSTS.61.1.S1_at	3070	HELLS	4.18E-04	1.235	9.66E-01
MmuSTS.1534.1.S1_at	10112	KIF20A	1.93E-04	1.242	9.89E-01
MmugDNA.4224.1.S1_at	1979	EIF4EBP2	7.67E-04	1.247	9.61E-01
MmuSTS.4533.1.S1_at	672	BRCA1	6.19E-04	1.256	9.76E-01

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmuSTS.4111.1.S1_at	1950	EGF	4.36E-04	1.258	9.77E-01
MmugDNA.43521.1.S1_at	143888		4.92E-04	1.263	9.66E-01
MmugDNA.10587.1.S1_at	4690	NCK1	3.47E-04	1.272	9.82E-01
MmuSTS.3374.1.S1_at	969	CD69	4.77E-05	1.276	9.93E-01
MmugDNA.34027.1.S1_at	9754	STARD8	8.81E-04	1.277	9.60E-01
MmuSTS.3471.1.S1_s_at	663	BNIP2	1.02E-03	1.278	9.61E-01
MmuSTS.4250.1.S1_at	10635	RAD51AP1	6.23E-05	1.281	9.95E-01
MmugDNA.32954.1.S1_at	79980	C20orf172	7.15E-04	1.281	9.47E-01
MmuSTS.1558.1.S1_at	9833	MELK	1.20E-06	1.285	1.00E+00
MmugDNA.13029.1.S1_at	121506	FLJ32115	9.44E-04	1.287	9.59E-01
MmuSTS.4145.1.S1_at	22944	KIN	3.54E-04	1.291	9.74E-01
MmugDNA.23209.1.S1_at	64581	CLEC7A	5.49E-04	1.296	9.58E-01
MmugDNA.15250.1.S1_at	1058	CENPA	6.16E-05	1.299	9.96E-01
MmuSTS.4420.1.S1_at	22974	TPX2	8.13E-04	1.300	9.55E-01
MmugDNA.14909.1.S1_at	9	NAT1	3.67E-05	1.301	9.96E-01
MmugDNA.511.1.S1_at	5687	PSMA6	1.57E-04	1.309	9.85E-01
MmugDNA.34057.1.S1_at	55848	C9orf46	3.45E-04	1.310	9.65E-01
MmugDNA.34652.1.S1_at	51163	DBR1	1.89E-05	1.313	9.96E-01
MmugDNA.22342.1.S1_at	5683	PSMA2	4.56E-04	1.318	9.72E-01
MmugDNA.22602.1.S1_at	81619	TSPAN14	7.89E-04	1.320	9.58E-01
MmugDNA.12109.1.S1_at	8412	BCAR3	5.32E-04	1.324	9.70E-01
MmugDNA.13933.1.S1_at	84525	HOP	5.28E-04	1.324	9.46E-01
MmunewRS.843.1.S1_at	388325		4.69E-05	1.325	9.94E-01
MmugDNA.23177.1.S1_at	81788	NUAK2	1.55E-04	1.327	9.87E-01
MmugDNA.29322.1.S1_at	134701		1.22E-04	1.327	9.83E-01
MmuSTS.4671.1.S1_at	944	TNFSF8	1.38E-05	1.331	9.95E-01
MmugDNA.41491.1.S1_at	400077		3.84E-04	1.331	9.73E-01
MmugDNA.3232.1.S1_at	57452	GALNTL1	6.72E-04	1.333	9.47E-01
MmugDNA.3469.1.S1_at	441108		7.50E-04	1.334	9.69E-01
MmugDNA.30433.1.S1_at	79801		1.47E-05	1.337	9.99E-01
MmugDNA.13218.1.S1_at	64065	PERP	1.11E-04	1.340	9.92E-01
MmuSTS.514.1.S1_at	286410	ATP11C	5.37E-04	1.343	9.68E-01
MmugDNA.22343.1.S1_at	5684	PSMA3	1.51E-04	1.343	9.69E-01
MmugDNA.11386.1.S1_at	51479	ANKFY1	4.84E-04	1.343	9.50E-01
MmuSTS.4458.1.S1_at	2268	FGR	1.64E-04	1.344	9.70E-01
MmuSTS.948.1.S1_at	8790	FPGT	2.23E-04	1.345	9.73E-01
MmugDNA.36685.1.S1_at	83852	SETDB2	7.30E-04	1.346	9.60E-01
MmugDNA.19553.1.S1_at	1643	DDB2	1.02E-03	1.347	9.49E-01
MmugDNA.19644.1.S1_at	83464	APH1B	1.89E-03	1.348	9.49E-01
MmugDNA.35965.1.S1_at	80231	CXorf21	2.50E-04	1.349	9.91E-01
MmugDNA.29523.1.S1_s_at	646890		6.86E-04	1.349	9.48E-01
MmuSTS.4796.1.S1_at	2329	FMO4	1.56E-04	1.349	9.72E-01
MmugDNA.36612.1.S1_s_at	51009	DERL2	1.10E-03	1.350	9.64E-01

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmugDNA.18835.1.S1_at	158376		7.09E-05	1.350	9.90E-01
MmugDNA.14092.1.S1_s_at	55062		3.33E-04	1.351	9.80E-01
MmugDNA.29197.1.S1_at	223082	ZNRF2	1.78E-06	1.351	9.99E-01
MmuSTS.911.1.S1_at	2146	EZH2	1.75E-06	1.353	9.99E-01
MmugDNA.13953.1.S1_at	55250	STATIP1	9.93E-04	1.356	9.57E-01
MmugDNA.20073.1.S1_at	122769	PPIL5	3.27E-05	1.357	9.96E-01
MmuSTS.3321.1.S1_at	900	CCNG1	5.53E-04	1.358	9.67E-01
MmugDNA.18603.1.S1_at	3117	HLADQA1;	1.94E-04	1.359	9.92E-01
MmuSTS.4707.1.S1_at	7508	XPC	9.73E-04	1.361	9.49E-01
MmuSTS.489.1.S1_s_at	389	RHOC	1.14E-03	1.361	9.50E-01
MmugDNA.22391.1.S1_at	165631	PARP15	1.43E-04	1.362	9.88E-01
MmugDNA.7816.1.S1_s_at	1536	CYBB	7.97E-06	1.362	9.98E-01
MmugDNA.29494.1.S1_at	2149	F2R	2.41E-05	1.365	9.96E-01
MmuSTS.2598.1.S1_at	29904	EEF2K	7.50E-04	1.366	9.61E-01
MmugDNA.13455.1.S1_at	26018	LRIG1	4.47E-04	1.369	9.61E-01
MmugDNA.31283.1.S1_at	79866		9.56E-06	1.370	9.98E-01
MmugDNA.3208.1.S1_at	54918	CKLFSF6	6.44E-05	1.372	9.88E-01
MmugDNA.9295.1.S1_at	11145	HRASLS3	8.53E-04	1.372	9.59E-01
MmugDNA.35654.1.S1_at	4175	MCM6	1.23E-03	1.373	9.51E-01
MmuSTS.4534.1.S1_at	695	BTK	1.13E-03	1.374	9.58E-01
MmugDNA.19285.1.S1_s_at	27230	SERP1	3.02E-04	1.375	9.68E-01
MmugDNA.23860.1.S1_at	8339	HIST1H2BG8343	4.86E-05	1.375	9.93E-01
MmuSTS.4321.1.S1_at	30061	SLC40A1	1.36E-04	1.376	9.88E-01
MmugDNA.865.1.S1_at	5111	PCNA	1.20E-04	1.376	9.74E-01
MmuSTS.999.1.S1_at	4609	MYC	9.31E-04	1.376	9.66E-01
MmugDNA.9744.1.S1_at	154141		3.51E-05	1.378	9.99E-01
MmugDNA.13670.1.S1_at	23549	DNPEP	3.36E-04	1.378	9.77E-01
MmuSTS.3274.1.S1_at	26750	RPS6KC1	3.15E-04	1.378	9.68E-01
MmugDNA.7779.1.S1_at	9766	KIAA0247	2.04E-04	1.379	9.81E-01
MmugDNA.11830.1.S1_at	64927	HCC8;	5.11E-04	1.380	9.80E-01
MmugDNA.32914.1.S1_at	92667		4.89E-04	1.380	9.46E-01
MmugDNA.2373.1.S1_at	5337	PLD1	3.15E-04	1.380	9.82E-01
MmugDNA.35207.1.S1_at	57761	TRIB3	3.60E-05	1.381	9.97E-01
Mmu.6471.1.S1_x_at	6348	CCL3	1.58E-04	1.382	9.87E-01
MmuSTS.1539.1.S1_at	4001	LMNB1	3.22E-04	1.385	9.78E-01
MmugDNA.2615.1.S1_at	732322		3.00E-04	1.386	9.86E-01
MmugDNA.3622.1.S1_at	11170	TU3A	3.04E-05	1.387	9.94E-01
Mmu.12708.1.S1_at	51411	BIN2	6.20E-04	1.390	9.77E-01
MmugDNA.36576.1.S1_at	58511	DNASE2B	4.48E-05	1.390	9.96E-01
MmugDNA.31400.1.S1_at	204	AK2	1.50E-03	1.392	9.50E-01
MmugDNA.7768.1.S1_at	9780		7.44E-04	1.393	9.51E-01
MmugDNA.24411.1.S1_at	728913		7.81E-05	1.394	9.88E-01
MmugDNA.19621.1.S1_at	643079		3.40E-04	1.395	9.73E-01

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmugDNA.4964.1.S1_at	9517	SPTLC2	2.92E-04	1.395	9.80E-01
MmugDNA.35537.1.S1_at	1514	CTSL	1.67E-04	1.396	9.88E-01
MmugDNA.19377.1.S1_at	10577		3.57E-04	1.396	9.62E-01
MmuSTS.3561.1.S1_at	10390	CEPT1	3.18E-04	1.399	9.71E-01
MmugDNA.30725.1.S1_at	10261	IGSF6	1.27E-04	1.400	9.89E-01
Mmu.11727.1.S1_at	5583	PRKCH	4.50E-05	1.400	9.98E-01
MmugDNA.27645.1.S1_at	120376		4.94E-04	1.401	9.68E-01
MmugDNA.34671.1.S1_at	824	CAPN2	1.67E-04	1.402	9.84E-01
MmuSTS.2327.1.S1_at	9491	PSMF1	1.36E-04	1.402	9.88E-01
MmugDNA.8940.1.S1_at	729659		2.15E-04	1.404	9.91E-01
Mmu.7083.1.S1_at	51267	CLEC1A	5.11E-05	1.404	9.92E-01
MmugDNA.28580.1.S1_at	307	ANXA4	5.92E-04	1.405	9.64E-01
MmugDNA.35783.1.S1_at	64114	PP1201	3.44E-05	1.405	9.96E-01
MmuSTS.4738.1.S1_at	10205	EVA1	7.75E-05	1.406	9.94E-01
MmugDNA.23709.1.S1_at	5872	RAB13	5.64E-04	1.406	9.70E-01
MmuSTS.1554.1.S1_at	1326	MAP3K8	7.80E-08	1.408	1.00E+00
MmuSTS.3203.1.S1_at	25897	RNF19	4.74E-04	1.409	9.53E-01
MmugDNA.20111.1.S1_at	10627	MRCL3	2.17E-04	1.409	9.82E-01
MmuSTS.4575.1.S1_at	29969	MDFIC	9.87E-04	1.410	9.61E-01
MmugDNA.11328.1.S1_at	467	ATF3	3.71E-04	1.410	9.56E-01
Mmu.737.1.S1_s_at	8992	ATP6V0E	7.61E-04	1.411	9.67E-01
MmugDNA.546.1.S1_at	23193	GANAB	1.97E-04	1.412	9.83E-01
MmuSTS.682.1.S1_at	3570	IL6R	7.43E-04	1.413	9.60E-01
MmuSTS.1444.1.S1_at	4005	LMO2	3.79E-04	1.415	9.73E-01
MmuSTS.84.1.S1_at	3112	HLADOB;	1.20E-05	1.416	9.97E-01
MmuSTS.1535.1.S1_at	11004	KIF2C	1.13E-04	1.417	9.91E-01
MmugDNA.35225.1.S1_at	642393		7.61E-04	1.420	9.66E-01
MmugDNA.16753.1.S1_at	152137		1.13E-04	1.420	9.85E-01
MmuSTS.2169.1.S1_at	29957	SLC25A24	1.36E-04	1.422	9.90E-01
MmuSTS.3525.1.S1_at	11118	BTN3A2	1.91E-05	1.423	9.99E-01
MmugDNA.36055.1.S1_at	80301		1.50E-03	1.423	9.50E-01
MmugDNA.12065.1.S1_at	25878	MXRA5	2.05E-05	1.423	9.97E-01
MmugDNA.19464.1.S1_at	29128	UHRF1	1.40E-05	1.423	9.97E-01
Mmu.16124.1.S1_at	202	AIM1	5.17E-05	1.429	9.94E-01
MmugDNA.27747.1.S1_s_at	10783	NEK6	4.25E-05	1.429	9.89E-01
MmugDNA.30847.1.S1_at	127544	IBRDC3	1.32E-04	1.430	9.88E-01
MmugDNA.23069.1.S1_at	4616	GADD45B	7.25E-04	1.430	9.49E-01
MmugDNA.27004.1.S1_at	11167	FSTL1	1.23E-04	1.430	9.85E-01
MmugDNA.23728.1.S1_at	153768		1.06E-04	1.430	9.96E-01
MmugDNA.30565.1.S1_at	5031	P2RY6	8.41E-04	1.431	9.58E-01
Mmu.116.1.S1_s_at	51065	RPS27L	6.21E-05	1.432	9.88E-01
MmugDNA.38.1.S1_at	29916	SNX11	8.09E-05	1.432	9.90E-01
MmuSTS.4469.1.S1_s_at	7408	VASP	1.18E-04	1.432	9.82E-01

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmugDNA.17811.1.S1_at	3554	IL1R1	1.74E-05	1.433	9.97E-01
MmuSTS.3184.1.S1_at	8569	MKNK1	6.23E-04	1.433	9.83E-01
MmugDNA.43623.1.S1_s_at	1601	DAB2	8.22E-04	1.436	9.79E-01
MmuSTS.2182.1.S1_at	282890	ZNF311	1.57E-04	1.436	9.87E-01
MmuSTS.4034.1.S1_at	51311	TLR8	2.53E-06	1.437	1.00E+00
MmugDNA.12091.1.S1_at	578	BAK1	3.49E-04	1.438	9.78E-01
MmugDNA.18586.1.S1_at	10592	SMC2L1	1.74E-03	1.439	9.51E-01
MmugDNA.8062.1.S1_at	94120	SYTL3	6.60E-04	1.441	9.77E-01
MmuSTS.577.1.S1_at	57514	CDGAP	1.12E-03	1.442	9.63E-01
MmugDNA.19385.1.S1_at	23585	TMEM50A	2.92E-04	1.442	9.67E-01
MmugDNA.36391.1.S1_at	143903	LOC143903	5.87E-04	1.443	9.61E-01
MmugDNA.27473.1.S1_s_at	8837	CFLAR	1.51E-04	1.443	9.77E-01
MmugDNA.5460.1.S1_at	55920	RCC2	5.26E-05	1.444	9.92E-01
MmugDNA.38606.1.S1_at	84270	C9orf89	5.64E-04	1.444	9.64E-01
MmuSTS.144.1.S1_at	4171	MCM2	5.03E-05	1.445	9.94E-01
MmugDNA.34348.1.S1_at	80067		2.99E-04	1.447	9.69E-01
MmugDNA.10980.1.S1_at	79621		2.42E-04	1.447	9.83E-01
MmuSTS.3029.1.S1_at	5547	PRCP	3.55E-04	1.448	9.81E-01
MmugDNA.34150.1.S1_s_at	112597		3.45E-04	1.448	9.60E-01
MmugDNA.36470.1.S1_at	10403	KNTC2	1.84E-06	1.449	1.00E+00
MmugDNA.14810.1.S1_at	10906	TRAFD1	5.73E-05	1.451	9.93E-01
MmugDNA.8146.1.S1_at	79993	ELOVL7	2.12E-04	1.451	9.87E-01
MmuSTS.355.1.S1_at	10630	PDPN	1.54E-03	1.452	9.63E-01
MmugDNA.14995.1.S1_at	1003	CDH5	5.54E-04	1.452	9.53E-01
MmugDNA.3106.1.S1_at	5476	PPGB	8.37E-04	1.452	9.64E-01
MmugDNA.30611.1.S1_at	5106	PCK2	2.65E-04	1.453	9.73E-01
MmugDNA.15282.1.S1_at	9204	ZNF258	1.23E-03	1.453	9.58E-01
MmugDNA.23022.1.S1_s_at	55233	MOBK1B	3.10E-05	1.455	9.96E-01
MmugDNA.40059.1.S1_at	148932	MOBKL2C	3.69E-04	1.455	9.85E-01
MmuSTS.379.1.S1_at	3965	LGALS9	4.96E-04	1.456	9.67E-01
MmugDNA.14828.1.S1_at	10972	TMP21	1.75E-05	1.460	9.93E-01
MmugDNA.27555.1.S1_at	134359		7.25E-04	1.461	9.59E-01
MmugDNA.24713.1.S1_at	64318	C10orf117	1.01E-04	1.461	9.83E-01
MmugDNA.5873.1.S1_at	55165		3.57E-06	1.462	1.00E+00
MmuSTS.1581.1.S1_at	10788	IQGAP2	2.00E-04	1.463	9.85E-01
MmugDNA.1988.1.S1_at	1317	SLC31A1	6.25E-05	1.465	9.93E-01
MmugDNA.32991.1.S1_at	133121	ENPP6	3.07E-04	1.465	9.82E-01
MmuSTS.3533.1.S1_at	811	CALR	2.38E-04	1.465	9.71E-01
MmugDNA.14380.1.S1_at	967	CD63	3.94E-05	1.467	9.96E-01
MmugDNA.42902.1.S1_at	114885	OSBPL11	8.43E-05	1.469	9.81E-01
MmuSTS.2363.1.S1_at	29940	SART2	1.13E-03	1.471	9.52E-01
MmugDNA.14404.1.S1_at	8832	CD84	1.18E-03	1.471	9.67E-01
MmugDNA.17943.1.S1_at	3955	LFNG	3.91E-04	1.472	9.71E-01

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmugDNA.30164.1.S1_at	54962		5.34E-06	1.472	9.99E-01
MmugDNA.5841.1.S1_at	85025	TMEM60	3.29E-04	1.474	9.70E-01
MmugDNA.6277.1.S1_at	699	BUB1	1.40E-07	1.475	1.00E+00
MmugDNA.32603.1.S1_at	3017	HIST1H2BD	1.36E-07	1.476	1.00E+00
MmugDNA.42920.1.S1_at	3021		6.59E-06	1.477	9.95E-01
MmugDNA.5165.1.S1_at	1894	ECT2	6.05E-04	1.478	9.80E-01
MmuSTS.3946.1.S1_at	84541	TAKRP;	1.07E-04	1.480	9.78E-01
MmugDNA.29934.1.S1_at	1052	CEBPD	2.42E-04	1.481	9.80E-01
MmugDNA.5783.1.S1_at	7098	TLR3	1.43E-05	1.481	9.99E-01
MmugDNA.20899.1.S1_at	1030	CDKN2B	7.70E-05	1.482	9.97E-01
MmugDNA.21162.1.S1_at	3707	ITPKB	2.81E-04	1.482	9.68E-01
MmugDNA.2781.1.S1_at	55843	ARHGAP15	1.27E-04	1.484	9.92E-01
MmugDNA.2376.1.S1_at	5341	PLEK	1.47E-03	1.489	9.49E-01
MmugDNA.12415.1.S1_at	79713		5.74E-04	1.489	9.75E-01
MmugDNA.8651.1.S1_at	51312	SLC25A37	2.18E-04	1.489	9.88E-01
MmugDNA.39626.1.S1_at	284217	LAMA1	4.06E-04	1.490	9.70E-01
MmuSTS.3762.1.S1_at	5992	RFX4	2.36E-04	1.492	9.83E-01
MmugDNA.5028.1.S1_at	285195	SLC9A9	2.73E-04	1.494	9.80E-01
MmugDNA.15181.1.S1_at	222171		1.17E-03	1.495	9.53E-01
MmugDNA.13533.1.S1_at	6892	TAPBP	1.37E-03	1.496	9.52E-01
MmuSTS.2860.1.S1_at	6678	SPARC	3.97E-06	1.498	9.97E-01
MmugDNA.40144.1.S1_at	7127	TNFAIP2	2.31E-04	1.498	9.82E-01
MmugDNA.33773.1.S1_at	4234	METTL1	9.68E-05	1.500	9.93E-01
Mmu.13961.1.S1_at	25934		2.40E-04	1.501	9.69E-01
MmugDNA.3480.1.S1_at	145389	SLC38A6	1.04E-04	1.501	9.93E-01
MmugDNA.12386.1.S1_at	79711	IPO4	6.86E-04	1.502	9.82E-01
Mmu.12285.1.S1_at	25816	TNFAIP8	5.26E-04	1.505	9.80E-01
MmugDNA.1679.1.S1_at	5266	PI3	1.44E-03	1.506	9.65E-01
MmuSTS.2509.1.S1_at	55766	H2AFJ	4.66E-04	1.508	9.85E-01
MmugDNA.5017.1.S1_at	4488	MSX2	4.61E-04	1.508	9.87E-01
MmugDNA.32180.1.S1_at	2313	FLI1	3.97E-04	1.509	9.72E-01
MmugDNA.27238.1.S1_at	5947	RBP1	6.32E-04	1.510	9.76E-01
MmugDNA.41093.1.S1_at	168537	GIMAP7	1.85E-05	1.513	9.98E-01
MmugDNA.25779.1.S1_at	66002	CYP4F12	2.22E-04	1.514	9.87E-01
MmugDNA.5059.1.S1_at	7042	TGFB2	2.61E-05	1.517	9.96E-01
MmugDNA.19727.1.S1_at	57007	CMKOR1	1.67E-04	1.519	9.93E-01
MmugDNA.27037.1.S1_at	51338	MS4A4A	8.52E-07	1.519	1.00E+00
MmugDNA.13217.1.S1_at	121457		7.19E-06	1.520	9.99E-01
MmuSTS.165.1.S1_at	3399	ID3	2.96E-05	1.523	9.96E-01
MmugDNA.13285.1.S1_at	5292	PIM1	6.34E-04	1.524	9.85E-01
MmuSTS.2912.1.S1_at	3566	IL4R	1.48E-04	1.524	9.79E-01
MmuSTS.513.1.S1_at	57205	ATP10D	2.71E-04	1.524	9.77E-01
MmugDNA.41170.1.S1_at	4258	MGST2	1.88E-04	1.525	9.81E-01

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmuSTS.1445.1.S1_at	10019	LNK	3.45E-05	1.525	9.96E-01
MmugDNA.31966.1.S1_at	79887		3.67E-04	1.527	9.70E-01
MmunewRS.500.1.S1_at	728641		8.30E-04	1.527	9.75E-01
MmugDNA.25625.1.S1_at	79626		7.88E-04	1.528	9.62E-01
MmugDNA.14327.1.S1_at	114294	LACTB	2.59E-06	1.531	9.99E-01
Mmu.4440.1.S1_at	55300	PI4K2B	1.94E-06	1.531	9.99E-01
MmuSTS.4364.1.S1_at	6646	SOAT1	7.90E-08	1.536	1.00E+00
MmuSTS.3165.1.S1_at	25825	BACE2	1.06E-04	1.536	9.92E-01
Mmu.13804.1.S1_s_at	5052	PRDX1	1.91E-06	1.539	9.98E-01
Mmu.2018.1.S1_at	3109	HLADMB;	3.85E-05	1.539	9.93E-01
MmugDNA.11557.1.S1_at	64151	HCAPG;	6.81E-07	1.542	1.00E+00
MmugDNA.16295.1.S1_s_at	1029	CDKN2A	5.77E-04	1.544	9.85E-01
MmugDNA.7381.1.S1_s_at	7940	LST1	1.11E-03	1.544	9.65E-01
MmugDNA.43133.1.S1_at	8382	NME5	6.58E-06	1.545	9.99E-01
MmuSTS.3358.1.S1_at	947	CD34	2.83E-05	1.548	9.97E-01
MmuSTS.249.1.S1_at	3689	ITGB2	1.61E-03	1.551	9.51E-01
MmuSTS.676.1.S1_at	50615	IL21R	2.88E-04	1.553	9.93E-01
MmuAffx.161.1.S1_at	341	APOC1	6.78E-04	1.554	9.61E-01
MmugDNA.32458.1.S1_at	55577	NAGK	5.17E-06	1.555	9.99E-01
MmuSTS.737.1.S1_at	3717	JAK2	3.97E-04	1.555	9.78E-01
MmugDNA.34184.1.S1_at	114908	PORIMIN	4.85E-05	1.558	9.94E-01
MmunewRS.400.1.S1_at	4129	MAOB	5.12E-07	1.559	1.00E+00
MmugDNA.26934.1.S1_at	253635		5.89E-06	1.560	9.98E-01
MmugDNA.12374.1.S1_at	51251	NT5C3	3.27E-05	1.560	9.95E-01
MmugDNA.40313.1.S1_at	4017	LOXL2	8.48E-05	1.561	9.95E-01
MmugDNA.22836.1.S1_s_at	3006	HIST1H1C	3.27E-05	1.562	9.96E-01
MmugDNA.33437.1.S1_s_at	55905	ZNF313	2.83E-04	1.563	9.70E-01
MmugDNA.37853.1.S1_at	91947		1.78E-04	1.565	9.87E-01
MmugDNA.9854.1.S1_at	241	ALOX5AP	2.13E-05	1.566	9.92E-01
MmuSTS.3531.1.S1_s_at	6351	CCL4	1.57E-04	1.567	9.77E-01
MmugDNA.41036.1.S1_at	26509	FER1L3	3.94E-04	1.567	9.96E-01
MmugDNA.42759.1.S1_at	64757		1.08E-03	1.574	9.53E-01
MmuSTS.1556.1.S1_at	4148	MATN3	7.47E-08	1.578	1.00E+00
MmugDNA.11858.1.S1_at	60484	HAPLN2	1.05E-03	1.578	9.62E-01
MmugDNA.24513.1.S1_s_at	8741		2.29E-05	1.579	9.97E-01
MmugDNA.32060.1.S1_s_at	60559	SPCS3	3.22E-05	1.581	9.97E-01
MmugDNA.6959.1.S1_at	10039	PARP3	3.81E-04	1.583	9.83E-01
MmugDNA.152.1.S1_at	966	CD59	2.43E-04	1.585	9.60E-01
MmugDNA.28820.1.S1_at	342035	COLM	4.38E-04	1.590	9.73E-01
MmugDNA.23099.1.S1_at	5777	PTPN6	1.50E-04	1.592	9.94E-01
MmugDNA.21562.1.S1_at	5450	POU2AF1	1.10E-05	1.593	9.98E-01
MmugDNA.1828.1.S1_at	3071	HEM1	5.47E-04	1.593	9.79E-01
Mmu.10973.1.S1_at	8673	VAMP8	2.54E-04	1.594	9.85E-01

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmugDNA.40799.1.S1_at	8737	RIPK1	5.35E-05	1.594	9.98E-01
MmuSTS.1537.1.S1_at	10184		2.65E-05	1.597	9.95E-01
MmuSTS.4663.1.S1_at	7133	TNFRSF1B	3.69E-04	1.606	9.88E-01
MmugDNA.25554.1.S1_at	11067		2.02E-04	1.607	9.86E-01
MmugDNA.18235.1.S1_at	60436	TGIF2	2.00E-04	1.607	9.89E-01
MmugDNA.29436.1.S1_at	8935	SCAP2	2.87E-04	1.608	9.82E-01
MmugDNA.732.1.S1_at	3074	HEXB	8.25E-06	1.609	9.99E-01
MmugDNA.11758.1.S1_at	22809	ATF5	2.31E-04	1.609	9.88E-01
MmugDNA.41579.1.S1_at	340205		5.26E-04	1.610	9.65E-01
MmugDNA.22360.1.S1_s_at	5699	PSMB10	2.08E-04	1.612	9.92E-01
MmugDNA.11781.1.S1_at	79686		6.78E-05	1.614	9.92E-01
Mmu.8637.1.S1_at	347	APOD	1.53E-05	1.616	9.97E-01
MmuSTS.88.1.S1_at	8349	HIST2H2BE	1.51E-05	1.619	9.99E-01
MmugDNA.6706.1.S1_at	26586		2.33E-05	1.620	9.98E-01
MmuSTS.2246.1.S1_at	5336	PLCG2	4.47E-04	1.621	9.82E-01
MmuSTS.4658.1.S1_at	4953	ODC1	2.21E-06	1.622	9.99E-01
MmugDNA.28015.1.S1_at	23325		1.14E-05	1.625	9.99E-01
MmugDNA.3160.1.S1_s_at	58527		2.43E-05	1.627	9.96E-01
MmunewRS.203.1.S1_x_at	8970	HIST1H2BJ	3.02E-06	1.629	1.00E+00
MmugDNA.42101.1.S1_at	196410	MGC17301	1.53E-05	1.630	9.97E-01
MmugDNA.6653.1.S1_at	23424	TDRD7	5.50E-06	1.630	9.98E-01
MmuSTS.4012.1.S1_at	11046	SLC35D2	1.87E-04	1.637	9.92E-01
MmugDNA.29473.1.S1_at	54440		1.07E-03	1.639	9.74E-01
MmuSTS.2849.1.S1_at	397	ARHGDIB	4.09E-05	1.639	9.94E-01
Mmu.629.2.S1_at	1075	CTSC	1.01E-04	1.640	9.94E-01
MmugDNA.1040.1.S1_at	375287	FLJ45645	1.43E-04	1.640	9.94E-01
MmugDNA.2122.1.S1_at	2040	STOM	6.82E-06	1.641	9.98E-01
Mmu.11959.1.A1_at	2634	GBP2	2.01E-07	1.644	1.00E+00
MmugDNA.38956.1.S1_at	51203	NUSAP1	2.74E-07	1.645	1.00E+00
MmugDNA.15037.1.S1_at	5937	RBMS1	1.74E-05	1.645	9.98E-01
MmugDNA.39211.1.S1_at	639	PRDM1	1.61E-05	1.645	9.99E-01
MmugDNA.34948.1.S1_at	2799	GNS	8.03E-07	1.646	9.99E-01
MmugDNA.12530.1.S1_at	9871	SEC24D	1.68E-06	1.646	1.00E+00
MmuSTS.1881.1.S1_at	1293	COL6A3	2.20E-05	1.647	9.97E-01
MmuSTS.3988.1.S1_at	1520	CTSS	1.50E-04	1.648	9.88E-01
Mmu.1161.1.S1_at	3553	IL1B	2.35E-06	1.648	1.00E+00
MmugDNA.35226.1.S1_at	80162		3.18E-04	1.649	9.84E-01
MmuSTS.3819.1.S1_at	25805	BAMBI	1.04E-03	1.650	9.76E-01
MmugDNA.167.1.S1_at	2876	GPX1	4.62E-04	1.652	9.85E-01
MmugDNA.28497.1.S1_at	9473	C1orf38	2.86E-04	1.653	9.86E-01
MmugDNA.207.1.S1_at	55640	C14orf58	3.18E-06	1.653	1.00E+00
MmugDNA.34778.1.S1_s_at	7764	ZNF217	1.61E-06	1.654	9.99E-01
MmugDNA.14691.1.S1_at	1193	CLIC2	5.56E-06	1.655	9.99E-01

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmugDNA.5661.1.S1_at	23070	KIAA0082	4.78E-07	1.655	1.00E+00
MmugDNA.23424.1.S1_at	4792	NFKBIA	5.86E-06	1.658	9.99E-01
MmugDNA.32059.1.S1_at	316	AOX1	1.44E-04	1.658	9.90E-01
MmugDNA.27664.1.S1_at	9535	GMFG	8.22E-05	1.659	9.92E-01
MmugDNA.37069.1.S1_at	1807	DPYS	3.98E-04	1.666	9.93E-01
MmunewRS.633.1.S1_s_at	164668	ARP10	2.94E-06	1.667	9.99E-01
MmuSTS.1377.1.S1_at	79370	BCL2L14	1.69E-07	1.668	1.00E+00
MmugDNA.28260.1.S1_at	8644	AKR1C3	5.09E-04	1.673	9.77E-01
Mmu.14249.1.S1_at	727797		1.14E-07	1.673	1.00E+00
MmugDNA.28488.1.S1_at	340348	MGC50844	1.23E-06	1.674	1.00E+00
MmugDNA.22603.1.S1_at	11314	CD300A	1.99E-04	1.677	9.95E-01
MmugDNA.15589.1.S1_at	56944		2.08E-03	1.678	9.61E-01
MmugDNA.16581.1.S1_at	1266	CNN3	8.13E-06	1.681	9.98E-01
MmuSTS.1798.1.S1_at	220963	SLC16A9	1.94E-05	1.683	9.98E-01
MmugDNA.21105.1.S1_at	3978	LIG1	1.19E-06	1.684	1.00E+00
MmuSTS.1106.1.S1_at	5175	PECAM1	4.08E-05	1.684	9.98E-01
MmugDNA.5880.1.S1_at	7132	TNFRSF1A	3.96E-06	1.687	1.00E+00
MmugDNA.1120.1.S1_at	2202	EFEMP1	2.98E-04	1.687	9.96E-01
MmuSTS.3655.1.S1_at	5327	PLAT	1.47E-05	1.687	9.99E-01
MmugDNA.6224.1.S1_at	56344	CABP5	5.07E-05	1.688	9.96E-01
MmugDNA.40394.1.S1_at	4061	LY6E	3.37E-06	1.689	9.98E-01
MmugDNA.23317.1.S1_at	1992	SERPINB1	3.70E-06	1.691	9.99E-01
MmugDNA.13531.1.S1_at	822	CAPG	9.64E-05	1.693	9.90E-01
MmugDNA.43514.1.S1_at	167410	LIX1	1.23E-04	1.695	9.91E-01
MmuSTS.4526.1.S1_at	6850	SYK	5.96E-04	1.698	9.81E-01
MmugDNA.33195.1.S1_at	10161	P2RY5	2.97E-05	1.701	9.98E-01
MmuSTS.681.1.S1_at	3557	IL1RN	9.47E-05	1.702	9.98E-01
MmugDNA.11430.1.S1_at	139818	DOCK11	1.22E-05	1.703	9.99E-01
MmuSTS.3575.1.S1_at	6354	CCL7	3.03E-03	1.703	9.82E-01
MmugDNA.11007.1.S1_at	79630		2.48E-06	1.704	1.00E+00
MmugDNA.21820.1.S1_at	257106	ARHGAP30	2.21E-05	1.704	9.99E-01
MmugDNA.4052.1.S1_at	54507	TSRC1	1.11E-05	1.707	1.00E+00
MmuSTS.2402.1.S1_at	861	RUNX1	1.19E-04	1.707	9.97E-01
MmugDNA.34019.1.S1_at	93082		1.74E-05	1.709	9.99E-01
MmugDNA.28244.1.S1_at	2120	ETV6	1.96E-05	1.710	9.98E-01
MmugDNA.16474.1.S1_s_at	56925	LXN	2.47E-04	1.711	9.86E-01
MmugDNA.23693.1.S1_at	5836	PYGL	6.82E-05	1.712	9.97E-01
MmugDNA.5794.1.S1_at	441019		1.89E-08	1.715	1.00E+00
MmuSTS.3635.1.S1_at	51474	EPLIN	3.92E-06	1.719	9.99E-01
MmugDNA.13787.1.S1_at	126969		2.65E-05	1.721	9.99E-01
MmugDNA.6561.1.S1_at	146722	CD300LF	1.30E-04	1.722	9.98E-01
MmuSTS.1347.1.S1_at	3148	HMGB2	2.60E-06	1.723	9.99E-01
Mmu.12235.2.S1_at	221178	SPATA13338872	3.30E-05	1.728	9.93E-01

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmuSTS.174.1.S1_at	3426	IF	2.78E-05	1.728	9.98E-01
MmugDNA.7797.1.S1_at	2219	FCN1	2.55E-03	1.729	9.66E-01
MmugDNA.40742.1.S1_at	9133	CCNB2	3.93E-08	1.729	1.00E+00
MmugDNA.41518.1.S1_s_at	7431	VIM	5.96E-06	1.730	9.99E-01
MmugDNA.16011.1.S1_at	285440	CYP4V2	1.32E-05	1.730	9.97E-01
MmugDNA.37813.1.S1_at	57561		1.07E-05	1.732	9.97E-01
MmugDNA.35730.1.S1_at	64092		4.00E-11	1.736	1.00E+00
MmugDNA.27959.1.S1_at	2533	FYB	1.16E-04	1.737	9.94E-01
MmuSTS.3400.1.S1_at	1026	CDKN1A	5.66E-04	1.745	9.88E-01
MmugDNA.22331.1.S1_s_at	6503	SLA	7.54E-06	1.747	9.99E-01
MmuSTS.1343.1.S1_at	123	ADFP	2.85E-06	1.747	1.00E+00
MmuSTS.3224.1.S1_at	4542	MYO1F	1.39E-03	1.750	9.82E-01
MmuSTS.4674.1.S1_at	7157	TP53	9.01E-06	1.752	9.99E-01
MmuSTS.3937.1.S1_at	6778	STAT6	3.14E-05	1.752	9.99E-01
MmuSTS.3830.1.S1_at	9120	SLC16A6	2.88E-05	1.756	9.99E-01
MmugDNA.8456.1.S1_at	11015	KDELR3	3.74E-07	1.757	1.00E+00
MmugDNA.24231.1.S1_at	23607	CD2AP	1.07E-05	1.760	9.99E-01
MmuSTS.549.1.S1_at	580	BARD1	7.47E-07	1.761	1.00E+00
MmugDNA.30129.1.S1_at	8743	TNFSF10	2.44E-06	1.762	9.99E-01
MmugDNA.26254.1.S1_at	3383	ICAM1	1.38E-05	1.762	9.99E-01
Mmu.8945.1.S1_at	3075	CFH	9.46E-05	1.762	9.82E-01
MmuSTS.4659.1.S1_at	7128	TNFAIP3	7.04E-07	1.765	1.00E+00
MmuSTS.4058.1.S1_at	8406	SRPX	1.51E-05	1.772	9.96E-01
MmuSTS.3971.1.S1_at	1508	CTSB	5.31E-08	1.772	1.00E+00
MmugDNA.26598.1.S1_at	1316	KLF6	1.43E-05	1.776	9.99E-01
MmugDNA.6430.1.S1_at	9308	CD83	1.61E-05	1.777	9.99E-01
MmugDNA.7359.1.S1_s_at	7305	TYROBP	1.30E-03	1.778	9.59E-01
MmugDNA.23207.1.S1_at	160364	CLEC12A	6.65E-05	1.779	9.98E-01
MmuSTS.1925.1.S1_at	2152	F3	3.27E-07	1.779	1.00E+00
MmugDNA.17365.1.S1_at	6275	S100A4	1.09E-04	1.781	9.95E-01
MmuSTS.4661.1.S1_at	4982	TNFRSF11B	1.26E-04	1.782	9.98E-01
MmugDNA.25436.1.S1_at	4172	MCM3	1.53E-07	1.782	1.00E+00
MmuSTS.2821.1.S1_at	9404	LPXN	8.19E-06	1.783	9.99E-01
MmugDNA.22158.1.S1_at	8407	TAGLN2	2.34E-05	1.787	9.97E-01
MmugDNA.3746.1.S1_at	56670	SUCNR1	1.20E-03	1.790	9.93E-01
MmugDNA.19512.1.S1_at	3371	TNC	4.03E-06	1.791	1.00E+00
MmugDNA.28918.1.S1_at	196527		4.75E-06	1.791	9.99E-01
MmugDNA.10692.1.S1_at	51246	SCOTIN	1.32E-06	1.796	1.00E+00
MmuSTS.4567.1.S1_at	9450	LY86	5.24E-04	1.800	9.73E-01
MmugDNA.17682.1.S1_at	6563	SLC14A1	7.60E-04	1.802	9.84E-01
MmugDNA.17986.1.S1_s_at	401494		1.15E-05	1.802	9.98E-01
MmugDNA.14342.1.S1_at	951	CD37	1.81E-04	1.804	9.89E-01
MmuSTS.3742.1.S1_at	5806	PTX3	5.00E-04	1.811	9.95E-01

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmuSTS.3129.1.S1_at	6004	RGS16	1.57E-02	1.814	9.47E-01
MmugDNA.40137.1.S1_at	1997	ELF1	3.90E-05	1.816	9.96E-01
MmugDNA.1096.1.S1_at	9961	MVP	5.59E-05	1.818	9.98E-01
MmugDNA.35565.1.S1_at	7805	LAPTM5	8.44E-05	1.825	9.92E-01
MmugDNA.16536.1.S1_at	3597	IL13RA1	1.08E-05	1.828	9.96E-01
MmugDNA.22096.1.S1_at	3956	LGALS1	1.18E-04	1.830	9.91E-01
MmugDNA.38698.1.S1_at	22918	C1QR1	3.98E-06	1.836	1.00E+00
MmuSTS.700.1.S1_at	23601	CLEC5A	2.16E-04	1.836	9.91E-01
MmugDNA.43509.1.S1_at	388389		4.22E-04	1.843	9.95E-01
MmuSTS.3582.1.S1_at	4267	CD994267	1.02E-08	1.847	1.00E+00
MmuSTS.2130.1.S1_s_at	6280	S100A9	1.05E-03	1.848	9.69E-01
MmugDNA.17159.1.S1_s_at	9603	NFE2L3	9.47E-07	1.848	1.00E+00
MmuSTS.2579.1.S1_at	10791	VAMP5	4.25E-06	1.848	1.00E+00
MmuSTS.3015.1.S1_at	1728	NQO1	3.09E-05	1.849	9.99E-01
MmugDNA.33529.1.S1_at	642061		9.07E-07	1.860	1.00E+00
MmugDNA.16377.1.S1_at	4671	BIRC1	4.16E-06	1.861	1.00E+00
MmuSTS.4816.1.S1_at	1647	GADD45A	2.44E-05	1.863	9.97E-01
MmuSTS.721.1.S1_at	80896	NPL	5.20E-09	1.871	1.00E+00
MmugDNA.26355.1.S1_at	837	CASP4	5.04E-07	1.875	1.00E+00
MmugDNA.21905.1.S1_at	170954		4.29E-06	1.876	9.99E-01
Mmu.13018.1.S1_at	85363	TRIM5	1.04E-06	1.877	1.00E+00
MmugDNA.21377.1.S1_at	4343	MOV10	5.23E-06	1.877	1.00E+00
MmugDNA.5064.1.S1_at	7045	TGFBI	3.98E-03	1.878	9.59E-01
Mmu.8766.1.S1_at	51667	NYREN18	2.08E-06	1.878	1.00E+00
MmugDNA.30721.1.S1_at	1545	CYP1B1	2.85E-05	1.882	9.98E-01
MmugDNA.13403.1.S1_at	728	C5R1	1.12E-04	1.884	9.96E-01
MmuSTS.3349.1.S1_at	727811		4.68E-08	1.884	1.00E+00
MmugDNA.18224.1.S1_at	8477	GPR65	9.19E-10	1.886	1.00E+00
MmugDNA.10824.1.S1_at	145270	PRIMA1	7.45E-05	1.886	9.98E-01
MmugDNA.4698.1.S1_at	10241	NDP52	1.51E-08	1.891	1.00E+00
MmuSTS.2433.1.S1_at	81844	TRIM56	9.65E-07	1.898	1.00E+00
MmuSTS.646.1.S1_at	3601	IL15RA	9.26E-07	1.902	1.00E+00
MmugDNA.22091.1.S1_at	3684	ITGAM	4.38E-06	1.902	9.99E-01
MmugDNA.41794.1.S1_at	4478	MSN	4.72E-07	1.909	1.00E+00
MmugDNA.15515.1.S1_at	2745	GLRX	2.28E-08	1.910	1.00E+00
MmuSTS.1849.1.S1_at	952	CD38	3.61E-04	1.922	9.84E-01
MmuSTS.4740.1.S1_at	2162	F13A1	2.95E-08	1.924	1.00E+00
MmugDNA.32019.1.S1_at	401433		8.15E-08	1.931	1.00E+00
MmugDNA.9266.1.S1_at	1890	ECGF1	2.60E-04	1.931	9.95E-01
MmugDNA.14897.1.S1_at	558	AXL	1.03E-05	1.934	9.99E-01
Mmu.3058.1.S1_at	6774	STAT3	1.72E-04	1.942	9.61E-01
MmugDNA.7695.1.S1_at	24138	IFIT5	7.10E-07	1.943	1.00E+00
MmugDNA.13690.1.S1_at	51303	FKBP11	8.76E-07	1.947	1.00E+00

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmuSTS.2906.1.S1_at	7450	VWF	5.18E-04	1.950	9.94E-01
MmugDNA.10481.1.S1_at	9595	PSCDBP	8.62E-07	1.959	1.00E+00
Mmu.2883.1.S1_at	730	C7	7.18E-05	1.964	9.97E-01
MmuSTS.1991.1.S1_at	9180	OSMR	7.43E-07	1.966	1.00E+00
MmugDNA.1987.1.S1_at	6518	SLC2A5	3.54E-06	1.979	9.99E-01
MmugDNA.2786.1.S1_at	6688	SPI1	3.34E-06	1.980	1.00E+00
MmugDNA.9809.1.S1_s_at	1051	CEBPB	1.77E-08	1.989	1.00E+00
MmuSTS.4521.1.S1_at	6773	STAT2	2.59E-06	1.998	1.00E+00
MmuSTS.1770.1.S1_at	6281	S100A10	5.55E-06	1.999	1.00E+00
MmuSTS.3453.1.S1_at	29108	PYCARD	2.00E-04	2.000	9.96E-01
MmuSTS.2144.1.S1_at	9531	BAG3	7.96E-09	2.000	1.00E+00
MmugDNA.9727.1.S1_at	222256	FLJ23834	2.00E-09	2.010	1.00E+00
MmugDNA.38481.1.S1_at	2012	EMP1	1.75E-07	2.011	1.00E+00
MmugDNA.328.1.S1_at	26471	P8	1.94E-05	2.018	9.99E-01
MmuSTS.617.1.S1_at	8870	IER3	3.04E-07	2.021	1.00E+00
MmugDNA.11623.1.S1_at	24145	PANX1	5.22E-07	2.021	1.00E+00
MmugDNA.13846.1.S1_at	51330	TNFRSF12A	1.13E-04	2.022	1.00E+00
MmugDNA.9235.1.S1_s_at	103	ADAR	8.93E-12	2.026	1.00E+00
MmuSTS.664.1.S1_at	840	CASP7	4.63E-09	2.036	1.00E+00
MmugDNA.42026.1.S1_at	155038	GIMAP8	2.74E-06	2.036	1.00E+00
MmuSTS.4685.1.S1_at	7298	TYMS	3.04E-07	2.039	1.00E+00
MmugDNA.32051.1.S1_at	7048	TGFBR2	1.26E-08	2.039	1.00E+00
MmugDNA.5964.1.S1_at	197135		4.75E-11	2.044	1.00E+00
MmugDNA.1832.1.S1_at	4688	NCF2	7.77E-07	2.048	1.00E+00
MmugDNA.28886.1.S1_at	115004		9.83E-06	2.055	1.00E+00
MmuSTS.3060.1.S1_at	55647	RAB20	2.79E-06	2.055	1.00E+00
MmugDNA.5610.1.S1_at	152007	C9orf19	3.57E-05	2.055	9.99E-01
MmugDNA.5827.1.S1_at	3937	LCP2	6.89E-08	2.065	1.00E+00
MmugDNA.23080.1.S1_at	83937	RASSF4	7.65E-07	2.065	1.00E+00
MmugDNA.6970.1.S1_at	116441	TM4SF18	1.88E-09	2.067	1.00E+00
MmugDNA.12043.1.S1_at	219285		1.38E-05	2.069	1.00E+00
MmugDNA.18189.1.S1_at	3487	IGFBP4	3.89E-05	2.081	9.99E-01
MmugDNA.30554.1.S1_at	10095	ARPC1B	3.90E-06	2.086	1.00E+00
MmuSTS.486.1.S1_at	27350	APOBEC3C	2.51E-05	2.090	1.00E+00
MmuSTS.4836.1.S1_at	2760	GM2A	1.57E-05	2.109	9.99E-01
MmuSTS.92.1.S1_at	3162	HMOX1	5.39E-06	2.115	1.00E+00
MmugDNA.35760.1.S1_at	64108	IFRG28	3.85E-09	2.119	1.00E+00
MmugDNA.18099.1.S1_at	653879		1.07E-04	2.120	9.97E-01
MmugDNA.20554.1.S1_at	4067	LYN	3.18E-08	2.124	1.00E+00
MmugDNA.38506.1.S1_at	2014	EMP3	3.97E-05	2.131	9.99E-01
MmugDNA.10526.1.S1_at	301	ANXA1	5.62E-07	2.141	1.00E+00
MmugDNA.6929.1.S1_s_at	654816		1.20E-05	2.145	1.00E+00
MmugDNA.31245.1.S1_at	10385	BTN2A2	7.25E-09	2.145	1.00E+00

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
Mmu.10941.2.S1_at	9056	SLC7A7	1.01E-08	2.148	1.00E+00
MmugDNA.26925.1.S1_s_at	2207	FCER1G	1.31E-06	2.149	1.00E+00
MmugDNA.40170.1.S1_s_at	10537	UBD	1.04E-03	2.155	9.96E-01
MmuSTS.1500.1.S1_at	4170	MCL1	1.60E-08	2.158	1.00E+00
MmugDNA.38530.1.S1_at	6614	SN	3.67E-06	2.163	1.00E+00
MmuSTS.1895.1.S1_at	2920	CXCL2	3.65E-10	2.166	1.00E+00
MmugDNA.35794.1.S1_at	133	ADM	3.16E-06	2.176	1.00E+00
MmugDNA.5901.1.S1_at	355	FAS	6.55E-09	2.188	1.00E+00
MmuSTS.3928.1.S1_at	1282	COL4A1	2.23E-07	2.195	1.00E+00
MmugDNA.25634.1.S1_at	7453	WARS	2.69E-05	2.211	1.00E+00
MmugDNA.4914.1.S1_s_at	302	ANXA2	3.17E-07	2.219	1.00E+00
MmugDNA.33372.1.S1_at	162394		1.00E-04	2.221	9.96E-01
MmugDNA.24291.1.S1_s_at	5265	SERPINA1	5.12E-05	2.234	1.00E+00
MmugDNA.17370.1.S1_s_at	6277	S100A6	1.02E-06	2.237	1.00E+00
MmuSTS.4005.1.S1_at	7058	THBS2	7.17E-06	2.239	9.93E-01
MmugDNA.41156.1.S1_at	56245	C21orf62	1.41E-04	2.242	1.00E+00
MmugDNA.10434.1.S1_at	55603		3.38E-09	2.245	1.00E+00
MmugDNA.38723.1.S1_at	10863	ADAM28	1.10E-06	2.246	1.00E+00
MmugDNA.1812.1.S1_at	81030	ZBP1	2.49E-06	2.249	1.00E+00
MmuSTS.1890.1.S1_at	1522	CTSZ	9.23E-05	2.258	9.97E-01
MmugDNA.9213.1.S1_at	55337		5.01E-09	2.274	1.00E+00
MmuSTS.3489.1.S1_at	968	CD68	5.40E-07	2.278	1.00E+00
MmuSTS.3532.1.S1_at	963	CD53	2.39E-08	2.279	1.00E+00
MmugDNA.3190.1.S1_at	80216	ALPK1	1.04E-09	2.281	1.00E+00
MmuSTS.4595.1.S1_at	7096	TLR1	1.93E-08	2.283	1.00E+00
MmuSTS.3275.1.S1_at	6241	RRM2	2.22E-08	2.292	1.00E+00
MmugDNA.33060.1.S1_x_at	10288	LILRB2	2.43E-04	2.310	9.97E-01
MmugDNA.11765.1.S1_at	118932		1.17E-04	2.323	1.00E+00
Mmu.3782.2.S1_a_at	3588	IL10RB	1.53E-08	2.331	1.00E+00
MmuSTS.2862.1.S1_at	6696	SPP1	9.45E-10	2.338	1.00E+00
MmugDNA.17376.1.S1_at	6279	S100A8	2.11E-03	2.366	9.70E-01
MmugDNA.30800.1.S1_at	25939	SAMHD1	1.13E-09	2.370	1.00E+00
MmugDNA.26746.1.S1_at	83706	URP2	9.30E-07	2.372	1.00E+00
MmuSTS.1982.1.S1_at	929	CD14	4.24E-05	2.372	1.00E+00
MmuSTS.4308.1.S1_s_at	3059	HCLS1	3.78E-06	2.378	1.00E+00
MmugDNA.20313.1.S1_at	3587	IL10RA	2.76E-07	2.384	1.00E+00
MmugDNA.4421.1.S1_at	9935	MAFB	1.89E-07	2.386	1.00E+00
MmuSTS.2126.1.S1_s_at	445	ASS	7.51E-07	2.394	1.00E+00
Mmu.16393.1.S1_at	330	BIRC3	2.32E-07	2.411	1.00E+00
MmugDNA.40598.1.S1_at	2532	FY	1.86E-06	2.415	1.00E+00
MmugDNA.7314.1.S1_at	85441	PRIC285	7.47E-08	2.425	1.00E+00
MmugDNA.25492.1.S1_s_at	9582	APOBEC3B	1.23E-09	2.427	1.00E+00
MmugDNA.13395.1.S1_at	3428	IFI16	1.60E-08	2.436	1.00E+00

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmuSTS.991.1.S1_at	5880	RAC2	1.82E-06	2.439	1.00E+00
MmugDNA.40434.1.S1_at	143630	MGC20470	2.19E-07	2.443	1.00E+00
MmugDNA.7109.1.S1_at	6672	SP100	1.72E-08	2.446	1.00E+00
MmugDNA.36497.1.S1_at	58472	SQRDL	3.49E-08	2.453	1.00E+00
MmugDNA.11614.1.S1_at	23643	LY96	1.26E-06	2.457	1.00E+00
MmugDNA.5658.1.S1_at	597	BCL2A1	5.32E-06	2.466	1.00E+00
MmugDNA.11631.1.S1_s_at	2919	CXCL1	5.45E-10	2.469	1.00E+00
MmugDNA.34145.1.S1_at	6993	TCTEL1	3.23E-09	2.469	1.00E+00
MmugDNA.21695.1.S1_at	5720	PSME1	1.57E-10	2.489	1.00E+00
MmugDNA.25916.1.S1_at	5359	PLSCR1	5.14E-07	2.492	1.00E+00
MmugDNA.32094.1.S1_at	3958	LGALS3	6.48E-06	2.499	1.00E+00
MmugDNA.19109.1.S1_at	3600	IL15	6.79E-06	2.500	1.00E+00
Mmu.10556.1.S1_s_at	9447	AIM2	2.44E-07	2.519	1.00E+00
MmugDNA.10394.1.S1_at	6041	RNASEL	2.04E-09	2.523	1.00E+00
MmugDNA.3046.1.S1_at	1192	CLIC1	1.30E-10	2.533	1.00E+00
MmugDNA.25384.1.S1_at	11065	UBE2C	2.26E-07	2.533	1.00E+00
MmugDNA.40602.1.S1_at	170575	GIMAP1	4.86E-06	2.535	1.00E+00
MmuSTS.615.1.S1_at	55340	GIMAP5	1.17E-06	2.556	1.00E+00
MmugDNA.4113.1.S1_at	26353	HSPB8	4.69E-11	2.562	1.00E+00
MmugDNA.1554.1.S1_at	11177	BAZ1A	1.36E-09	2.582	1.00E+00
MmuSTS.1193.1.S1_at	5366		2.56E-07	2.585	1.00E+00
MmugDNA.12273.1.S1_at	7153	TOP2A	6.24E-08	2.600	1.00E+00
MmuSTS.4059.1.S1_at	6737	TRIM21	1.96E-11	2.602	1.00E+00
MmugDNA.33823.1.S1_s_at	9768		1.25E-07	2.604	1.00E+00
MmugDNA.27352.1.S1_at	57169	KIAA1404	2.05E-09	2.609	1.00E+00
MmugDNA.4866.1.S1_at	493869	LOC493869	6.29E-08	2.616	1.00E+00
MmugDNA.35023.1.S1_s_at	116496	C1orf24	3.78E-07	2.616	1.00E+00
Mmu.10025.1.S1_at	1512	CTSH	1.44E-08	2.666	1.00E+00
MmuSTS.2359.1.S1_at	6282	S100A11	6.15E-08	2.669	1.00E+00
Mmu.13989.1.S1_at	51056	LAP3	9.99E-06	2.674	1.00E+00
MmugDNA.35261.1.S1_at	2212	FCGR2A	2.04E-08	2.694	1.00E+00
MmuSTS.1158.1.S1_at	51131	PHF11	1.07E-09	2.699	1.00E+00
MmugDNA.36434.1.S1_at	3001	GZMA	6.61E-04	2.702	9.98E-01
MmuSTS.20.1.S1_at	7462	LAT2	9.08E-08	2.704	1.00E+00
MmugDNA.34.1.S1_at	29760	BLNK	2.96E-08	2.709	1.00E+00
MmugDNA.14551.1.S1_at	51816	CECR1	4.13E-07	2.712	1.00E+00
MmugDNA.8496.1.S1_at	6363	CCL19	5.71E-05	2.740	9.99E-01
MmugDNA.7390.1.S1_at	7318	UBE1L	6.48E-07	2.754	1.00E+00
MmugDNA.11115.1.S1_s_at	79651		1.07E-08	2.772	1.00E+00
MmugDNA.38197.1.S1_at	440836	LOC440836	1.28E-06	2.786	1.00E+00
MmugDNA.9155.1.S1_s_at	55332		9.28E-09	2.792	1.00E+00
MmugDNA.36182.1.S1_at	358	AQP1	1.46E-05	2.795	9.99E-01
MmugDNA.24747.1.S1_at	4973	OLR1	1.72E-08	2.818	1.00E+00

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmugDNA.13638.1.S1_at	51296	SLC15A3	2.75E-07	2.840	1.00E+00
MmugDNA.18647.1.S1_s_at	57817	HAMP	1.53E-07	2.865	1.00E+00
MmuSTS.3131.1.S1_at	5996	RGS1	1.25E-06	2.872	1.00E+00
MmugDNA.10110.1.S1_at	6039	RNASE6	6.44E-06	2.883	1.00E+00
MmugDNA.15798.1.S1_s_at	3118	HLADQA2;	3.87E-05	2.905	1.00E+00
MmugDNA.31716.1.S1_s_at	55034		1.34E-08	2.907	1.00E+00
MmuSTS.1204.1.S1_at	7097	TLR2	7.67E-11	2.921	1.00E+00
MmugDNA.10008.1.S1_at	4615	MYD88	5.81E-10	2.942	1.00E+00
MmuSTS.2833.1.S1_at	80832	APOL4	4.17E-07	2.950	1.00E+00
MmuSTS.3702.1.S1_at	5610	EIF2AK2	1.69E-10	2.951	1.00E+00
MmugDNA.424.1.S1_at	120329	CASP12P1	2.88E-08	2.980	1.00E+00
MmuSTS.1560.1.S1_at	116071	MGC20410	5.64E-06	2.983	1.00E+00
MmugDNA.3247.1.S1_at	8363		4.22E-08	2.984	1.00E+00
MmugDNA.5628.1.S1_at	567	B2M	4.88E-15	2.986	1.00E+00
MmugDNA.34787.1.S1_at	2209	FCGR1A	8.85E-10	2.993	1.00E+00
MmugDNA.28808.1.S1_at	5788	PTPRC	3.94E-05	3.012	9.99E-01
MmugDNA.19485.1.S1_s_at	960	CD44	6.68E-07	3.062	1.00E+00
MmugDNA.19588.1.S1_at	92610		3.54E-06	3.079	1.00E+00
MmugDNA.30026.1.S1_at	10894	XLKD1	6.79E-05	3.129	9.99E-01
MmugDNA.12881.1.S1_at	9023	CH25H	4.36E-07	3.165	1.00E+00
MmugDNA.41339.1.S1_at	2643	GCH1	2.11E-07	3.165	1.00E+00
MmugDNA.40745.1.S1_at	197259	MLKL	4.89E-10	3.170	1.00E+00
MmugDNA.37154.1.S1_at	3113	HLADPA1;	1.86E-06	3.202	1.00E+00
MmugDNA.8529.1.S1_at	5552		8.35E-10	3.202	1.00E+00
MmugDNA.29625.1.S1_at	3665	IRF7	1.96E-07	3.220	1.00E+00
MmugDNA.39974.1.S1_at	26157	GIMAP2	5.84E-11	3.227	1.00E+00
MmuSTS.3375.1.S1_at	942	CD86	2.68E-08	3.242	1.00E+00
MmugDNA.5541.1.S1_at	50856	CLEC4A	1.81E-08	3.317	1.00E+00
MmuSTS.946.1.S1_at	57664	PLEKHA4	1.17E-06	3.333	1.00E+00
Mmu.8608.1.S2_at	6352	CCL5	7.63E-09	3.338	1.00E+00
MmugDNA.12153.1.S1_at	629	BF	3.31E-05	3.358	1.00E+00
MmugDNA.42599.1.S1_at	51703	ACSL5	2.70E-10	3.381	1.00E+00
MmuSTS.4842.1.S1_at	53831	GPR84	6.51E-08	3.433	1.00E+00
MmugDNA.41728.1.S1_at	4332	MNDA	2.12E-10	3.445	1.00E+00
MmuSTS.1399.1.S1_at	717	C2	2.52E-08	3.447	1.00E+00
MmugDNA.2176.1.S1_at	3669	ISG20	9.22E-12	3.471	1.00E+00
MmuSTS.1185.1.S1_at	51365	PLA1A	2.47E-13	3.474	1.00E+00
MmugDNA.27315.1.S1_at	54809	SAMD9	1.01E-10	3.488	1.00E+00
MmugDNA.18925.1.S1_at	1356	CP	2.22E-03	3.537	9.99E-01
MmugDNA.20604.1.S1_at	4071	TM4SF1	6.77E-09	3.561	1.00E+00
MmugDNA.43273.1.S1_at	1117	CHI3L2	1.49E-05	3.611	1.00E+00
MmunewRS.55.1.S1_at	3936	LCP1	2.72E-07	3.676	1.00E+00
MmugDNA.22991.1.S1_at	5721	PSME2	4.70E-11	3.679	1.00E+00

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
Mmu.865.1.S1_at	4938	OAS1	6.36E-07	3.696	1.00E+00
MmugDNA.23367.1.S1_at	441168		1.18E-04	3.718	1.00E+00
MmugDNA.30287.1.S1_at	474344	GIMAP6	4.54E-10	3.730	1.00E+00
MmugDNA.5715.1.S1_at	7076	TIMP1	2.14E-06	3.739	1.00E+00
MmugDNA.30526.1.S1_at	155465		1.83E-04	3.772	1.00E+00
MmuSTS.3251.1.S1_at	719	C3AR1	1.39E-10	3.775	1.00E+00
MmugDNA.35003.1.S1_at	63901		1.95E-10	3.776	1.00E+00
Mmu.12364.1.S1_at	11119	BTN3A1	1.81E-08	3.778	1.00E+00
MmuSTS.2890.1.S1_at	9830	TRIM14	3.96E-10	3.873	1.00E+00
MmuSTS.2320.1.S1_at	5630	PRPH	1.33E-03	3.910	9.98E-01
MmuSTS.2856.1.S1_at	6648	SOD2	3.01E-05	3.940	1.00E+00
MmugDNA.15862.1.S1_at	1164	CKS2	1.85E-07	3.953	1.00E+00
MmuSTS.4590.1.S1_at	7052	TGM2	8.36E-10	4.015	1.00E+00
MmugDNA.32666.1.S1_s_at	84790	TUBA6	5.08E-08	4.024	1.00E+00
MmuSTS.558.1.S1_at	26228	BRDG1	2.25E-06	4.115	1.00E+00
MmugDNA.13052.1.S1_at	118788	PIK3AP1	1.23E-10	4.142	1.00E+00
MmuSTS.1447.1.S1_at	4065	LY75	9.95E-07	4.190	1.00E+00
MmugDNA.2478.1.S1_at	972	CD74	4.59E-09	4.198	1.00E+00
MmuSTS.2540.1.S1_at	10068	IL18BP	1.03E-06	4.203	1.00E+00
MmugDNA.34954.1.S1_at	10410	IFITM3	3.63E-11	4.228	1.00E+00
MmuSTS.1612.1.S1_at	8638	OASL	1.31E-09	4.280	1.00E+00
MmuSTS.2736.1.S1_at	710	SERPING1	1.06E-07	4.320	1.00E+00
MmugDNA.30463.1.S1_at	9111	NMI	5.50E-10	4.365	1.00E+00
MmuSTS.2830.1.S1_at	23780	APOL2	9.89E-09	4.373	1.00E+00
MmugDNA.31841.1.S1_at	10673	TNFSF13B	3.60E-07	4.432	1.00E+00
Mmu.2142.1.S1_at	10346	TRIM22	2.15E-07	4.441	1.00E+00
MmugDNA.35854.1.S1_at	714	C1QG	5.30E-11	4.476	1.00E+00
MmugDNA.1533.1.S1_at	3108	HLADMA;	6.79E-06	4.536	1.00E+00
MmuSTS.1513.1.S1_at	3430	IFI35	1.62E-10	4.563	1.00E+00
MmugDNA.19209.1.S1_s_at	3934	LCN2	1.83E-04	4.596	1.00E+00
MmugDNA.36516.1.S1_at	10379	ISGF3G	3.44E-15	4.647	1.00E+00
MmugDNA.18601.1.S1_at	3115	HLADPB1;	8.41E-09	4.705	1.00E+00
MmuSTS.704.1.S1_at	3659	IRF1	8.85E-07	4.749	1.00E+00
MmuSTS.1397.1.S1_at	713	C1QB	4.07E-08	4.772	1.00E+00
MmugDNA.42778.1.S1_at	64761	PARP12	1.98E-12	4.772	1.00E+00
MmuSTS.690.1.S1_at	3576	IL8	1.47E-07	4.804	1.00E+00
MmugDNA.8251.1.S1_at	55281		6.02E-09	4.836	1.00E+00
MmugDNA.8389.1.S1_at	55303	GIMAP4	1.00E-10	4.846	1.00E+00
MmunewRS.436.1.S1_s_at	3123	HLADRB1;	1.87E-06	4.884	1.00E+00
MmugDNA.30775.1.S1_at	54979		4.18E-07	4.940	1.00E+00
MmugDNA.36369.1.S1_at	922	CD5L	5.94E-04	4.943	9.68E-01
MmuSTS.1957.1.S1_at	712	C1QA	6.97E-10	4.946	1.00E+00
MmugDNA.7038.1.S1_at	2214	FCGR3A	4.05E-07	4.961	1.00E+00

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmuSTS.1168.1.S1_at	1116	CHI3L1	2.12E-06	5.113	1.00E+00
MmugDNA.208.1.S1_at	6398	SECTM1	2.75E-06	5.253	1.00E+00
MmugDNA.7052.1.S1_at	9246	UBE2L6	3.90E-12	5.275	1.00E+00
MmuSTS.4428.1.S1_at	2537	G1P3	0.00E+00	5.308	1.00E+00
MmugDNA.22615.1.S1_s_at	93349	LOC93349	6.87E-10	5.335	1.00E+00
MmunewRS.536.1.S1_at	10437	IFI30	1.14E-07	5.475	1.00E+00
MmuSTS.4003.1.S1_at	4283	CXCL9	1.16E-07	5.507	1.00E+00
MmugDNA.15476.1.S1_at	94240		3.16E-09	5.513	1.00E+00
MmuSTS.793.1.S1_at	715	C1R	1.38E-09	5.554	1.00E+00
MmuSTS.19.1.S1_at	11326	VSIG4	5.56E-07	5.622	1.00E+00
MmuSTS.3917.1.S1_at	6772	STAT1	0.00E+00	5.796	1.00E+00
MmugDNA.41706.1.S1_at	91543	RSAD2	7.13E-12	5.940	1.00E+00
MmuSTS.4697.1.S1_at	7412	VCAM1	1.31E-11	6.078	1.00E+00
MmugDNA.41065.1.S1_at	721		2.51E-12	6.151	1.00E+00
MmugDNA.43116.1.S1_at	10457	GPNMB	2.23E-05	6.246	1.00E+00
MmugDNA.32616.1.S1_at	54739	BIRC4BP	4.51E-07	6.338	1.00E+00
MmugDNA.17449.1.S1_at	22797	TFEC	9.99E-10	6.353	1.00E+00
MmugDNA.14204.1.S1_at	3002	GZMB	1.03E-05	6.386	1.00E+00
MmunewRS.936.1.S1_at	163351	GBP6	8.13E-06	6.552	1.00E+00
MmuSTS.1398.1.S1_at	716	C1S	6.19E-10	6.640	1.00E+00
MmugDNA.32570.1.S1_at	4069	LYZ	2.28E-08	6.740	1.00E+00
MmugDNA.1046.1.S1_s_at	3122	HLADRA;	9.82E-12	6.747	1.00E+00
MmugDNA.20819.1.S1_s_at	4940	OAS3	2.26E-12	6.781	1.00E+00
MmuSTS.2737.1.S1_at	27299	ADAMDEC1	4.32E-04	6.917	1.00E+00
MmugDNA.19189.1.S1_at	23586	DDX58	9.29E-08	6.977	1.00E+00
MmugDNA.20898.1.S1_at	128346		1.24E-08	7.073	1.00E+00
MmuSTS.4713.1.S1_at	56829	ZC3HAV1	2.44E-15	7.523	1.00E+00
MmugDNA.15286.1.S1_at	54625	PARP14	8.12E-13	7.740	1.00E+00
MmugDNA.1081.1.S1_at	6890	TAP1	2.90E-11	8.176	1.00E+00
MmugDNA.5697.1.S1_s_at	8519	IFITM1	3.11E-13	8.638	1.00E+00
MmugDNA.34622.1.S1_at	83666	PARP9	1.22E-14	9.200	1.00E+00
MmugDNA.34026.1.S1_at	8605	PLA2G4C	1.71E-12	9.215	1.00E+00
MmuSTS.1927.1.S1_at	4600	MX2	7.27E-11	9.232	1.00E+00
MmugDNA.29742.1.S1_s_at	9332	CD163	2.80E-06	9.326	1.00E+00
MmuSTS.1439.1.S1_at	3959	LGALS3BP	9.35E-11	9.375	1.00E+00
MmunewRS.816.1.S1_s_at	3512	IGJ	7.58E-05	9.627	1.00E+00
MmugDNA.22932.1.S1_at	5698	PSMB9	3.35E-12	9.985	1.00E+00
MmugDNA.36047.1.S1_at	5696	PSMB8	3.30E-09	10.439	1.00E+00
MmugDNA.20634.1.S1_at	79132	LGP2	1.11E-16	10.614	1.00E+00
Mmu.11584.1.S1_at	684	BST2	9.53E-13	10.685	1.00E+00
MmugDNA.13757.1.S1_at	51316		7.27E-12	11.247	1.00E+00
MmugDNA.18432.1.S1_at	3620	INDO	1.07E-06	11.707	1.00E+00
MmugDNA.35740.1.S1_at	3433	IFIT2	2.61E-10	12.113	1.00E+00

Table S2.2 Differential Gene Expression in Monkeys with SIVE. (cont)

Affymetrix Probe Set ID	Human Entrez ID	Human Gene Name	Bayes.p	fold	PPDE(p)
MmuSTS.1579.1.S1_at	3434	IFIT1	4.44E-16	12.346	1.00E+00
MmugDNA.8510.1.S1_s_at	5920	RARRES3	8.83E-07	13.097	1.00E+00
MmugDNA.30921.1.S1_at	55008	HERC6	2.32E-14	13.243	1.00E+00
MmugDNA.38910.1.S1_at	51191	HERC5	2.51E-10	14.991	1.00E+00
MmugDNA.15618.1.S1_s_at	6373	CXCL11	4.17E-07	15.181	1.00E+00
Mmu.11912.1.S1_at	6355	CCL8	5.00E-10	16.796	1.00E+00
Mmu.3779.1.S1_at	3437	IFIT3	7.88E-12	16.979	1.00E+00
MmugDNA.36297.1.S1_at	64135	IFIH1	6.55E-14	17.012	1.00E+00
MmugDNA.27193.1.S1_at	55601	FLJ20035	3.12E-12	17.149	1.00E+00
Mmu.10083.1.S1_s_at	12	SERPINA3	4.77E-12	17.213	1.00E+00
MmunewRS.254.1.S1_at	3429	IFI27	0.00E+00	21.864	1.00E+00
MmugDNA.12892.1.S1_at	10561	IFI44	1.56E-13	22.087	1.00E+00
MmuSTS.3006.1.S1_at	4599	MX1	1.11E-16	22.342	1.00E+00
MmugDNA.41315.1.S1_at	2633	GBP1	2.72E-10	23.028	1.00E+00
MmuSTS.3317.1.S1_at	6347	CCL2	3.93E-08	23.658	1.00E+00
MmuSTS.1893.1.S1_s_at	3627	CXCL10	8.34E-08	36.293	1.00E+00
MmugDNA.17118.1.S1_at	9636	G1P2	1.89E-15	42.157	1.00E+00
MmuSTS.4748.1.S1_at	11274	USP18	1.74E-13	42.663	1.00E+00

Table S2.3 Mapping of Rat Genes to Human Entrez IDs.
NCBI/Gene and RGD (Rat Genome Database) were used to map rat gene names to human Entrez Gene IDs via Entrez Gene Symbols and aliases. Only genes reported to have at least 1 EGR1 response element (ERE) are included. Lgals5 is a discontinued gene. DE – Differentially expressed.

Rat Gene ID	Effect of EGR1 Over-expression	Rat Gene Symbol	Human Entrez	p <0.05, James *et al*[93]	Present in SIVE dataset	DE in SIVE dataset
Agrin	Down	Agrn	375790	Y		
Alcam	Down	Alcam	214			
Adenosine deaminase 3	Down	Ampd3	272			
Ap2a2	Down	Ap2a2	161	Y		
Ap2S1	Down	Ap2s1	1175			
Ap3m1	Down	Ap3m1	26985			
Apba3	Down	Apba3	9546			
Apol-III	Down	Apol3	80833	Y	N	
Arf6	Down	Arf6	382			
Arfrp1	Down	Arfrp1	10139			
ATF3	Down	Atf3	467			Y
ATP5c1	Down	Atp5c1	509			
AIM-1	Down	Aurkb	9212	Y		
CCL2	Down	Ccl2	6347	Y		Y
CCL5	Down	Ccl5	6352			Y
Cyclin B1	Down	Ccnb1	891			
IAP	Down	CD47	961	Y		
CD48	Down	CD48	962			
cdc20	Down	Cdc20	991			
cdc5	Down	Cdc5l	988			
cnp1	Down	Cnp	1267			
Col1a2	Down	Col1a2	1278			
csn7a	Down	Cops7A	50813			
carboxypeptidase E	Down	Cpe	1363			
cst3	Down	Cst3	1471			
CXCL10	Down	Cxcl10	3627			Y
drebrin	Down	Dbnl	28988			
eef2	Down	Eef2	1938			
EGR1	Up	Egr1	1958			Y
elF2c1	Down	Eif2c1	26523			
ferredoxin	Down	Fdx1	2230			
fra1	Down	Fosl1	8061			

Table S2.3 Mapping of Rat Genes to Human Entrez IDs. (cont)

Rat Gene ID	Effect of EGR1 Over-expression	Rat Gene Symbol	Human Entrez	p <0.05, James *et al*[93]	Present in SIVE dataset	DE in SIVE dataset
Gadd45a	Down	Gadd45a	1647			Y
GAK	Down	Gak	2580			
mir16	Down	Gde1	51573			
EF-G	Down	Gfm1	85476			
CFR-1	Down	Glg1	2734			
gnb1	Down	Gnb1	2782			
gnb2	Down	Gnb2	2783		N	
gephyrin	Down	Gphn	10243			
csn1/gps1	Down	Gps1	2873			
GSK3b	Down	Gsk3b	2932			
gtf2f2	Down	Gtf2f2	2963			
hnrpa1	Down	Hnrnpa1	3178		N	
hnrpm	Down	Hnrpm	4670			
Irf1	Down	Irf1	3659	Y		Y
Irf7	Down	Irf7	3665	Y		Y
c-jun	Down	Jun	3725			
KLF10	Down	Klf10	7071			
Importin B	Down	Kpnb1	3837			
Lgals3 BP	Down	Lgals3bp	3959			Y
Lgals5	Down	Lgals5	3961		N	
Lgals9	Down	Lgals9	3965			Y
Mdm2	Down	Mdm2	4193			
Mx1	Down	Mx1	4599			Y
Mx2	Down	Mx2	4600			Y
myr2	Down	Myo1c	4641			
neuropilin	Down	NRP1	8829			
NUB1	Down	Nub1	51667			Y
oas1	Down	Oas1	4938	Y		Y
PDGF	Up	Pdgfa	5154			
SAK	Down	Plk4	10733			
prolyl endopeptidase	Down	Prep	5550	Y		
psma5	Down	Psma5	5686			
psmb8	Down	Psmb8	5696	Y		Y
psmb9	Down	Psmb9	5698	Y		Y
psmc4	Down	Psmc4	5704		N	
psme2	Down	Psme2	5721	Y		Y

Table S2.3 Mapping of Rat Genes to Human Entrez IDs. (cont)

Rat Gene ID	Effect of EGR1 Over-expression	Rat Gene Symbol	Human Entrez	p <0.05, James *et al*[93]	Present in SIVE dataset	DE in SIVE dataset
poliovirus receptor	Down	Pvr	5817			
Rab13	Down	Rab13	5872			Y
rabaptin-5	Down	Rabep1	9135			
ZNF313	Down	Rnf114	55905	Y		Y
ROCK II	Down	Rock2	9475			
p70S6ribosomal kinase	Down	Rps6kb1	6198			
Raga	Down	Rraga	10670			
Best5/cig5/viperin	Down	Rsad2	91543	Y		Y
SCFD1	Down	ScfD1	23256			
PAI1	Down	Serpine1	5054			
Srfs10	Down	Sfrs10	6434			
SGK	Down	SGK1	6446			
small glutamine-rich tetratricopeptide	Down	Sgta	6449		N	
CHOT1	Down	Slc6a8	6535			
SMARCB1/SNF1	Down	Smarcb1	6598			
SOCS	Down	Socs1	8651			
syntaxin 4a	Down	Stx4	6810			
Succinate CoA ligase	Down	Suclg1	8802			
synapsin II	Down	Syn2	6854	Y		
synaptotagmin 3	Up	Syt3	84258			
Tap1 transporter	Down	Tap1	6890	Y		Y
Tap2 transporter	Down	Tap2	6891			
Tapasin	Down	Tapbp	6892			Y
Tyrosine hydroxylase	Down	Th	7054			
transketolase	Down	Tkt	7086			
NIPL/SKIP3/tribbles	Down	Trib3	57761			Y
ubiquitin D	Down	Ubd	10537			Y
UBC7	Down	Ube2G2	7327			
UBE2L6/UBCH8	Down	Ube2L6	9246			Y
Ubiquilin 1	Down	Ubqln1	29979			
vhl	Down	Vhl	7428			
Hbeta58-vacuolar protein sorting 26	Down	Vps26a	9559			

Acknowledgements

We thank Han-Yu Chuang for assistance with the PinnacleZ software, Samad Lotia for implementing PinnacleZ as a plug-in to Cytoscape, and James Buescher for assistance with immunohistochemistry.

Chapter 2, in full, is a re-editing of the materials published in Gersten, M., Alirezaei, M., Marcondes, M.C., Flynn, C., Ravasi, T., Ideker, T. & Fox, H.S. An integrated systems analysis implicates EGR1 downregulation in simian immunodeficiency virus encephalitis-induced neural dysfunction. J Neurosci 29, 12467-76 (2009). The dissertation author was the primary investigator and author of this paper.

3. GENOMIC SEQUENCING OF HYPOXIA-ADAPTED FLIES

Introduction

Oxygen is essential for metazoan growth, development and maintenance[47]. Hypoxia, a condition in which a tissue or organism is deprived of an adequate supply of oxygen, may occur with exposure to low O_2 tension (at high altitude), or as a result of insufficient air/blood flow, either physiological (within strenuously exercising muscle or rapidly growing tumors) or pathological (*eg*, airway obstruction, stroke, myocardial infarction). The major consequence of hypoxia is a reduction in the rate of ATP production as less oxygen is available for the more efficient generation of ATP *via* oxidative phosphorylation. Organisms employ adaptive responses to survive hypoxic challenge, including increased ATP production via anaerobic respiration and stabilization of HIF1A (Hypoxia Inducible Factor 1A)[113-115]; as oxygen availability decreases, adaptive mechanisms eventually become overwhelmed and an insufficiency of ATP leads to ion pump failure and cellular demise[116].

While all animals require oxygen, organisms vary with respect to the degree and duration of hypoxia that can be tolerated[117,118]. Furthermore, natural (*ie*, residence at high altitude) or experimental selection over many generations may allow the evolution of individuals with increased tolerance to reduced oxygen supply. Understanding the mechanisms that allow more hypoxia-tolerant organisms to thrive at lower oxygen tension would be expected to provide clues to improved approaches

for the treatment of disorders in which hypoxia plays a major or contributory role in morbidity and mortality.

Among the model organisms used to study hypoxia, *Drosophila melanogaster* offers several advantages. Like other insects, its tissues are ventilated with the ambient atmospheric O_2 tension, by means of a branching tracheal system, so that the actual O_2 level in tissue is known[47]. In addition, it has a short generation time, its development and genetics are well-characterized, and there is a plethora of available genetic tools[119].

D. melanogaster was used in a long-term selection strategy, starting with a pool of 27 isogenic strains, to generate a population of flies that is able to reproduce and thrive at 4% O_2, a level which is lethal to the original parental lines[49,50]. This was achieved by a stepwise reduction in ambient oxygen, starting at 8% O_2. Tolerance to 5% O_2 was relatively easily achieved by generation 13; the population then experienced a bottleneck of 19 generations before reaching 4% O_2 tolerance (Figure 3.1). Control pools were maintained in parallel at 21% O_2. The hypoxia-adapted fly population (AF) continues to survive at 4% O_2 at well over 200 generations.

The observation that tolerance to 4% O_2 persists upon reexposure to hypoxia after several generations in a normoxic environment suggested that genetic and/or epigenetic changes likely play a role in AF hypoxia tolerance. We asked, therefore, what does evolution select for in adaptation to hypoxia, as revealed by study of hypoxia-tolerant flies? In particular, is there evidence for (i) genomic rearrangements creating new genes or novel regulatory mechanisms for gene expression; (ii)

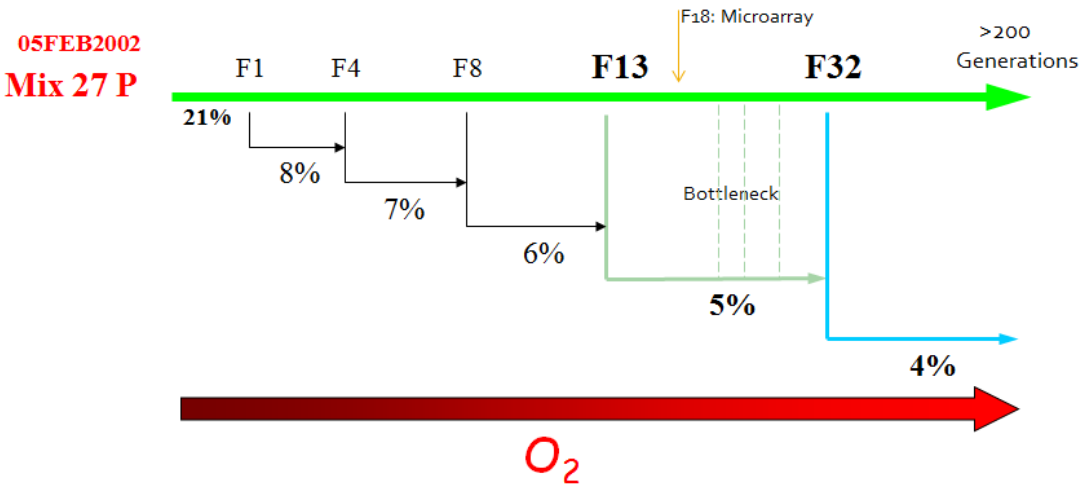


Figure 3.1 Selection of hypoxia-tolerant flies.
Three replicate pools of 27 isogenic *D. melanogaster* strains were exposed to a stepwise reduction in ambient oxygen starting at 8% O_2. Tolerance to 4% O_2 was achieved after 32 generations. Control pools were maintained in parallel at 21% O_2.

coordinated polymorphisms leading to changes in development and/or adaptive pathways; and/or (iii) modification of genomic regulatory regions driving epigenetic or transcriptomic changes? To further understand genetic mechanisms underlying AF adaptation to hypoxia, we used Illumina short read sequencing technology to resequence control and AF flies that had been under hypoxia selection for 180 generations.

Analysis of the resequencing data was complicated by two considerations. First, use of a diverse starting population, while facilitating more rapid selection of hypoxia-tolerant flies, presented a challenge for identifying genomic differences that contribute to hypoxia tolerance because the control population consisted not of a single genome, but of a population of bits and pieces reassembled from all 27 parental

lines. Second, processing limitations existing at the time the study was conducted necessitated our use of fly pools, rather than individual flies, for sequencing. As a result, it could not be determined whether multiple reads came from the same or different individual flies, and therefore neither linkage disequilibrium nor copy number variation could be ascertained. I approached the data, therefore, with the assumption that the reproductive bottleneck experienced between generations 13 and 32 led to a contraction of genetic variation within the adapted populations. I then assumed that (i) many of the important genetic differences *vs* the control population (mutations and/or allele selection) would be widespread within an adapted population, and (ii) at least some of these differences would be shared by independently adapted populations arising from the same parental pool. Based on these assumptions I devised a fine-grained approach to detecting genomic differences that distinguished control and adapted populations, which allowed for analysis of approximately half of all euchromatin loci. Given the inability to utilize population genetics methods to identify genes likely to be important in hypoxia tolerance, I opted for a systems approach to evaluate the polymorphic differences identified, looking for pathways or processes containing a disproportionate number of components harboring one or more polymorphisms. This approach to the data is based on the hypothesis that the presence of many polymorphisms in a pathway would enable fine-tuning of signaling changes in a cost-benefit tradeoff, and therefore would be more likely to participate in the adaptive response.

Methods and Results

Notch Signaling Pathway

Sample preparation and genomic resequencing using the Illumina GA II sequencer was performed by D. Zhou, as described in Zhou *et al*[48]. Maq v.0.7.1 software[120] was used under its default parameters to map reads from two control (C1, C2) and two hypoxia-tolerant (H1, H2) populations to the *Drosophila melanogaster* reference genome release 5.16 downloaded from FlyBase[121]. Since neither C1 nor C2 was under selective pressure, these two populations would be expected to differ only as the result of genetic drift. The C2 data was found to have much lower high-quality coverage (as defined below) and was not used in subsequent analyses. Between 120 million and 200 million 54-bp paired-end reads per population were aligned per population, corresponding to >60× coverage.

In my first analysis of the resequencing data, I implemented an approach to detect loci with a definitive base call that differed between control and adapted populations. I found that loci with high-quality base calls (base quality ≥20, best read quality ≥40, calls limited to A, C, G, or T) showed similar high read consistency for coverages between 20 and 40 (Table 3.1), whether or not the base call was the same or different from the reference base. I therefore examined euchromatin loci that had ≥20x coverage and high-quality base calls in sample pools C1, H1, and H2, which covered approximately 45% of euchromatin loci. I disregarded pool C2 since it had much lower 20x coverage than the other three pools (36,594,586 loci in C2 *vs* 67,933,396 in C1); furthermore, since neither C1 nor C2 was under selective pressure,

Table 3.1 Consistency of reads at high-quality loci.
Analysis limited to loci with base quality ≥ 20, best read quality ≥ 40, A/C/G/T base call. Percent match refers to the percent of reads matching the base called.

Category	≥40 reads		≥30 reads		≥20 reads	
	C1 (%)	H1 (%)	C1 (%)	H1 (%)	C1 (%)	H1 (%)
Consensus = Ref All	47719617	61279655	52394684	66695997	57892280	72832544
Consensus = Ref Match ≥ 90%	47526958 (99.60%)	60894190 (99.37%)	52178712 (99.59%)	66229769 (99.30%)	57644129 (99.57%)	72227182 (99.17%)
Consensus = Ref Match ≥ 80%	47683168 (99.92%)	61230135 (99.92%)	52352451 (99.92%)	66638390 (99.91%)	57841907 (99.91%)	72761296 (99.90%)
Consensus = Ref Match ≥ 70%	47709140 (99.98%)	61269128 (99.98%)	52381910 (99.98%)	66682928 (99.98%)	57876313 (99.97)	72815769 (99.98%)
Consensus ≠ Ref All	78467	86496	87538	96456	99225	108959
Consensus ≠ Ref Match ≥ 90%	64575 (82.30%)	69657 (80.53%)	72914 (83.29%)	78476 (81.36%)	83661 (84.31%)	89532 (82.17%)
Consensus ≠ Ref Match ≥ 80%	76545 (97.55%)	84881 (98.13%)	85503 (97.68%)	94682 (98.16%)	96961 (97.72%)	106924 (98.13%)
Consensus ≠ Ref Match ≥ 70%	78035 (99.45%)	86092 (99.53%)	87035 (99.42%)	95972 (99.50%)	98581 (99.35%)	108357 (99.45%)

only differences due to genetic drift would be anticipated. Scripts were written to parse Maq view output files to identify loci meeting the above constraints in which both the H1 and H2 base calls differed from both reference and C1 bases. In particular, both the first and second best Maq calls for H1 and H2 were required to differ from both the first and second best base calls for C1. In all these cases, the base calls for H1 and H2 were found to agree. Indels were identified using Maq indelpe output files corrected for homopolymer tracts. After eliminating indel positions shared with C1, the set of putative hypoxia-tolerance related indels was further filtered to remove loci with <20x coverage in either the H1 or H2 pool or in which the percent of reads mapped with the indel was <50% in either the H1 or H2 pool. Identified SNPs and indel loci were mapped to genes/gene regions using fasta files downloaded from FlyBase[121]. An overview of this methodology is presented in Figure 3.2.

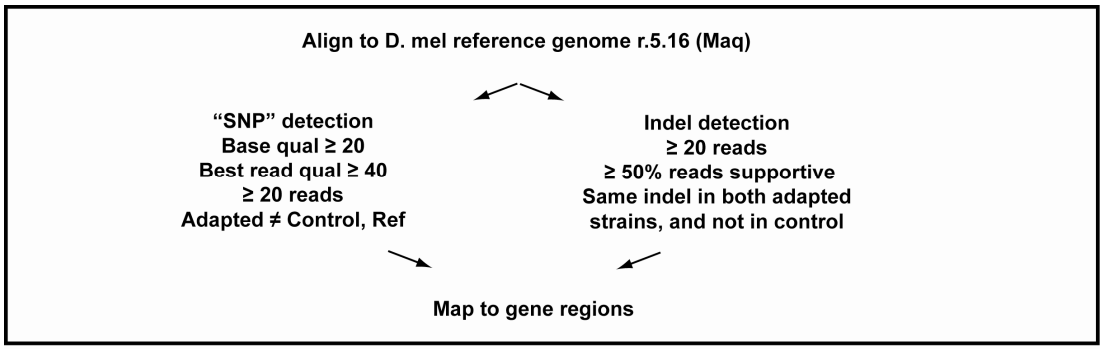


Figure 3.2 Overview of SNP/indel identification procedure.

A total of 2440 euchromatin SNPs and 1505 small indels distinguished the adapted *vs* control populations, of which 1846 SNPs and 1043 indels could be mapped

to a total of 1817 FlyBase extended genes (gene plus 2kB up- and downstream). Polymorphisms were identified in all gene regions, with a preponderance appearing in the X chromosome (Figures 3.3A and 3.3B). Many of observations based on the fine-grained analysis were corroborated by a coarse-grained analysis done by N. Udpa designed to identify hypoxia-selected regions[48].

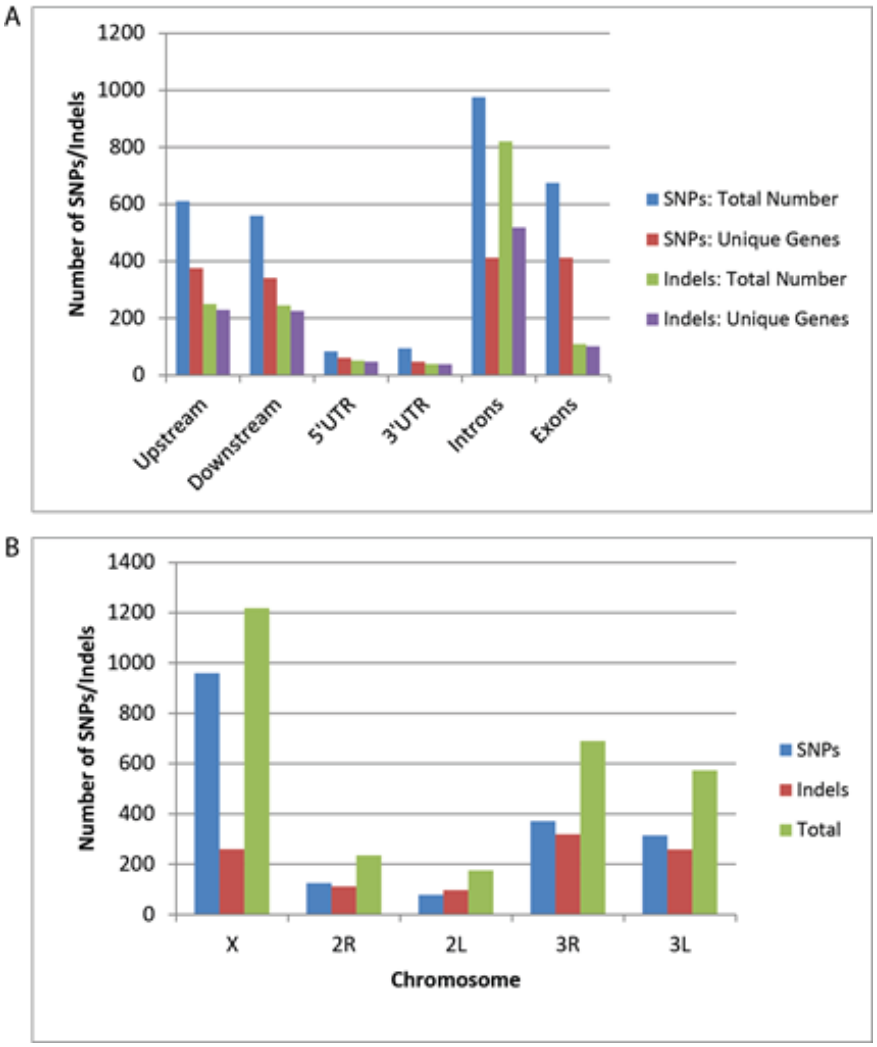


Figure 3.3 Distribution of hypoxia-tolerance polymorphisms.
Bar graphs indicating the distribution of SNPs and small indels in extended gene regions by gene region (A) and chromosome (B). Loci falling in the same gene region of more than one gene were counted for each gene.

Based on gene expression data obtained from 5% O_2-adapted flies which suggested that the Notch pathway, among others, might be involved in hypoxia tolerance[50], we focused our initial analysis on determining whether there was evidence of a genetic basis for Notch signaling in hypoxia tolerance. KEGG[14,15] pathway analysis using DAVID[122,123] identified five Notch pathway genes among those with one or more polymorphisms: *Notch*, *Delta*, *Pcaf*, *nejire*, and *fringe*. A search of the literature identified additional polymorphism-containing Notch pathway interactors, several of which mapped to a hypoxia-selected region identified by the coarse-grained analysis. These data, together with differential gene expression are presented in Figure 3.4, illustrating the position of each affected gene in the Notch signaling pathway.

I found additional evidence suggesting a role for Notch signaling in hypoxia-tolerance by integrating the set of polymorphic genes with a high-confidence *D. melanogaster* network of functional relationships, constructed using genetic interaction, protein interaction and gene expression datasets[124], containing 5021 genes (nodes) and 20,000 interactions (edges), which I visualized using Cytoscape[26]. Of the 1817 genes with ≥1 polymorphism, 519 were present in the network. In the subnetwork containing only the 519 polymorphic genes, 232 genes had no connections and the remaining 287 had a total of 423 interactions. Notch (N) itself was directly connected to 25 genes, including fringe (fng) and Delta (Dl) (Figure 3.5). I next identified non-polymorphic genes highly connected (≥5 interactions) to genes containing polymorphisms, and added them, generating a network of 765 nodes and

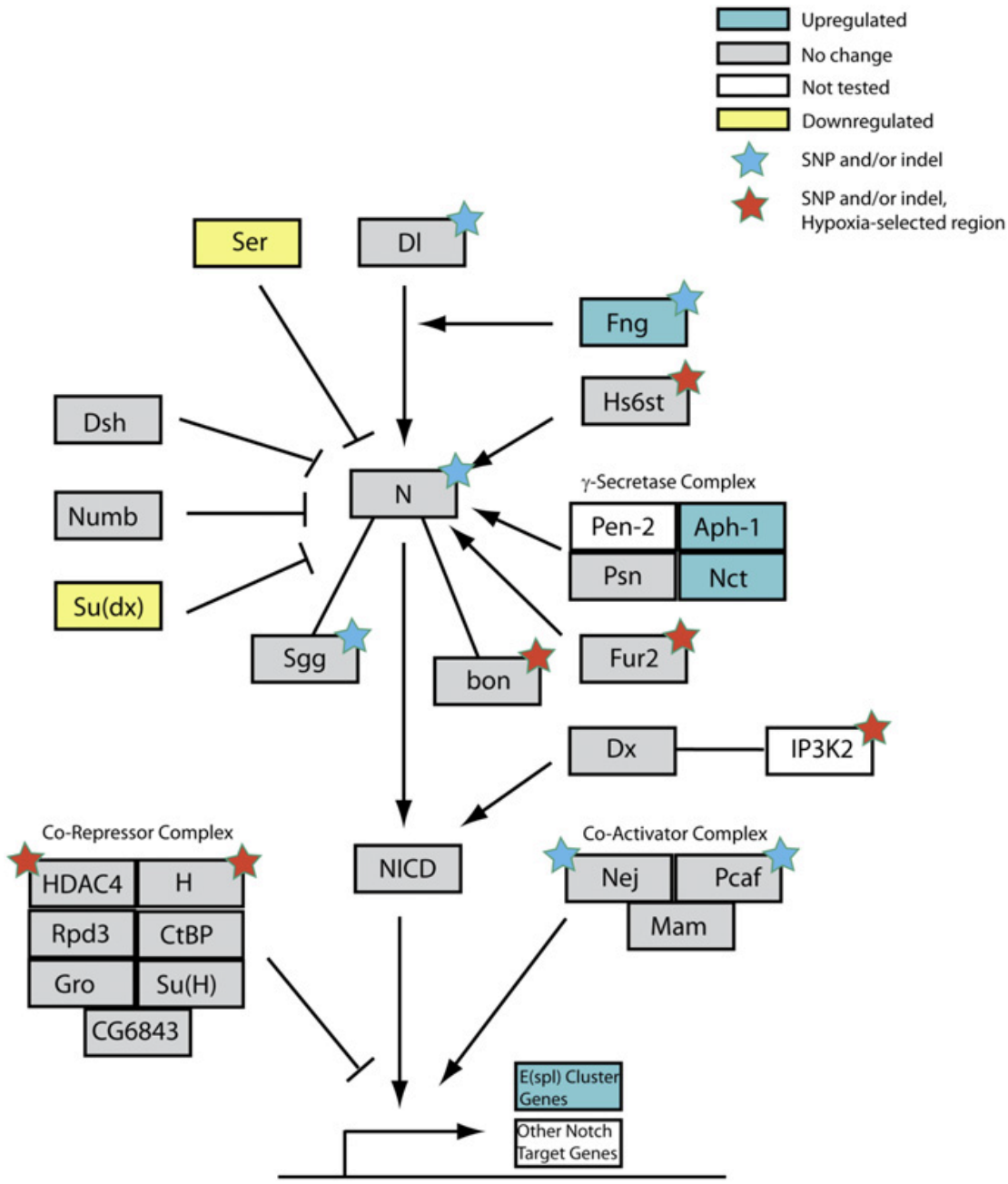


Figure 3.4 Enrichment of fixed SNPs and indels in an extended Notch pathway.
The Notch pathway was adapted from KEGG[14,15] by adding Notch interactors from the literature to create an expanded Notch signaling pathway. Genes differentially expressed in larva (expression levels from Zhou *et al*[50]) are cyan (upregulated) or yellow (downregulated), genes showing no change in expression are gray, and untested genes are white. Genes for which one more SNP and/or indels became fixed are indicated with stars: red for genes located within a hypoxia-selected region and blue for all others.

4323 interactions. Within this expanded network, Notch was directly connected to 74 genes, nearly 10% of the entire network (Figure 3.6). This set of genes included an additional seven KEGG Notch pathway genes, as well as members of other KEGG pathways: Dorso-ventral axis formation, MAPK signaling, Wnt signaling and Hedgehog signaling. This analysis suggests that adaptation of the AF populations to 4% O_2 involves not only changes in Notch signaling, but also its interaction/crosstalk with other developmental pathways.

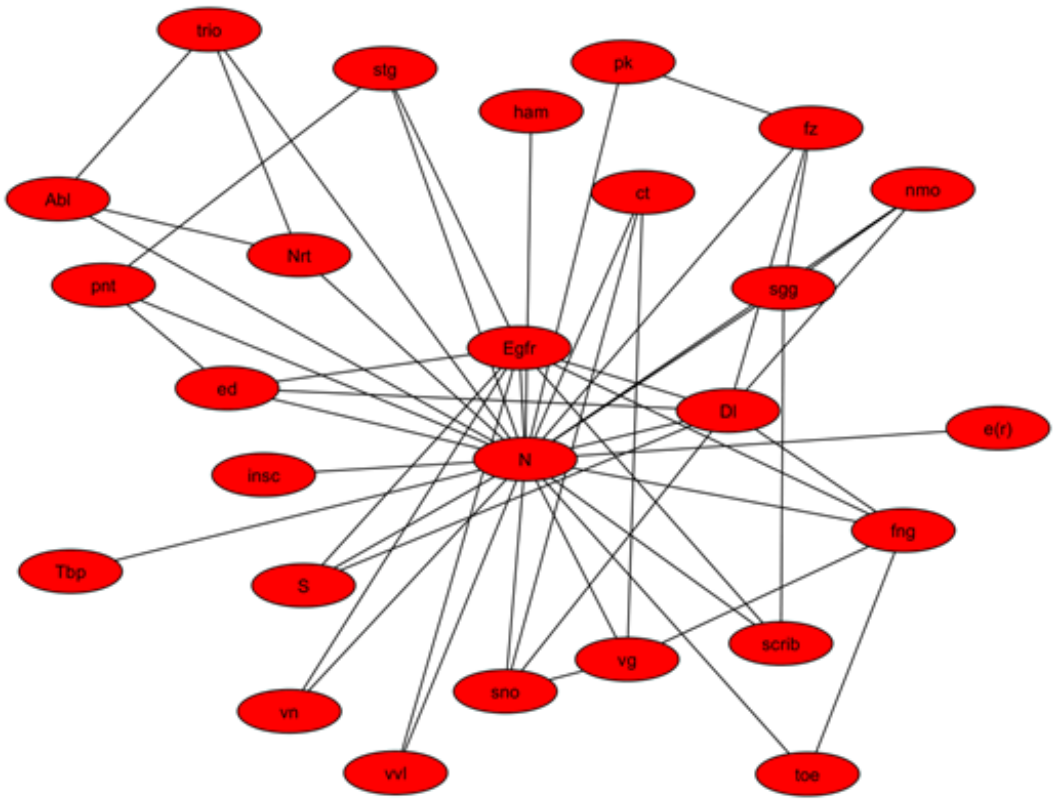


Figure 3.5 Polymorphism-containing nearest neighbors of Notch.
Interrogation of the Costello high confidence 20K network of functional relationships[124] revealed 25 polymorphism-containing genes that directly interact with Notch.

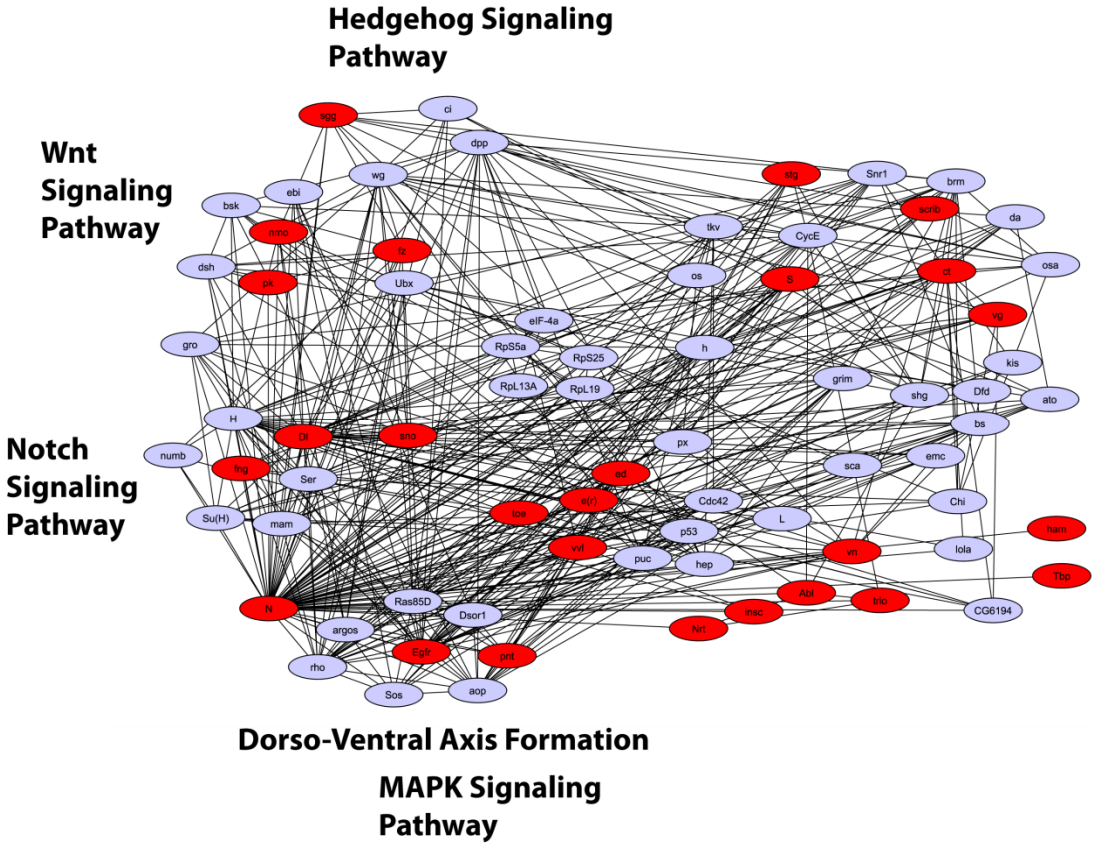


Figure 3.6 Expanded network of Notch interactors.
Within the Costello 20K network[124], Notch interacts with 25 polymorphism-containing genes (colored red) as well as 49 additional genes (colored blue) that are highly connected (≥5 interactions) to polymorphism-containing genes.

Experimental validation of a role for Notch pathway activation in hypoxia tolerance was provided by the work of collaborating scientists[48] using Notch LOF (loss of function) and GOF (gain of function) fly strains, chemical inhibition of Notch cleavage, and genetic crosses using the Gal4-UAS expression system[125-127]. Measures of hypoxia tolerance included rate of adult eclosion and adult survival under hypoxic conditions (Supplemental Figures S3.1 and S3.2).

Wnt Signaling Pathway

Noting the evolutionary bottleneck encountered between achieving 5% vs 4% O_2 tolerance, and the network evidence of Notch interaction with other polymorphism-containing pathways, I searched for additional mechanisms that contribute to the tolerant phenotype. I reanalyzed the sequencing data using a refinement of the fine-grained approach. Specifically, I looked for polymorphisms fixed (≥90%) in H1 and H2 and rare (≤10%) in C at high-quality loci, defined in terms of coverage (≥10X, after duplicate read removal) and Maq[120] reported base quality (≥20) and best read quality (≥40). Figure 3.7 provides an overview of the reanalysis methodology.

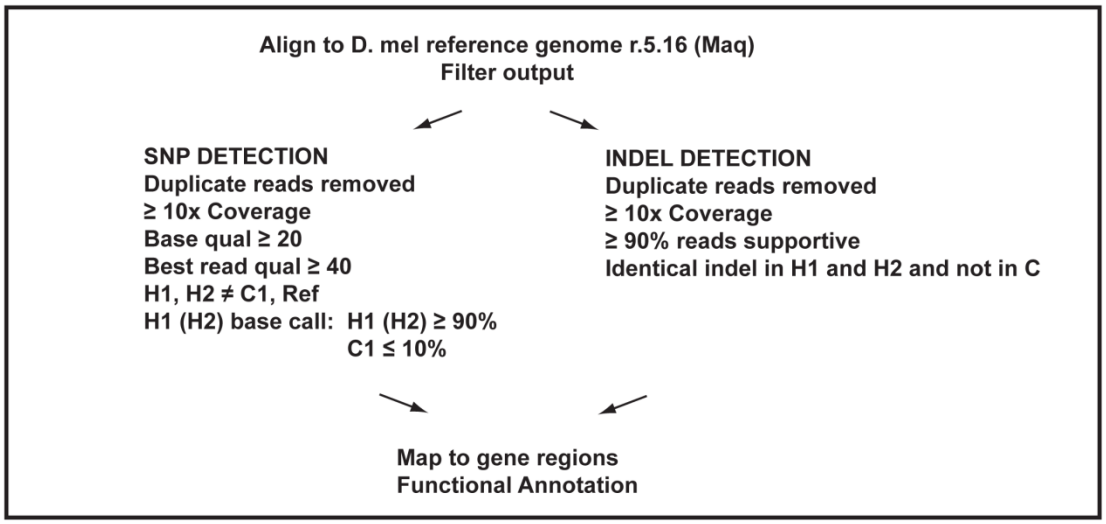


Figure 3.7 Overview of SNP/indel reanalysis procedure.

Again, approximately 50% of the euchromatin portion of the five major gene-bearing chromosomes (X, 2L, 2R, 3L, 3R; range 45.4-53.4%) met the selection criteria in the C dataset as well as in both H1 and H2 datasets and could be analyzed for SNPs

and small indels. A total of 2514 SNPs and 405 small indels distinguished H1 and H2 from C, of which 1940 SNPs and 283 indels could be mapped to a total of 1072 FlyBase16 extended genes (gene plus 2kB up- and downstream; Table 3.2). Many

Table 3.2 Fixed SNPs and Indels Distinguishing AF from Control Flies

Category	SNP	Indel
Total number of loci	2514	405
Loci mapped to genes: No. of loci (No. unique FlyBase genes)		
Extended gene	1940 (921)	283 (320)
Upstream (2000 bp)	507 (315)	67 (71)
Downstream (2000 bp)	471 (291)	57 (58)
5' UTR	80 (55)	15 (14)
3' UTR	63 (48)	13 (16)
Intron	971 (337)	216 (185)
Exon	743 (333)	29 (29)
CDS	625 (288)	3 (3)
NS change	135 (99)	2 (2)

genes contained more than one polymorphism, and due to physical overlap of genes in the *Drosophila* genome, some polymorphic loci mapped to more than one gene. Applying the same SNP selection criteria to the analyzable loci I detected only a single locus fixed in one H population and rare in the other, for an estimated FDR of 4.0×10^{-4}. Polymorphisms were identified in all gene regions (Figures 3.8), though relatively fewer mapped to 3'UTR (SNPs, indels), and exons (indels) after normalizing for gene region size. Approximately half the indels and nearly two-thirds

of SNPs mapped to the X chromosome after normalization for (euchromatin) chromosome size (Figure 3.8). At all these polymorphic loci the H1 and H2 sequences matched each other, suggesting that allele selection rather than mutation may have dominated during adaptation to hypoxia.

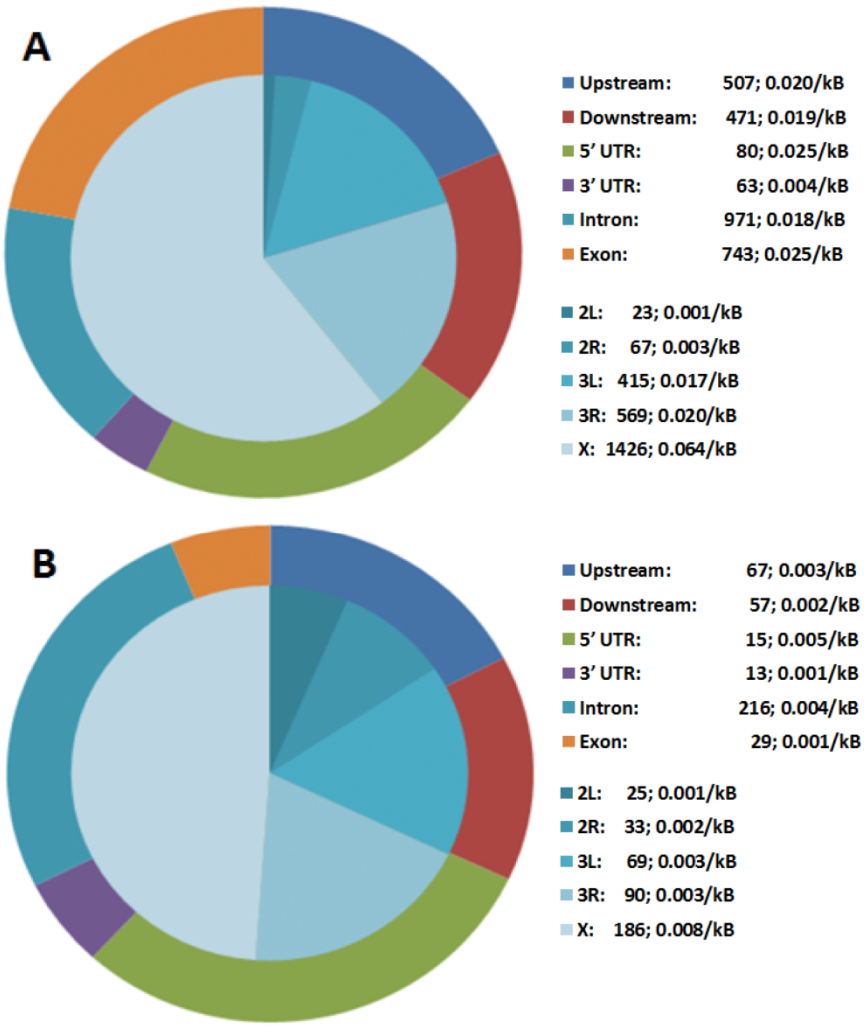


Figure 3.8 Normalized distributions of polymorphisms identified hypoxia-adapted flies. Chromosome and gene region distributions of SNPs (A) or indels (B) were normalized to (euchromatin) chromosome size or number of total gene region loci, respectively. The legends specify the number of polymorphisms detected and polymorphisms/kB (euchromatin) chromosome or gene region. Not included are 14 SNPs mapping to XHet, YHet or U/Uextra and 2 indels mapping to U and Chrom 4.

DAVID[122,123] analysis of genes with ≥1 polymorphism revealed most top-scoring GO-BP annotations to relate to development and morphogenesis (Supplemental Table S3.1). Of 66 GO terms with a corrected Benjamini p-value of ≤ 0.05, 57 were associated with one of these processes, including ten terms relating to the nervous system. Lower scoring annotations included 11 genes related to oxidative phosphorylation, 45 genes related to oogenesis and 52 to cell cycle, including ATR homolog, *mei-41*, with 26 SNPs, which regulates a meiotic checkpoint during *Drosophila* oogenesis[128].

Pathway analysis of genes with ≥1 polymorphism revealed the Wnt signaling pathway to be at or near the top of both KEGG[14,15] (N=15, p=0.002) and Panther[16] (N=18, p=0.055) reported pathways (Figure 3.9, Supplemental Table S3.2). A union of the two sets of Wnt pathway genes yielded a total of 28 genes across all three major modes of Wnt signaling: canonical (*armadillo*/*β-catenin* mediated), planar cell polarity, and Wnt/calcium pathways. In this analysis, other major developmental pathways were far less enriched for polymorphic genes, several of which also participate in Wnt signaling. Additional Wnt-pathway related genes were identified by comparing the list of polymorphic genes with (i) the set of *D. melanogaster* genes assigned a Wnt association by Gene Ontology[17], which added four genes, and (ii) the set of putative canonical Wnt-pathway regulatory genes identified by DasGupta in a genomic RNAi screen[129]. Of the 238 putative regulators identified in the latter screen, 207 could be mapped to a total of 212 FlyBase genes, adding another 29 genes for a total of 56 genes among those identified as containing a polymorphism. The overlap

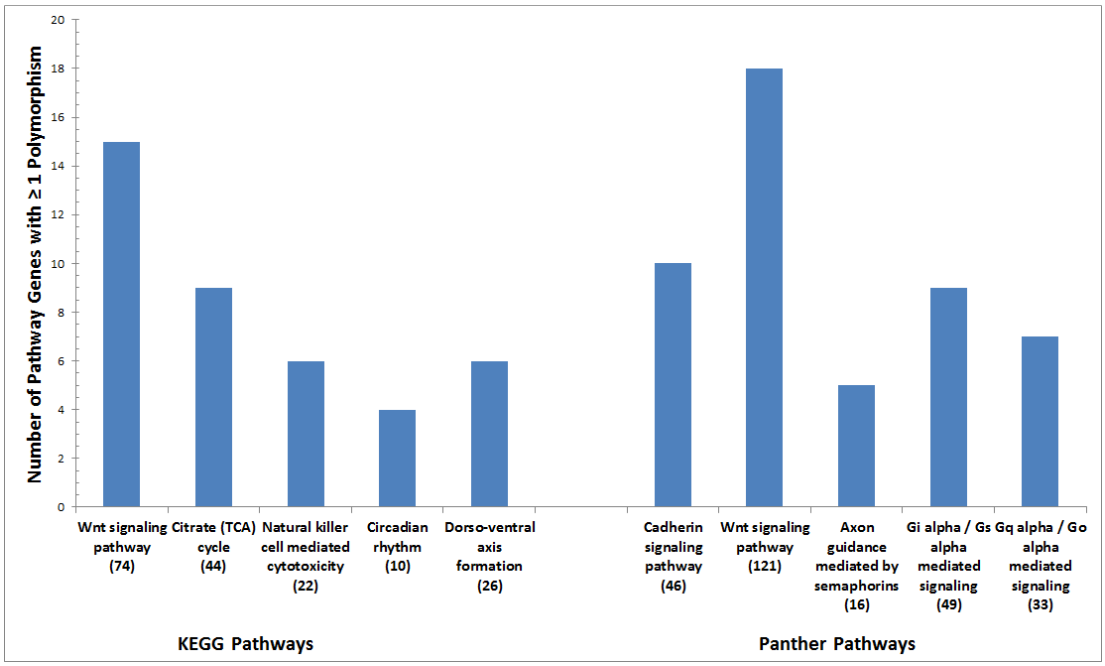


Figure 3.9 Pathway enrichment of hypoxia-tolerance polymorphisms
DAVID[122,123]-identified KEGG[14,15] and Panther[16] pathway enrichment of polymorphisms. The number of pathway genes with ≥ 1 polymorphism is plotted and the total number of pathway genes is indicated below each pathway name.

between the polymorphic gene set and the DasGupta set was highly significant (hypergeometric p-value of 4.51E-04).

Supplemental Tables S3.3 and S3.4 provide information regarding all SNPs and indels detected in Wnt pathway-associated genes and Supplemental Table S3.5 summarizes data on coding region polymorphisms. Table 3.3 summarizes the chromosomal location and number of polymorphisms observed for the 56 Wnt pathway-associated genes and their association (when known) with Wnt signaling. The large majority of genes are located on chromosomes X, 3R and 3L; however the large overall skew towards X-chromosome location (Figure 3.8) is not seen here. The majority of genes (38 of 56) had three or fewer polymorphisms, although the

Table 3.3 Polymorphism-Containing Wnt Pathway-Associated Genes
Compiled using KEGG[14,15], Panther[16], DasGupta[129], GO[17] and Flybase[121].

Symbol	Enrichment Source	Wnt Pathway	Function	SNP	Indel	CDS SNP	NS SNP
arr	KEGG, DasGupta, GO	Canonical	Co-Receptor; RNAi: Activator		1		
Axn	KEGG, DasGupta, GO	Canonical	β-catenin binding; RNAi: Inhibitor	7		1	
bab1	DasGupta	Canonical	RNAi:Inhibitor	2	1		
Bap60	Panther	Canonical	Brahma associated protein; Activator	3			
bon	DasGupta	Canonical	RNAi:Inhibitor		2		
Cad99C	Panther	?	Cadherin	3		2	
CanA1	KEGG, Panther	Calcium	Dephosphorylates NFAT; Activator	3		2	
CanA-14F	KEGG	Calcium	Dephosphorylates NFAT; Activator	8	1		
Cby	GO	Canonical	Inhibitor	6		1	
CG10107	DasGupta	Canonical	RNAi:Inhibitor	7	1		
CG11873	DasGupta	Canonical	RNAi:Inhibitor	6		3	3
CG14107	DasGupta	Canonical	RNAi:Activator	1			
CG32105	DasGupta	Canonical	RNAi:Inhibitor	1			
CG32432	DasGupta	Canonical	RNAi:Inhibitor		1		
CG32575/hang	DasGupta	Canonical	RNAi:Inhibitor	49	1	26	2
CG32683	Panther	Canonical	Arrestin-related; β-arrestin is an Activator		1		
CG4328	DasGupta	Canonical	RNAi:Inhibitor	1			
CG6606/Rip11	DasGupta	Canonical	RNAi:Activator	1			
CG6834	DasGupta	Canonical	RNAi:Activator	1			
CG7628/CG42575	DasGupta	Canonical	RNAi:Inhibitor	3	1	1	
CG7837	DasGupta	Canonical	RNAi:Activator	7		3	
CG7990	DasGupta	Canonical	RNAi:Inhibitor		1		
cpo	DasGupta	Canonical	RNAi:Inhibitor	1			
dalao	Panther	Canonical	Transcription regulation; Activator	2			
dally	KEGG, GO	Canonical	Co-Receptor		1		
dco	Panther	Canonical	Casein kinase Iε; Activator	5	1	4	
dm	Panther	Canonical	(Myc) Transcription factor; Target Gene	1			

Table 3.3 Polymorphism-Containing Wnt Pathway-Associated Genes. (cont)

Symbol	Enrichment Source	Wnt Pathway	Function	SNP	Indel	CDS SNP	NS SNP
dom	Panther, DasGupta	Canonical	Helicase; RNAi:Inhibitor	2		2	
ds	Panther	PCP	Cadherin-related	1			
fru	DasGupta	Canonical	RNAi:Inhibitor		1		
ft	Panther, GO	PCP	Cadherin-related	2		2	
fz	Panther, DasGupta, GO	PCP > Canonical	Receptor; RNAi: Activator	1			
gskt	Panther	Canonical	Putative GSK3Bhomolog; Inhibitor	1			
hdc	DasGupta	Canonical	RNAi:Inhibitor	6	2		
lilli	DasGupta	Canonical	RNAi:Activator		1		
Med	KEGG, Panther	Canonical	Binds TCF; Activator		1		
NFAT	KEGG	Calcium	Transcription factor; Activator	1			
nmo	KEGG, Panther, GO	Canonical	Kinase; Inhibitor	8	1		
Pkcdelta	Panther	Calcium	Kinase	2			
pnt	DasGupta	Canonical	RNAi:Inhibitor	6			
Pp2B-14D	KEGG, Panther	Calcium	Dephosphorylates NFAT; Activator	7	1	3	
pygo	Panther, DasGupta, GO	Canonical	β-catenin nuclear targeting; RNAi:Activator	1		1	1
Rac1	KEGG	PCP	Activator	2			
Rho1	KEGG, GO	PCP	Activator		1		
rok	KEGG, GO	PCP	Activator	14	4	2	
sfl	GO	Canonical	Sulfotransferase; Activator	4	1		
sgg	KEGG, Panther, GO	Canonical	GSK3B homolog; β-catenin binding, Inhibitor	5	1	2	
sgl	GO	Canonical	UDP Glucose Dehydrogenase; Activator		1		
sima	DasGupta	Canonical	RNAi:Inhibitor	14	5	1	1
skpC	KEGG	Canonical	β-catenin proteolysis; Inhibitor	1	1		
skpD	KEGG	Canonical	β-catenin proteolysis; Inhibitor	1			
Smr	DasGupta	Canonical	RNAi:Inhibitor	15		6	
tna	DasGupta	Canonical	RNAi:Inhibitor	1	1		

Table 3.3 Polymorphism-Containing Wnt Pathway-Associated Genes. (cont)

Symbol	Enrichment Source	Wnt Pathway	Function	SNP	Indel	CDS SNP	NS SNP
ttv	GO		Heparan sulfate proteoglycan biosynthesis; Activator		1		
vn	DasGupta	Canonical	RNAi:Inhibitor	1			
wts	DasGupta	Canonical	RNAi:Inhibitor	1	1	1	

observed polymorphisms may well underestimate the true number, since only approximately half of the eukaryotic sequence had sufficient coverage across all three datasets (C, H1, H2) to permit polymorphism detection. Only four genes had ten or more polymorphisms: *sima* (19), *Smr* (15), *rok* (18) and *hang* (50). When adjusted for extended gene size (gene plus 2kB up- and downstream), five genes had ≥0.9 polymorphisms/kB: *Cby* (1.26), *CG7837* (0.92), *hang* (2.68), *Pp2B-14D* (0.93), and *rok* (1.10). Eighteen genes had one or more coding region SNPs, the majority of which were synonymous changes. Four genes had one or more nonsynonymous (NS) SNPs: *CG11873* (3 NS), *sima* (1 NS), *pygo* (1 NS) and *hang* (2 NS). *Hang* had an additional 24 synonymous SNPs as well as one CDS indel which was not frame-shifting.

Discussion

As a result of its tracheal system, the tissues of the 4% O_2-adapted flies are exposed at sea-level to the ambient oxygen tension of approximately 30 mm Hg, slightly lower than the roughly 38 mm Hg tension in capillary blood of high-altitude dwellers living at 4500 m[130]. The achievement of this degree of hypoxia tolerance required 32 generations of selection, corresponding to 800-100 years in human years.

Given that the Tibetan and Andean highlanders have been evolving to high altitude residence for perhaps as long as 25,000 and 11,000 years, respectively[130,131], the fly adaptation seems remarkably rapid, and may reflect the graded selection protocol and in particular the diverse starting genetic pool of 27 isogenic parental lines.

Although use of a diverse starting population was important for rapid selection of hypoxia-tolerant flies, it presented a challenge for identifying genomic differences that contribute to hypoxia tolerance, which was compounded by the technical requirement to use fly pools rather than individual flies for sequencing. Assuming that the reproductive bottleneck experienced between generations 13 and 32 led to a contraction of genetic variation within the adapted populations, I devised a fine-grained approach to detecting genomic differences fixed in the adapted and rare in the control populations, which allowed for analysis of approximately half of all euchromatin loci. My method enabled the reliable detection, despite the use of pooled samples, of homozygous polymorphisms with sufficiently high coverage and quality present in the majority of the adapted populations. Recognizing that many, perhaps key, polymorphisms may have been missed as a result of only 50% genome evaluability and the inability to identify heterozygous polymorphisms, and unable to utilize classic population genetics methods, I used a systems approach to identify pathways or processes enriched for genes containing polymorphisms. I reasoned that the presence of many polymorphic components would allow for the flexibility and fine-tuning of signaling changes in pathways normally used in diverse cellular,

developmental and homeostatic contexts that would enable a favorable cost-benefit adaptive outcome.

Remarkably, analyses were able to identify two highly conserved, key developmental pathways, Notch and Wnt, with many pathway members containing one or more polymorphisms. Notch is a highly conserved pathway that participates in developmental and physiological processes involving a wide spectrum of cell types and stages of development[132,133]. Transmembrane receptor *Notch* interacts with a transmembrane ligand (*Delta*, *Serrate*/*Jagged*) on a contacting cell, which leads to proteolytic cleavage of Notch and translocation of the resulting Notch intracellular domain (NICD) to the nucleus where it interacts with the DNA-binding protein *CSL* (*Su(H)*) and coactivator *Mastermind* to promote target gene transcription[132,133]. Wnt signaling is more complex, comprising several highly conserved pathways which respond to activation by one of several secreted Wnt ligands binding to one of several frizzled receptors. These pathways control a variety of developmental processes, including cell fate specification, polarity and stem cell renewal[134-138]. The canonical pathway, which activates transcription *via* stabilization of *β-catenin* (*arm*), is best studied. In the absence of Wnt signaling cytoplasmic *β-catenin* is phosphorylated and targeted for proteosomal degdradation by a complex consisting of *Apc*, *Axin* and *GSK3B* (*sgg*) which leads to proteosomal degradation. Wnt signaling, acting *via Dsh*, modulates the destruction complex, leading to *β-catenin* stabilization and entry into the nucleus where it binds to and derepresses *TCF*/*LEF* transcription factors[135,139-141]. The non-canonical pathways, which are less well-studied, include the Planar Cell

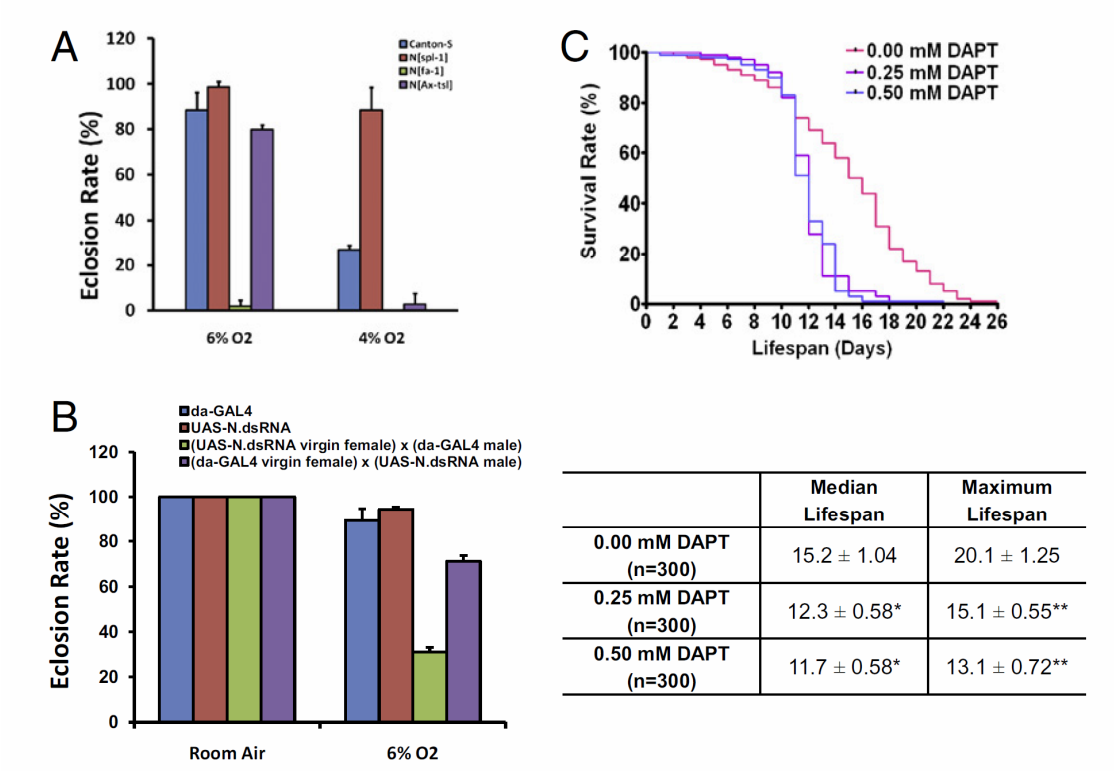
Polarity (PCP) pathway and the Wnt/Ca^{++} pathway. The PCP pathway regulates structure orientation within the epithelial plane and convergent extension in vertebrate gastrulation[135,142]. The Wnt/Ca^{++} pathway is activated *via* Wnt-induced intracellular Ca^{++} release in which calcium-sensitive proteins activate the NFAT transcription factor, and it may also modulate activity of the other Wnt pathways[135,141,143].

The identification of Notch and Wnt pathway polymorphisms in hypoxia-tolerant flies begs the question not only of how, mechanistically, alterations in each of these pathways may contribute to hypoxia tolerance, but also of how these two pathways might interact with each other in the adaptation to hypoxia. Regarding the latter question, many articles have noted that Notch-canonical Wnt cross-talk may be antagonistic or cooperative, depending on the cellular/developmental context[132,139,144-146]. A variety of mechanisms have been described to explain many of the observations. Canonical Wnt signaling upregulates *Notch2* in colon cancer[147], during oncogenic transformation of mammary epithelial cells[148], and may upregulate *Notch* in the maintenance of hematopoietic stem cells[149]; it induces transcription of the *Notch* ligand *Jag1* in hair follicles[150] and during intestinal tumorigenesis[151]. Phosphorylation of *Notch* by *GSK-3β*, part of the *β-catenin* destruction complex, inhibits *Notch* transcriptional activation of target genes in kidney and fibroblast tissue culture cells[145], whereas LiCl treatment, which inhibits *GSK-3β*, increases *Notch* levels and target gene transcription in lung cancer[152] and tissue culture cells[153]. *GSK-3β* also interacts with the *Notch* transcriptional coactivator, *MAML1* to decrease its activity[154], and *NLK*, which suppresses canonical Wnt signaling[155], also impairs *Notch* activity[156].

Conversely, *dishevelled*, a key mediator of Wnt signaling binds *Notch* and blocks *Notch* signaling[157,158], and Wnt downregulates *Notch* to promote proliferation/maintenance of prostate epithelial progenitors[159]. Presenilins, necessary for *Notch* cleavage and activation promote *β-catenin* degradation[160,161], *Notch* associates with activated *arm* to decrease its concentration and activity[162], and *Notch* may inhibit Wnt signaling by maintaining the *dTCf* repressive state in fly visceral mesoderm[163]. In contrast to its effect on *β-catenin*, *GSK3β* increases NICD stability and signaling in fibroblasts[164], and it is subject to reciprocal effects on its activation status by Notch (activation maintained) and Wnt (*GSK3β* is inactivated) signaling during the temporal switch from Notch-mediated proliferative phase to Wnt-mediated differentiation phase during myogenesis[165]. Given the diversity and ubiquity of interactions between Notch and Wnt in development and disease, it has been proposed that Wnt and Notch signaling comprise an integrated cellular system (Wntch) to exert mutual control in the determination of cell fate[139,166]. Depending on the tissues and developmental stages during which changes in Notch and/or canonical Wnt signaling act to promote hypoxia tolerance, conceivably, either pathway might be activated or repressed, with the two pathways cooperating or antagonistic. How non-canonical Wnt signaling might interact with Notch signaling is even less clear.

Experimental validation of Notch pathway activation in hypoxia tolerance in flies was performed by scientific collaborators and is described in Zhou *et al*[48]. Further exploration of the contribution of Wnt signaling to hypoxia tolerance in flies was performed as part of the current dissertation.

Supplemental Figures



	Median Lifespan	Maximum Lifespan
0.00 mM DAPT (n=300)	15.2 ± 1.04	20.1 ± 1.25
0.25 mM DAPT (n=300)	12.3 ± 0.58*	15.1 ± 0.55**
0.50 mM DAPT (n=300)	11.7 ± 0.58*	13.1 ± 0.72**

Figure S3.1 Hypoxia tolerance in Notch mutants and γ-secretase inhibitor-treated flies. This work was performed by D. Zhou[48] . (A) Contribution of Notch to hypoxia tolerance was determined by culturing the homozygous-viable Notch mutants in 6% O2 (relatively mild) or 4% O2 (severe) hypoxic conditions. Canton-S was used as a control. N[spl-1], which is a gain-of-function Notch allele, has a dramatically increased survival rate in hypoxic conditions. In contrast, the N[fa-1] and N[Ax-tsl] loss-of-function mutants had little or no eclosion in severe hypoxic condition and showed a reduced survival rate even in mild hypoxic condition (*P < 0.01, compared with Canton-S control). (B) RNAi-mediated Notch knockdown induces increased sensitivity to hypoxia in flies. Flies carrying a UAS-N.dsRNA transgene on the X chromosome were used to determine the function of Notch in hypoxia tolerance. Two crosses were used to generate flies that had N.dsRNA expression either in all progeny [cross A: (UAS-N.dsRNA virgin female) × (da-GAL4 male)] or only in the female progeny [cross B: (da-GAL4) × (UAS-N.dsRNA)]. Cross A progeny showed an increased sensitivity to hypoxia (*P < 0.01 compared with the cross B control). This hypersensitivity was rescued in cross B progeny in which the function of Notch was knocked down only in females. Both male and female progeny were included in the scoring of A and B crosses. (C) γ-Secretase activation plays important role in hypoxia tolerance in hypoxia-selected flies. Five-day-old adult hypoxia-selected flies were collected and treated with 0.25 or 0.50 mM DAPT and their life span was determined under 1.5% O2. Median life span was the time when 50% of death occurred in the sample, and maximum life span was the time when 90% of sample were dead. Compared with the vehicle-treated controls, flies treated with DAPT showed a significant reduction of both median (*P < 0.01) and maximum life span (**P < 0.01). The median and maximum life spans were calculated using GraphPad Prism 4, and the statistical significance was calculated by Student's t test.

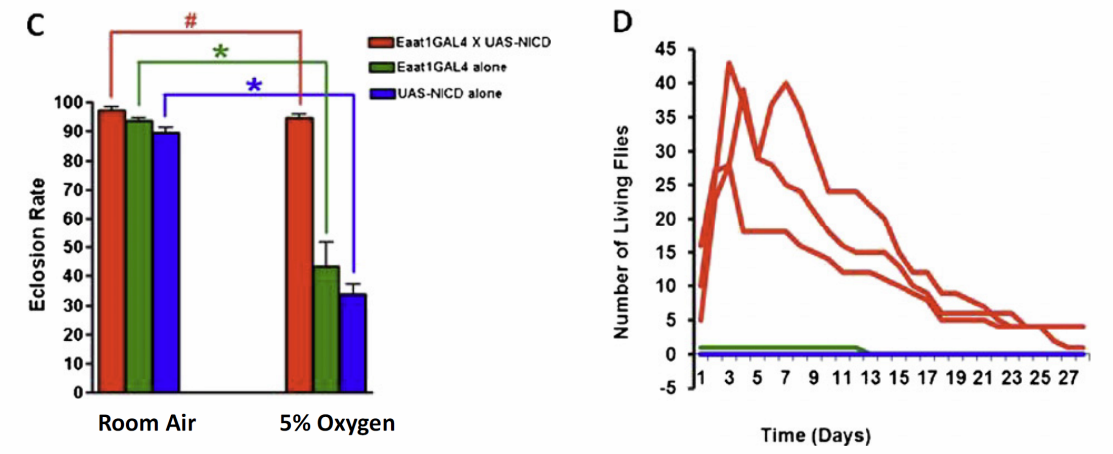


Figure S3.2 Increased hypoxia tolerance in flies with NICD overexpression in specific glial cells. This work was performed by D. Visk[48]. Figures A and B from the published paper are not presented here. (C) Glial-specific overexpression of NICD enhances the eclosion rate in hypoxia. #P value = 0.205; *P value <0.001 (unpaired Student's t test). Bars represent the mean ± SEM for six vials (n = 6). (D) Glial-specific overexpression of NICD enhances adult survival in hypoxia. Adult survival during hypoxia was evaluated by counting the number of newly eclosed experimental and control flies and transferring them to a new individual vial. The following day, the number of dead flies in the adult-only vial was subtracted from the previous day's total and the newly eclosed flies from the original vial were counted and added to the adult-only vial to avoid a cumbersome number of vials for all tests done. Hence, the number of flies in an adult-only vial could go down (due to adult death), but later increase (due to addition of newly eclosed adults from the original vial to the adult-only vial).

Supplemental Tables

Table S3.1 Biological Process Enrichment of Hypoxia-tolerance Polymorphisms.
Annotation was done using DAVID[122,123]-defined Gene Ontology-Biological Process "FAT" categories; enrichment assumed a *D. melanogaster* background. All annotations with Benjamini-corrected p <= 0.05 are included. All non-bolded annotations relate to development or morphogenesis.

Term	Count	PValue	Fold Enrichment	Benjamini	FDR
GO:0022610~biological adhesion	41	9.95E-10	2.899	1.91E-06	1.70E-06
GO:0030182~neuron differentiation	66	3.19E-09	2.145	3.06E-06	5.46E-06
GO:0048666~neuron development	58	9.53E-09	2.222	6.10E-06	1.63E-05
GO:0000904~cell morphogenesis involved in differentiation	52	2.64E-08	2.282	1.27E-05	4.52E-05
GO:0007155~cell adhesion	36	5.17E-08	2.751	1.98E-05	8.84E-05
GO:0030030~cell projection organization	58	6.08E-08	2.113	1.95E-05	1.04E-04
GO:0048667~cell morphogenesis involved in neuron differentiation	49	9.48E-08	2.262	2.60E-05	1.62E-04
GO:0007552~metamorphosis	57	1.12E-07	2.093	2.69E-05	1.92E-04
GO:0048707~instar larval or pupal morphogenesis	55	1.55E-07	2.107	3.31E-05	2.66E-04
GO:0000902~cell morphogenesis	65	1.65E-07	1.955	3.16E-05	2.82E-04
GO:0048812~neuron projection morphogenesis	48	2.00E-07	2.231	3.50E-05	3.43E-04
GO:0031175~neuron projection development	48	2.23E-07	2.224	3.57E-05	3.82E-04
GO:0048569~post-embryonic organ development	49	2.49E-07	2.193	3.67E-05	4.25E-04
GO:0009886~post-embryonic morphogenesis	55	2.74E-07	2.071	3.76E-05	4.69E-04
GO:0009791~post-embryonic development	63	4.22E-07	1.930	5.41E-05	7.23E-04
GO:0007409~axonogenesis	37	4.72E-07	2.484	5.67E-05	8.08E-04
GO:0035120~post-embryonic appendage morphogenesis	40	9.01E-07	2.322	1.02E-04	0.0015
GO:0048858~cell projection morphogenesis	50	9.81E-07	2.077	1.05E-04	0.0017
GO:0035114~imaginal disc-derived appendage morphogenesis	41	1.04E-06	2.281	1.05E-04	0.0018

Table S3.1 Biological Process Enrichment of Hypoxia-tolerance Polymorphisms. (cont)

Term	Count	PValue	Fold Enrichment	Benjamini	FDR
GO:0048563~post-embryonic organ morphogenesis	46	1.06E-06	2.153	1.01E-04	0.0018
GO:0007560~imaginal disc morphogenesis	46	1.06E-06	2.153	1.01E-04	0.0018
GO:0002165~instar larval or pupal development	60	1.23E-06	1.908	1.13E-04	0.0021
GO:0035107~appendage morphogenesis	41	1.44E-06	2.252	1.25E-04	0.0025
GO:0048737~imaginal disc-derived appendage development	41	1.60E-06	2.243	1.34E-04	0.0027
GO:0048736~appendage development	41	2.20E-06	2.216	1.76E-04	0.0038
GO:0032990~cell part morphogenesis	50	2.66E-06	2.008	2.04E-04	0.0046
GO:0007476~imaginal disc-derived wing morphogenesis	37	4.97E-06	2.256	3.67E-04	0.0085
GO:0007472~wing disc morphogenesis	37	6.16E-06	2.236	4.38E-04	0.0105
GO:0007411~axon guidance	27	7.34E-06	2.639	5.03E-04	0.0126
GO:0007444~imaginal disc development	57	7.58E-06	1.839	5.01E-04	0.0130
GO:0032989~cellular component morphogenesis	66	1.97E-05	1.694	0.0013	0.0337
GO:0016337~cell-cell adhesion	16	2.03E-05	3.605	0.0013	0.0348
GO:0035220~wing disc development	42	2.53E-05	1.987	0.0015	0.0433
GO:0007166~cell surface receptor linked signal transduction	**76**	**2.60E-05**	**1.609**	**0.0015**	**0.0445**
GO:0035239~tube morphogenesis	19	3.88E-05	3.007	0.0022	0.0664
GO:0006793~phosphorus metabolic process	**66**	**1.34E-04**	**1.592**	**0.0073**	**0.2286**
GO:0006796~phosphate metabolic process	**66**	**1.34E-04**	**1.592**	**0.0073**	**0.2286**
GO:0007186~G-protein coupled receptor protein signaling pathway	**43**	**1.55E-04**	**1.821**	**0.0083**	**0.2656**
GO:0006928~cell motion	42	1.67E-04	1.831	0.0086	0.2847
GO:0008038~neuron recognition	13	2.01E-04	3.527	0.0101	0.3433
GO:0008037~cell recognition	13	2.47E-04	3.457	0.0121	0.4213
GO:0035218~leg disc development	13	3.01E-04	3.389	0.0143	0.5139

Table S3.1 Biological Process Enrichment of Hypoxia-tolerance Polymorphisms. (cont)

Term	Count	PValue	Fold Enrichment	Benjamini	FDR
GO:0035295~tube development	19	3.55E-04	2.552	0.0165	0.6058
GO:0048813~dendrite morphogenesis	21	3.99E-04	2.386	0.0181	0.6804
GO:0016358~dendrite development	21	3.99E-04	2.386	0.0181	0.6804
GO:0060446~branching involved in open tracheal system development	11	4.56E-04	3.750	0.0202	0.7783
GO:0048754~branching morphogenesis of a tube	11	4.56E-04	3.750	0.0202	0.7783
GO:0042067~establishment of ommatidial polarity	12	4.60E-04	3.468	0.0199	0.7847
GO:0001736~establishment of planar polarity	14	4.82E-04	3.051	0.0204	0.8220
GO:0030166~proteoglycan biosynthetic process	**7**	**4.86E-04**	**6.204**	**0.0201**	**0.8285**
GO:0007164~establishment of tissue polarity	14	5.69E-04	3.002	0.0230	0.9696
GO:0006413~translational initiation	**13**	**6.31E-04**	**3.142**	**0.0249**	**1.0738**
GO:0006029~proteoglycan metabolic process	**7**	**7.29E-04**	**5.816**	**0.0282**	**1.2397**
GO:0001763~morphogenesis of a branching structure	11	8.61E-04	3.482	0.0325	1.4628
GO:0060562~epithelial tube morphogenesis	11	8.61E-04	3.482	0.0325	1.4628
GO:0030431~sleep	**5**	**9.17E-04**	**9.496**	**0.0339**	**1.5576**
GO:0007424~open tracheal system development	24	9.98E-04	2.085	0.0362	1.6948
GO:0060541~respiratory system development	24	9.98E-04	2.085	0.0362	1.6948
GO:0006470~protein amino acid dephosphorylation	**16**	**0.0010111**	**2.594**	**0.0360**	**1.7161**
GO:0001737~establishment of imaginal disc-derived wing hair orientation	7	0.00105545	5.474	0.0369	1.7907
GO:0007157~heterophilic cell adhesion	6	0.00119165	6.647	0.0408	2.0196
GO:0048859~formation of anatomical boundary	11	0.00126658	3.324	0.0425	2.1453
GO:0046530~photoreceptor cell differentiation	21	0.00129095	2.181	0.0426	2.1862
GO:0060429~epithelium development	32	0.00139916	1.810	0.0453	2.3673
GO:0007398~ectoderm development	14	0.00141723	2.737	0.0451	2.3976

Table S3.1 Biological Process Enrichment of Hypoxia-tolerance Polymorphisms. (cont)

Term	Count	PValue	Fold Enrichment	Benjamini	FDR
GO:0016203~muscle attachment	8	0.00145333	4.432	0.0455	2.4579

GO:0007049~cell cycle	**52**	**0.2462632**	**1.122**	**0.844**	**99.2074**

GO:0048477~oogenesis	**45**	**0.2985956**	**1.110**	**0.885**	**99.7686**

GO:0006119~oxidative phosphorylation	**11**	**0.536**	**1.108**	**0.972**	**99.9998**

Table S3.2 Pathway Enrichment of Hypoxia-tolerance Polymorphisms.
Annotation was done using DAVID[122,123]; enrichment assumed a *D. melanogaster* background.

Category	Term	Count	PVal	Genes	Fold Enriched
KEGG	dme04310:Wnt signaling pathway	15	0.002	FBGN0026597, FBGN0011655, FBGN0010015, FBGN0011826, FBGN0010333, FBGN0000119, FBGN0011817, FBGN0030505, FBGN0030758, FBGN0026174, FBGN0026181, FBGN0011577, FBGN0026175, FBGN0003371, FBGN0014020	2.449
KEGG	dme00020:Citrate cycle (TCA cycle)	9	0.024	FBGN0035240, FBGN0038922, FBGN0030975, FBGN0035239, FBGN0029722, FBGN0027291, FBGN0052026, FBGN0028325, FBGN0036162, FBGN0001248	2.471
KEGG	dme04650:Natural killer cell mediated cytotoxicity	6	0.030	FBGN0030758, FBGN0010015, FBGN0011826, FBGN0040068, FBGN0010333, FBGN0030505	3.295
KEGG	dme04711:Circadian rhythm	4	0.043	FBGN0016694, FBGN0002413, FBGN0014396, FBGN0003371	4.833
KEGG	dme04320:Dorso-ventral axis formation	6	0.057	FBGN0000256, FBGN0000810, FBGN0003731, FBGN0004647, FBGN0001404, FBGN0003118	2.788
…					
KEGG	dme04350:TGF-beta signaling pathway	6	0.142	FBGN0026174, FBGN0026181, FBGN0011655, FBGN0003463, FBGN0026175, FBGN0014020	2.132
KEGG	dme04070:Phosphatidylinositol signaling system	6	0.244	FBGN0051140, FBGN0085388, FBGN0085413, FBGN0085373, FBGN0030761, FBGN0040335	1.768
KEGG	dme04150:mTOR signaling pathway	4	0.386	FBGN0010715, FBGN0015542, FBGN0035709, FBGN0035860	1.790
KEGG	dme04013:MAPK signaling pathway	3	0.443	FBGN0003731, FBGN0003720, FBGN0003118	2.014
KEGG	dme04330:Notch signaling pathway	2	0.850	FBGN0020388, FBGN0004647	1.098
KEGG	dme04340:Hedgehog signaling pathway	2	0.874	FBGN0002413, FBGN0003371	1.007

Table S3.2 Pathway Enrichment of Hypoxia-tolerance Polymorphisms. (cont)

Category	Term	Count	PVal	Genes	Fold Enriched
PANTHER	P00012:Cadherin signaling pathway	10	0.024	FBGN0000723, FBGN0011742, FBGN0001085, FBGN0046332, FBGN0003731, FBGN0003138, FBGN0001075, FBGN0003371, FBGN0000497, FBGN0039709	2.272
PANTHER	P00057:Wnt signaling pathway	18	0.055	FBGN0259680, FBGN0043900, FBGN0001085, FBGN0046332, FBGN0011655, FBGN0000472, FBGN0010015, FBGN0011826, FBGN0001075, FBGN0011817, FBGN0000497, FBGN0052683, FBGN0020306, FBGN0025463, FBGN0002413, FBGN0030093, FBGN0003371, FBGN0039709	1.555
PANTHER	P00007:Axon guidance mediated by semaphorins	5	0.057	FBGN0000723, FBGN0025743, FBGN0035574, FBGN0010333, FBGN0014020	3.266
PANTHER	P00026:Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway	9	0.084	FBGN0004573, FBGN0039747, FBGN0250910, FBGN0004834, FBGN0046332, FBGN0038063, FBGN0036789, FBGN0003371, FBGN0030087	1.919
PANTHER	P00027:Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway	7	0.084	FBGN0039747, FBGN0030444, FBGN0259680, FBGN0053517, FBGN0035574, FBGN0036789, FBGN0014020	2.217
…					
PANTHER	P00033:Insulin/IGF pathway-protein kinase B signaling cascade	4	0.621	FBGN0031086, FBGN0046332, FBGN0045759, FBGN0003371	1.267
PANTHER	P00025:Hedgehog signaling pathway	3	0.744	FBGN0002413, FBGN0046332, FBGN0003371	1.161
PANTHER	P00052:TGF-beta signaling pathway	4	0.916	FBGN0031086, FBGN0015789, FBGN0011655, FBGN0045759	0.746
PANTHER	P00018:EGF receptor signaling pathway	4	0.932	FBGN0259680, FBGN0038603, FBGN0003731, FBGN0010333	0.709
PANTHER	P00045:Notch signaling pathway	2	0.946	FBGN0020388, FBGN0004647	0.721

Table S3.3 SNPs in Wnt-Pathway Associated Genes.
Location of SNPs within extended gene (gene + 2000 bases up- and downstream) for Wnt-associated genes. The base call at each SNP locus is provided for the reference, control and AF genomes. Asterisk denotes that the SNP also maps to an additional gene.

SNP-Chrom		SNP-Location	FBgn	Gene	Ref	Control	H-1	H-2
3R		25855268	FBgn0026597	Axn	A	A	G	G
3R		25856260	FBgn0026597	Axn	C	C	G	G
3R		25857098	FBgn0026597	Axn	C	C	G	G
3R		25857108	FBgn0026597	Axn	A	A	G	G
3R		25857149	FBgn0026597	Axn	T	T	C	C
3R		25857899	FBgn0026597	Axn	T	T	A	A
3R		25858547	FBgn0026597	Axn	C	C	A	A
3L		1060268	FBgn0004870	bab1	A	A	G	G
3L		1066588	FBgn0004870	bab1	T	T	A	A
X	*	13051724	FBgn0025463	Bap60	C	C	T	T
X	*	13051921	FBgn0025463	Bap60	G	G	C	C
X	*	13051929	FBgn0025463	Bap60	G	G	A	A
3R		25678061	FBgn0039709	Cad99C	C	C	T	T
3R	*	25678724	FBgn0039709	Cad99C	T	T	C	C
3R	*	25680582	FBgn0039709	Cad99C	A	A	T	T
3R	*	26868343	FBgn0010015	CanA1	C	C	T	T
3R	*	26868364	FBgn0010015	CanA1	T	T	C	C
3R	*	26868453	FBgn0010015	CanA1	T	T	C	C
X		16468264	FBgn0030758	CanA-14F	G	G	T	T
X		16470744	FBgn0030758	CanA-14F	T	T	C	C
X		16471027	FBgn0030758	CanA-14F	T	T	G	G
X	*	16480873	FBgn0030758	CanA-14F	G	G	A	A
X	*	16480901	FBgn0030758	CanA-14F	G	G	A	A
X	*	16482157	FBgn0030758	CanA-14F	G	G	A	A
X	*	16482356	FBgn0030758	CanA-14F	G	G	A	A
X	*	16482494	FBgn0030758	CanA-14F	T	T	A	A
3R	*	17691971	FBgn0067317	Cby	T	T	A	A
3R	*	17691989	FBgn0067317	Cby	A	A	G	G
3R	*	17692079	FBgn0067317	Cby	T	T	C	C
3R	*	17692121	FBgn0067317	Cby	T	T	C	C
3R	*	17692411	FBgn0067317	Cby	A	A	G	G
3R	*	17694243	FBgn0067317	Cby	C	C	T	T
3L	*	6736271	FBgn0035713	CG10107	G	G	T	T

Table S3.3 SNPs in Wnt-Pathway Associated Genes. (cont)

SNP-Chrom		SNP-Location	FBgn	Gene	Ref	Control	H-1	H-2
3L	*	6736715	FBgn0035713	CG10107	A	A	G	G
3L	*	6736925	FBgn0035713	CG10107	C	C	T	T
3L	*	6736983	FBgn0035713	CG10107	C	C	T	T
3L	*	6737096	FBgn0035713	CG10107	A	A	T	T
3L	*	6737119	FBgn0035713	CG10107	A	A	C	C
3L		6743650	FBgn0035713	CG10107	C	C	G	G
3R		24898471	FBgn0039633	CG11873	T	T	C	C
3R		24898541	FBgn0039633	CG11873	A	A	G	G
3R		24898572	FBgn0039633	CG11873	C	C	A	A
3R		24924998	FBgn0039633	CG11873	G	G	C	C
3R		24926626	FBgn0039633	CG11873	A	A	T	T
3R		24926645	FBgn0039633	CG11873	G	G	A	A
3L	*	13373224	FBgn0036351	CG14107	G	G	C	C
3L		12325640	FBgn0052105	CG32105	G	G	C	C
X		16315878	FBgn0026575	hang	T	T	C	C
X		16315939	FBgn0026575	hang	G	G	T	T
X		16318068	FBgn0026575	hang	A	A	G	G
X		16318530	FBgn0026575	hang	G	G	T	T
X		16318819	FBgn0026575	hang	G	G	C	C
X		16318865	FBgn0026575	hang	A	A	G	G
X		16318943	FBgn0026575	hang	C	C	G	G
X		16318984	FBgn0026575	hang	G	G	C	C
X		16318993	FBgn0026575	hang	T	T	C	C
X		16319584	FBgn0026575	hang	C	C	T	T
X		16320246	FBgn0026575	hang	T	T	A	A
X		16321056	FBgn0026575	hang	C	C	T	T
X		16321537	FBgn0026575	hang	T	T	C	C
X		16322274	FBgn0026575	hang	C	C	G	G
X		16322483	FBgn0026575	hang	A	A	G	G
X		16323014	FBgn0026575	hang	T	T	G	G
X		16323338	FBgn0026575	hang	C	C	T	T
X		16323466	FBgn0026575	hang	A	A	C	C
X		16323533	FBgn0026575	hang	T	T	A	A
X		16323941	FBgn0026575	hang	A	A	T	T
X		16324767	FBgn0026575	hang	C	C	T	T
X		16324823	FBgn0026575	hang	G	G	C	C

Table S3.3 SNPs in Wnt-Pathway Associated Genes. (cont)

SNP-Chrom		SNP-Location	FBgn	Gene	Ref	Control	H-1	H-2
X		16324886	FBgn0026575	hang	C	C	A	A
X		16324923	FBgn0026575	hang	T	T	C	C
X		16325000	FBgn0026575	hang	T	T	C	C
X		16325027	FBgn0026575	hang	G	G	A	A
X		16325138	FBgn0026575	hang	A	A	G	G
X		16325270	FBgn0026575	hang	G	G	A	A
X		16325480	FBgn0026575	hang	A	A	G	G
X		16325894	FBgn0026575	hang	T	T	C	C
X		16325939	FBgn0026575	hang	C	C	T	T
X		16325942	FBgn0026575	hang	A	A	G	G
X		16325975	FBgn0026575	hang	T	T	C	C
X		16326152	FBgn0026575	hang	C	C	T	T
X		16326182	FBgn0026575	hang	G	G	A	A
X		16326212	FBgn0026575	hang	A	A	G	G
X		16326458	FBgn0026575	hang	C	C	T	T
X		16327253	FBgn0026575	hang	T	T	C	C
X		16327985	FBgn0026575	hang	T	T	C	C
X		16327994	FBgn0026575	hang	A	A	T	T
X		16328249	FBgn0026575	hang	G	G	T	T
X		16328265	FBgn0026575	hang	A	A	C	C
X		16329558	FBgn0026575	hang	G	G	A	A
X		16329702	FBgn0026575	hang	C	C	A	A
X		16329714	FBgn0026575	hang	A	A	G	G
X		16329972	FBgn0026575	hang	C	C	T	T
X		16330252	FBgn0026575	hang	A	A	C	C
X	*	16333134	FBgn0026575	hang	G	G	A	A
X	*	16333145	FBgn0026575	hang	A	A	G	G
3L		12284712	FBgn0036274	CG4328	A	A	C	C
X		18537604	FBgn0027335	Rip11	T	T	C	C
3R	*	7514334	FBgn0037935	CG6834	G	G	T	T
3L		11120635	FBgn0260795	CG42575	T	T	C	C
3L		11120643	FBgn0260795	CG42575	T	T	G	G
3L	*	11126284	FBgn0260795	CG42575	C	C	T	T
3R	*	25621714	FBgn0039696	CG7837	T	T	C	C
3R	*	25621888	FBgn0039696	CG7837	A	A	G	G
3R	*	25623948	FBgn0039696	CG7837	A	A	G	G

Table S3.3 SNPs in Wnt-Pathway Associated Genes. (cont)

SNP-Chrom		SNP-Location	FBgn	Gene	Ref	Control	H-1	H-2
3R	*	25623950	FBgn0039696	CG7837	A	A	C	C
3R	*	25624104	FBgn0039696	CG7837	A	A	G	G
3R	*	25624556	FBgn0039696	CG7837	T	T	A	A
3R	*	25625952	FBgn0039696	CG7837	T	T	C	C
3R	*	13825334	FBgn0000363	cpo	G	G	A	A
X	*	9047536	FBgn0030093	dalao	T	T	C	C
X	*	9047809	FBgn0030093	dalao	G	G	A	A
3R		26882410	FBgn0002413	dco	G	G	A	A
3R		26882515	FBgn0002413	dco	T	T	A	A
3R		26882761	FBgn0002413	dco	T	T	C	C
3R		26882794	FBgn0002413	dco	G	G	A	A
3R		26886394	FBgn0002413	dco	T	T	A	A
X		3278089	FBgn0000472	dm	G	G	A	A
2R	*	17219902	FBgn0020306	dom	T	T	C	C
2R	*	17220640	FBgn0020306	dom	C	C	T	T
2L		659638	FBgn0000497	ds	T	T	C	C
2L		4214127	FBgn0001075	ft	T	T	G	G
2L		4214130	FBgn0001075	ft	A	A	T	T
3L		14299897	FBgn0001085	fz	T	T	G	G
3R	*	27196173	FBgn0046332	gskt	C	C	G	G
3R		26156142	FBgn0010113	hdc	T	T	G	G
3R		26157633	FBgn0010113	hdc	G	G	C	C
3R		26174967	FBgn0010113	hdc	T	T	A	A
3R		26177602	FBgn0010113	hdc	A	A	G	G
3R		26184449	FBgn0010113	hdc	C	C	T	T
3R		26184459	FBgn0010113	hdc	C	C	T	T
X		13543214	FBgn0030505	NFAT	C	C	T	T
3L		7976664	FBgn0011817	nmo	C	C	T	T
3L	*	8023055	FBgn0011817	nmo	C	C	T	T
3L	*	8023278	FBgn0011817	nmo	A	A	G	G
3L		8030970	FBgn0011817	nmo	G	G	A	A
3L		8032191	FBgn0011817	nmo	G	G	A	A
3L		8035281	FBgn0011817	nmo	T	T	C	C
3L		8044386	FBgn0011817	nmo	T	T	A	A
3L		8044395	FBgn0011817	nmo	A	A	T	T
X		12324763	FBgn0259680	Pkcdelta	C	C	A	A

Table S3.3 SNPs in Wnt-Pathway Associated Genes. (cont)

SNP-Chrom		SNP-Location	FBgn	Gene	Ref	Control	H-1	H-2
X		12344985	FBgn0259680	Pkcdelta	G	G	C	C
3R		19126196	FBgn0003118	pnt	T	T	C	C
3R		19169700	FBgn0003118	pnt	C	C	T	T
3R		19169703	FBgn0003118	pnt	A	A	C	C
3R		19171816	FBgn0003118	pnt	A	A	T	T
3R	*	19171970	FBgn0003118	pnt	T	T	C	C
3R	*	19173445	FBgn0003118	pnt	C	C	G	G
X		16451640	FBgn0011826	Pp2B-14D	C	C	T	T
X		16451663	FBgn0011826	Pp2B-14D	C	C	G	G
X		16453000	FBgn0011826	Pp2B-14D	A	A	G	G
X		16453114	FBgn0011826	Pp2B-14D	A	A	G	G
X		16453381	FBgn0011826	Pp2B-14D	G	G	A	A
X		16453645	FBgn0011826	Pp2B-14D	G	G	A	A
X		16453711	FBgn0011826	Pp2B-14D	C	C	T	T
3R		27403752	FBgn0043900	pygo	A	A	T	T
3L	*	1300237	FBgn0010333	Rac1	G	G	A	A
3L	*	1303255	FBgn0010333	Rac1	A	A	G	G
X		16522505	FBgn0026181	rok	C	C	T	T
X	*	16523825	FBgn0026181	rok	A	A	G	G
X	*	16530148	FBgn0026181	rok	G	G	T	T
X	*	16530173	FBgn0026181	rok	G	G	A	A
X	*	16530435	FBgn0026181	rok	T	T	C	C
X	*	16530805	FBgn0026181	rok	G	G	C	C
X	*	16530963	FBgn0026181	rok	T	T	G	G
X	*	16531499	FBgn0026181	rok	A	A	C	C
X	*	16531593	FBgn0026181	rok	G	G	A	A
X	*	16531595	FBgn0026181	rok	A	A	C	C
X	*	16532133	FBgn0026181	rok	G	G	C	C
X	*	16532262	FBgn0026181	rok	G	G	A	A
X	*	16532401	FBgn0026181	rok	C	C	T	T
X	*	16532402	FBgn0026181	rok	C	C	A	A
3L		6532545	FBgn0020251	sfl	A	A	G	G
3L		6534030	FBgn0020251	sfl	A	A	G	G
3L		6534065	FBgn0020251	sfl	T	T	G	G
3L		6540733	FBgn0020251	sfl	C	C	G	G
X		2530444	FBgn0003371	sgg	T	T	G	G

Table S3.3 SNPs in Wnt-Pathway Associated Genes. (cont)

SNP-Chrom		SNP-Location	FBgn	Gene	Ref	Control	H-1	H-2
X		2532290	FBgn0003371	sgg	A	A	C	C
X		2553538	FBgn0003371	sgg	C	C	T	T
X		2554975	FBgn0003371	sgg	T	T	C	C
X		2560726	FBgn0003371	sgg	G	G	A	A
3R	*	25883225	FBgn0015542	sima	T	T	C	C
3R	*	25888163	FBgn0015542	sima	A	A	G	G
3R	*	25890435	FBgn0015542	sima	C	C	G	G
3R		25894002	FBgn0015542	sima	A	A	G	G
3R		25894059	FBgn0015542	sima	C	C	T	T
3R		25897834	FBgn0015542	sima	C	T	A	A
3R		25897835	FBgn0015542	sima	G	G	T	T
3R		25920200	FBgn0015542	sima	G	G	A	A
3R		25925062	FBgn0015542	sima	T	T	C	C
3R		25925100	FBgn0015542	sima	T	T	C	C
3R		25928248	FBgn0015542	sima	T	T	G	G
3R		25934509	FBgn0015542	sima	C	C	T	T
3R		25940704	FBgn0015542	sima	A	A	T	T
3R	*	25948527	FBgn0015542	sima	C	C	G	G
X		19711825	FBgn0026175	skpC	G	G	A	A
X		19705790	FBgn0026174	skpD	T	T	C	C
X		12584535	FBgn0024308	Smr	T	T	C	C
X		12584544	FBgn0024308	Smr	T	T	C	C
X		12584820	FBgn0024308	Smr	T	T	C	C
X		12584838	FBgn0024308	Smr	A	A	G	G
X		12584850	FBgn0024308	Smr	C	C	A	A
X		12586881	FBgn0024308	Smr	T	T	C	C
X		12594631	FBgn0024308	Smr	C	C	G	G
X	*	12600108	FBgn0024308	Smr	T	T	C	C
X	*	12600167	FBgn0024308	Smr	C	C	T	T
X	*	12601629	FBgn0024308	Smr	T	T	C	C
X	*	12601631	FBgn0024308	Smr	C	C	A	A
X	*	12603364	FBgn0024308	Smr	T	T	G	G
X	*	12603450	FBgn0024308	Smr	A	A	G	G
X		12612904	FBgn0024308	Smr	A	A	C	C
X		12621985	FBgn0024308	Smr	G	G	C	C
3L		10853267	FBgn0026160	tna	A	A	C	C

Table S3.3 SNPs in Wnt-Pathway Associated Genes. (cont)

SNP-Chrom		SNP-Location	FBgn	Gene	Ref	Control	H-1	H-2
3L		5838466	FBgn0003984	vn	G	G	A	A
3R		26618196	FBgn0011739	wts	C	C	T	T

Table S3.4 Indels in Wnt-Pathway Associated Genes.
Location of small indels within extended gene (gene + 2000 bases up- and downstream) for Wnt-associated genes. The size (bases) of each indel is given. Asterisk denotes that the SNP also maps to an additional gene.

Indel-Chrom		Indel-Location	FBgn	Gene	H-1 Indel Size	H-2 Indel Size
2R		9361436	FBgn0000119	arr	-4	-4
3L		1067902	FBgn0004870	bab1	-12	-12
3R	*	16419218	FBgn0023097	bon	7	7
3R	*	16419269	FBgn0023097	bon	-3	-3
X	*	16480046	FBgn0030758	CanA-14F	1	1
3L	*	6736946	FBgn0035713	CG10107	-1	-1
3L		20843272	FBgn0052432	CG32432	-1	-1
X		16324021	FBgn0026575	hang	-3	-3
X	*	10310448	FBgn0052683	CG32683	1	1
3L		11120212	FBgn0260795	CG42575	1	1
X		19092498	FBgn0030997	CG7990	1	1
3L	*	8833831	FBgn0011577	dally	1	1
3R		26881704	FBgn0002413	dco	-1	-1
3R		14292859	FBgn0004652	fru	1	1
3R		26152119	FBgn0010113	hdc	1	1
3R		26155075	FBgn0010113	hdc	-1	-1
2L		2918409	FBgn0041111	lilli	-4	-4
3R	*	27437643	FBgn0011655	Med	2	2
3L		8030913	FBgn0011817	nmo	1	1
X		16454010	FBgn0011826	Pp2B-14D	1	1
2R	*	11992045	FBgn0014020	Rho1	-2	-1
X	*	16530163	FBgn0026181	rok	-3	-3
X	*	16531620	FBgn0026181	rok	2	2
X	*	16531813	FBgn0026181	rok	1	1
X	*	16532424	FBgn0026181	rok	-1	-1
3L		6531103	FBgn0020251	sfl	-1	-1
X		2538552	FBgn0003371	sgg	1	1
3L	*	6957020	FBgn0010851	sgl	-3	-3
3R	*	25890695	FBgn0015542	sima	-1	-1
3R		25921624	FBgn0015542	sima	-1	-1
3R		25922644	FBgn0015542	sima	-2	-2
3R		25925309	FBgn0015542	sima	-1	-1
3R		25928578	FBgn0015542	sima	-3	-3

Table S3.4 Indels in Wnt-Pathway Associated Genes. (cont)

Indel-Chrom		Indel-Location	FBgn	Gene	H-1 Indel Size	H-2 Indel Size
X		19711833	FBgn0026175	skpC	2	2
3L		10858407	FBgn0026160	tna	-1	-1
2R		10428290	FBgn0020245	ttv	-1	-1
3R		26628332	FBgn0011739	wts	4	4

Table S3.5 Coding Region Polymorphisms in Wnt pathway-Associated Genes.
For each coding region SNP inWnt-associated genes, the location, strand, reading frame, and reference and H1/H2 codon and amino acid (AA) is provided and the SNP is identified as synonymous (S) or non-synonymous (NS).

Gene	SNP-Chrom	SNP-Location	Strand	Frame	Ref-Codon	Ref-AA	H1/H2-Codon	H1/H2-AA	Type
Axn	3R	25857149	+	1	TTG	L	CTG	L	S
Cad99C	3R	25678061	+	0	GAC	D	GAT	D	S
Cad99C	3R	25678724	+	0	CGT	R	CGC	R	S
CanA1	3R	26868343	+	0	AGC	S	AGT	S	S
CanA1	3R	26868364	+	0	TTT	F	TTC	F	S
Cby	3R	17694243	+	0	CTC	L	CTT	L	S
CG11873	3R	24924998	+	0	AGC	S	ACC	T	NS
CG11873	3R	24926626	+	0	ACT	T	TCT	S	NS
CG11873	3R	24926645	+	0	AGC	S	AAC	N	NS
hang	X	16323338	+	0	CTC	L	CTT	L	S
hang	X	16323466	+	0	CAA	Q	CCA	P	NS
hang	X	16323533	+	0	GCT	A	GCA	A	S
hang	X	16323941	+	0	ACA	T	ACT	T	S
hang	X	16324767	+	0	CGC	R	CGT	R	S
hang	X	16325000	+	0	GCT	A	GCC	A	S
hang	X	16325027	+	0	CCG	P	CCA	P	S
hang	X	16325138	+	0	CTA	L	CTG	L	S
hang	X	16325270	+	0	CGG	R	CGA	R	S
hang	X	16325480	+	0	AAA	K	AAG	K	S
hang	X	16325894	+	0	ACT	T	ACC	T	S
hang	X	16325939	+	0	GTC	V	GTT	V	S
hang	X	16325942	+	0	CAA	Q	CAG	Q	S
hang	X	16325975	+	0	TTT	F	TTC	F	S
hang	X	16326152	+	0	AAC	N	AAT	N	S
hang	X	16326182	+	0	AGG	R	AGA	R	S
hang	X	16326212	+	0	AAA	K	AAG	K	S
hang	X	16326458	+	0	AAC	N	AAT	N	S
hang	X	16327253	+	2	AGT	S	AGC	S	S
hang	X	16327985	+	2	TCT	S	TCC	S	S
hang	X	16327994	+	2	GCA	A	GCT	A	S
hang	X	16329558	+	0	GAG	E	GAA	E	S
hang	X	16329702	+	0	GTC	V	GTA	V	S
hang	X	16329714	+	0	GTA	V	GTG	V	S
hang	X	16329972	+	0	GAC	D	GAT	D	S

Table S3.5 Coding Region Polymorphisms in Wnt pathway-Associated Genes. (cont)

Gene	SNP-Chrom	SNP-Location	Strand	Frame	Ref-Codon	Ref-AA	H1/H2-Codon	H1/H2-AA	Type
hang	X	16330252	+	0	ACG	T	CCG	P	NS
CG42575	3L	11126284	+	0	AGC	S	AGT	S	S
CG7837	3R	25621714	-	0	GAA	E	GAG	E	S
CG7837	3R	25621888	-	0	CAT	H	CAC	H	S
CG7837	3R	25624104	-	0	TTT	F	TTC	F	S
dco	3R	26882410	-	0	GAC	D	GAT	D	S
dco	3R	26882515	-	0	GCA	A	GCT	A	S
dco	3R	26882761	-	0	AAA	K	AAG	K	S
dco	3R	26882794	-	0	TTC	F	TTT	F	S
dom	2R	17219902	+	0	GGT	G	GGC	G	S
dom	2R	17220640	+	0	CAC	H	CAT	H	S
ft	2L	4214127	-	0	TCA	S	TCC	S	S
ft	2L	4214130	-	0	GGT	G	GGA	G	S
Pp2B-14D	X	16453000	-	0	ATT	I	ATC	I	S
Pp2B-14D	X	16453114	-	0	CAT	H	CAC	H	S
Pp2B-14D	X	16453381	-	0	GTC	V	GTT	V	S
pygo	3R	27403752	-	0	ATC	I	AAC	N	NS
rok	X	16522505	-	0	GAG	E	GAA	E	S
rok	X	16523825	-	0	GGT	G	GGC	G	S
sgg	X	2553538	+	0	AGC	S	AGT	S	S
sgg	X	2554975	+	0	GCT	A	GCC	A	S
sima	3R	25894059	+	1	TCG	S	TTG	L	NS
Smr	X	12584535	-	2	CAA	Q	CAG	Q	S
Smr	X	12584544	-	2	CAA	Q	CAG	Q	S
Smr	X	12584820	-	2	CCA	P	CCG	P	S
Smr	X	12584838	-	2	GGT	G	GGC	G	S
Smr	X	12584850	-	2	TCG	S	TCT	S	S
Smr	X	12586881	-	2	CAA	Q	CAG	Q	S
wts	3R	26618196	-	0	AAG	K	AAA	K	S

Acknowledgements

Chapter 3, in part, is a re-editing of materials published in Zhou, D., Udpa, N., Gersten, M., Visk, D.W., Bashir, A., Xue, J., Frazer, K.A., Posakony, J.W., Subramaniam, S., Bafna, V. & Haddad, G.G. Experimental selection of hypoxia-tolerant *Drosophila melanogaster*. Proc Natl Acad Sci U S A 108, 2349-54 (2011). The dissertation author was a co-second author responsible for analyzing data and contributing to the writing of the paper. Chapter 3 is also, in part, a re-editing of materials currently being prepared for submission for publication: Gersten, M., Zhou, D., Haddad, G.G., & Subramaniam, S. Wnt pathway activation increases hypoxia tolerance in *Drosophila melanogaster*, in preparation. The dissertation author is the primary investigator and author of this paper.

4. TRANSCRIPTIONAL ANALYSIS OF HYPOXIA-ADAPTED FLIES

Introduction

The functional state of a biological system is reflected in the nature and quantities of its current parts list – the transcripts, proteins, and metabolites that exist at the instant the system is studied. High throughput gene expression studies quantitate transcript abundance of thousands of genes simultaneously, providing a description of the transcriptional state of the system with clues to the proteomic state of the system. Divergent phenotypes or responses to environmental stress are reflected in changes to the system's functional state. The concomitant transcriptional changes can be analyzed to gain insight into the pathways and processes associated with these changes in functional state.

A previous study of gene expression in hypoxia-adapted 3rd instar larvae and adult flies, using cDNA arrays, was performed at generation 18 when the flies were adapted to 5% O_2[49,50]. Significant differential expression relative to control flies maintained in parallel was assessed using Significance Analysis of Microarray[167] (SAM) software, with a cutoff of >1.5-fold and FDR < 0.05. Approximately 2750 differentially expressed genes were identified in larvae, but fewer than 150 in adult flies. The two stages shared 23 upregulated genes, mostly related to immunity, and seven downregulated genes mostly related to metabolism. Multiple components of several signaling pathways were differentially expressed in larvae, including EGF, insulin, Notch, Toll/IMD and JNK. In addition, many metabolism genes, and

particularly those related to carbohydrate metabolism, were altered in larvae. Many genes related to energy production were downregulated, including genes in the TCA cycle and respiratory chain complexes, and perhaps surprisingly, many glycolytic enzymes as well. Evidence for coordinated regulation of the downregulated TCA genes was suggested by the presence of binding sites for the transcriptional suppressor, *hairy*, in the regulatory regions of these genes, upregulation of *hairy* itself, and ChIP binding of *hairy* to a target gene promoter under hypoxic, but not normoxic conditions.

Since the adapted fly population had undergone a reproductive bottleneck between the time that the original microarray and the genome sequencing were performed, I revisited gene expression in hypoxia-adapted flies, now tolerant to 4% O_2. The objectives of this second microarray experiment were the following. First, I was interested to determine whether there was physiologic evidence for Wnt pathway involvement in hypoxia tolerance, as suggested by the genomic sequencing analysis which was performed in 4% O_2-adapted flies. Second, I wanted to determine whether gene expression differences between control and adapted flies primarily reflect inherited genomic differences or whether they require hypoxic conditions to become manifest, and this required analysis of a third population, hypoxia-tolerant flies grown in normoxia. Third, I was curious what differences in gene expression had occurred in association with, and likely as a result of genetic changes contributing to, emergence from the reproductive bottleneck. Finally, since eclosion is an energetically demanding activity, requiring the fly to forcefully and rapidly extrude its adult body

from the confining pupal sac, I asked whether there are detectable eclosion-related expression differences between control and hypoxia-adapted flies not observed in other developmental stages. This latter question required inclusion of a third developmental stage in addition to larvae and adults. Since pupa stages vary in gene expression and are difficult to time between populations, I used early post-eclosion flies, collected 0.5-3 hours after emergence from the pupa.

Methods

Affymetrix microarray analysis

Twenty-seven samples from hypoxia-selected flies maintained in 4% O_2 (H) or reverted for five generations to normoxia (HR) and control flies (C) were used for microarray experiments. In each case, three replicate pools of flies were collected from three developmental stages: larva (L: 25/pool); 0.5-3 hours post-eclosion (Ec: 25 females and 25 males/pool); and 7-9 day adults (A: 25 females and 25 males/pool). Total RNA was extracted using TRIzol (Invitrogen, Carlsbad, CA) followed by a clean-up with RNeasy kit (Qiagen, Valencia, CA). Affymetrix *Drosophila* Genome 2.0 arrays were used and probe labeling, array hybridization, and image scanning were performed following the standard protocol according to manufacturer's instruction (Affymetrix, Santa Clara, CA). Gene expression values (probeset summary values) were generated using the Plier algorithm, an improved method developed by Affymetrix[168,169], which was implemented using the Affymetrix Expression Console

software (Affymetrix, Santa Clara, CA). Differential expression between experimental groups was determined using VAMPIRE[170] software. VAMPIRE models the global error structure of microarray data to obtain a better measure of variability for each gene's expression than provided by the variance of the measured data. It does this by optimizing two parameters that describe the expression-dependent and –independent variances. It then integrates this variance model with a Bayesian framework that converts the measured mean intensity into a probability density that more accurately describes the true expression level of the gene. The computed expression density is then used in statistical tests to determine whether or not a gene is differentially expressed between two conditions at a given threshold for significance. VAMPIRE has been particularly effective at identifying significant differential expression in experiments with few replicate measures. Samples with a mean Control expression <64 were discarded due to poor correlation among biological replicate samples, as shown in Figure 4.1.

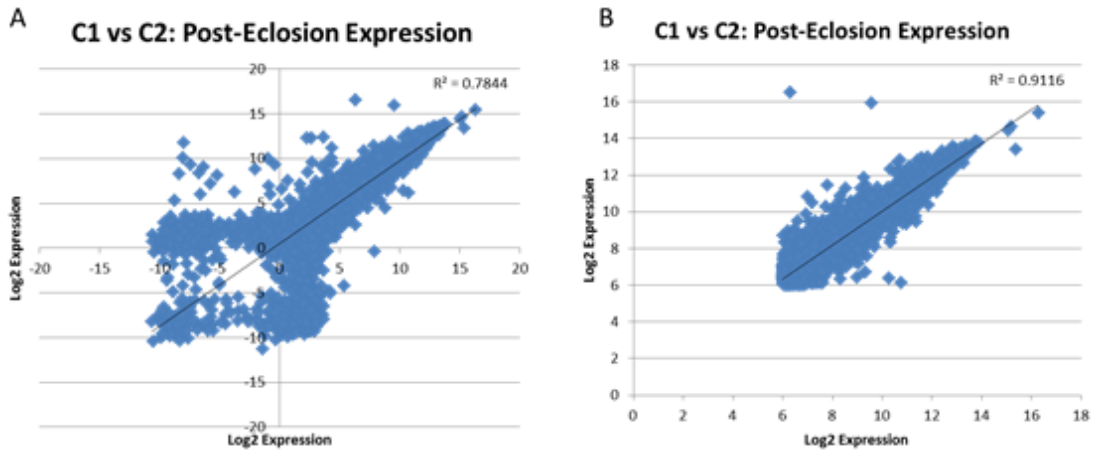


Figure 4.1 Correlation between two log2-transformed control biological replicate datasets. (A) All data and (B) data with expression values >= 64. Elimination of expression values < 64 greatly improves the correlation between sets, increasing confidence in the results; therefore, data with control mean expression < 64 was excluded fromfurther analysis.

RT-PCR

First strand cDNA was synthesized using SuperScript III reverse transcriptase and Oligo-(dT) primers (Invitrogen, Carlsbad, CA). Real-time PCR amplification was performed using ABI Prism 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA). For each reaction, 10 µl of 2× SYBR green PCR master mix (Applied Biosystems, Foster City, CA) and 0.15 µM of both forward and reverse primers along with 100 ng of each appropriate cDNA sample were mixed (total reaction volume: 20 µl). The relative expression level was calculated using $2^{-\Delta\Delta Ct}$ method, as described previously[171,172]. *Drosophila melanogaster* actin was used as internal control. The final results were presented as fold change of H or HR over C. All experiments were done in triplicate. The primers used are shown in Table 4.1.

Table 4.1 Primers Used in RT-PCR Validation Experiments

Oligo	Sequence (5'->3')
pim Forward	TGGAAAACAGATGCCTTCCGCCC
pim Reverse	TCAGGCTCTGCAGGCTCCATGT
MDR50 Forward	GCAGAAGGGAGTTTCGGCAGCC
MDR50 Reverse	ACACTCCCGGTCGGTCCACA
CKIIB Forward	TGCTCTTCATGGTGCATCCCGA
CKIIB Reverse	CTGCTGCCTGCAGCTGAATTTGAT
pk Forward	TCCGGACACACTGCCCCGAT
pk Reverse	TGTCCCGATCCCGCTCCCTC
CG13422 Forward	GCCGCCGATGTGGTCAGTTCA
CG13422 Reverse	GACTCCATGGTAACGCGCCGT
Act88F Forward	CCCAACAACCTCGGCTCGGAC
Act88F Reverse	TGAGCACCGACAACCGGAGGT
Act5C Forward	CTAACCTCGCCCTCTCCTCT
Act5C Reverse	GCAGCCAAGTGTGAGTGTGT

STEM analysis

To identify processes that might have a genetic contribution but that were not identified by VAMPIRE analysis to be shared by H and HR flies I explored the microarray data using Short Time-series Expression Miner software[173] (STEM), treating the three developmental stages as a short time series. The algorithm clusters time-dependent expression profiles and allows for a comparison of two experimental groups with respect to profiles with shared genes. I compared H *vs* C, HR *vs* C and H

vs HR, in each case using the excluded dataset as the denominator to generate the required ratio data input. I then manually evaluated the comparison figures with respect to Gene Ontology of genes whose direction of change differed at the post-eclosion data point.

Results

Table 4.2 summarizes the distribution of differentially expressed genes relative to control across developmental stages. There are relatively few differences in adult and larva samples, except for downregulated genes in adult AF grown in normoxia, and these largely mapped to annotations related to peptidase activity, oxidation reduction and host defense. Among the 48 downregulated oxidation-reduction genes in AF maintained at 21% O2 were seven genes related to the electron transport chain, and an additional two genes that mapped to “generation of precursor metabolites and energy”. This downregulation in hypoxia-adapted flies grown in normoxia may reflect a reduced need for oxidative metabolism, suggesting a possible response to perceived hyperoxia in 21% O_2 for these 4% O_2-adapted flies. Supporting this hypothesis, this population also had five upregulated genes annotated as “response to oxidative stress”.

Table 4.2 Summary of Gene Expression Differences By Developmental Stage
Plier[168] expression values, excluding probesets with mean Control expression < 64; VAMPIRE[170] significance analysis, cutoff FDR 1%.

Comparison	Adult	Post-Eclosion	Larva
Number Upregulated			
H vs C	339	2149	198
HR vs C	207	56	142
H-HR Common	117	45	83
Number Downregulated			
H vs C	342	1992	233
HR vs C	616	75	151
H-HR Common	200	52	102

Table 4.3 lists KEGG pathway and GO Biological Process (BP) annotations for up- and downregulated genes shared by H and HR. Most of the annotations relate to metabolism of drugs and xenobiotics and defense responses, with both up- and downregulated genes in each category. Between 6-18 downregulated genes annotated for "oxidation reduction" were shared by H and HR across the three developmental stages. There was also a shared reduction in oogenesis-related genes in the post-eclosion stage.

Most striking was the large number of both up- and downregulated genes observed for post-eclosion hypoxia-adapted flies maintained in hypoxia (Table 4.2). Similar numbers of differentially expressed genes were obtained using other methods to summarize expression values (RMA[56]) and to detect significance (CyberT[57]). RT-PCR was performed on several genes, confirming their differential expression

Table 4.3 KEGG and GO BP Annotations for Differentially Expressed Genes Shared by H and HR Flies.
Annotations shared across two or more developmental stages are italicized and underlined.

	Adult		Post-Eclosion		Larva	
UPREGULATED						
KEGG	*Limonene/ pinene degradation*	5	*Limonene/ pinene degradation*	3	*Metabolism xenobiotics (P450)*	5
	Lysosome	3			*Drug metabolism*	5
	Glutathione metabolism	2			Tyrosine metabolism	3
	Metabolism xenobiotics (P450)	2			Retinol metabolism	3
	Drug metabolism	2			Fatty acid metabolism	2
GO-BP	*Defense response*	10	Sperm individualization	2	Oxidation reduction	11
	…				Proteolysis	10
	Antimicrobial humoral resp	5			*Defense response*	4
DOWNREGULATED						
KEGG	*Limonene/ pinene degradation*	4	*Limonene/pinene degradation*	2	*Limonene/ pinene degradation*	4
	Metabolism xenobiotics (P450)	3			Glycerophospholipid metab	2
	Drug metabolism	3			*Glutathione metabolism*	2
	Lysosome	3			*Metabolism xenobiotics (P450)*	2
	…				*Drug metabolism*	2
	Starch & sucrose metabolism	2				
	Glutathione metabolism	2				
GO-BP	Peptidoglycan metab process	4	Chorion eggshell form'n	6	*Oxidation reduction*	11
	…		…		*Defense response*	6
	Oxidation reduction	18	Sexual reproduction / Oogenesis	7	Biogenic amine metab process	3
	Defense response	8	…		Locomotory behavior	5
			Oxidation reduction	6		
			Proteolysis	6		
			Defense response to fungus	2		

(Figure 4.2). Table 4.4 summarizes the most highly enriched GO-BP annotations for differentially expressed genes in post-eclosion AF grown in 4% O_2. Upregulated genes relate primarily to development (particularly respiratory system development) and metabolism, including 14 genes related to glycolysis. Downregulated genes are highly enriched for DNA replication, cell cycle and DNA repair, which may

contribute to the reduced size observed in AF[49]. These, as well as the 117 downregulated oogenesis genes (not shown), may also reflect a reduced energetic investment in gametocyte production, as observed in a model of mitochondrial disease[174].

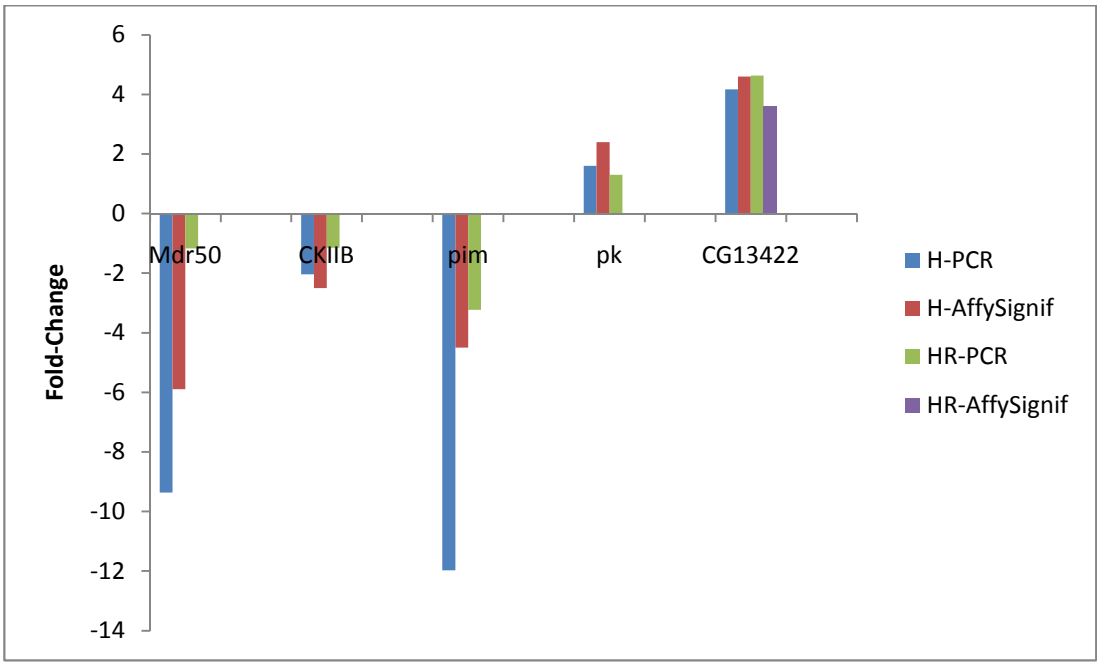


Figure 4.2 PCR confirmation of post-eclosion differential expression in hypoxia-adapted flies. Actin-normalized Ct values for all H *vs* C comparisons were significant at $p<0.05$. None of the PCR results were significant of HR *vs* C. *Actin5C* was used in all experiments except for *pim*, where *Act88F* was used.

Although reduction in cell cycle genes was not detected in AF grown in normoxia by the VAMPIRE analysis, the shared reduction in oogenesis-related genes in H and HR post-eclosion flies suggested that there might be a genetic contribution to this observation. I explored this possibility by evaluating the post-eclosion gene expression data using Short Time-series Expression Miner[173] software (STEM),

Table 4.4 Top-Scoring GO[17]-BP Annotations for Differentially Expressed Genes in Post-Eclosion Hypoxia-Tolerant Flies Grown in 4% O_2

Downregulated Genes				
Category	Term	Count	PValue	Fold Enrichment
GOBP_FAT	GO:0006259~DNA metabolic process	117	1.69E-42	3.719
GOBP_FAT	GO:0051276~chromosome organization	120	9.32E-32	2.989
GOBP_FAT	GO:0007049~cell cycle	195	9.50E-31	2.223
GOBP_FAT	GO:0006260~DNA replication	63	5.30E-30	4.609
GOBP_FAT	GO:0000278~mitotic cell cycle	135	2.45E-29	2.649
GOBP_FAT	GO:0022403~cell cycle phase	161	1.58E-26	2.280
GOBP_FAT	GO:0022402~cell cycle process	171	4.13E-26	2.196
GOBP_FAT	GO:0000279~M phase	153	1.80E-24	2.248
GOBP_FAT	GO:0051726~regulation of cell cycle	71	5.12E-19	3.004
GOBP_FAT	GO:0000087~M phase of mitotic cell cycle	66	7.53E-19	3.132
GOBP_FAT	GO:0006325~chromatin organization	71	2.44E-18	2.934
GOBP_FAT	GO:0007052~mitotic spindle organization	77	2.59E-18	2.788
GOBP_FAT	GO:0000280~nuclear division	65	3.68E-18	3.085
GOBP_FAT	GO:0007067~mitosis	64	7.70E-18	3.079
GOBP_FAT	GO:0006974~response to DNA damage stimulus	55	1.91E-17	3.359
Upregulated genes				
Category	Term	Count	PValue	Fold Enrichment
GOBP_FAT	GO:0007155~cell adhesion	51	1.26E-09	2.451
GOBP_FAT	GO:0022610~biological adhesion	51	2.17E-08	2.269
GOBP_FAT	GO:0006928~cell motion	63	1.34E-05	1.728
GOBP_FAT	GO:0060541~respiratory system development	38	1.63E-05	2.077
GOBP_FAT	GO:0007424~open tracheal system development	38	1.63E-05	2.077
GOBP_FAT	GO:0035150~regulation of tube size	12	2.30E-05	4.364
GOBP_FAT	GO:0035151~regulation of tube size, open tracheal system	11	3.37E-05	4.600
GOBP_FAT	GO:0007156~homophilic cell adhesion	13	3.84E-05	3.883
GOBP_FAT	GO:0035152~regulation of tube architecture, open tracheal system	15	6.62E-05	3.301

Table 4.4 Top-Scoring GO17-BP Annotations for Differentially Expressed Genes in Post-Eclosion Hypoxia-Tolerant Flies Grown in 4% O2 (cont)

Category	Term	Count	PValue	Fold Enrichment
GOBP_FAT	GO:0016052~carbohydrate catabolic process	24	7.14E-05	2.418
GOBP_FAT	GO:0008277~regulation of G-protein coupled receptor protein signaling pathway	11	9.42E-05	4.182
GOBP_FAT	GO:0006096~glycolysis	14	1.09E-04	3.345
GOBP_FAT	GO:0006006~glucose metabolic process	19	1.80E-04	2.605
GOBP_FAT	GO:0005976~polysaccharide metabolic process	35	2.13E-04	1.913
GOBP_FAT	GO:0006022~aminoglycan metabolic process	33	2.48E-04	1.944

treating the three developmental stages as a short time series. The algorithm clusters time-dependent expression profiles and allows for a comparison of two experimental groups with respect to profiles with shared genes. I compared H vs C, HR vs C and H vs HR, in each case using the excluded dataset as the denominator to generate the required ratio data input. I then evaluated the comparison figures with respect to Gene Ontology of genes whose direction of change differed at the post-eclosion data point. The results of this analysis, illustrated in Figure 4.3 and summarized in Table 4.5, suggest that reduced ribosome biogenesis and increased development-related gene expression seen in hypoxia-tolerant flies depends on a hypoxic environment; normal ribosome biogenesis in HR flies is consistent with the observation that these flies are of normal size. In contrast, decreased DNA replication appears to be at least partially genetic/epigenetic in origin. Among the genes identified using STEM, 41 probesets had significantly reduced expression in H flies, as assessed by VAMPIRE, 40 with fold change 0.14-0.60 and p-value < 10-4, and there were 28 probesets with reduced

expression in HR flies, 11 with fold change 0.65-0.777 and p-value 0.0014-0.033. Inherited changes in cell cycle genes, if true, might account for the reduction in oogenesis gene expression seen in both AF populations (Table 4.3) and reflect the polymorphisms detected in oogenesis and cell cycle genes (Supplemental Table S3.1).

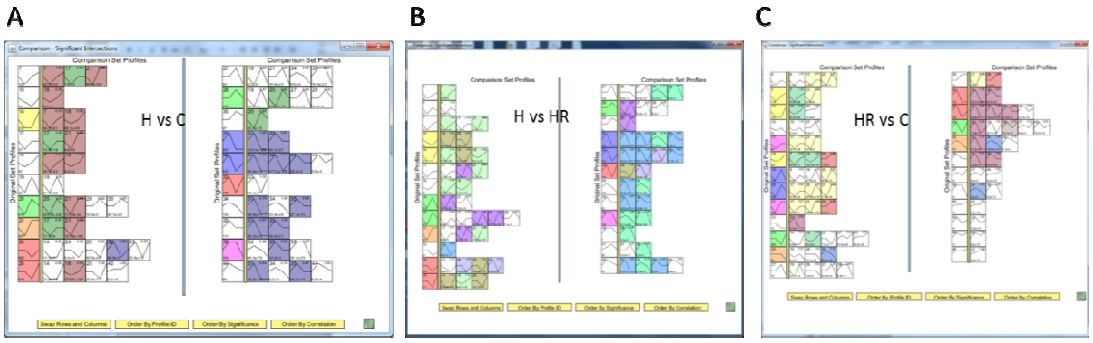


Figure 4.3 STEM[173] analysis of gene expression in post-eclosion flies.
Each panel is a comparison of STEM-generated microarray time-series profiles identified for two conditions, original and comparator, here using a third condition as the denominator to generate ratio data. In each case, the profiles identified for the first condition (original) appear in the left-hand column. Profiles for the second condition (comparator) which contain genes appearing in an original profile are positioned in the same row, to the right of the original profile. (A) H (original) *vs* C (comparator); (B) H (original) *vs* HR (comparator); (C) HR (original) *vs* C (comparator).

Table 4.5 Summary of STEM[173] Analysis of Gene Expression in Post-Eclosion Flies.
Summary of results from STEM analysis illustrated in Figure 4.3. Processes dependent on a hypoxic environment are italicized; processes that may have a genetic/epigenetic component are bolded.

	H *vs* HR (Control=C)	HR *vs* C (Control=H)	H *vs* C (Control=HR)
O_2 Exposure Difference	*YES*	No	*YES*
Genetic Difference	No	**YES**	**YES**
Ribosome biogenesis	*Decreased in H*		*Decreased in H*
Development (various)	*Increased in H*		*Increased in H*
DNA replication		**Decreased in HR**	**Decreased in H**
Cell cycle		**Decreased in HR**	**Decreased in H**
EtOH, hexose metab; oxidoreductase		Increased in HR	

Similar to the GO analysis, pathway analysis of the set of differentially expressed genes in post-eclosion H flies also identified upregulated genes relating to energy-producing pathways, including 15 pyruvate, 14 glycolysis/gluconeogenesis and 12 citric acid cycle (TCA) pathway genes. There were also 29 downregulated DNA replication pathway genes, as well as several annotations for downregulated DNA repair pathways (Supplemental Table S4.1). Combining up- and downregulated genes, pathway analysis identified several developmental/signaling pathways showing multiple gene involvement, among which Wnt signaling figured prominently (Figure 4.4, Supplemental Table S4.2). Several of these pathways other than Wnt were identified as affected in larvae in the earlier microarray study[50]. KEGG and Panther identified a total of 30 upregulated and 22 downregulated Wnt pathway genes; GO-BP identified an additional 10 upregulated and 6 downregulated genes. Table 4.6 summarizes the expression changes observed and the Wnt pathway association for these 68 genes. Combining data from genome sequencing and gene expression, Figure 4.5 highlights the genes in the Wnt signaling pathway that were differentially expressed in post-eclosion H flies and/or which contain polymorphisms. It is clear that both genetic and expression changes are detected in all three major Wnt pathways.

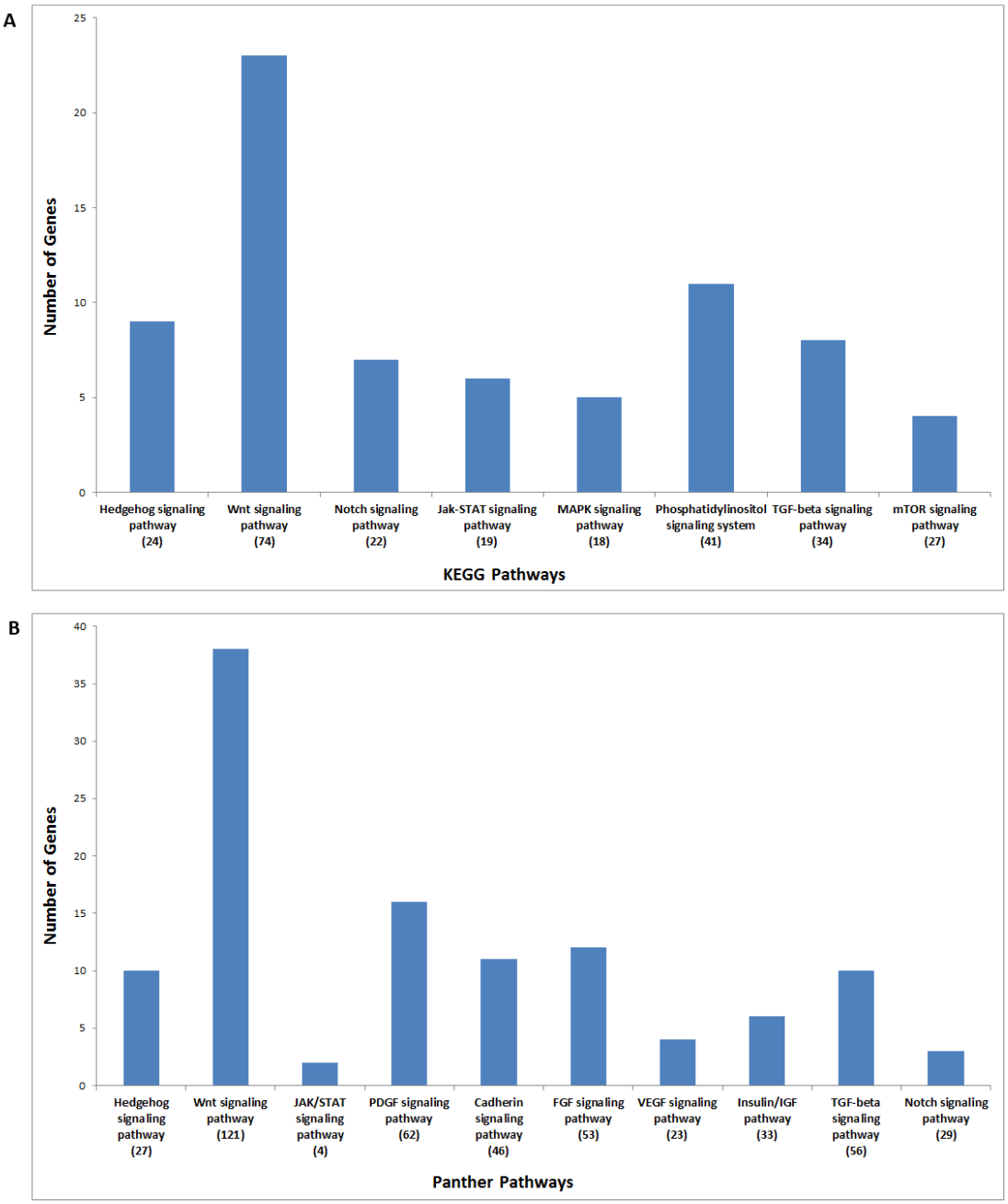


Figure 4.4 Pathway analysis of differentially expressed genes in post-eclosion hypoxia-tolerant flies grown in 4% O_2.
Analysis done using DAVID[122,123] software. (A) KEGG[14,15] and (B) Panther[16] signaling pathways with ≥2 differentially expressed genes. The number of differentially expressed pathway genes is plotted and the total number of pathway genes is indicated below each pathway name.

Table 4.6 Wnt Pathway-Associated Genes Differentially Expressed in Post-Eclosion Hypoxia Tolerant Flies Grown at 4% O_2

Affymetrix Probeset IDs were converted to FlyBase gene symbols using DAVID. Asterisks indicate that the gene identified by DAVID could not be confirmed at the Panther (*) or AMIGO (**) website.

Symbol	Enrich-ment Source	Wnt Pathway	Function	Affy Probeset	Mean Control	Fold-Change	p-value
abd-A	Panther	Canonical		1637813_at	231.56	1.5667	5.42E-06
Antp	Panther	Canonical		1640208_s_at	145.34	1.5141	3.30E-05
Antp	Panther	Canonical		1624759_s_at	119.1	1.5018	5.55E-05
arr	KEGG, GO-BP	Canonical	Co-Receptor	1634211_at	211.76	1.3682	0.001353
CadN	Panther		Cadherin	1633000_a_at	82.89	1.5962	6.88E-06
CaMKII	KEGG	Calcium	Kinase	1640534_s_at	375.61	1.4346	2.25E-04
CanB	Panther	Calcium	Dephosphorylates NFAT	1641460_at	524.45	1.5138	2.32E-05
CG11533 /Asator	Panther	Canonical	Kinase	1627955_a_at	460.5	1.605	1.42E-06
CG11848 /Bili	GO-BP	Canonical	Negative regulator	1640291_at	77.1	1.5769	1.42E-05
CG12227 /skpF	KEGG	Canonical	β-catenin proteolysis	1639319_at	324.66	1.4468	1.64E-04
CG15800	KEGG	Canonical	β-catenin proteolysis	1624073_at	118.46	1.6831	2.42E-07
CG31009 /Cad99C	Panther		Cadherin	1630399_at	192.22	1.5929	2.62E-06
CG6445 /Cad74A	Panther		Cadherin	1635742_s_at	149.93	1.8881	1.67E-10
CG7913 /PP2A-B'	KEGG, Panther	Canonical	Protein phosphatase, regulatory subunit	1628476_at	164.5	1.4113	4.99E-04
CG8881 /skpB	KEGG	Canonical	β-catenin proteolysis	1636591_at	321.97	1.3734	0.001116
crol	GO-BP	Canonical	Negative regulator	1626018_a_at	337.78	1.4093	4.45E-04
DAAM	KEGG	PCP	Actin binding	1629015_a_at	538.64	1.9488	9.59E-12
dally	KEGG, GO-BP	Canonical	Co-Receptor	1628489_at	929.18	1.482	5.74E-05
drl	GO-BP		Wnt binding protein	1624297_at	140.04	1.6969	1.23E-07

Table 4.6 Wnt Pathway-Associated Genes Differentially Expressed in Post-Eclosion Hypoxia Tolerant Flies Grown at 4% O2 (cont)

Symbol	Enrich-ment Source	Wnt Pathway	Function	Affy Probeset	Mean Control	Fold-Change	p-value
ds	Panther	PCP	Cadherin-related	1640627_at	100.8	1.7088	1.44E-07
fj	GO-BP	PCP	Wnt binding protein, protein kinase	1636091_at	98.27	1.4948	8.46E-05
ft	Panther, GO-BP		Cadherin-related	1624125_at	195.12	1.9448	1.65E-11
Galpha49B	Panther		GTPase	1639773_s_at	1586.75	1.7539	9.47E-09
gish	Panther, GO-BP	Canonical	Protein kinase	1626805_s_at	922.42	1.6424	4.10E-07
gskt	Panther	Canonical	Putative GSK3B homolog, β-catenin binding	1632217_at	283.67	1.5701	4.59E-06
Hel89B *	Panther		Helicase	1627970_at	295.96	1.5865	2.74E-06
Itp-r83A	Panther	Calcium	Inositol triphosphate receptor	1639461_a_at	101.1	1.4801	1.21E-04
nimC1	GO-BP		Wnt protein binding	1639502_at	304.1	1.5185	2.15E-05
nkd	KEGG, GO-BP	Canonical, PCP	Binds dsh	1630361_at	68	1.56	3.16E-05
pan	GO-BP	Canonical	Transcription factor	1629467_at	67.68	2.1253	7.14E-13
pk	KEGG	PCP	Interacts with Vang	1636612_a_at	89	2.3881	1.54E-17
Pka-C2	KEGG	Canonical	cAMP-dependent protein kinase	1632787_a_at	142.47	1.4964	5.52E-05
Pkc53E	KEGG, Panther	Calcium	Protein kinase C	1632478_a_at	192	1.5593	7.34E-06
Pkcdelta	Panther	Calcium	Protein kinase C	1628677_at	383.79	2.0954	4.75E-14
Roc1b	KEGG	Canonical	β-catenin proteolysis	1639817_at	214.85	1.533	1.54E-05
sfl	GO-BP	Canonical	Sulfotransferase	1631534_at	170.23	1.4608	1.33E-04

Table 4.6 Wnt Pathway-Associated Genes Differentially Expressed in Post-Eclosion Hypoxia Tolerant Flies Grown at 4% O2 (cont)

Symbol	Enrich-ment Source	Wnt Pathway	Function	Affy Probeset	Mean Control	Fold-Change	p-value
sgg	KEGG, Panther, GO-BP	Canonical	GSK3B homolog, β-catenin binding	1623084_at	193.58	1.7096	6.03E-08
sgg	KEGG, Panther, GO-BP	Canonical	GSK3B homolog, β-catenin binding	1630774_s_at	1131.65	1.5943	1.92E-06
sinah	KEGG, Panther	Canonical	β-catenin proteolysis	1638838_at	708.75	1.3782	9.50E-04
SoxN	GO-BP	Canonical	Chromatin DNA binding	1631408_at	162.67	2.0163	1.54E-12
spen	GO-BP	Canonical	Nucleic acid binding	1641518_a_at	833.94	1.4392	1.91E-04
tsh **	GO-BP		Transcriptio n factor	1626150_at	121.64	1.5331	2.25E-05
babo	Panther	Canonical	Kinase	1623424_a_at	210.1	0.5842	3.53E-05
Bap60	Panther	Canonical	Brahma associated protein; Activator	1634157_at	306.34	0.5179	1.36E-06
CG1240	Panther	Canonical	Chromatin binding	1636036_at	639.65	0.5681	1.34E-05
CG3226	KEGG	Canonical	β-catenin proteolysis	1623702_at	669.89	0.6697	7.87E-04
CG5899/ Etl1 *	Panther		Helicase	1624673_at	182.55	0.5165	1.74E-06
CG7055/ dalao	Panther	Canonical	Chromatin binding	1631677_at	516.13	0.5839	2.76E-05
CkIIbeta	KEGG, Panther, GO-BP	Canonical	Protein kinase regulator	1637769_s_at	1221.23	0.6537	4.40E-04
CkIIbeta	KEGG, Panther, GO-BP	Canonical	Protein kinase regulator	1628344_at	118.53	0.404	9.15E-09
CycD	KEGG, Panther	Canonical	Cyclin-dependent protein kinase regulator; Target gene	1627295_s_at	212.13	0.5176	1.62E-06

Table 4.6 Wnt Pathway-Associated Genes Differentially Expressed in Post-Eclosion Hypoxia Tolerant Flies Grown at 4% O2 (cont)

Symbol	Enrich-ment Source	Wnt Pathway	Function	Affy Probeset	Mean Control	Fold-Change	p-value
dom	Panther	Canonical	Helicase	1633331_at	542.9	0.5932	4.12E-05
dom	Panther	Canonical	Helicase	1636034_at	276.23	0.3471	5.11E-11
dx	GO-BP		Protein binding	1626617_at	334.52	0.4934	3.68E-07
ebi	KEGG, Panther	Canonical	β-catenin proteolysis	1638600_at	268.11	0.5704	1.73E-05
fz4	Panther, GO-BP		Receptor for Wnt	1631777_a_at	212.88	0.5555	9.87E-06
Iswi	Panther	Canonical	Helicase	1641309_s_at	334.47	0.6721	8.98E-04
krz	Panther		β-Arrestin, binds to dsh	1623793_at	133.71	0.5872	5.81E-05
lin **	GO-BP		Catalytic activity	1623324_at	253.51	0.6672	7.99E-04
mor	Panther	Canonical	Chromatin binding	1640418_at	504.19	0.572	1.64E-05
nej	KEGG, Panther, GO-BP	Canonical	CREB binding protein	1622925_at	320.87	0.53	2.46E-06
pont	KEGG, GO-BP	Canonical	Helicase, transcription cofactor	1635279_at	442.56	0.5846	2.91E-05
por	KEGG, GO-BP	Canonical	Acyl transferase	1636527_at	80.47	0.6406	8.70E-04
Pp2B-14D	KEGG, Panther	Calcium	Dephosphorylates NFAT	1626781_at	575.74	0.6251	1.52E-04
pygo	Panther, GO-BP	Canonical	β-catenin nuclear targeting	1640730_at	433.99	0.6681	7.61E-04
rept **	GO-BP		Helicase	1630792_at	557.71	0.6441	3.16E-04
Rfabg	GO-BP		Retinoid and Fatty Acid Binding	1637843_at	5214.94	0.314	2.99E-12
Rpd3	Panther	Canonical	Histone deacetylase	1633700_at	363.93	0.5692	1.51E-05
skd	GO-BP	Canonical	Transcription cofactor; Positive regulator	1630269_s_at	118.27	0.44	6.92E-08

Table 4.6 Wnt Pathway-Associated Genes Differentially Expressed in Post-Eclosion Hypoxia Tolerant Flies Grown at 4% O2 (cont)

Symbol	Enrich-ment Source	Wnt Pathway	Function	Affy Probeset	Mean Control	Fold-Change	p-value
smo	Panther, GO-BP		G protein-coupled receptor activity	1634442_at	330.12	0.5642	1.23E-05
Snr1	Panther	Canonical	DNA binding	1625112_at	362.13	0.5896	3.73E-05
tum	GO-BP	Canonical	Small GTPase activator; Negative regulator	1634690_at	201.05	0.3178	9.59E-12

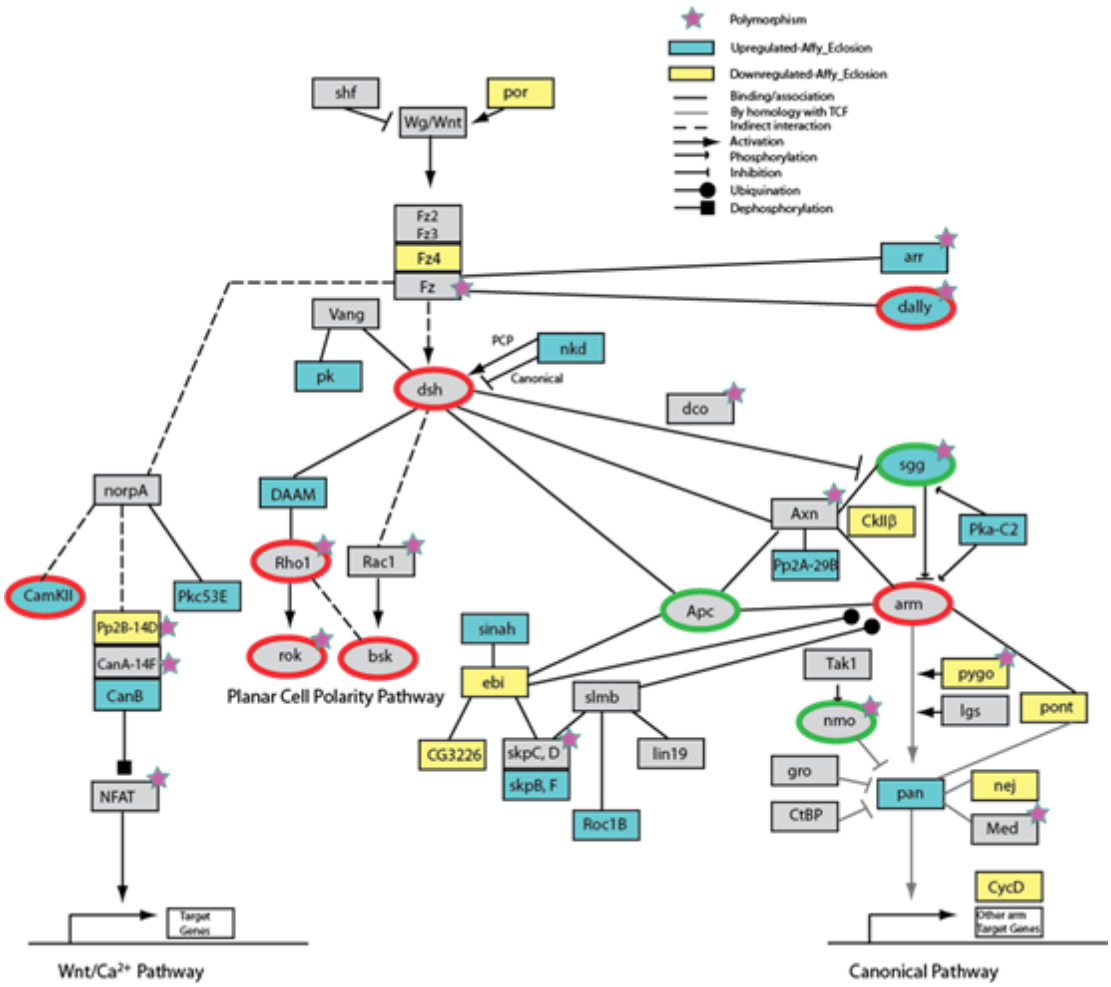


Figure 4.5 Consolidated map of Wnt pathway polymorphisms and genes differentially expressed in AF flies grown in 4% O_2.
The Wnt pathway was adapted from KEGG[14,15] with additional interactors added. Genes with one or more fixed polymorphism are indicated with a magenta star. Genes differentially expressed in post-eclosion AF maintained in 4% O2 are colored cyan (upregulated) or yellow (downregulated). Wnt pathway genes tested in genetic crosses are elliptical in shape (see Chapter 5); pathway activators are indicated by a red border and pathway inhibitors by a green border.

I also looked for further evidence that Wnt signaling contributed to the hypoxia-tolerant phenotype by determining whether any Wnt pathway target genes were differentially expressed. As shown in Table 4.7, several target genes identified by Nusse (Wnt Homepage, http://www.stanford.edu/~rnusse/pathways/targets.html) were regulated in post-eclosion H flies, sometimes, but not always in the direction expected for Wnt pathway activation.

Table 4.7 Wnt Pathway Target Genes Differentially Expressed in Post-eclosion Hypoxia-tolerant (H) Flies.
Target gene identification and expected effect of Wnt activation are from Nusse (http://www.stanford.edu/~rnusse/pathways/targets.html). Target genes not specifically identified in *Drosophila* were mapped to *Drosophila* and the original identified gene is indicated in parentheses. Significant expression differences determined by VAMPIRE[170], FDR 1% cutoff.

	Mean Control Expression	Fold-Change	Expected Change
Downregulated			
ovo/shavenbaby	707.06	(-3.3)	down
Deterin (survivin)	181.39	(-2.8)	up
pim (PTTG)	378.61	(-4.5)	?
Mdr50 (MDR1)	128.76	(-5.9)	up
Upregulated			
bnl (FGF)	81.77	2	up
pan (Tcf)	67.68	2.1	up
wengen (Tnfrsf19)	732.44	1.7	up
sr	168.25	2.2	down
mirror (Irx3)	83.65	1.7	?
nkd	68	1.6	up
dally	929.18	1.5	down

Discussion

Gene expression analysis of flies adapted to 4% O_2 was performed to address four questions: 1) Is a hypoxic environment necessary for differential gene expression by the hypoxia-selected flies? 2) Has the scope of differential gene expression changed in these flies coincident with their emergence from a 4% O_2-imposed reproductive bottleneck? 3) Is the sensitivity to hypoxia normally seen in adult eclosion reflected in a transcriptional response by hypoxia-adapted flies around the time of eclosion? 4) Is there evidence in post-bottleneck adapted flies of differential expression of Wnt pathway genes that would provide physiological support for the identification of Wnt pathway polymorphisms in these flies? Briefly stated, the results of the gene expression analysis indicated that the large majority of differential gene expression occurred in hypoxia-tolerant flies early after eclosion, but only in the context of a hypoxic environment; they provided evidence of transcriptional changes in several Wnt pathway genes in these post-eclosion flies, and they differed in several respects from the earlier microarray study[50].

Compared with the earlier study, the most striking difference is the relatively small number of gene expression differences seen here in 3rd instar larvae. The large number of differentially expressed genes in the post-eclosion flies led me, therefore, to consider that perhaps larval and post-eclosion samples had been interchanged. However, the detection of very high expression of larva-specific proteins FBp1 and Lsp1α only in the larva- labeled samples indicated that the samples were intact. Since post-eclosion flies were not studied previously, no direct comparisons can be made.

However, comparing post-eclosion differential expression here with 3rd instar differential expression in the earlier study revealed involvement of some common processes, including upregulation of developmental processes and host defense, though the latter were less pronounced in the current study. Some signaling pathways were also affected in both studies, including Notch and JAK/STAT, and upon reanalysis of the earlier study, Wnt as well. Most striking, however, were some of the differences between the two studies. In particular, the downregulation of DNA replication/cell cycle genes seen here was absent in the larva study, and the large decrease in gene expression related to glycolysis, TCA cycle and oxidative phosphorylation seen in the larva study was absent in the post-eclosion flies; in fact, there was upregulation of many glycolysis and TCA cycle genes in these flies.

There are several possible contributors to the differences between the two gene expression studies enumerated above, including differences in developmental stage studied, differences in array type (cDNA *vs* Affymetrix oligonucleotide), and analysis method (SAM[167] *vs* VAMPIRE[170]). Most intriguing is the possibility that genetic/epigenetic changes arising during adaptation to 4% O_2 replaced compensatory transcriptional responses previously necessary for survival under hypoxia, but perhaps otherwise maladaptive[175].

The latter conjecture may find support in studies comparing acute physiological responses to high altitude with the physiology of evolutionarily-adapted high-altitude dwellers. Table 4.8 summarizes acute lowlander responses to high-altitude hypoxia as well as physiological adaptations of three highlander populations

in Ethiopia, Tibet and the Andes, who have evolved different strategies for successful survival at 3500-4000 m. Four observations stand out. First, several lowlander responses can be viewed as maladaptive. These include pulmonary vasoconstriction,

Table 4.8 Strategies For Adaptation to High-Altitude Hypoxia.
Data adapted from Beall[130,176]. Responses are relative to lowlanders living at sea-level. For lowlanders, BMR, HVR, and Maximal O_2 uptake responses to hypoxia last days-weeks; changes in Hb and red cells take weeks to plateau.

		Altitude 3500-4000 m		
	Lowlander Response	Ethiopia	Tibet	Andes
O_2 saturation		Normal	Very Decreased	Decreased
Hemoglobin (Hb) concentration	Increased	Normal	Normal	Increased
Calculated Arterial O_2 Content		Normal	Decreased	Increased
Red Cell Mass		Normal	Normal	Increased
Resting Ventilation	Increased		Increased	Normal
Hypoxic Ventilatory Response (HVR)	Increased		Normal	Decreased
Muscle Capillary Density			Increased	Normal
Pulmonary Hypertension	Pulmonary vasoconstriction		Absent or slight	Present
Blood vessel nitric oxide synthesis	Decreased		Increased	?Decreased
Basal Metabolic Rate (BMR)	Increased		Normal	Normal
Maximal O_2 uptake	Decreased		Normal	Normal

designed to shunt blood away from underperfused lung, but here potentially leading to pulmonary hypertension and impaired O_2 diffusion; and increased red cells which increase blood viscosity and may reduce cardiac output[175]. Second, all three adapted

populations have evolved away from one or more aspects of the typical lowlander response to high-altitude hypoxia, perhaps to reduce any associated untoward effects. Third, all three groups have adopted different strategies for successful survival at high altitude, which likely reflects, at least in part, the different complement of genetic diversity within each respective founder population. Finally, the most "ancient" group, those living in the Ethiopian highlands, exhibit physiological characteristics that belie their chronic residence at high altitude, whereas the "youngest" group, those residing in the Andes, maintain several maladaptive responses, likely contributing to their higher incidence of chronic mountain sickness[177].

I also compared the gene expression changes in hypoxia-tolerant post-eclosion flies with results reported for 1 day old adult male and female *tko*^{25t} *Drosophila* with impaired mitochondrial oxidative phosphorylation[174]. There were three interesting commonalities between the two analyses: 1) differential expression of genes involved in detoxification and defense, largely upregulated in *tko*^{25t} flies; 2) evidence of reduced oogenesis, confirmed for *tko*^{25t} flies by measuring egg counts; and 3) upregulation of larva-specific genes including Lsp1 and Fbp1, which the authors suggest may reflect developmental delay. They also note that these genes have been proposed to play a role in multiple processes, including nutrient transport and oxygen diffusion. These genes are annotated for oxygen transport (Gene Ontology[17]) and contain hemocyanin protein domains (Interpro)[178], raising the question whether in hypoxia-adapted flies they may have been co-opted to facilitate oxygen transport as a compensatory response to the reduced oxygen level. The data for these two genes

from the present study are presented in Figure 4.6. Upregulation of these genes in both H and HR flies suggests that there may be a genetic/epigenetic component to the upregulation.

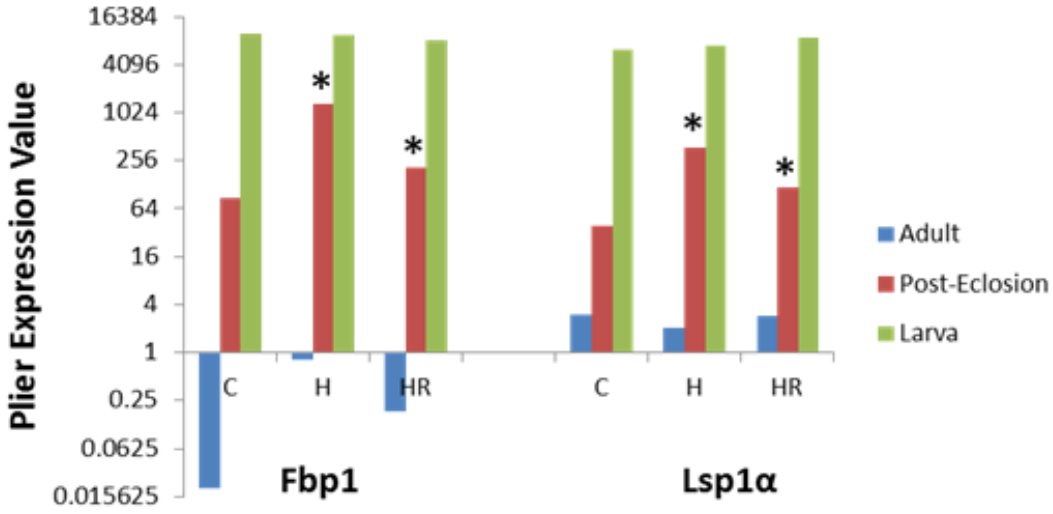


Figure 4.6 Expression of Fbp1 and Lsp1α in control and hypoxia-adapted flies.
Affymetrix expression values for larva-specific genes Fbp1 and Lsp1α. C - Control, H – adapted flies grown at 4% O_2, HR – adapted flies grown at 21% O_2. Asterisk (*) denotes significant difference from respective control value (VAMPIRE[170], FDR 1% cutoff).

Finally, pathway analysis of differential expression in post-eclosion adapted flies maintained in hypoxia showed Wnt signaling to figure prominently in both KEGG and Panther databases. Combining pathway and GO BP annotations yielded a total of 68 Wnt pathway-associated genes that were differentially expressed relative to control flies. Thus, both genomic sequencing (Chapter 3) and gene expression analysis

identified Wnt signaling as a candidate pathway involved in adaptation to hypoxia. These observations provide the rationale for performing validation experiments, which are described in Chapter 5.

Supplemental Tables

Table S4.1 Top-Scoring KEGG[14,15] Annotations for Differentially Expressed Genes in Post-Eclosion Hypoxia-Tolerant Flies Grown in 4% O_2.

Term	Count	PValue	Fold Enrichment
Upregulated			
dme00620:Pyruvate metabolism	15	0.002	2.474
dme00052:Galactose metabolism	10	0.008	2.673
dme00600:Sphingolipid metabolism	9	0.010	2.790
dme00564:Glycerophospholipid metabolism	16	0.011	1.968
dme00520:Amino sugar and nucleotide sugar metabolism	14	0.012	2.087
dme00010:Glycolysis / Gluconeogenesis	14	0.012	2.087
dme00680:Methane metabolism	5	0.012	4.844
dme00020:Citrate cycle (TCA cycle)	12	0.020	2.114
dme00051:Fructose and mannose metabolism	9	0.037	2.250
dme00565:Ether lipid metabolism	8	0.040	2.385
dme00561:Glycerolipid metabolism	11	0.047	1.938
dme00430:Taurine and hypotaurine metabolism	4	0.049	4.429
dme04070:Phosphatidylinositol signaling system	10	0.071	1.890
dme04310:Wnt signaling pathway	15	0.084	1.571
dme04711:Circadian rhythm	4	0.127	3.100
dme00910:Nitrogen metabolism	6	0.142	2.114
dme00230:Purine metabolism	20	0.212	1.271
dme00410:beta-Alanine metabolism	5	0.221	2.040
dme00500:Starch and sucrose metabolism	11	0.235	1.421
dme00380:Tryptophan metabolism	5	0.250	1.938
Downregulated			
dme03030:DNA replication	29	0.000	4.283
dme03420:Nucleotide excision repair	27	0.000	3.874
dme03430:Mismatch repair	17	0.000	4.493
dme03040:Spliceosome	46	0.000	2.081
dme00240:Pyrimidine metabolism	36	0.000	2.260
dme03440:Homologous recombination	12	0.000	3.545
dme03410:Base excision repair	12	0.000	3.013
dme04914:Progesterone-mediated oocyte maturation	17	0.003	2.082

Table S4.1 Top-Scoring KEGG14,15 Annotations for Differentially Expressed Genes in Post-Eclosion Hypoxia-Tolerant Flies Grown in 4% O2. (cont)

Term	Count	PValue	Fold Enrichment
dme01040:Biosynthesis of unsaturated fatty acids	9	0.006	2.825
dme03018:RNA degradation	18	0.008	1.883
dme03020:RNA polymerase	12	0.010	2.232
dme03450:Non-homologous end-joining	5	0.016	4.185
dme00563:Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	7	0.029	2.704
dme00280:Valine, leucine and isoleucine degradation	12	0.047	1.826
dme00230:Purine metabolism	33	0.048	1.358
dme00071:Fatty acid metabolism	13	0.048	1.764
dme00903:Limonene and pinene degradation	23	0.113	1.343
dme00310:Lysine degradation	10	0.123	1.674
dme00410:beta-Alanine metabolism	7	0.158	1.850
dme00592:alpha-Linolenic acid metabolism	6	0.160	2.009

Table S4.2 Signaling Pathway Analysis of Differential Expression in Post-Eclosion Hypoxia-Tolerant Flies Maintained in 4% O_2
KEGG[14,15] and Panther[16] signaling pathways with ≥2 differentially expressed genes, as determined by DAVID[122,123], are included.

Category	Term	Count	PValue	Affymetrix Probesets
KEGG	dme04340:Hedgehog signaling pathway	9	0.549	1633554_AT, 1626805_S_AT, 1632787_A_AT, 1634442_AT, 1635643_AT, 1630774_S_AT, 1635117_AT, 1641272_AT, 1623144_AT, 1623084_AT
KEGG	dme04310:Wnt signaling pathway	23	0.748	1640534_S_AT, 1627295_S_AT, 1632478_A_AT, 1628344_AT, 1622925_AT, 1639319_AT, 1628489_AT, 1639817_AT, 1635279_AT, 1636591_AT, 1632787_A_AT, 1638838_AT, 1623084_AT, 1630361_AT, 1626781_AT, 1636527_AT, 1638600_AT, 1628476_AT, 1629015_A_AT, 1636612_A_AT, 1624073_AT, 1637769_S_AT, 1634211_AT, 1630774_S_AT, 1623702_AT
KEGG	dme04330:Notch signaling pathway	7	0.776	1632457_S_AT, 1626617_AT, 1638568_S_AT, 1633700_AT, 1622925_AT, 1623200_AT, 1639144_A_AT
KEGG	dme04630:Jak-STAT signaling pathway	6	0.796	1627295_S_AT, 1637703_A_AT, 1627362_A_AT, 1626708_AT, 1622925_AT, 1635677_A_AT, 1632783_A_AT, 1635580_AT
KEGG	dme04013:MAPK signaling pathway	5	0.889	1627394_S_AT, 1635325_A_AT, 1628369_AT, 1641143_S_AT, 1638587_AT
KEGG	dme04070:Phosphatidylinositol signaling system	11	0.908	1641349_AT, 1639095_AT, 1632797_A_AT, 1635502_AT, 1632478_A_AT, 1633654_S_AT, 1639461_A_AT, 1630829_AT, 1633184_AT, 1641006_S_AT, 1627129_AT
KEGG	dme04350:TGF-beta signaling pathway	8	0.960	1636591_AT, 1622925_AT, 1624073_AT, 1627736_AT, 1633556_S_AT, 1623424_A_AT, 1639319_AT, 1639817_AT
KEGG	dme04150:mTOR signaling pathway	4	0.998	1627886_AT, 1624043_AT, 1625051_AT, 1640539_A_AT, 1634082_AT

Table S4.2 Signaling Pathway Analysis of Differential Expression in Post-Eclosion Hypoxia-Tolerant Flies Maintained in 4% O2 (cont)

Category	Term	Count	PValue	Affymetrix Probesets
PANTHER	P00025:Hedgehog signaling pathway	10	0.438	1632217_AT, 1626805_S_AT, 1622925_AT, 1634442_AT, 1625885_AT, 1630774_S_AT, 1635117_AT, 1632565_AT, 1628979_AT, 1623144_AT, 1623084_AT
PANTHER	P00057:Wnt signaling pathway	38	0.514	1627295_S_AT, 1640418_AT, 1633000_A_AT, 1625112_AT, 1633700_AT, 1633331_AT, 1632478_A_AT, 1626805_S_AT, 1628344_AT, 1622925_AT, 1640208_S_AT, 1627970_AT, 1639773_S_AT, 1637813_AT, 1632217_AT, 1638838_AT, 1627955_A_AT, 1623084_AT, 1634157_AT, 1626781_AT, 1639461_A_AT, 1636034_AT, 1634442_AT, 1640627_AT, 1631777_A_AT, 1623424_A_AT, 1628476_AT, 1638600_AT, 1641460_AT, 1624125_AT, 1624759_S_AT, 1628677_AT, 1635742_S_AT, 1624673_AT, 1640730_AT, 1623793_AT, 1637769_S_AT, 1631677_AT, 1636036_AT, 1630774_S_AT, 1641309_S_AT, 1630399_AT
PANTHER	P00038:JAK/STAT signaling pathway	2	0.765	1626730_S_AT, 1627362_A_AT, 1632783_A_AT
PANTHER	P00047:PDGF signaling pathway	16	0.892	1628552_AT, 1634440_S_AT, 1632478_A_AT, 1635639_A_AT, 1639461_A_AT, 1637828_A_AT, 1636614_A_AT, 1628677_AT, 1632217_AT, 1629859_S_AT, 1625516_AT, 1641021_A_AT, 1632787_A_AT, 1635325_A_AT, 1635603_AT, 1630774_S_AT, 1635054_AT, 1623084_AT
PANTHER	P00012:Cadherin signaling pathway	11	0.932	1633000_A_AT, 1628344_AT, 1638438_A_AT, 1631049_AT, 1631777_A_AT, 1640627_AT, 1624125_AT, 1635742_S_AT, 1632217_AT, 1637769_S_AT, 1630774_S_AT, 1623084_AT, 1630399_AT

Table S4.2 Signaling Pathway Analysis of Differential Expression in Post-Eclosion Hypoxia-Tolerant Flies Maintained in 4% O2 (cont)

Category	Term	Count	PValue	Affymetrix Probesets
PANTHER	P00021:FGF signaling pathway	12	0.960	1628677_AT, 1633088_AT, 1629859_S_AT, 1632478_A_AT, 1635325_A_AT, 1633665_AT, 1624979_AT, 1636375_AT, 1635677_A_AT, 1641006_S_AT, 1629546_AT, 1628476_AT
PANTHER	P00056:VEGF signaling pathway	4	0.986	1628677_AT, 1632478_A_AT, 1641006_S_AT, 1640486_S_AT
PANTHER	P00033:Insulin/IGF pathway-protein kinase B signaling cascade	6	0.988	1632217_AT, 1631242_AT, 1641006_S_AT, 1630774_S_AT, 1625727_AT, 1630305_AT, 1623084_AT
PANTHER	P00052:TGF-beta signaling pathway	10	0.996	1641053_S_AT, 1622925_AT, 1631242_AT, 1625141_AT, 1627736_AT, 1633556_S_AT, 1626766_S_AT, 1625727_AT, 1623424_A_AT, 1630305_AT
PANTHER	P00045:Notch signaling pathway	3	1.000	1639900_AT, 1623200_AT, 1639144_A_AT

Acknowledgements

Chapter 4 is, in part, a re-editing of materials currently being prepared for submission for publication: Gersten, M., Zhou, D., Haddad, G.G., & Subramaniam, S. Wnt pathway activation increases hypoxia tolerance in *Drosophila melanogaster*, in preparation. The dissertation author is the primary investigator and author of this paper.

5. EXPERIMENTAL VALIDATION OF WNT PATHWAY ACTIVATION IN TOLERANCE TO HYPOXIA IN FLIES

Introduction

Just over a decade ago a new approach to doing biology was proposed[179-181], which represented a paradigm shift from the reductionist approach that had dominated during the twentieth century. Rather than a bottom-up approach, focusing on intensive investigation of single molecules and small-scale modeling, "systems biology" would leverage the ever-increasing availability of high-throughput genome-scale data and powerful computing capabilities to integrate different types of data, formulate system-level models and propose testable hypotheses. Results of simulations and wet lab experiments would be fed back into the proposed model to initiate the next cycle in an iterative process of model refinement (Figure 5.1). Although systems biology is challenged by the noise, complexity and multiple timescales of biological systems and is still establishing its place in biological research[182], it is clear that the value of this integrated, multi-disciplinary methodology is its promise of revealing the structure and dynamics of biological systems, an understanding not afforded by detailed descriptions of components in isolation[181].

As described in Chapter 3 and 4, I analyzed two system-level datasets, genome sequencing and gene expression, to better understand how the hypoxia-tolerant *Drosophila* adapted to life at 4% O_2. Focusing on pathways rather than individual

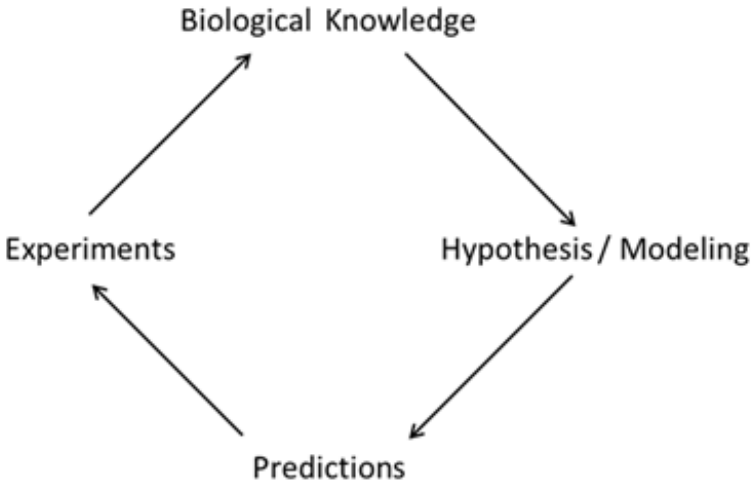


Figure 5.1 Systems biology research cycle.
Adapted from Kitano[181].

genes, both analyses identified Wnt signaling as significantly affected in the adapted flies, predicting that this pathway plays a role in fly adaptation to hypoxia. I sought to test this prediction by perturbing the system, using genetic manipulation of the Wnt pathway, to determine whether doing so would affect the ability of unadapted flies to tolerate a hypoxic environment. To do so, I utilized the GAL4-UAS system to overexpress or knock-down pathway genes. GAL4 is the yeast transcriptional activator of genes induced by galactose, regulating transcription upon binding to a defined DNA sequence (Upstream Activating Sequence, UAS), but GAL4 expression can stimulate transcription in a variety of systems, including *Drosophila*, where the gene of interest (GOI) is under UAS control[126,127]. The GAL4 driver/UAS system was developed by Brand and Perrimon[125] into a genetic tool that allows cell/tissue-specific expression in *Drosophila*, and currently there exist fly lines that express GAL4 in many different cells/tissues, as well as fly lines with a variety of genes under UAS

control. Although the system can be used in many ways[126], the most basic application (Figure 5.2) mates homozygous strains to produce F1 progeny which express the GOI in the cell/tissue in which the inserted regulatory element is active.

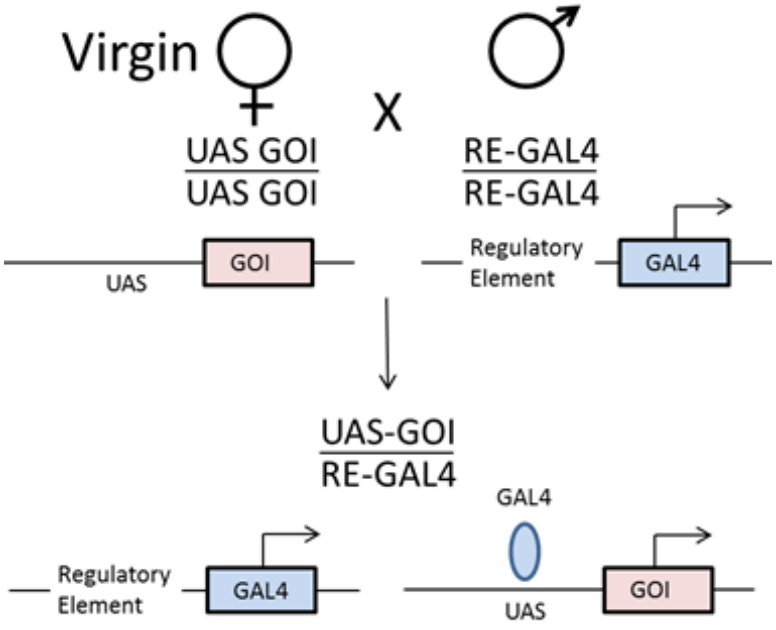


Figure 5.2 GAL4-UAS system.
GOI-gene of interest; RE- regulatory element. Adapted from Duffy[126] and St Johnston[127].

I evaluated hypoxia tolerance based on adult eclosion rate since in several forms of environmental stress, including hypoxia, flies are able develop up to the point of emergence from the pupa, an activity that requires highly coordinated neural/muscular activity[183]. Thus, the ability to successfully eclose from pupae is a key achievement in adaptation to chronic hypoxia. I used a hypoxic stress of 5% O_2, approximately equal to the roughly 38 mm Hg tension reported in capillary blood of high-altitude dwellers[130]; this level of hypoxia is severe enough to restrict eclosion of

control flies but not so severe as to preclude an impact by single gene effects[48,49]. I used the GAL4-UAS system to both overexpress and knockdown selected Wnt pathway genes. The genes were selected based on availability of a homozygous UAS strain as well as strategic location within the Wnt pathway; fly strain information is provided in Supplemental Table S5.1.

Methods

Drosophila stocks and cultures.

Elav and *Hml* GAL4 driver lines and UAS-expressor lines (*arm*, *dsh*, *dally*, *Rho1*, *bsk*, *rok*, *CaMKII*) were obtained from Bloomington *Drosophila* Stock Center (http://flystocks.bio.indiana.edu/). UAS-RNAi lines were obtained from the Vienna Stock Collection (http://stockcenter.vdrc.at/control/main; *sgg*, *nmo*) or Bloomington (*Apc*). Stocks were cultured on standard media.

Hypoxia tolerance testing.

The impact of Wnt-pathway activation on survival in hypoxia was assessed by the ability of flies to eclose after culture in a hypoxic environment. UAS(-RNAi) stocks were crossed with *elav* or *Hml* GAL4 driver lines to produce F1 flies that express the UAS insert in neurons or larval hemocytes, respectively. Each cross contained 5-10 virgin females and 5-10 males; except for UAS-*arm* and UAS-*dally* (with inserts on the X-chromosome), GAL4 virgin females and UAS males were used. UAS-*arm* and UAS-*dally* virgin females were crossed with males containing the *elav*-

Gal4 insert on the 3rd chromosome (driver 8760); other neuron-specific crosses used *elav*-Gal4 driver 458 with insert on the X chromosome. Flies were allowed to lay eggs for approximately 48 hours in normoxia; adults were then transferred to fresh vials and the vials containing eggs were moved to a room temperature chamber, computer-controlled to maintain a 5% oxygen atmosphere. Three replicate vials were set up for each cross and adults from each cross laid one or more clutches of eggs over 48-hour periods. Parental lines for each cross were tested in parallel as controls. After 3-4 weeks, vials from both crosses and parental lines were evaluated for eclosion rate. Experiments were excluded if a parental (control) rate of eclosion was > 0.70. Only vials with ≥15 total pupae were included. Numbers of eclosed (empty pupal case) and non-eclosed pupae were counted and the ratio between eclosed and total number of pupae was calculated. Each datapoint presented reflects 5-23 replicate tubes, each with ≥ 15 pupae, derived from 3-8 experiments; except for elav-GAL4/UAS-Apc results, experiments reflect at least two independent matings. Number of tubes, total number of pupae, and statistical details for each genetic cross experiment are provided in Supplemental Table S5.2. Statistical significance was calculated using two-tailed unpaired Student t-test and data are presented as mean ± 1.96 SEM (95% C.I.)[184]. An overview of this protocol is presented in Figure 5.3.

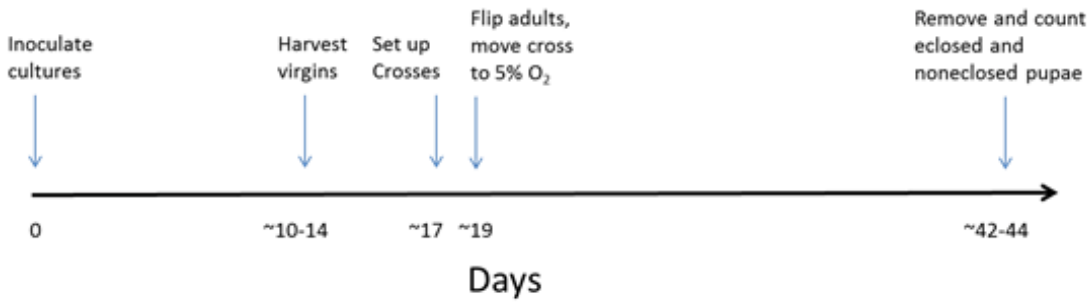


Figure 5.3 Hypoxia tolerance testing of genetically manipulated flies.
Crosses were made using 5-10 virgin females and 5-10 male flies with a minimum of three replicate tubes/cross. Flies were allowed to lay eggs for 2 days at 21% O_2 before adults were removed. See text for additional details.

Results

Neurons

Although the most straightforward approach to begin investigating the effects of Wnt pathway perturbation would be to overexpress or knockdown pathway genes ubiquitously, using a universal GAL4 driver, I considered, given the breadth of effects of the Wnt pathway in normal development, that such a strategy might be deleterious to the developing fly. Noting that the nervous system was a target tissue for hypoxia rescue in *C elegans*[185], I asked whether Wnt signaling modulation in the nervous system would impact hypoxia tolerance in flies. I used GAL4 drivers expressing *elav*, with inserts on chromosome 1 (X chromosome) and 3, respectively.

As shown in Figure 4.5, polymorphisms and gene expression differences were seen in all three Wnt pathways; however the canonical pathway is best studied and many target genes have been described across eukaryotic species. The gene expression changes observed in the post-eclosion AF (Figure 4.5), as well as expression changes

seen in Wnt target genes (Table 4.7), offer evidence for both activation and suppression of the Wnt canonical pathway. This is not surprising given that regulation of Wnt signaling is tissue- as well as time-dependent and the expression changes observed represent a mixture of fly tissues; furthermore, cross-talk with other pathways and feedback loops may affect the observed expression. I noted that the canonical pathway co-receptors *arr* and *dally* were both upregulated and considered that in the presence of Wnt signaling, upregulation of the *Tcf* homolog, *pang*, which otherwise may inhibit Wnt target gene expression[137], would potentiate the pathway. Consequently, I looked to determine whether Wnt pathway activation would affect the ability of unadapted flies to tolerate a 5% O_2 environment.

I looked at overexpression of several canonical Wnt pathway activators (*arm*, *dsh*, *dally*) and knockdown of pathway inhibitors (*GSK3β/sgg*, *Apc*) using two *elav* GAL4 driver lines, 8760 and 458. The *elav* 458 line has been used extensively in the Haddad laboratory and was the preferred driver, but could not be used with *arm* and *dally* which also have inserts on the X chromosome. The results of these experiments are presented in Figures 5.4 and 5.5. These experiments show that activation of canonical Wnt signaling in neurons leads to significantly increased adult eclosion. Mean eclosion rates for parental lines are significantly lower than rates for their respective crosses, with p-values ranging from $1.7x10^{-3}$ to $1.9x10^{-30}$. A closer examination of these results raises two interesting points. First, it is evident from Figure 5.5 that the mean eclosion rate for parental strain *elav* 458 differs between the

two experiments. This variability, which was often apparent for strains tested many times, may reflect variables including how early eggs were laid during the 2-day egg-

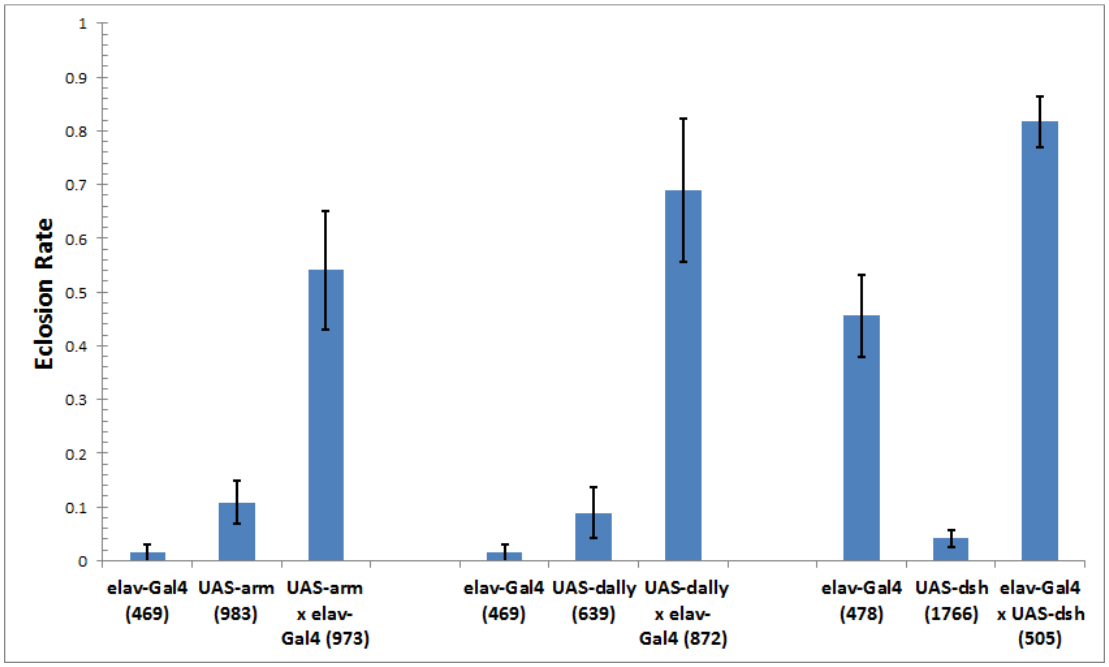


Figure 5.4 Overexpression of Wnt canonical pathway activators in neurons.
Neuron-specific overexpression of canonical pathway activators arm, dsh or dally driven by *elav*-GAL4. Highly efficient elav driver 458 (insert on X chromosome) was used with UAS-dsh; elav driver 8760 (insert on chromosome 3) was used with UAS-arm and UAS-dally, with inserts on the X chromosome. Eclosion rate for each cross is compared with that of its respective, concurrently tested parental strains using two-tailed unpaired Student's t-test. Bars represent the mean ± 1.96 SEM (95% C.I). Eclosion rate of cross compared with its respective parental strains in all cases yielded $p < 1x10^{-6}$. Total number of pupae counted is indicated in parentheses for each strain. (See Table S5.2 for details.).

laying period, and how many times and when in the developmental cycle the hypoxia chamber was accessed to add or remove experiments. Second, as seen in Figure 5.4, the mean eclosion rate for *elav* 8760 was much lower than for *elav* 458; this was a consistent observation for these two GAL4 driver lines. Nevertheless, improved eclosion rates for crosses could be seen using both of these parental lines, which apparently differ in their inherent ability to survive at 5% O_2, though parental

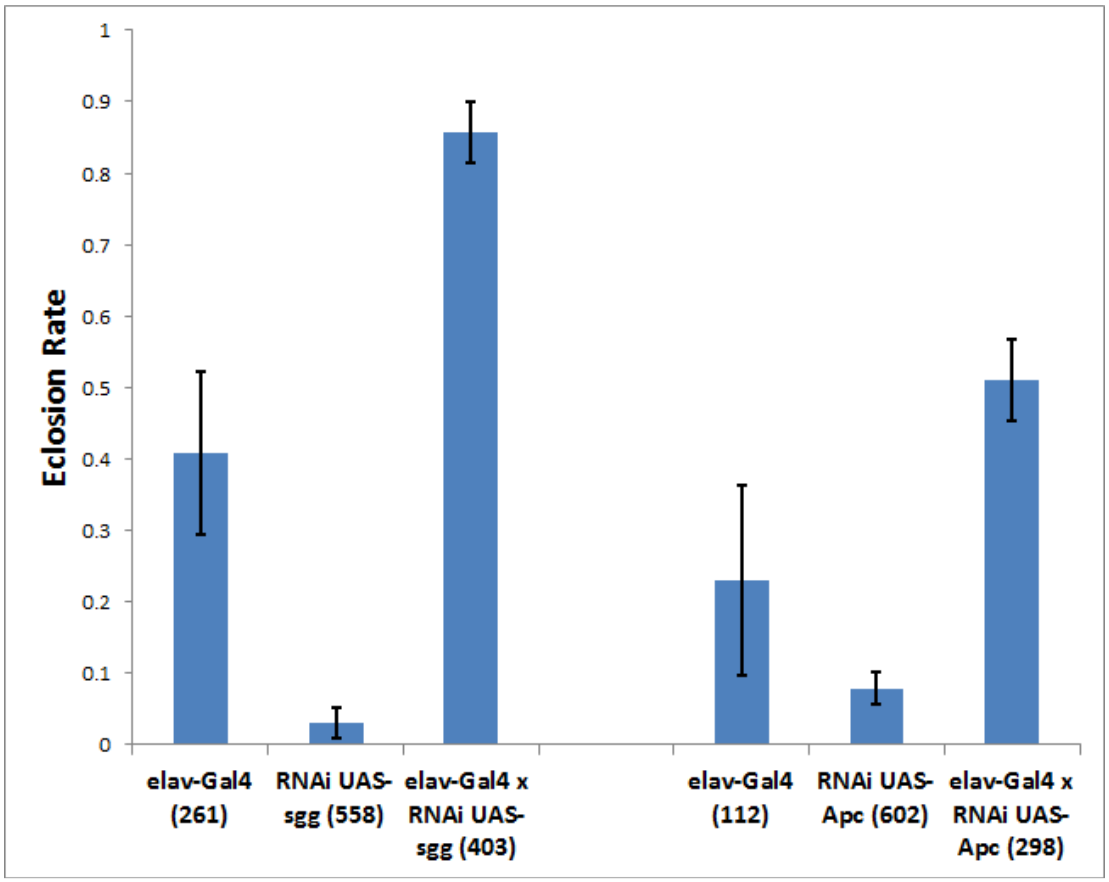


Figure 5.5 Knockdown of Wnt canonical pathway inhibitors in neurons.
Neuron-specific knockdown of canonical pathway inhibitors *sgg* or *Apc* driven by *elav*-GAL4 driver 458. Eclosion rate for each cross is compared with that of its respective, concurrently tested parental strains using two-tailed unpaired Student's t-test. Bars represent the mean ± 1.96 SEM (95% C.I). Eclosion rate of cross compared with its respective parental strains in all cases yielded $p < 0.002$. Total number of pupae counted is indicated in parentheses for each strain. (See Table S5.2 for details.) The different mean *elav*-Gal4 eclosion rates for the two experiments reflects variability observed in experiments conducted at different times.

background may have affected the absolute level of eclosion rate achieved by the respective crosses. Robustness to inherent parental hypoxia-tolerance increases confidence in the results observed.

Hemocytes

In selecting a second target tissue to explore, I noted that defense-related genes were among those differentially expressed by hypoxia-tolerant flies whether or not they were grown under hypoxic conditions, suggesting there might be a genetic component to the differential expression. This led me to ask whether hemocytes, which are prominently involved in *Drosophila* host defense, might play a role in hypoxia tolerance, and specifically, whether Wnt activation in hemocytes might impact hypoxia tolerance. These cells, the *Drosophila* counterpart of mammalian phagocytes, produce several antimicrobial peptides (AMP) as well as release *Spz* to activate Toll signaling in the fat body[186], and also play a key role in tissue remodeling during metamorphosis by removing apoptosing larval tissues[187]. In addition to the genes studied in neurons, I was also able to test the effect of a *nmo* knockdown line (which yielded few progeny when crossed with *elav*-Gal4) using the larval hemocyte driver, *Hml*-Gal4. Figures 5.6 and 5.7 show, respectively, that canonical pathway activator overexpression and inhibitor knockdown in hemocytes also leads to significantly increased adult eclosion, with p-values of crosses *vs* parental strains ranging from $2.7x10^{-2}$ to $5.7x10^{-21}$.

Since polymorphisms and differential expression were also detected in genes within the Wnt noncanonical pathways, I used the Hml-Gal4 driver to look at the effect of overexpression of several PCP pathway genes and a Wnt Ca^{2+} pathway gene. As shown in Figure 5.8, overexpression of *Rho1* and *bsk* (*Jnk*), in addition to *dsh*, but not *rok* were associated with significantly increased adult eclosion, with p-values

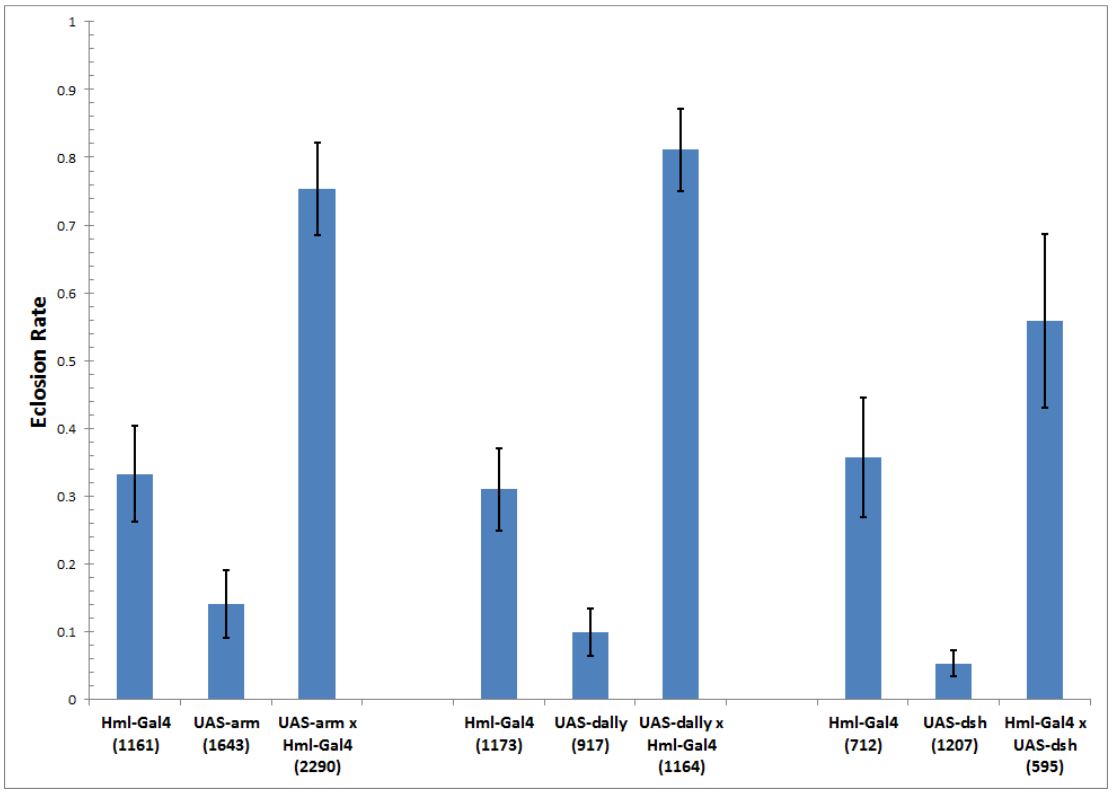


Figure 5.6 Overexpression of Wnt canonical pathway activators in hemocytes.
Hemocyte-specific overexpression of canonical pathway activators *arm*, *dsh* or *dally* driven by *Hml*-GAL4. Eclosion rate for each cross is compared with that of its respective, concurrently tested parental strains using two-tailed unpaired Student's t-test. Bars represent the mean ± 1.96 SEM (95% C.I). Eclosion rate of cross compared with its respective parental strains in all cases yielded $p < 1x10^{-7}$, except *Hml*-Gal4 x UAS-*dsh vs Hml*-Gal4 (p=0.016). Total number of pupae counted is indicated in parentheses for each strain. (See Table S5.2 for details.)

ranging from $1.6x10^{-2}$ to $2.3x10^{-15}$. In addition, overexpression of CaMKII (Figure 5.9) increased adult eclosion in 5% O_2 ($p < 2x10^{-6}$). Figure 4.5 indicates the location within the Wnt pathways of the genes studied, and also identifies Wnt pathway genes with polymorphisms and/or differential expression in post-eclosion AF. Although several Wnt pathway genes participate directly or *via* cross-talk with other signaling pathways, the fact that eclosion rate is increased by activation of these pathways at

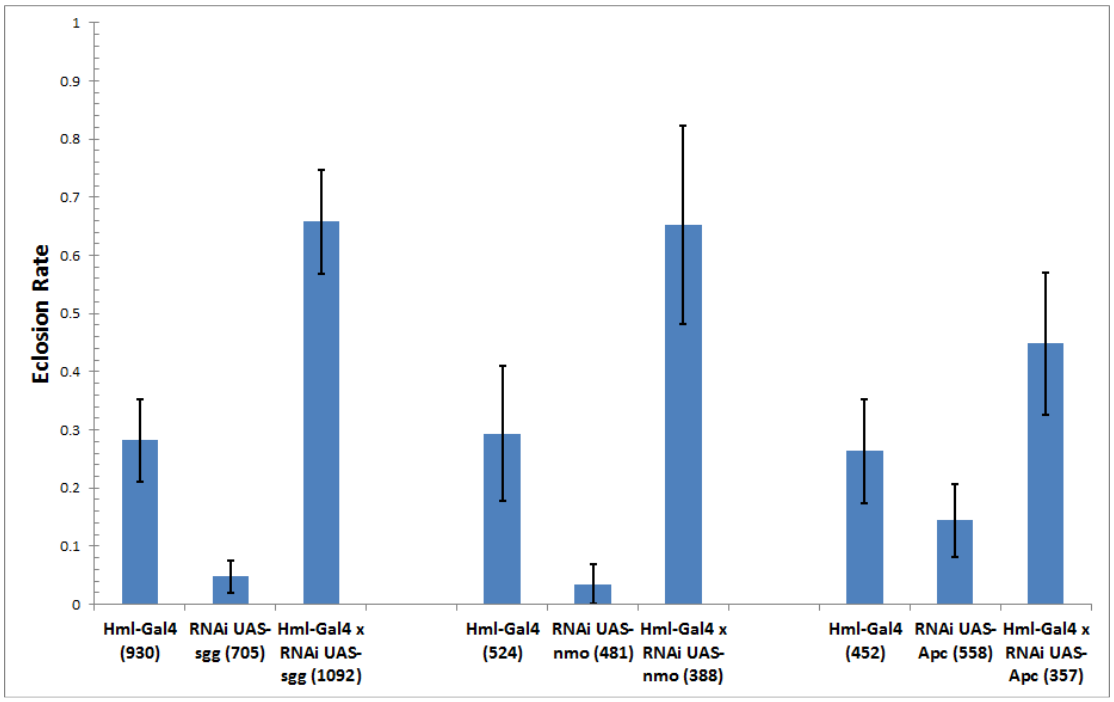


Figure 5.7 Knockdown of Wnt canonical pathway inhibitors in hemocytes.
Hemocyte-specific knockdown of canonical pathway inhibitors *sgg*, *nmo* or *Apc* driven by *Hml*-GAL4. Eclosion rate for each cross is compared with that of its respective, concurrently tested parental strains using two-tailed unpaired Student's t-test. Bars represent the mean ± 1.96 SEM (95% C.I). Eclosion rate of cross compared with its respective parental strains in all cases yielded $p < 0.004$, except *Hml*-Gal4 x RNAi UAS-*Apc vs Hml*-Gal4 ($p=0.027$). Total number of pupae counted is indicated in parentheses for each strain. (See Table S5.2 for details.)

multiple points provides support that activation of Wnt signaling *per se* contributes to increased hypoxia tolerance.

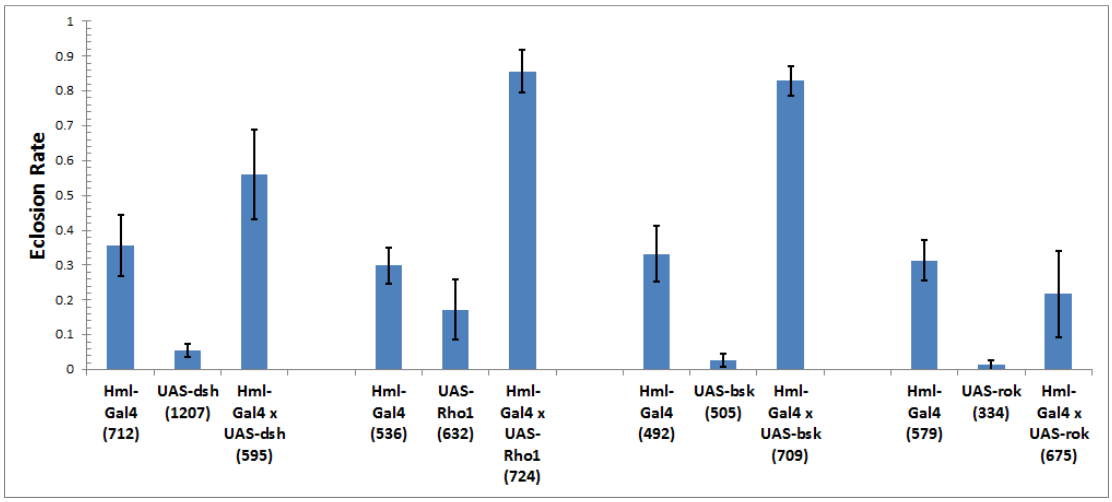


Figure 5.8 Overexpression of planar cell polarity pathway activators in hemocytes.
Hemocyte-specific expression of PCP pathway activators *dsh* (also a canonical pathway activator), *Rho1*, *bsk*, and *rok* driven by *Hml*-GAL4. Eclosion rate for each cross is compared with that of its respective, concurrently tested parental strains using two-tailed unpaired Student's t-test. Bars represent the mean ± 1.96 SEM (95% C.I). Eclosion rate of cross compared with its respective parental strains in all cases yielded $p < 1x10^{-7}$, except *Hml*-Gal4 x UAS-*dsh vs Hml*-Gal4 (p=0.016), *Hml*-Gal4 x UAS-*rok vs Hml*-Gal4 (p=0.184) and *Hml*-Gal4 x UAS-*rok vs* UAS-*rok* (p=0.0055). Total number of pupae counted is indicated in parentheses for each strain. (See Table S5.2 for details.)

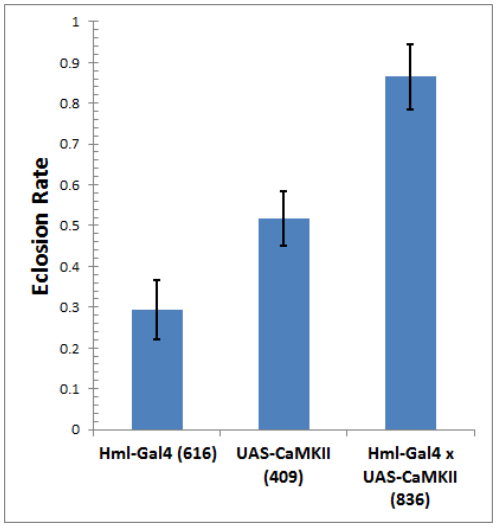


Figure 5.9 Overexpression of Wnt calcium pathway activator in hemocytes.
Hemocyte-specific expression of Wnt calcium pathway activator *CaMKII* driven by *Hml*-GAL4. Eclosion rate of cross is compared with that of its concurrently tested parental strains using two-tailed unpaired Student's t-test. Bars represent the mean ± 1.96 SEM (95% C.I). Eclosion rate of cross compared with its parental strains in both cases yielded $p < 2x10^{-6}$. Total number of pupae counted is indicated in parentheses for each strain. (See Table S5.2 for details.)

Discussion

Given its extensive participation in diverse developmental and transcriptional programs, it is not surprising that adaptation to hypoxia, which widely affects physiological processes, should involve Wnt signaling pathways. Recent work describes at least two points of intersection between hypoxia and Wnt signaling. In colon carcinoma lines, hypoxia (1% O_2) causes *HIF-1α*, the major regulator of the transcriptional response to hypoxia[113-115], to compete with *TCF-4* for binding to *β-catenin*; this causes an increase in *HIF-1α*-mediated transcription while decreasing *TCF-4* mediated transcription[188], which the authors suggest may promote survival and adaptation to hypoxia and contribute to the cell cycle arrest induced by hypoxia, respectively. In contrast to its effect on differentiated cells, hypoxia, acting via *HIF-1α,* is reported to enhance canonical Wnt signaling in embryonic stem cells and neural stem cells, promoting stem cell proliferation and Wnt-regulated differentiation[189,190]. Given the vulnerability of brain tissue to an inadequate oxygen supply, alterations in the number and type of neurons produced as a result of hypoxia stimulation of Wnt signaling, might increase an organism's ability to survive the reduced oxygen environment. The observed increase in hypoxia tolerance by activating canonical Wnt signaling exclusively in neurons supports a neurologic mechanism as at least contributory to hypoxia adaptation. Considering the criticality of coordinated neuronal-muscular activity in adult eclosion, it is possible that neuromuscular junction (NMJ) innervation may play a role. In this regard it is of interest that Wnt signaling in neurons contributes to synaptogenesis and bouton formation at NMJs, although it

appears to require only upstream elements of the canonical pathway[191]. In contrast, *hang*, a Wnt signaling inhibitor[129], and the gene containing the largest number of polymorphisms (Table 3.3), negatively regulates NMJ bouton formation[192]. Increased bouton formation may strengthen the NMJ allowing the adult to eclose at O_2 concentrations lower than normally required.

I have also shown that activation of the Wnt canonical or PCP pathway in hemocytes promotes increased adult eclosion in 5% O_2. Although host defense gene expression may be affected in the course of hypoxia adaptation simply as part of the stress response, it is possible, given their critical role in morphogenesis[187], that hemocyte physiology is specifically targeted. In particular, an increased number of circulating hemocytes might increase the efficiency of tissue remodeling essential for successful emergence of the adult fly. In this regard, canonical Wnt signaling has been shown to maintain hemocyte precursors both directly by preventing their differentiation, and indirectly by promoting proliferation and maintenance of cells in the hematopoietic microenvironment that maintain precursor stemness[193]; more prolonged precursor status might allow for increased rounds of replication and eventual hemocyte numbers. Noncanonical Wnt signaling could also play a role, as *Rac1* overexpression, acting via *bsk*, was shown to increase the number of circulating hemocytes by mobilizing the sessile hemocyte population[194]; *Rac* may also increase canonical Wnt signaling by promoting *β-catenin* nuclear translocation, either directly or by activated *Jnk* phosphorylation of *β-catenin*[195,196]. Finally, *Rac1* and *Rho1* GTPases, which both help regulate shape and migration of *Drosophila* hemocytes[197],

can reciprocally induce each other: *Rac1* acting *via bsk*, and *Rho1* acting *via dia*[198]. Thus, overexpression of either *Rac1*, *Rho1*, or *bsk* could potentially lead to induction of all three genes.

A role for Wnt signaling in hypoxia tolerance has been demonstrated here. Future work should address identifying the mechanism(s) through which Wnt pathway activation in neurons and hemocytes, and perhaps other tissues, promotes adaptation to hypoxia.

Supplemental Tables

Table S5.1 Fly strains used in genetic cross experiments.
Comments are from Bloomington Stock Center (http://flystocks.bio.indiana.edu/), except for *arm*, which was extracted from FlyBase[121]. B – Bloomington Stock Center; V – Vienna *Drosophila* RNAi Center (http://stockcenter.vdrc.at/control/main).

Gene	Source	Stock / transformant number	Genotype	Chrom(s) affected	Insertion chrom(s)	Comments
UAS Insert						
dsh	B	9453	w[*]; P{w[+mC]=UAS-dsh.myc}1-16	1;3	3	Expresses wild type dsh under UAS control
arm	B	4782	P{w[+mC]=UAS-arm.S10}C, y[1] w[1118]	1	1	Constitutively active
dally	B	5397	w[*] P{w[+mC]=UAS-dally.J}SJ1	1	1	
Rho1	B	7334	w[*]; P{w[+mC]=UAS-Rho1.Sph}2.1	1;3	3	Expresses wild type Rho1 under the control of UAS
bsk	B	6407	y[1] w[1118]; P{w[+mC]=UAS-bsk.A-Y}1	1;2	2	
rok	B	6669	y[1] w[*]; P{w[+mC]=UAS-rok.CAT}3.1	1;3	3	
CaMKII	B	29662	w[*]; P{w[+mC]=UAS-CaMKII.R3}2	1;2	2	Expresses the wild type R3 isoform of CaMKII under UAS control

Table S5.1 Fly strains used in genetic cross experiments. (cont)

Gene	Source	Stock / transformant number	Genotype	Chrom(s) affected	Insertion chrom(s)	Comments
RNAi-UAS Insert						
Apc	B	28582	y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HM05070}attP2	1;3	3	Expresses dsRNA for RNAi of Apc (FBgn0015589) under UAS control, TRiP
sgg	V	101538/KK			2	
nmo	V	104885/KK			2	
Gal4 Insert						
elav	B	8760	w[*]; P{w[+mC]=GAL4-elav.L}3	1;3	3	Expresses GAL4 in the nervous system.
elav	B	458	P{w[+mW.hs]=GawB}elav[C155]	1	1	GAL4 expressed in all tissues of the embryonic nervous system beginning at stage 12
Hml	B	6396	w[1118]; P{w[+mC]=Hml-GAL4.G}6-4	1;2	2	GAL4 expressed in larval hemocytes; Hml normally expressed at low levels

Table S5.2 Experimental Details for Genetic Cross Experiments.
Included are experiments in which the control (parental) eclosion rate was <70% and replicate tubes that contained ≥15 pupae. SD – Standard Deviation. Crosses are denoted as Virgin Female x Male.

		Eclosion Rate					
Targeted Tissue	Fly Strain	Mean	SD	p *vs* cross	# Expts	# Tubes	Total Pupae
Neuron	elav-Gal4	0.014	0.026	4.777E-09	4	12	469
	UAS-arm	0.108	0.070	3.051E-07		12	983
	UAS-arm x elav-Gal4	0.541	0.195			12	973
	elav-Gal4	0.014	0.026	1.741E-09	4	12	469
	UAS-dally	0.089	0.079	8.536E-08		11	639
	UAS-dally x elav-Gal4	0.689	0.237			12	872
	elav-Gal4	0.456	0.165	2.237E-09	8	18	478
	UAS-dsh	0.041	0.039	1.935E-30		23	1766
	elav-Gal4 x UAS-dsh	0.817	0.106			19	505
	elav-Gal4	0.409	0.202	1.012E-07	5	12	261
	RNAi UAS-sgg	0.031	0.040	3.322E-23		14	558
	elav-Gal4 x RNAi UAS-sgg	0.857	0.081			14	403
	elav-Gal4	0.230	0.151	1.711E-03	3	5	112
	RNAi UAS-Apc	0.079	0.107	3.148E-07		9	602
	elav-Gal4 x RNAi UAS-Apc	0.511	0.097			8	298
Larval Hemocyte	Hml-Gal4	0.333	0.165	2.368E-10	7	21	1161
	UAS-arm	0.141	0.118	3.890E-17		21	1643
	UAS-arm x Hml-Gal4	0.754	0.160			21	2290
	Hml-Gal4	0.310	0.144	1.050E-13	7	21	1173
	UAS-dally	0.099	0.077	5.724E-21		19	917
	UAS-dally x Hml-Gal4	0.811	0.136			19	1164

Table S5.2 Experimental Details for Genetic Cross Experiments. (cont)

		Eclosion Rate					
Targeted Tissue	Fly Strain	Mean	SD	p *vs* cross	# Expts	# Tubes	Total Pupae
	Hml-Gal4	0.358	0.174	1.603E-02	5	15	712
	UAS-dsh	0.054	0.038	1.688E-08		15	1207
	Hml-Gal4 x dsh	0.559	0.245			14	595
	Hml-Gal4	0.282	0.140	5.259E-07	5	15	930
	RNAi UAS-sgg	0.048	0.056	3.389E-13		15	705
	Hml-Gal4 x RNAi UAS-sgg	0.658	0.176			15	1092
	Hml-Gal4	0.294	0.178	3.374E-03	3	9	524
	RNAi UAS-nmo	0.035	0.053	1.691E-06		9	481
	Hml-Gal4 x RNAi UAS-nmo	0.652	0.230			7	388
	Hml-Gal4	0.263	0.136	2.734E-02	3	9	452
	RNAi UAS-Apc	0.144	0.097	4.291E-04		9	558
	Hml-Gal4 x RNAi UAS-Apc	0.448	0.176			8	357
	Hml-Gal4	0.299	0.079	3.081E-10	3	9	536
	UAS-Rho1	0.172	0.131	8.517E-10		9	632
	Hml-Gal4 x UAS-Rho1	0.855	0.093			9	724
	Hml-Gal4	0.332	0.122	9.212E-09	3	9	492
	UAS-bsk	0.026	0.028	2.273E-15		8	505
	Hml-Gal4 x UAS-bsk	0.829	0.064			9	709
	Hml-Gal4	0.313	0.088	1.838E-01	3	9	579
	UAS-rok	0.014	0.022	5.504E-03		9	334
	Hml-Gal4 x UAS-rok	0.217	0.189			9	675
	Hml-Gal4	0.294	0.126	5.281E-10	4	12	616
	UAS-CaMKII	0.517	0.113	1.902E-06		11	409
	Hml-Gal4 x UAS-CaMKII	0.865	0.141			12	836

Acknowledgements

Chapter 5 is, in part, a re-editing of materials currently being prepared for submission for publication: Gersten, M., Zhou, D., Haddad, G.G., & Subramaniam, S. Wnt pathway activation increases hypoxia tolerance in *Drosophila melanogaster*, in preparation. The dissertation author is the primary investigator and author of this paper.

6. CONCLUSIONS

High-throughput genome-scale datasets and computational tools to effectively analyze them are providing the means to study complex phenotypes from a systemic perspective. Such a perspective is needed to address the multiple and diverse interacting pathways and processes that produce complex traits and diseases. Animal models, which were productively used in conventional methods in biology, remain useful in the current systems biology paradigm, offering the same advantages for control of genetic and environmental variables and more extensive *in vivo* experimental manipulation.

I studied two complex phenotypes with relevance to human diseases utilizing two very different animal models which provided different kinds of benefit for the study of their respective human diseases. The rhesus macaque is phylogenetically very close to human, and SIV encephalopathy closely reproduces many of the clinico-pathologic aspects of human HIV Associated Dementia. More distant in the phylogenetic spectrum is *Drosophila melanogaster*, which nonetheless has been extremely useful in studying human conditions, having orthologs to approximately 75% of human disease genes[44,45]. Its particular advantages in the current study were its short developmental cycle, which permitted the experimental selection of a population of hypoxia-adapted flies over years rather than centuries or millenia, and the many genetic tools available for testing hypotheses.

In Chapter 2, I integrated genome-scale hippocampus expression data from SIVE monkeys with a human protein-protein interaction network and applied a

module-finding algorithm[32] to detect differentially expressed, connected groups of genes that distinguished control from infected monkeys. This allowed me to identify EGR1, among the set of downregulated genes, as a candidate gene in the pathogenesis of SIV-related neurodysfunction. Experimental work, coupled with a synthesis of data reported in the literature, led to the proposal that EGR1 is downregulated in SIV encephalitis as a consequence of the host defense response, leading to changes in proteosomal subunit selection that contribute to cognitive deficits. This conjecture highlights the complexity of host defense responses, which may produce deleterious, as well as beneficial effects, and illustrates how this predicament may arise from cross-talk between different signaling pathways.

In Chapters 3 and 4, I analyzed genomic sequencing and genome-scale gene expression data in flies that were selected for tolerance to 4% O_2, a level normally prohibitive to successful completion of the fly developmental cycle. Data from both analyses pointed to involvement of the Wnt signaling pathway in the adaptation to hypoxia. Like Notch signaling, which has a demonstrated effect on tolerance to hypoxia[48], Wnt signaling comprises a highly conserved developmental pathway, one which impacts axis specification, morphogenesis, axon guidance, stem cell maintenance and differentiation, and in vertebrates, neural tube formation and carcinogenesis. In Chapter 5, I utilized the GAL4-UAS system[125-127] to show that activation of Wnt signaling in neurons or hemocytes (the *Drosophila* counterpart of mammalian macrophages) promoted hypoxia tolerance, as measured by increased rates of adult eclosion in 5% O_2.

While the implications of improved fly eclosion rates for human hypoxia-related disorders are not straightforward, it is possible that an understanding of the mechanisms underlying the Wnt effect in flies may suggest avenues for parallel exploration in, for example, hypoxia-tolerant tumors. As a start, future studies in flies might determine the Wnt target genes that are up- or downregulated in Wnt-activating genetic crosses demonstrating increased hypoxia tolerance.

REFERENCES

1. Thomas, P.D. & Kejariwal, A. Coding sing le-nucleotide polymorphisms associated with complex vs. Mendelian disease: Evolutionary evidence for differences in molecular effects. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15398-15403 (2004).

2. Lesnick, T.G., Papapetropoulos, S., Mash, D.C., Ffrench-Mullen, J., Shehadeh, L., de Andrade, M., Henley, J.R., Rocca, W.A., Ahlskog, J.E. & Maraganore, D.M. A genomic pathway approach to a complex disease: axon guidance and Parkinson disease. *PLoS Genet* **3**, e98 (2007).

3. Schork, N.J. Genetics of complex disease: approaches, problems, and solutions. *Am J Respir Crit Care Med* **156**, S103-9 (1997).

4. Noorbakhsh, F., Overall, C.M. & Power, C. Deciphering complex mechanisms in neurodegenerative diseases: the advent of systems biology. *Trends Neurosci* **32**, 88-100 (2009).

5. Villoslada, P., Steinman, L. & Baranzini, S.E. Systems biology and its application to the understanding of neurological diseases. *Ann Neurol* **65**, 124-39 (2009).

6. Murphy, D. Gene expression studies using microarrays: Principles, problems, and prospects. *Advances in Physiology Education* **26**, 256-270 (2002).

7. Allison, D.B., Cui, X.Q., Page, G.P. & Sabripour, M. Microarray data analysis: from disarray to consolidation and consensus. *Nature Reviews Genetics* **7**, 55-65 (2006).

8. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* **10**, 57-63 (2009).

9. Hutchison, C.A., 3rd. DNA sequencing: bench to bedside and beyond. *Nucleic Acids Res* **35**, 6227-37 (2007).

10. Metzker, M.L. Sequencing technologies - the next generation. *Nat Rev Genet* **11**, 31-46 (2010).

11. Chait, B.T. Mass spectrometry in the postgenomic era. *Annu Rev Biochem* **80**, 239-46 (2011).

12. Yates, J.R., Ruse, C.I. & Nakorchevsky, A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng* **11**, 49-79 (2009).

13. Massie, C.E. & Mills, I.G. ChIPping away at gene regulation. *EMBO Rep* **9**, 337-43 (2008).

14. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30 (2000).

15. Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M. & Hirakawa, M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res* **38**, D355-60 (2010).

16. Mi, H. & Thomas, P. PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. *Methods Mol Biol* **563**, 123-40 (2009).

17. 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., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M. & Sherlock, G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-9 (2000).

18. Carter, S.L., Brechbuhler, C.M., Griffin, M. & Bond, A.T. Gene co-expression network topology provides a framework for molecular characterization of cellular state. *Bioinformatics* **20**, 2242-50 (2004).

19. Giot, L., Bader, J.S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y.L., Ooi, C.E., Godwin, B., Vitols, E., Vijayadamodar, G., Pochart, P., Machineni,

H., Welsh, M., Kong, Y., Zerhusen, B., Malcolm, R., Varrone, Z., Collis, A., Minto, M., Burgess, S., McDaniel, L., Stimpson, E., Spriggs, F., Williams, J., Neurath, K., Ioime, N., Agee, M., Voss, E., Furtak, K., Renzulli, R., Aanensen, N., Carrolla, S., Bickelhaupt, E., Lazovatsky, Y., DaSilva, A., Zhong, J., Stanyon, C.A., Finley, R.L., Jr., White, K.P., Braverman, M., Jarvie, T., Gold, S., Leach, M., Knight, J., Shimkets, R.A., McKenna, M.P., Chant, J. & Rothberg, J.M. A protein interaction map of Drosophila melanogaster. *Science* **302**, 1727-36 (2003).

20. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M. & Sakaki, Y. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A* **98**, 4569-74 (2001).

21. Li, S., Armstrong, C.M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P.O., Han, J.D., Chesneau, A., Hao, T., Goldberg, D.S., Li, N., Martinez, M., Rual, J.F., Lamesch, P., Xu, L., Tewari, M., Wong, S.L., Zhang, L.V., Berriz, G.F., Jacotot, L., Vaglio, P., Reboul, J., Hirozane-Kishikawa, T., Li, Q., Gabel, H.W., Elewa, A., Baumgartner, B., Rose, D.J., Yu, H., Bosak, S., Sequerra, R., Fraser, A., Mango, S.E., Saxton, W.M., Strome, S., Van Den Heuvel, S., Piano, F., Vandenhaute, J., Sardet, C., Gerstein, M., Doucette-Stamm, L., Gunsalus, K.C., Harper, J.W., Cusick, M.E., Roth, F.P., Hill, D.E. & Vidal, M. A map of the interactome network of the metazoan C. elegans. *Science* **303**, 540-3 (2004).

22. Uetz, P., Giot, L., Cagney, G., Mansfield, T.A., Judson, R.S., Knight, J.R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S. & Rothberg, J.M. A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature* **403**, 623-7 (2000).

23. Ewing, R.M., Chu, P., Elisma, F., Li, H., Taylor, P., Climie, S., McBroom-Cerajewski, L., Robinson, M.D., O'Connor, L., Li, M., Taylor, R., Dharsee, M., Ho, Y., Heilbut, A., Moore, L., Zhang, S., Ornatsky, O., Bukhman, Y.V., Ethier, M., Sheng, Y., Vasilescu, J., Abu-Farha, M., Lambert, J.P., Duewel, H.S., Stewart, II, Kuehl, B., Hogue, K., Colwill, K., Gladwish, K., Muskat, B., Kinach, R., Adams, S.L., Moran, M.F., Morin, G.B., Topaloglou, T. & Figeys, D. Large-scale mapping of human protein-protein interactions by mass spectrometry. *Mol Syst Biol* **3**, 89 (2007).

24. Rual, J.F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G.F., Gibbons, F.D., Dreze, M., Ayivi-Guedehoussou, N., Klitgord, N., Simon, C., Boxem, M., Milstein, S., Rosenberg, J., Goldberg, D.S., Zhang, L.V., Wong, S.L., Franklin, G., Li, S., Albala, J.S., Lim, J., Fraughton, C., Llamosas, E., Cevik, S., Bex, C., Lamesch, P., Sikorski, R.S., Vandenhaute, J., Zoghbi, H.Y., Smolyar, A., Bosak, S., Sequerra, R., Doucette-Stamm, L., Cusick, M.E., Hill, D.E., Roth, F.P. & Vidal, M. Towards a proteome-scale map of the human protein-protein interaction network. *Nature* **437**, 1173-8 (2005).

25. Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F.H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., Timm, J., Mintzlaff, S., Abraham, C., Bock, N., Kietzmann, S., Goedde, A., Toksoz, E., Droege, A., Krobitsch, S., Korn, B., Birchmeier, W., Lehrach, H. & Wanker, E.E. A human protein-protein interaction network: a resource for annotating the proteome. *Cell* **122**, 957-68 (2005).

26. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. & Ideker, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-504 (2003).

27. Fields, S. & Sternglanz, R. The two-hybrid system: an assay for protein-protein interactions. *Trends Genet* **10**, 286-92 (1994).

28. Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G. & Superti-Furga, G. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141-7 (2002).

29. Sharan, R., Ulitsky, I. & Shamir, R. Network-based prediction of protein function. *Mol Syst Biol* **3**, 88 (2007).

30. Vazquez, A., Flammini, A., Maritan, A. & Vespignani, A. Global protein function prediction from protein-protein interaction networks. *Nat Biotechnol* **21**, 697-700 (2003).

31. Chen, J. & Yuan, B. Detecting functional modules in the yeast protein-protein interaction network. *Bioinformatics* **22**, 2283-90 (2006).

32. Chuang, H.Y., Lee, E., Liu, Y.T., Lee, D. & Ideker, T. Network-based classification of breast cancer metastasis. *Mol Syst Biol* **3**, 140 (2007).

33. Ideker, T., Ozier, O., Schwikowski, B. & Siegel, A.F. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* **18 Suppl 1**, S233-40 (2002).

34. Segal, E., Wang, H. & Koller, D. Discovering molecular pathways from protein interaction and gene expression data. *Bioinformatics* **19 Suppl 1**, i264-72 (2003).

35. Aitman, T.J., Boone, C., Churchill, G.A., Hengartner, M.O., Mackay, T.F. & Stemple, D.L. The future of model organisms in human disease research. *Nat Rev Genet* **12**, 575-82 (2011).

36. Ances, B.M. & Clifford, D.B. HIV-associated neurocognitive disorders and the impact of combination antiretroviral therapies. *Curr Neurol Neurosci Rep* **8**, 455-61 (2008).

37. Joska, J.A., Gouse, H., Paul, R.H., Stein, D.J. & Flisher, A.J. Does highly active antiretroviral therapy improve neurocognitive function? A systematic review. *J Neurovirol* **16**, 101-14 (2010).

38. Roberts, E.S., Burudi, E.M., Flynn, C., Madden, L.J., Roinick, K.L., Watry, D.D., Zandonatti, M.A., Taffe, M.A. & Fox, H.S. Acute SIV infection of the brain leads to upregulation of IL6 and interferon-regulated genes: expression patterns throughout disease progression and impact on neuroAIDS. *J Neuroimmunol* **157**, 81-92 (2004).

39. Roberts, E.S., Zandonatti, M.A., Watry, D.D., Madden, L.J., Henriksen, S.J., Taffe, M.A. & Fox, H.S. Induction of pathogenic sets of genes in macrophages and neurons in NeuroAIDS. *Am J Pathol* **162**, 2041-57 (2003).

40. Kaul, M., Zheng, J., Okamoto, S., Gendelman, H.E. & Lipton, S.A. HIV-1 infection and AIDS: consequences for the central nervous system. *Cell Death Differ* **12 Suppl 1**, 878-92 (2005).

41. Bozon, B., Davis, S. & Laroche, S. Regulated transcription of the immediate-early gene Zif268: mechanisms and gene dosage-dependent function in synaptic plasticity and memory formation. *Hippocampus* **12**, 570-7 (2002).

42. Jones, M.W., Errington, M.L., French, P.J., Fine, A., Bliss, T.V., Garel, S., Charnay, P., Bozon, B., Laroche, S. & Davis, S. A requirement for the immediate early gene Zif268 in the expression of late LTP and long-term memories. *Nat Neurosci* **4**, 289-96 (2001).

43. Knapska, E. & Kaczmarek, L. A gene for neuronal plasticity in the mammalian brain: Zif268/Egr-1/NGFI-A/Krox-24/TIS8/ZENK? *Prog Neurobiol* **74**, 183-211 (2004).

44. Pandey, U.B. & Nichols, C.D. Human disease models in Drosophila melanogaster and the role of the fly in therapeutic drug discovery. *Pharmacol Rev* **63**, 411-36 (2011).

45. Reiter, L.T., Potocki, L., Chien, S., Gribskov, M. & Bier, E. A systematic analysis of human disease-associated gene sequences in Drosophila melanogaster. *Genome Res* **11**, 1114-25 (2001).

46. Zhou, D., Visk, D.W. & Haddad, G.G. Drosophila, a golden bug, for the dissection of the genetic basis of tolerance and susceptibility to hypoxia. *Pediatr Res* **66**, 239-47 (2009).

47. Van Voorhies, W.A. Metabolic function in Drosophila melanogaster in response to hypoxia and pure oxygen. *J Exp Biol* **212**, 3132-41 (2009).

48. Zhou, D., Udpa, N., Gersten, M., Visk, D.W., Bashir, A., Xue, J., Frazer, K.A., Posakony, J.W., Subramaniam, S., Bafna, V. & Haddad, G.G. Experimental selection of hypoxia-tolerant Drosophila melanogaster. *Proc Natl Acad Sci U S A* **108**, 2349-54 (2011).

49. Zhou, D., Xue, J., Chen, J., Morcillo, P., Lambert, J.D., White, K.P. & Haddad, G.G. Experimental selection for Drosophila survival in extremely low O(2) environment. *PLoS One* **2**, e490 (2007).

50. Zhou, D., Xue, J., Lai, J.C., Schork, N.J., White, K.P. & Haddad, G.G. Mechanisms underlying hypoxia tolerance in Drosophila melanogaster: hairy as a metabolic switch. *PLoS Genet* **4**, e1000221 (2008).

51. Diesing, T.S., Swindells, S., Gelbard, H. & Gendelman, H.E. HIV-1-associated dementia: a basic science and clinical perspective. *AIDS Read* **12**, 358-68 (2002).

52. Mattson, M.P., Haughey, N.J. & Nath, A. Cell death in HIV dementia. *Cell Death Differ* **12 Suppl 1**, 893-904 (2005).

53. Rausch, D.M., Murray, E.A. & Eiden, L.E. The SIV-infected rhesus monkey model for HIV-associated dementia and implications for neurological diseases. *J Leukoc Biol* **65**, 466-74 (1999).

54. Masliah, E., Roberts, E.S., Langford, D., Everall, I., Crews, L., Adame, A., Rockenstein, E. & Fox, H.S. Patterns of gene dysregulation in the frontal cortex of patients with HIV encephalitis. *J Neuroimmunol* **157**, 163-75 (2004).

55. Roberts E.S.; Fox, H.S. Transcriptional analysis of inflammatory and glial responses in the central nervous system: development of RNA fingerprints for neuroAIDS and other central nervous system disorders. in *The Neurology of AIDS, Second Edition* (ed. Gendelman H.E.; Grant, I.E., I.P.; Lipton, S.A.; Swindells, S.) 225-237 (Oxford University Press, 2005).

56. Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B. & Speed, T.P. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* **31**, e15 (2003).

57. Baldi, P. & Long, A.D. A Bayesian framework for the analysis of microarray expression data: regularized t -test and statistical inferences of gene changes. *Bioinformatics* **17**, 509-19 (2001).

58. Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C. & Lempicki, R.A. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* **4**, P3 (2003).

59. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society B* **57**, 289-300 (1995).

60. Mak, H.C., Daly, M., Gruebel, B. & Ideker, T. CellCircuits: a database of protein network models. *Nucleic Acids Res* **35**, D538-45 (2007).

61. Maere, S., Heymans, K. & Kuiper, M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* **21**, 3448-9 (2005).

62. Peles, E., Nativ, M., Lustig, M., Grumet, M., Schilling, J., Martinez, R., Plowman, G.D. & Schlessinger, J. Identification of a novel contactin-associated transmembrane receptor with multiple domains implicated in protein-protein interactions. *Embo J* **16**, 978-88 (1997).

63. Chen, T.C., Lai, Y.K., Yu, C.K. & Juang, J.L. Enterovirus 71 triggering of neuronal apoptosis through activation of Abl-Cdk5 signalling. *Cell Microbiol* **9**, 2676-88 (2007).

64. Taga, T., Suzuki, A., Gonzalez-Gomez, I., Gilles, F.H., Stins, M., Shimada, H., Barsky, L., Weinberg, K.I. & Laug, W.E. alpha v-Integrin antagonist EMD 121974 induces apoptosis in brain tumor cells growing on vitronectin and tenascin. *Int J Cancer* **98**, 690-7 (2002).

65. Linding, R., Jensen, L.J., Ostheimer, G.J., van Vugt, M.A., Jorgensen, C., Miron, I.M., Diella, F., Colwill, K., Taylor, L., Elder, K., Metalnikov, P., Nguyen, V., Pasculescu, A., Jin, J., Park, J.G., Samson, L.D., Woodgett, J.R., Russell, R.B., Bork, P., Yaffe, M.B. & Pawson, T. Systematic discovery of in vivo phosphorylation networks. *Cell* **129**, 1415-26 (2007).

66. Lock, P., Abram, C.L., Gibson, T. & Courtneidge, S.A. A new method for isolating tyrosine kinase substrates used to identify fish, an SH3 and PX domain-containing protein, and Src substrate. *Embo J* **17**, 4346-57 (1998).

67. Lieser, S.A., Shindler, C., Aubol, B.E., Lee, S., Sun, G. & Adams, J.A. Phosphoryl transfer step in the C-terminal Src kinase controls Src recognition. *J Biol Chem* **280**, 7769-76 (2005).

68. Bhat, M.A., Rios, J.C., Lu, Y., Garcia-Fresco, G.P., Ching, W., St Martin, M., Li, J., Einheber, S., Chesler, M., Rosenbluth, J., Salzer, J.L. & Bellen, H.J. Axon-glia interactions and the domain organization of myelinated axons requires neurexin IV/Caspr/Paranodin. *Neuron* **30**, 369-83 (2001).

69. von Hertzen, L.S. & Giese, K.P. Memory reconsolidation engages only a subset of immediate-early genes induced during consolidation. *J Neurosci* **25**, 1935-42 (2005).

70. Matilla, A., Roberson, E.D., Banfi, S., Morales, J., Armstrong, D.L., Burright, E.N., Orr, H.T., Sweatt, J.D., Zoghbi, H.Y. & Matzuk, M.M. Mice lacking ataxin-1 display learning deficits and decreased hippocampal paired-pulse facilitation. *J Neurosci* **18**, 5508-16 (1998).

71. Marchionini, D.M., Lehrmann, E., Chu, Y., He, B., Sortwell, C.E., Becker, K.G., Freed, W.J., Kordower, J.H. & Collier, T.J. Role of heparin binding growth factors in nigrostriatal dopamine system development and Parkinson's disease. *Brain Res* **1147**, 77-88 (2007).

72. Pavlov, I., Voikar, V., Kaksonen, M., Lauri, S.E., Hienola, A., Taira, T. & Rauvala, H. Role of heparin-binding growth-associated molecule (HB-GAM) in hippocampal LTP and spatial learning revealed by studies on overexpressing and knockout mice. *Mol Cell Neurosci* **20**, 330-42 (2002).

73. Kruger, R. The role of synphilin-1 in synaptic function and protein degradation. *Cell Tissue Res* **318**, 195-9 (2004).

74. Crocker, S.F., Costain, W.J. & Robertson, H.A. DNA microarray analysis of striatal gene expression in symptomatic transgenic Huntington's mice (R6/2)

reveals neuroinflammation and insulin associations. *Brain Res* **1088**, 176-86 (2006).

75. Blalock, E.M., Chen, K.C., Sharrow, K., Herman, J.P., Porter, N.M., Foster, T.C. & Landfield, P.W. Gene microarrays in hippocampal aging: statistical profiling identifies novel processes correlated with cognitive impairment. *J Neurosci* **23**, 3807-19 (2003).

76. Booth, S., Bowman, C., Baumgartner, R., Sorensen, G., Robertson, C., Coulthart, M., Phillipson, C. & Somorjai, R.L. Identification of central nervous system genes involved in the host response to the scrapie agent during preclinical and clinical infection. *J Gen Virol* **85**, 3459-71 (2004).

77. Garden, G.A. & Morrison, R.S. The multiple roles of p53 in the pathogenesis of HIV associated dementia. *Biochem Biophys Res Commun* **331**, 799-809 (2005).

78. Fowler, M.A., Sidiropoulou, K., Ozkan, E.D., Phillips, C.W. & Cooper, D.C. Corticolimbic expression of TRPC4 and TRPC5 channels in the rodent brain. *PLoS ONE* **2**, e573 (2007).

79. Caffrey, T.M., Joachim, C., Paracchini, S., Esiri, M.M. & Wade-Martins, R. Haplotype-specific expression of exon 10 at the human MAPT locus. *Hum Mol Genet* **15**, 3529-37 (2006).

80. Ambrosini, E. & Aloisi, F. Chemokines and glial cells: a complex network in the central nervous system. *Neurochem Res* **29**, 1017-38 (2004).

81. Bajetto, A., Bonavia, R., Barbero, S. & Schettini, G. Characterization of chemokines and their receptors in the central nervous system: physiopathological implications. *J Neurochem* **82**, 1311-29 (2002).

82. Darbinian-Sarkissian, N., Czernik, M., Peruzzi, F., Gordon, J., Rappaport, J., Reiss, K., Khalili, K. & Amini, S. Dysregulation of NGF-signaling and Egr-1 expression by Tat in neuronal cell culture. *J Cell Physiol* **208**, 506-15 (2006).

83. Cho, C. & Miller, R.J. Chemokine receptors and neural function. *J Neurovirol* **8**, 573-84 (2002).

84. Ragozzino, D. CXC chemokine receptors in the central nervous system: Role in cerebellar neuromodulation and development. *J Neurovirol* **8**, 559-72 (2002).

85. Moelling, K., Schad, K., Bosse, M., Zimmermann, S. & Schweneker, M. Regulation of Raf-Akt Cross-talk. *J Biol Chem* **277**, 31099-106 (2002).

86. Zimmermann, S. & Moelling, K. Phosphorylation and regulation of Raf by Akt (protein kinase B). *Science* **286**, 1741-4 (1999).

87. O'Boyle, G., Brain, J.G., Kirby, J.A. & Ali, S. Chemokine-mediated inflammation: Identification of a possible regulatory role for CCR2. *Mol Immunol* **44**, 1944-53 (2007).

88. Moosavi, M., Maghsoudi, N., Zahedi-Asl, S., Naghdi, N., Yousefpour, M. & Trounce, I.A. The role of PI3/Akt pathway in the protective effect of insulin against corticosterone cell death induction in hippocampal cell culture. *Neuroendocrinology* **88**, 293-8 (2008).

89. Fukunaga, K. & Kawano, T. Akt is a molecular target for signal transduction therapy in brain ischemic insult. *J Pharmacol Sci* **92**, 317-27 (2003).

90. Limatola, C., Ciotti, M.T., Mercanti, D., Santoni, A. & Eusebi, F. Signaling pathways activated by chemokine receptor CXCR2 and AMPA-type glutamate receptors and involvement in granule cells survival. *J Neuroimmunol* **123**, 9-17 (2002).

91. Gosselin, R.D., Varela, C., Banisadr, G., Mechighel, P., Rostene, W., Kitabgi, P. & Melik-Parsadaniantz, S. Constitutive expression of CCR2 chemokine receptor and inhibition by MCP-1/CCL2 of GABA-induced currents in spinal cord neurones. *J Neurochem* **95**, 1023-34 (2005).

92. Xiong, H., Boyle, J., Winkelbauer, M., Gorantla, S., Zheng, J., Ghorpade, A., Persidsky, Y., Carlson, K.A. & Gendelman, H.E. Inhibition of long-term

potentiation by interleukin-8: implications for human immunodeficiency virus-1-associated dementia. *J Neurosci Res* **71**, 600-7 (2003).

93. James, A.B., Conway, A.M. & Morris, B.J. Genomic profiling of the neuronal target genes of the plasticity-related transcription factor -- Zif268. *J Neurochem* **95**, 796-810 (2005).

94. James, A.B., Conway, A.M. & Morris, B.J. Regulation of the neuronal proteasome by Zif268 (Egr1). *J Neurosci* **26**, 1624-34 (2006).

95. Ozato, K., Tailor, P. & Kubota, T. The interferon regulatory factor family in host defense: mechanism of action. *J Biol Chem* **282**, 20065-9 (2007).

96. Paun, A. & Pitha, P.M. The IRF family, revisited. *Biochimie* **89**, 744-53 (2007).

97. Namiki, S., Nakamura, T., Oshima, S., Yamazaki, M., Sekine, Y., Tsuchiya, K., Okamoto, R., Kanai, T. & Watanabe, M. IRF-1 mediates upregulation of LMP7 by IFN-gamma and concerted expression of immunosubunits of the proteasome. *FEBS Lett* **579**, 2781-7 (2005).

98. Lopez-Salon, M., Alonso, M., Vianna, M.R., Viola, H., Mello e Souza, T., Izquierdo, I., Pasquini, J.M. & Medina, J.H. The ubiquitin-proteasome cascade is required for mammalian long-term memory formation. *Eur J Neurosci* **14**, 1820-6 (2001).

99. Speese, S.D., Trotta, N., Rodesch, C.K., Aravamudan, B. & Broadie, K. The ubiquitin proteasome system acutely regulates presynaptic protein turnover and synaptic efficacy. *Curr Biol* **13**, 899-910 (2003).

100. Peri, S., Navarro, J.D., Amanchy, R., Kristiansen, T.Z., Jonnalagadda, C.K., Surendranath, V., Niranjan, V., Muthusamy, B., Gandhi, T.K., Gronborg, M., Ibarrola, N., Deshpande, N., Shanker, K., Shivashankar, H.N., Rashmi, B.P., Ramya, M.A., Zhao, Z., Chandrika, K.N., Padma, N., Harsha, H.C., Yatish, A.J., Kavitha, M.P., Menezes, M., Choudhury, D.R., Suresh, S., Ghosh, N., Saravana, R., Chandran, S., Krishna, S., Joy, M., Anand, S.K., Madavan, V., Joseph, A., Wong, G.W., Schiemann, W.P., Constantinescu, S.N., Huang, L., Khosravi-Far, R., Steen, H., Tewari, M., Ghaffari, S., Blobe, G.C., Dang, C.V.,

Garcia, J.G., Pevsner, J., Jensen, O.N., Roepstorff, P., Deshpande, K.S., Chinnaiyan, A.M., Hamosh, A., Chakravarti, A. & Pandey, A. Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Res* **13**, 2363-71 (2003).

101. Bader, G.D., Donaldson, I., Wolting, C., Ouellette, B.F., Pawson, T. & Hogue, C.W. BIND--The Biomolecular Interaction Network Database. *Nucleic Acids Res* **29**, 242-5 (2001).

102. Joshi-Tope, G., Gillespie, M., Vastrik, I., D'Eustachio, P., Schmidt, E., de Bono, B., Jassal, B., Gopinath, G.R., Wu, G.R., Matthews, L., Lewis, S., Birney, E. & Stein, L. Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res* **33**, D428-32 (2005).

103. Xenarios, I., Rice, D.W., Salwinski, L., Baron, M.K., Marcotte, E.M. & Eisenberg, D. DIP: the database of interacting proteins. *Nucleic Acids Res* **28**, 289-91 (2000).

104. Ramani, A.K., Bunescu, R.C., Mooney, R.J. & Marcotte, E.M. Consolidating the set of known human protein-protein interactions in preparation for large-scale mapping of the human interactome. *Genome Biol* **6**, R40 (2005).

105. Bader, G.D. & Hogue, C.W. Analyzing yeast protein-protein interaction data obtained from different sources. *Nat Biotechnol* **20**, 991-7 (2002).

106. Bader, J.S., Chaudhuri, A., Rothberg, J.M. & Chant, J. Gaining confidence in high-throughput protein interaction networks. *Nat Biotechnol* **22**, 78-85 (2004).

107. Burdo, T.H., Marcondes, M.C., Lanigan, C.M., Penedo, M.C. & Fox, H.S. Susceptibility of Chinese rhesus monkeys to SIV infection. *Aids* **19**, 1704-6 (2005).

108. Watry, D., Lane, T.E., Streb, M. & Fox, H.S. Transfer of neuropathogenic simian immunodeficiency virus with naturally infected microglia. *Am J Pathol* **146**, 914-23 (1995).

109. Allison, D.B., Gadbury, G.L., Heo, M.S., Fernandez, J.R., Lee, C.K., Prolla, T.A. & Weindruch, R. A mixture model approach for the analysis of microarray gene expression data. *Computational Statistics & Data Analysis* **39**, 1-20 (2002).

110. Remm, M., Storm, C.E. & Sonnhammer, E.L. Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J Mol Biol* **314**, 1041-52 (2001).

111. Wikoff, W.R., Pendyala, G., Siuzdak, G. & Fox, H.S. Metabolomic analysis of the cerebrospinal fluid reveals changes in phospholipase expression in the CNS of SIV-infected macaques. *J Clin Invest* **118**, 2661-9 (2008).

112. Gersten, M., Alirezaei, M., Marcondes, M.C., Flynn, C., Ravasi, T., Ideker, T. & Fox, H.S. An integrated systems analysis implicates EGR1 downregulation in simian immunodeficiency virus encephalitis-induced neural dysfunction. *J Neurosci* **29**, 12467-76 (2009).

113. Kenneth, N.S. & Rocha, S. Regulation of gene expression by hypoxia. *Biochem J* **414**, 19-29 (2008).

114. Semenza, G.L. Targeting HIF-1 for cancer therapy. *Nat Rev Cancer* **3**, 721-32 (2003).

115. Weidemann, A. & Johnson, R.S. Biology of HIF-1alpha. *Cell Death Differ* **15**, 621-7 (2008).

116. Storey, K.B. & Storey, J.M. Metabolic rate depression and biochemical adaptation in anaerobiosis, hibernation and estivation. *Q Rev Biol* **65**, 145-74 (1990).

117. Hochachka, P.W. & Lutz, P.L. Mechanism, origin, and evolution of anoxia tolerance in animals. *Comp Biochem Physiol B Biochem Mol Biol* **130**, 435-59 (2001).

118. Ramirez, J.M., Folkow, L.P. & Blix, A.S. Hypoxia tolerance in mammals and birds: from the wilderness to the clinic. *Annu Rev Physiol* **69**, 113-43 (2007).

119. Haddad, G.G. Tolerance to low O2: lessons from invertebrate genetic models. *Exp Physiol* **91**, 277-82 (2006).

120. Li, H., Ruan, J. & Durbin, R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* **18**, 1851-8 (2008).

121. Tweedie, S., Ashburner, M., Falls, K., Leyland, P., McQuilton, P., Marygold, S., Millburn, G., Osumi-Sutherland, D., Schroeder, A., Seal, R. & Zhang, H. FlyBase: enhancing Drosophila Gene Ontology annotations. *Nucleic Acids Res* **37**, D555-9 (2009).

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

123. Huang da, W., Sherman, B.T. & Lempicki, R.A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**, 1-13 (2009).

124. Costello, J.C., Dalkilic, M.M., Beason, S.M., Gehlhausen, J.R., Patwardhan, R., Middha, S., Eads, B.D. & Andrews, J.R. Gene networks in Drosophila melanogaster: integrating experimental data to predict gene function. *Genome Biol* **10**, R97 (2009).

125. Brand, A.H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-15 (1993).

126. Duffy, J.B. GAL4 system in Drosophila: a fly geneticist's Swiss army knife. *Genesis* **34**, 1-15 (2002).

127. St Johnston, D. The art and design of genetic screens: Drosophila melanogaster. *Nat Rev Genet* **3**, 176-88 (2002).

128. Abdu, U., Brodsky, M. & Schupbach, T. Activation of a meiotic checkpoint during Drosophila oogenesis regulates the translation of Gurken through Chk2/Mnk. *Curr Biol* **12**, 1645-51 (2002).

129. DasGupta, R., Kaykas, A., Moon, R.T. & Perrimon, N. Functional genomic analysis of the Wnt-wingless signaling pathway. *Science* **308**, 826-33 (2005).

130. Beall, C.M. Two routes to functional adaptation: Tibetan and Andean high-altitude natives. *Proc Natl Acad Sci U S A* **104 Suppl 1**, 8655-60 (2007).

131. Aldenderfer, M.S. Moving up in the world. *American Scientist* **91**, 542-549 (2003).

132. Andersson, E.R., Sandberg, R. & Lendahl, U. Notch signaling: simplicity in design, versatility in function. *Development* **138**, 3593-612 (2011).

133. Bray, S.J. Notch signalling: a simple pathway becomes complex. *Nat Rev Mol Cell Biol* **7**, 678-89 (2006).

134. Cadigan, K.M. & Nusse, R. Wnt signaling: a common theme in animal development. *Genes Dev* **11**, 3286-305 (1997).

135. Komiya, Y. & Habas, R. Wnt signal transduction pathways. *Organogenesis* **4**, 68-75 (2008).

136. Logan, C.Y. & Nusse, R. The Wnt signaling pathway in development and disease. *Annu Rev Cell Dev Biol* **20**, 781-810 (2004).

137. Archbold, H.C., Yang, Y.X., Chen, L. & Cadigan, K.M. How do they do Wnt they do?: regulation of transcription by the Wnt/beta-catenin pathway. *Acta Physiol (Oxf)* (2011).

138. van Amerongen, R. & Nusse, R. Towards an integrated view of Wnt signaling in development. *Development* **136**, 3205-14 (2009).

139. Hayward, P., Kalmar, T. & Arias, A.M. Wnt/Notch signalling and information processing during development. *Development* **135**, 411-24 (2008).

140. Rao, T.P. & Kuhl, M. An updated overview on Wnt signaling pathways: a prelude for more. *Circ Res* **106**, 1798-806 (2010).

141. Miller, J.R., Hocking, A.M., Brown, J.D. & Moon, R.T. Mechanism and function of signal transduction by the Wnt/beta-catenin and Wnt/Ca2+ pathways. *Oncogene* **18**, 7860-72 (1999).

142. Seifert, J.R. & Mlodzik, M. Frizzled/PCP signalling: a conserved mechanism regulating cell polarity and directed motility. *Nat Rev Genet* **8**, 126-38 (2007).

143. Kohn, A.D. & Moon, R.T. Wnt and calcium signaling: beta-catenin-independent pathways. *Cell Calcium* **38**, 439-46 (2005).

144. Collu, G.M. & Brennan, K. Cooperation between Wnt and Notch signalling in human breast cancer. *Breast Cancer Res* **9**, 105 (2007).

145. Espinosa, L., Ingles-Esteve, J., Aguilera, C. & Bigas, A. Phosphorylation by glycogen synthase kinase-3 beta down-regulates Notch activity, a link for Notch and Wnt pathways. *J Biol Chem* **278**, 32227-35 (2003).

146. Hurlbut, G.D., Kankel, M.W., Lake, R.J. & Artavanis-Tsakonas, S. Crossing paths with Notch in the hyper-network. *Curr Opin Cell Biol* **19**, 166-75 (2007).

147. Ungerback, J., Elander, N., Grunberg, J., Sigvardsson, M. & Soderkvist, P. The Notch-2 gene is regulated by Wnt signaling in cultured colorectal cancer cells. *PLoS One* **6**, e17957 (2011).

148. Ayyanan, A., Civenni, G., Ciarloni, L., Morel, C., Mueller, N., Lefort, K., Mandinova, A., Raffoul, W., Fiche, M., Dotto, G.P. & Brisken, C. Increased Wnt signaling triggers oncogenic conversion of human breast epithelial cells by a Notch-dependent mechanism. *Proc Natl Acad Sci U S A* **103**, 3799-804 (2006).

149. Duncan, A.W., Rattis, F.M., DiMascio, L.N., Congdon, K.L., Pazianos, G., Zhao, C., Yoon, K., Cook, J.M., Willert, K., Gaiano, N. & Reya, T. Integration of Notch and Wnt signaling in hematopoietic stem cell maintenance. *Nat Immunol* **6**, 314-22 (2005).

150. Estrach, S., Ambler, C.A., Lo Celso, C., Hozumi, K. & Watt, F.M. Jagged 1 is a beta-catenin target gene required for ectopic hair follicle formation in adult epidermis. *Development* **133**, 4427-38 (2006).

151. Rodilla, V., Villanueva, A., Obrador-Hevia, A., Robert-Moreno, A., Fernandez-Majada, V., Grilli, A., Lopez-Bigas, N., Bellora, N., Alba, M.M., Torres, F., Dunach, M., Sanjuan, X., Gonzalez, S., Gridley, T., Capella, G., Bigas, A. & Espinosa, L. Jagged1 is the pathological link between Wnt and Notch pathways in colorectal cancer. *Proc Natl Acad Sci U S A* **106**, 6315-20 (2009).

152. Li, C., Zhang, Y., Lu, Y., Cui, Z., Yu, M., Zhang, S. & Xue, X. Evidence of the cross talk between Wnt and Notch signaling pathways in non-small-cell lung cancer (NSCLC): Notch3-siRNA weakens the effect of LiCl on the cell cycle of NSCLC cell lines. *J Cancer Res Clin Oncol* **137**, 771-8 (2011).

153. Jin, Y.H., Kim, H., Oh, M., Ki, H. & Kim, K. Regulation of Notch1/NICD and Hes1 expressions by GSK-3alpha/beta. *Mol Cells* **27**, 15-9 (2009).

154. Saint Just Ribeiro, M., Hansson, M.L., Lindberg, M.J., Popko-Scibor, A.E. & Wallberg, A.E. GSK3beta is a negative regulator of the transcriptional coactivator MAML1. *Nucleic Acids Res* **37**, 6691-700 (2009).

155. Li, M., Wang, H., Huang, T., Wang, J., Ding, Y., Li, Z., Zhang, J. & Li, L. TAB2 scaffolds TAK1 and NLK in repressing canonical Wnt signaling. *J Biol Chem* **285**, 13397-404 (2010).

156. Ishitani, T., Hirao, T., Suzuki, M., Isoda, M., Ishitani, S., Harigaya, K., Kitagawa, M., Matsumoto, K. & Itoh, M. Nemo-like kinase suppresses Notch signalling by interfering with formation of the Notch active transcriptional complex. *Nat Cell Biol* **12**, 278-85 (2010).

157. Axelrod, J.D., Matsuno, K., Artavanis-Tsakonas, S. & Perrimon, N. Interaction between Wingless and Notch signaling pathways mediated by dishevelled. *Science* **271**, 1826-32 (1996).

158. Munoz-Descalzo, S., Sanders, P.G., Montagne, C., Johnson, R.I., Balayo, T. & Arias, A.M. Wingless modulates the ligand independent traffic of Notch through Dishevelled. *Fly (Austin)* **4**(2010).

159. Shahi, P., Seethammagari, M.R., Valdez, J.M., Xin, L. & Spencer, D.M. Wnt and Notch pathways have interrelated opposing roles on prostate progenitor cell proliferation and differentiation. *Stem Cells* **29**, 678-88 (2011).

160. De Strooper, B. & Annaert, W. Where Notch and Wnt signaling meet. The presenilin hub. *J Cell Biol* **152**, F17-20 (2001).

161. Kang, D.E., Soriano, S., Frosch, M.P., Collins, T., Naruse, S., Sisodia, S.S., Leibowitz, G., Levine, F. & Koo, E.H. Presenilin 1 facilitates the constitutive turnover of beta-catenin: differential activity of Alzheimer's disease-linked PS1 mutants in the beta-catenin-signaling pathway. *J Neurosci* **19**, 4229-37 (1999).

162. Hayward, P., Brennan, K., Sanders, P., Balayo, T., DasGupta, R., Perrimon, N. & Martinez Arias, A. Notch modulates Wnt signalling by associating with Armadillo/beta-catenin and regulating its transcriptional activity. *Development* **132**, 1819-30 (2005).

163. Lawrence, N., Langdon, T., Brennan, K. & Arias, A.M. Notch signaling targets the Wingless responsiveness of a Ubx visceral mesoderm enhancer in Drosophila. *Curr Biol* **11**, 375-85 (2001).

164. Foltz, D.R., Santiago, M.C., Berechid, B.E. & Nye, J.S. Glycogen synthase kinase-3beta modulates notch signaling and stability. *Curr Biol* **12**, 1006-11 (2002).

165. Brack, A.S., Conboy, I.M., Conboy, M.J., Shen, J. & Rando, T.A. A temporal switch from notch to Wnt signaling in muscle stem cells is necessary for normal adult myogenesis. *Cell Stem Cell* **2**, 50-9 (2008).

166. Munoz-Descalzo, S., Tkocz, K., Balayo, T. & Arias, A.M. Modulation of the ligand-independent traffic of Notch by Axin and Apc contributes to the activation of Armadillo in Drosophila. *Development* **138**, 1501-6 (2011).

167. Tusher, V.G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* **98**, 5116-21 (2001).

168. Affymetrix. Technical Note: Guide to Probe Logarithmic Intensity Error (PLIER) Estimation. (2005).

169. Therneau, T.M. & Ballman, K.V. What does PLIER really do? *Cancer Inform* **6**, 423-31 (2008).

170. Hsiao, A., Worrall, D.S., Olefsky, J.M. & Subramaniam, S. Variance-modeled posterior inference of microarray data: detecting gene-expression changes in 3T3-L1 adipocytes. *Bioinformatics* **20**, 3108-27 (2004).

171. Ginzinger, D.G. Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. *Exp Hematol* **30**, 503-12 (2002).

172. Livak, K.J. & Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-8 (2001).

173. Ernst, J. & Bar-Joseph, Z. STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinformatics* **7**, 191 (2006).

174. Fernandez-Ayala, D.J., Chen, S., Kemppainen, E., O'Dell, K.M. & Jacobs, H.T. Gene expression in a Drosophila model of mitochondrial disease. *PLoS One* **5**, e8549 (2010).

175. Storz, J.F., Scott, G.R. & Cheviron, Z.A. Phenotypic plasticity and genetic adaptation to high-altitude hypoxia in vertebrates. *J Exp Biol* **213**, 4125-36 (2010).

176. Beall, C.M. Andean, Tibetan, and Ethiopian patterns of adaptation to high-altitude hypoxia. *Integr Comp Biol* **46**, 18-24 (2006).

177. Xing, G., Qualls, C., Huicho, L., Rivera-Ch, M., Stobdan, T., Slessarev, M., Prisman, E., Ito, S., Wu, H., Norboo, A., Dolma, D., Kunzang, M., Norboo, T.,

Gamboa, J.L., Claydon, V.E., Fisher, J., Zenebe, G., Gebremedhin, A., Hainsworth, R., Verma, A. & Appenzeller, O. Adaptation and mal-adaptation to ambient hypoxia; Andean, Ethiopian and Himalayan patterns. *PLoS One* **3**, e2342 (2008).

178. Hunter, S., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquenne, L., Finn, R.D., Gough, J., Haft, D., Hulo, N., Kahn, D., Kelly, E., Laugraud, A., Letunic, I., Lonsdale, D., Lopez, R., Madera, M., Maslen, J., McAnulla, C., McDowall, J., Mistry, J., Mitchell, A., Mulder, N., Natale, D., Orengo, C., Quinn, A.F., Selengut, J.D., Sigrist, C.J., Thimma, M., Thomas, P.D., Valentin, F., Wilson, D., Wu, C.H. & Yeats, C. InterPro: the integrative protein signature database. *Nucleic Acids Res* **37**, D211-5 (2009).

179. Ideker, T., Galitski, T. & Hood, L. A new approach to decoding life: Systems biology. *Annual Review of Genomics and Human Genetics* **2**, 343-372 (2001).

180. Kitano, H. Perspectives on systems biology. *New Generation Computing* **18**, 199-216 (2000).

181. Kitano, H. Systems biology: A brief overview. *Science* **295**, 1662-1664 (2002).

182. Macilwain, C. Systems Biology: Evolving into the Mainstream. *Cell* **144**, 839-841 (2011).

183. Yocum, G.D., Zdarek, J., Joplin, K.H., Lee, R.E., Jr., Smith, D.C., Manter, K.D. & Denlinger, D.L. Alteration of the eclosion rhythm and eclosion behavior in the flesh fly, Sarcophaga crassipalpis, by low and high temperature stress. *Journal of Insect Physiology* **40**, 13-21 (1994).

184. Streiner, D.L. Maintaining standards: differences between the standard deviation and standard error, and when to use each. *Can J Psychiatry* **41**, 498-502 (1996).

185. Scott, B.A., Avidan, M.S. & Crowder, C.M. Regulation of hypoxic death in C. elegans by the insulin/IGF receptor homolog DAF-2. *Science* **296**, 2388-91 (2002).

186. Irving, P., Ubeda, J.M., Doucet, D., Troxler, L., Lagueux, M., Zachary, D., Hoffmann, J.A., Hetru, C. & Meister, M. New insights into Drosophila larval haemocyte functions through genome-wide analysis. *Cellular Microbiology* **7**, 335-350 (2005).

187. Lanot, R., Zachary, D., Holder, F. & Meister, M. Postembryonic hematopoiesis in Drosophila. *Dev Biol* **230**, 243-57 (2001).

188. Kaidi, A., Williams, A.C. & Paraskeva, C. Interaction between beta-catenin and HIF-1 promotes cellular adaptation to hypoxia. *Nat Cell Biol* **9**, 210-7 (2007).

189. Mazumdar, J., O'Brien, W.T., Johnson, R.S., LaManna, J.C., Chavez, J.C., Klein, P.S. & Simon, M.C. O2 regulates stem cells through Wnt/beta-catenin signalling. *Nat Cell Biol* **12**, 1007-13 (2010).

190. Zhang, K., Zhu, L. & Fan, M. Oxygen, a Key Factor Regulating Cell Behavior during Neurogenesis and Cerebral Diseases. *Front Mol Neurosci* **4**, 5 (2011).

191. Miech, C., Pauer, H.U., He, X. & Schwarz, T.L. Presynaptic local signaling by a canonical wingless pathway regulates development of the Drosophila neuromuscular junction. *J Neurosci* **28**, 10875-84 (2008).

192. Schwenkert, I., Eltrop, R., Funk, N., Steinert, J.R., Schuster, C.M. & Scholz, H. The hangover gene negatively regulates bouton addition at the Drosophila neuromuscular junction. *Mech Dev* **125**, 700-11 (2008).

193. Sinenko, S.A., Mandal, L., Martinez-Agosto, J.A. & Banerjee, U. Dual role of wingless signaling in stem-like hematopoietic precursor maintenance in Drosophila. *Dev Cell* **16**, 756-63 (2009).

194. Williams, M.J., Wiklund, M.L., Wikman, S. & Hultmark, D. Rac1 signalling in the Drosophila larval cellular immune response. *J Cell Sci* **119**, 2015-24 (2006).

195. Wu, X., Tu, X., Joeng, K.S., Hilton, M.J., Williams, D.A. & Long, F. Rac1 activation controls nuclear localization of beta-catenin during canonical Wnt signaling. *Cell* **133**, 340-53 (2008).

196. Schlessinger, K., Hall, A. & Tolwinski, N. Wnt signaling pathways meet Rho GTPases. *Genes Dev* **23**, 265-77 (2009).

197. Fauvarque, M.O. & Williams, M.J. Drosophila cellular immunity: a story of migration and adhesion. *J Cell Sci* **124**, 1373-82 (2011).

198. Williams, M.J., Habayeb, M.S. & Hultmark, D. Reciprocal regulation of Rac1 and Rho1 in Drosophila circulating immune surveillance cells. *J Cell Sci* **120**, 502-11 (2007).