Identification of a Loss of Function Allele That Spontaneously Arose in a Widely Used C57BL/6 Substrain Reveals a Novel Role for Dock2 in CD8 T Cell Homeostasis and Differentiation

Identification of a loss of function allele that spontaneously arose in a widely used C57BL/6 substrain reveals a novel role for Dock2 in CD8 T cell homeostasis and differentiation

A dissertation presented

by

Ezana Demissie

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Immunology

Harvard University

Cambridge, Massachusetts

February 2018

Dissertation Advisor: Shiv Pillai Ezana Demissie

ABSTRACT

Commercially available C57BL/6 mice are widely in biomedical research laboratories. We used whole genome sequencing to identify a homozygous copy number variant that disrupts the function of DOCK2 in a commercially available C57BL/6 mouse substrain (C57BL/6NHsd). This strain of C57BL/6 mice, from Harlan-Sprague (acquired by Envigo Biosciences), exhibits several striking immune phenotypes that have also been described in the context of *Dock2* deficiency including the loss of marginal zone B cells, decrease in plasmacytoid dendritic cells and invariant NKT cells. To our knowledge this allele has been introduced into four independently generated knockout mice (*Siae*$^{-/-}$, *Cmah*$^{-/-}$, *Irf5*$^{-/-}$ and *Nod2*$^{-/-}$) during backcrossing into the C57BL/6 strain. Given the widespread use of the C57BL/6NHsd substrain across the world, it is likely that this allele has been introduced into other gene targeted mice. As a result, published studies where C57BL/6NHsd mice have been used as controls, as experimental animals, or for backcrossing into the C57BL/6 background will need to be reinterpreted as this mutation may be alter or be responsible for phenotypes ascribed to other genes.

We built on this finding by identifying a 3-5 fold expansion of memory phenotype CD8^{+} T cells in DOCK2 deficient mice that correlates with increased

resistance to intracellular infection. Bone marrow chimera and adoptive transfer studies indicate that these cells arise in a cell intrinsic manner following thymic egress. In addition, close inspection of their transcriptional profile, TCR repertoire and cell surface marker expression shows that *Dock2*$^{-/-}$ naive CD8^{+} T cells directly convert into “virtual memory” cells bypassing the effector T cell stage. This phenomenon also occurs in the context of a restricted TCR repertoire as DOCK2 deficient T cells expressing the OT1 transgene show a similarly dramatic expansion of their memory compartment. Direct conversion to memory is associated with TCR hypersensitivity to ex vivo weak agonist peptide and in vivo self-peptide triggering. Collectively, these findings suggest that DOCK2 sets the threshold for entry into the “virtual” memory compartment by negatively regulating tonic TCR triggering.

Acknowledgements

The best decision I made in my professional life is joining Shiv Pillai's lab. I owe him for the contagious enthusiasm and great science that made the laboratory feel more like a playground rather than a place of employment.

The best thing that happened to me in the Pillai lab was being given a desk close to Vinay Mahajan. I owe him my PhD!

Bruce Horwitz gave me my first immunology textbook! I owe him for over a decade's worth of guidance and mentorship.

I owe Hamid Mattoo for all those times he patiently trained a clumsy PhD student terrified of mice.

I owe the "older young" Pillai lab members, Grace Yuen, Faisal Alsufyani and Yesim Tuncay, for being my best friends in lab and my sanity.

I owe Adrian and Brian for the years of cohabitation, fun and lifelong friendship. Importantly, I owe them for teaching me how to be a "cool" coastal elite.

I owe the "cats" and all members of Pillai lab for the countless entertaining lunches, pictures and scientific discussion.

I owe my dissertation and thesis committee for their time and insightful suggestions.

I owe my classmates, Jon and Cat in particular, for teaching me how to smile despite the failed experiments.

I owe friends at the Harvard Immunology Program, Sydney in particular, and the Ragon Institute for the years of guidance and friendship.

I owe my extended family in Boston for all the years of delicious food!

I owe my siblings, Aida and Yoftahe for being there for me no matter what.

I owe my parents, Abebe Demissie and Akaleselassie Mekuria, for sacrificing their adult lives for us. I owe them for constantly working and not taking a vacation in 30 years so they can afford to send us to school in the US.

The best thing that has happened to me in life is meeting Hemen Solomon Muleta. I owe her for giving me the most important thing in my life, a future we can build together. I owe her for giving me something to constantly strive for, one day being good enough for her!

Table of contents

Chapter 1: Introduction

1.1 History of mouse models

Although dogs are man's best friend, the mouse's ubiquitous geographic distribution has earned it the title of man's closest companion (Masopust, Sivula, and Jameson 2017). Our closest companion, however, doesn't inspire the same level of affection humanity has bestowed upon dogs. In fact, the term mouse originates from the ancient Sanskrit word Mush which means "to steal" (Vandenbergh 2000) This isn't surprising as the presence of mice has enormous negative economic impact on civilization to this day. For instance, the rice lost to rodents can feed over half a billion people (Stenseth et al. 2003).

In biomedical science, the mouse has gone some way to repaying its debt to humanity. The winding path it took to the modern laboratory bench started in China, where it was bred by mouse fanciers as early as 307 AD. Western European scientists like Robert Hooke (early 17th century) and Louis-Theodore Colladon(19th century) also sporadically used mice for researching biological systems. Interestingly, Gregor Mendel initiated his experiments on the segregation of mouse coat colors before switching to peas as his monastery's management balked at breeding such foul animals in his small room (Vandenbergh 2000).

One of the laboratory mouse's biggest breaks arrived at the turn of the 20th century when mice from Japan were imported by English and American mouse fanciers (Vandenbergh 2000; Gondo 2008; Yang et al. 2011). Abbie Lathrop, who first tried her luck in the poultry business, started with pair of waltzing mice before eventually housing over 11,000 mice. Her small Granby, MA farm initially limited its target market to people interested in keeping "fancy" mice as pets. After sometime, however, she noticed that the number of orders she was getting dramatically exceeded the demand of such a small market. To her surprise, the biggest buyers of mice from her facility were scientists interested in using her carefully bred mice for experiments. Abbie Lathrop's contribution to biomedical science goes beyond just breeding mice. She kept meticulous records and closely monitored her colony for any phenotypic irregularities. Her observation that some mice were getting lesions, which were later diagnosed by Leo Loeb as malignant, led to her being listed as an author on 10 publications in prominent journals with Loeb. In fact, a collaborative effort between Lathrop and Loeb was the first to show that different strains of mice have varying rates of tumorigenesis and removal of both ovaries reduces the instances of mammary tumors in mice (Steensma, Kyle, and Shampo 2010).

William Castle, the founding director of the Busey Institute in Boston, was one of many prominent scientists who took advantage of mice supplied by the Abbie Lathrop's farm. Inspired by the rediscovery of Mendelian

genetics and building on the work Lucien Cuenot, Castle studied the inheritance of coat color in mice. His work was the first to experimentally show that the departure from Mendelian ratios seen when mice heterozygous for a lethal allele are crossed to each other is because a quarter of the progeny, i.e homozgous for the lethal allele, die during embryonic development (Paigen 2003; Snell and Reed 1993).

In addition to his scientific contributions, William Castle is probably best known for the caliber of spectacular trainees that came from his lab. Clarence Little, a Massachusetts native, trained at the Castle lab before becoming the President of the University of Maine. At only 33, he was the youngest university president at the time. After leaving Maine and spending four controversial years as president of the University of Michigan, Little went on to use funds he raised from wealthy Detroit area families (Ford and Jackson families) to establish the Jackson Laboratory in Bar Harbor, Maine. Here, he generated several homozygous inbred strains of mice, including the widely used C57BL strain in 1921 (Crow 2002; Linder and Davisson 2012). His efforts were immensely valuable as genetically identical inbred mice greatly aided the conducting of controlled and reproducible experiments. CC little, together with Ernest Tyzzer, further went on to use such mice to show that tumor rejection is a trait inherited in a manner consistent with it being controlled by multiple dominant and recessive alleles (Masopust, Sivula, and Jameson 2017; Paigen 2003).

George Snell, a Nobel laureate from Castle's lab, made enormous contributions by performing the earliest experiments demonstrating MHC restriction with his studies on histocompatibility. His most powerful tools were congenic strains that only differ at one H2 locus. Snell generated such mice by crossing two inbred strains and then iteratively backcrossing them to their parents. During backcrossing, he astutely observed that there was a strong linkage between a mutation that visibly deformed the tail and altered its the appearance and the ability to reject a tumor. By following the segregation of this genetic marker as opposed to technically cumbersome and laborious tumor rejection, Snell was able to rapidly narrow down and discover the Histocompatibility (H) locus. (Klein 2001; Paigen 2003)

1.2 Modern mouse models

The aforementioned experiments clearly show how much value can be extracted from the humble mouse by biomedical researchers. However, any aspiring mouse biologist should temper this with a through understanding of the limitations of modeling any human condition in mice.

1.2.1 Limitations of modern mouse models

The fact that there are no shortage of therapies that work spectacularly well in inbred mice but fail in equally grand fashion in humans is proof that mouse models are far from perfect. Indeed, the mouse is very different from humans. For instance, the temperature in which mice are housed, 26°C, although comfortable to humans, is much lower than their nesting temperature of 32°C. Not surprisingly, chronic exposure to such environmental stress has far reaching consequences for resting mouse physiology, metabolism and immune competence (Karp 2012). Consistent with this, recent studies have shown dramatic differences between mouse and human transcriptional response to inflammatory insult (Seok et al. 2013; Godec et al. 2016).

Diseases are frequently induced and their progression accelerated so that mice can rapidly show symptoms that are superficially similar to those seen in humans. For instance, the mouse models for multiple sclerosis (MS) faithfully reproduce CNS immune cell infiltration and demyelination (McCarthy, Richards, and Miller 2012). However, it seems unlikely that simple peptide/protein plus adjuvant injections can truly represent the complex etiology underlying human MS. The generation of atherosclerosis requires an aggressive genetic deletion that is not seen in humans, raising questions about the translatability of findings in this model to humans. Attempts to model tumorigenesis by transplanting tumors also suffers from such reductionist

thinking that doesn't take into account the unpredictable route incipient tumors take in humans (Masopust, Sivula, and Jameson 2017; Karp 2012). Modeling human infectious disease is also very challenging as pathogens have exquisitely adapted to their hosts. For example, mouse models of Tuberculosis infection don't cause latency or organized granulomas, two cardinal features of human disease, with mycobacterial counts progressively increasing until the mouse succumbs to infection (Barry et al. 2009).

The genetic homogeneity of inbred strains makes any colony particularly vulnerable to dramatic spread of the same set of communicable diseases. Fear of such an an event, combined with a desire to make experimental studies reproducible across institutions, has pushed researchers to house mice in ever more controlled facilities away from environmental microbes. Commensal and pathogenic microbes that colonize mammals play vital in immune system development and homeostasis. For instance, the NOD mouse develops diabetes in the context of certain microbial flora (M. S. Anderson and Bluestone 2005) and a single microbial species (SFB) is sufficient for the development of Th17 cells in the gut (Ivanov et al. 2009). Thus, it is not unreasonable to argue that such isolation has negatively affected the mouse's ability to model the human immune system. In fact, Stephen Jameson and David Masopust elegantly compared pet store and feral mice with commonly used lab strains. Not surprisingly, their studies showed dramatic differences in innate and adaptive immune cell

compartments that was reminiscent of the difference between neonates and adult humans (Beura et al. 2016).

1.2.2 Advantages of modern mouse models

Despite the aforementioned limitations, the laboratory mouse has an impressive resume. It is the most widely used model organism, with millions of them powering research in an impressive range of disciplines Scientists primarily working with mouse models have won over 2 dozen Nobel prizes. Key insights regarding MHC restriction, X-inactivation and immune cell differentiation/development have been gleaned from the mouse.

Compared to other larger mammals, housing mice is relatively inexpensive. As a possible consequence of their evolution in closely knit family groups, they also aren't susceptible to inbreeding depression (loss of viability due to inbreeding) (Masopust, Sivula, and Jameson 2017). This prolific breeding capacity makes it possible to rapidly obtain a sizable colony from a single breeding pair in a short amount of time. Their relatively short lifespan, about 2 years, allows the study of disease progression throughout their entire lives.

Mouse models are valuable in clinical settings when they are judiciously used with an understanding that they don't fully recapitulate human disease. Researchers in these environments must be careful to tailor their model to the

specific functional questions they are asking. For example, mice are widely used to validate drug targets and determine safe drug doses and combinations prior to entering clinical trials. This is especially valuable in rare diseases where small patient populations make large scale testing in humans unfeasible (Vandamme 2014).

The mouse is extremely valuable because of its remarkable similarity to human metabolism and anatomy, although there are some key differences. This isn't surprising as the mouse and human genome share >90% of the same genes, diverging from a common ancestor about 65 million to 75 million years ago (Masopust, Sivula, and Jameson 2017). The importance of the mouse to the biomedical world is underlined by the fact that it was the first mammal to have its genome sequenced, making studies in this animal genetically tractable (Mouse Genome Sequencing Consortium et al. 2002). As will be discussed below, the presence of numerous well characterized inbred strains, informative spontaneous mutations and our ability to easily modify the its genome has made the mouse an invaluable addition to biologists' tool box.

1.3 Inbred strains of mice

Table 1: Common inbred strains of mice (adapted from (Lutz, Linder, and Davisson 2012))

Strain	**Key Features**
129P3/J, 129/S1/SvImJ	Gene targeting, spontaneous testicular tumors
A/J	Widely used in immuno-oncology, cleft palate
AKR/J	High leukemia incidence
BALB/c	Th2 skewed immune response (Watanabe)
C57BL/6J	Background strain for mice carrying mutations/transgenes, Th1 skewed immune response (Watanabe)
C57BL/6N	ES cells used for gene targeting by the International Knockout Mouse Consortium
DBA/1J	Model for rheumatoid arthritis and immune mediated nephritis
DBA/2J	Less prone to getting atherosclerotic lesions, model for glaucoma research
NOD NZW/LacJ	Type 1 diabetes
NZB/B1NJ	Lupus like disease
SWR	Aged mice have high incidence of tumors

The presence of numerous strains that are very different from each other gives researchers a lot of flexibility in choosing the ones that best fit their experimental goals (See table 1). Immunologists, in particular, have taken advantage of mice that get polygenic diseases, allowing mapping of

quantitative trait loci to identify candidate susceptibility genes (Ermann and Glimcher 2012; M. S. Anderson and Bluestone 2005; Yamanouchi et al. 2007))

1.4 Informative spontaneous single mutations

Mutations arising randomly in inbred mouse colonies have provided valuable insights into human disease. These changes are rarely complete loss of function, usually generating hypomorphic or hypermorphic alleles that more closely mimic human disease (Bruneau et al. 2001). The high degree of gene conservation between mice and humans makes murine mutations that have phenotypic manifestations particularly valuable. For example deafness in mice with the pirouette (Odeh et al. 2010) and hurry-scurry (Longo-Guess et al. 2005) mutations led to the discovery of causative genes in human hearing loss while retinal degeneration 3 (RD3) (Friedman et al. 2006) and RD12 (Pang et al. 2005) mouse strains have pointed to genes responsible for early childhood blindness. A more detailed review is provided in (Davisson et al. 2012).

Random mutations in the mouse genome have also initiated studies that have showed how complex biological systems work. The scurfy mouse, a spontaneous mutant that arose at the Oak Ridge National Laboratory, provided the first clues about Y chromosome sex determination. Since it was

the first X-linked mutation discovered (Brunkow et al. 2001; Chatila et al. 2000), researchers saw that most of the affected mice were males. Importantly, karyotyping of the very few affected female mice showed that they were all XO, carrying only 39 chromosomes. The finding that XO and XX mice were female strongly suggested that the Y chromosome was sex determining. This argument was further bolstered by the later studies showing that XXY mice were all male (L. B. Russell 2013). Building on the discovery that XO mice are females, Mary Lyon hypothesized that a single X was enough for survival and went on to use the mottled mouse mosaicism of coat color pattern to confirm random x-inactivation (Gartler 2015; Lyon 1961).

The Nude mouse, discovered by a Scottish farmer who was alarmed at the sight of a hairless rodent, carries a spontaneous mutation in the *Forkhead Box N1* gene affecting thymic epithelial cell development (G. Anderson and McCarthy 2015). This immunocompromised mouse was the first recipient of xenografted human tissue. Xenografts of patient derived tumors are especially valuable for testing novel drugs because they maintain original tumor characteristics such as architecture, vasculature, gene expression and malignancy (Tentler et al. 2012).

Random germline mutations are an extremely rare event. A recent long term study that observed mice over 20 generations has shown that the germline mutations come up at a an extremely rare frequency of 1 in 5×10^9 bp per generation (Uchimura et al. 2015). In addition, severely deleterious

alleles will be rapidly selected against. Thus, genetic engineering of mice is critical to unlocking their full potential in the laboratory.

1.5 Classical genetic manipulation of mice

Genetic manipulation of mice has a very long history. X-rays were first shown to cause heritable changes in the genome with irradiated mice. The word “transgenic” was first coined to describe a mouse generated by pronuclear injection of DNA into fertilized mouse eggs (Gordon and Ruddle 1981). This was quickly followed by gene targeting approaches that utilized homologous recombination to target a specific locus or gene (Thomas and Capecchi 1987).

The systematic generation of engineered mice on industrial scale that followed this landmark paper has ushered in an era during which gene function has been uncovered at an unprecedented rate. It has allowed the production of recombinant products, such as anticoagulants and therapeutic antibodies, that have revolutionized the management of numerous diseases. It has also allowed scientists to study the effect of single genes in the disease pathogenesis with relatively minimal effort (Vandamme 2014).

1.5.1 Forward genetics

Generating randomly mutated mice in an unbiased manner followed by attempts to identify genetic basis of phenotypes seen in these mice has been very lucrative. This was initially done in a systematic manner at the Oak Ridge National Laboratory, motivated by studies to understand the effects of radiation on mammalian organisms (L. B. Russell 2013). However, X-rays cause complex chromosomal rearrangements making attempts to identify single gene driven contributions challenging. The discovery of the DNA alkylating agent N-ethyl-N-nitrosourea (ENU), also at Oak Ridge, opened the door for high efficiency genome wide in vivo screens (W. L. Russell et al. 1979; Cordes 2005). ENU, a potent mutagen with germline mutagenesis rates of approximately in 1 in 10^6 bp, is particularly valuable as it results in single base pair substitutions (Concepcion et al. 2004).

Sequencing of the human genome and the identification of ≈20,000 genes required the use of high throughput models to test the function of these genes. As a result, large scale ENU mutagenesis projects were initiated 20 years after the initial discovery of this approach. Dominant phenotypes were identified by administering ENU to generation 0 (G0) males and carefully screening G1 progeny (de Angelis et al. 2000; Nolan et al. 2000). However, because ENU causes 1000X fold more recessive mutations than dominant

ones, some studies have laboriously used a three generation crossing strategy to generate homozygous mutant G3 mice (Gondo 2008).

Such ENU based screens were used to identify a mouse model for familial adenomatous polyposis, a precursor of colon cancer (Su et al. 1992). *Clock*, a gene that controlled circadian rhythm, was first discovered via an ENU screen (Vitaterna et al. 1994; King et al. 1997). Immunologists have also used ENU screens to identify numerous genes including ones involved in innate sensing (Jiang et al. 2006, 2005), maintaining tolerance to self (Vinuesa et al. 2005) and immune cell development (Miosge et al. 2002). (See (Beutler, Du, and Xia 2007; Nelms and Goodnow 2001) for more details).

1.5.2 Reverse Genetics

The mouse is very valuable as it allows complimenting of random mutagenesis with designed genetic modifications followed by a careful examination of the resulting phenotype. This approach is so high-yield that it has led to the establishment of industrial scale efforts to target every single protein coding gene, generating thousands of knockout and conditional alleles (Gondo 2008; Skarnes et al. 2011).

The first efforts to introduce foreign DNA into the mouse used retroviruses to infect mouse embryos. Although foreign genetic material introduced in such a manner was successfully integrated into the genome and

passed on to progeny, its expression was silenced (Jaenisch 1976; Jaenisch and Mintz 1974). A few years later, germline transmission and stable expression of foreign genetic material was first accomplished by pronuclear injection of DNA into a fertilized mouse egg (Gordon and Ruddle 1981). Since then, several viable alternative gene delivery methods such as lentiviruses and transposons have resulted in significant improvements in frequency of transgenic founders and greater throughput (reviewed in (Rülicke 2012)). An important point to keep in mind when working with mice generated through random transgene insertion is that any observed phenotype can be influenced by disruption and/or increased function of an endogenous gene as well as differences in transgene copy number. In addition, transgene insertion can also cause large deletions and complex rearrangements at sites of integration. For instance, we have observed that transgenic mice expressing Cre under the control of the Lck promoter (Tg(Lck-cre)1Cwi) have an expansion of memory phenotype CD8 cells that is not seen in mice where the Cre transgene is integrated at a different locus (unpublished data).

Gene targeting via homologous recombination allows precise targeting of specific loci and isn't plagued by the undesired effects of random naked DNA integration. Key in establishing homologous recombination as way to precisely edit mammalian DNA was optimization of techniques for isolating and long-term culture of embryonic stem (ES) cells without loss of pluripotency (Evans and Kaufman 1981; Martin 1981). This was quickly

followed by another landmark study showing that cells cultured in such a way can be mixed into a new host's blastocyst followed by implantation into pseudopregnant females to give rise to somatic and germ cells (Bradley et al. 1984). Building on these findings, two labs independently made gene-targeted mice using homologous vectors injected to cultured ES cells (Thomas and Capecchi 1987; Doetschman et al. 1987). As rates of insertion through homologous recombination are relatively low in post-fertilization embryos (Hagmann et al. 1996), correctly targeted cells need to be selected for expression of drug resistance genes and screened by PCR and southern blotting. Care should also be taken so as to not affect adjacent genes and only make the smallest possible genetic alteration

The mouse geneticist's toolbox was greatly expanded by the discovery of site specific recombinase systems, the Cre-LoxP system being the most popular (Gu et al. 1994). In this system, loxP sites are engineered so that they flank a genetic region of interest. This allows the intervening region to either be excised or inverted by Cre recombinase, a 38 kDa enzyme from bacteriophage P1. Importantly, having carefully chosen promoters/control elements drive Cre expression allows conditional modification of the intervening region in a timely or spatially controlled manner (Bouabe and Okkenhaug 2013). This system is so powerful that more 500 different Cre transgenic mice have been generated.

This approaches for making genetically engineered mice have been widely used over the past 3 decades. Recently, however, newer methods of mutant generation have come online that have greatly reduced the cost and time of making such mice (see discussion, Chapter 4).

1.5.3 Backcrossing gene engineered mice

Until recently, most ES cells used for gene targeting were of 129 Origin because of relative ease of culturing them without loss of pluripotency and high efficiency germline transmission of mutant alleles (Rülicke 2012; van der Weyden et al. 2011; Kawase et al. 1994; Seong et al. 2004). Mutant 129 ES clones were usually combined with C57BL/6 blastocysts and implanted into pseudopregnant females. This results in chimeric mice, with contributions from ES cells of both strains. The chimeric mice with germline transmission of the mutant allele are then selected and backcrossed into the required background strain.

The simplest thing to do at this point would be to maintain the mutated allele in the 129 background. However, 129 mice aren't widely used in the lab because they don't breed well and have anatomical and behavioral problems.(Rivera and Tessarollo 2008). In addition, adoptive transfers into C57BL/6 mice, the most widely used strain in immunology, would not be possible because of immunological rejection. As a result, it is common

practice to backcross gene targeted and/or transgenic mice into the C57BL/6 background with each generation halving the genetic contribution of the parental ES 129 cell line, beyond the loci linked to mutant/selected allele. Consequently, 10 generations of backcrossing into the C57BL/6 background results in congenic mutant mice that are over 99.9% C57BL/6.

The dominance of C57BL/6 as the background mouse of choice for gene targeting has raised questions as to whether a single strain can accurately depict the breadth of genetic variation seen in human populations (Rivera and Tessarollo 2008). As will be further discussed below, the background context in which a mutant allele exists also has major consequences on the observed phenotype.

1.5.4 Effect of background

There are a lot of genetic differences underlying the phenotypic heterogeneity observed in various inbred strains. Thus, it isn't surprising that some genetically engineered mice show discordant phenotype:genotype pairings with genetic modifiers present in the background acting either independently or in epistasis with the mutant allele. For instance, flanking genes from parental ES cell line that are retained during backcrosses sometimes make interpretation of data from gene targeted mice challenging (Eisener-Dorman, Lawrence, and Bolivar 2009) Valor et al. elegantly

demonstrated this by using microarray data from several mutant lines to show that the genes with the most significant changes in expression are on the same chromosome as the knocked out genes (Valor and Grant 2007).

Background modifiers unlinked to mutant allele can also result in differences among distinct strains carrying the same mutation. The p53 knockout allele on the C57BL/6 background shows much lower rates of tumor incidence compared to the 129 background (Harvey et al. 1993). CD8α dendritic cell development, while defective in Batf3$^{-/-}$ 129 mice, is restored in knockout C57BL/6 mice because of a compensation by the related AP1 factor, Batf. (Tussiwand et al. 2012)

Even more dramatic strain-dependent phenotypes are seen in TGF-β and EGFR deficient mice. TGF-β loss is embryonic lethal after only one generation of backcrossing into the C57BL/6 background while BALB/c and mixed 129 X CF1 background TGF-β knockout mice phenocopy this 50% and 20% of the time, respectively (Kallapur, Ormsby, and Doetschman 1999). Similarly, loss of EGFR in 129/Sv background results in death mid-gestation while the same allele in a 129/Sv x C57BL/6 background results in some mice surviving until birth. Interestingly, such mice on a 129/Sv x C57BL/6 x MF1 background survive upto 20 days after birth (Sibilia and Wagner 1995; Threadgill et al. 1995).

Background dependent phenotypes can certainly be a concern for any biologist attempting to rapidly reap the rewards of publishing a novel mutant

allele. However, they are potentially very informative as they are reflective of what happens with most human mutations in diverse populations. Thus, attempts to discern the underlying basis of such discordance should not be discouraged.

1.5.5 Differences between C57BL strains

As discussed earlier, the C57BL strain was established by C.C. Little in the 1920s. In 1937, the C57BL strain was further split into two substrains C57BL/10 and, the currently widely used, C57BL/6. This substrain was maintained in the Jackson labs, hence the name C57BL/6J. The next fork in the history of this strain came when C57BL/6J mice were transferred to the NIH and maintained as separate colony (C57BL/6N) in 1951. Mice from this NIH colony were then transferred to several commercial vendors and independently maintained as C57BL/6NCrl, C57BL/6NHsd, C57BL/6NJ and C57BL/6NTac. (Song and Hwang 2017)

Although new mutations are rare and only a quarter of them are fixed by continuous sibling mating, the independent maintenance of C57BL/6 substrains has resulted substantial genetic drift (Zurita et al. 2011) (Mouse Phenome Database). Interestingly, such genetic drift has very important phenotypic implications. For example, C57BL substrains give differing results when tested in mouse models of autism and other neuropsychiatric disorders

(Lutz, Linder, and Davisson 2012). All C57BL/6J mice carry an arginine to leucine change in the leucine-rich repeat domain of *NLR Family Pyrin Domain Containing 12* which results in defective neutrophil recruitment to sites of inflammation. (Ulland et al. 2016) Most C57BL/6J mice also have in-frame deletion in *Nicotinamide Nucleotide Transhydrogenase,* which results in glucose intolerance in this substrain (Freeman et al. 2006).

C57BL/6N mice have also accumulated impactful mutations. The vision research community just recently became aware that this strain, widely used for crosses and gene targeting, has a single nucleotide deletion in the *Crumbs 1, Cell Polarity Complex Component* gene resulting in retinal degeneration (Mattapallil et al. 2012). C57BL6/N mice also have reduced responses to psychostimulant injection because of a single nucleotide nonsynonymous changed in the *Cytoplasmic FMR1 Interacting Protein 2* gene (Kumar et al. 2013). As part of this thesis, we have also identified a mutation that is restricted to mice from one commercial vendor of C57BL6N mice, Envigo Biosciences (previously Harlan Sparague Dawley) (Mahajan et al. 2016). This allele, a duplication around exon 28 and 29 of Dock2 in C57BL/6NHsd mice, has been backcrossed into multiple gene targeted mice over the past decades and results in dramatic immune phenotypes (see chapter 2).

1.6 Dedicator of Cytokinesis 2 (DOCK2)

DOCK2 activates the Rho GTPase Rac by allowing the transition from the inactive GDP-bound state to the active GTP-bound state. Published work has shown that DOCK2 permits close spatial control of Rac activation by localizing to cell membrane through interactions with PIP3 and phosphatidic acid (Fukui et al. 2001; Nishikimi et al. 2009, 2013). GTP bound RAC1 subsequently promotes actin polymerization aiding cytoskeletal rearrangements required for lymphocyte movement (Fukui et al. 2001; Terasawa et al. 2012; Gotoh et al. 2008), T cell interstitial motility (Nombela-Arrieta et al. 2007), plasmacytoid dendritic cell cytokine secretion (Gotoh et al. 2010), and TCR activation (Sanui et al. 2003; Le Floc'h et al. 2013). DOCK2's role in CD8+ TCR activation has been studied using relatively strong stimuli, i.e. anti-CD3 or cognate antigen. These studies have concluded that DOCK2 as a negative regulator of TCR clustering and immune synapse formation. A single study. used a weak agonist against the MHC-II restricted 2B4 TCR transgene to show a defect in the proliferative response of stimulated Dock2-/- CD4+ T-cells (Sanui et al. 2003). However, it is likely that DOCK2's role in regulating lymphocyte proliferation independent of the TCR complicates interpretation of this experiment (Wang et al. 2010). In contrast, work from our lab suggests that DOCK2 sets the threshold for TCR activation (see chapter 3).

References

Anderson, Graham, and Nicholas I. McCarthy. 2015. “Laying Bare the Nude Mouse Gene.” Journal of Immunology 194 (3): 847–48.

Anderson, Mark S., and Jeffrey A. Bluestone. 2005. “The NOD Mouse: A Model of Immune Dysregulation.” Annual Review of Immunology 23: 447–85.

Angelis, Martin Hrabé de, Heinrich Flaswinkel, Helmut Fuchs, Birgit Rathkolb, Dian Soewarto, Susan Marschall, Stephan Heffner, et al. 2000. “Genome-Wide, Large-Scale Production of Mutant Mice by ENU Mutagenesis.” Nature Genetics 25 (August). Nature America Inc.: 444.

Barry, Clifton E., 3rd, Helena I. Boshoff, Véronique Dartois, Thomas Dick, Sabine Ehrt, Joanne Flynn, Dirk Schnappinger, Robert J. Wilkinson, and Douglas Young. 2009. “The Spectrum of Latent Tuberculosis: Rethinking the Biology and Intervention Strategies.” Nature Reviews. Microbiology 7 (12): 845–55.

Beura, Lalit K., Sara E. Hamilton, Kevin Bi, Jason M. Schenkel, Oludare A. Odumade, Kerry A. Casey, Emily A. Thompson, et al. 2016. “Normalizing the Environment Recapitulates Adult Human Immune Traits in Laboratory Mice.” Nature 532 (7600): 512–16.

Beutler, Bruce, Xin Du, and Yu Xia. 2007. “Precis on Forward Genetics in Mice.” Nature Immunology 8 (July). Nature Publishing Group: 659.

Bouabe, Hicham, and Klaus Okkenhaug. 2013. “Gene Targeting in Mice: A Review.” Methods in Molecular Biology 1064: 315–36.

Bradley, Allan, Martin Evans, Matthew H. Kaufman, and Elizabeth Robertson. 1984. “Formation of Germ-Line Chimaeras from Embryo-Derived Teratocarcinoma Cell Lines.” Nature 309 (May). Nature Publishing Group: 255.

Bruneau, S., K. R. Johnson, M. Yamamoto, A. Kuroiwa, and D. Duboule. 2001. “The Mouse Hoxd13(spdh) Mutation, a Polyalanine Expansion Similar to Human Type II Synpolydactyly (SPD), Disrupts the Function but Not the Expression of Other Hoxd Genes.” Developmental Biology 237 (2): 345–53.

Brunkow, Mary E., Eric W. Jeffery, Kathryn A. Hjerrild, Bryan Paeper, Lisa B. Clark, Sue-Ann Yasayko, J. Erby Wilkinson, David Galas, Steven F.

Ziegler, and Fred Ramsdell. 2001. “Disruption of a New Forkhead/winged-Helix Protein, Scurfin, Results in the Fatal Lymphoproliferative Disorder of the Scurfy Mouse.” Nature Genetics 27 (January). Nature America Inc.: 68.

Chatila, T. A., F. Blaeser, N. Ho, H. M. Lederman, C. Voulgaropoulos, C. Helms, and A. M. Bowcock. 2000. “JM2, Encoding a Fork Head-Related Protein, Is Mutated in X-Linked Autoimmunity-Allergic Disregulation Syndrome.” The Journal of Clinical Investigation 106 (12): R75–81.

Concepcion, Dorothy, Kevin L. Seburn, Gen Wen, Wayne N. Frankel, and Bruce A. Hamilton. 2004. “Mutation Rate and Predicted Phenotypic Target Sizes in Ethylnitrosourea-Treated Mice.” Genetics 168 (2): 953–59.

Cordes, Sabine P. 2005. “N-Ethyl-N-Nitrosourea Mutagenesis: Boarding the Mouse Mutant Express.” Microbiology and Molecular Biology Reviews: MMBR 69 (3): 426–39.

Crow, James F. 2002. “C. C. Little, Cancer and Inbred Mice.” Genetics 161 (4): 1357–61.

Davisson, Muriel T., David E. Bergstrom, Laura G. Reinholdt, and Leah Rae Donahue. 2012. “Discovery Genetics - The History and Future of Spontaneous Mutation Research.” Current Protocols in Mouse Biology 2 (June): 103–18.

Doetschman, Thomas, Ronald G. Gregg, Nobuyo Maeda, Martin L. Hooper, David W. Melton, Simon Thompson, and Oliver Smithies. 1987. “Targetted Correction of a Mutant HPRT Gene in Mouse Embryonic Stem Cells.” Nature 330 (December). Nature Publishing Group: 576.

Eisener-Dorman, Amy F., David A. Lawrence, and Valerie J. Bolivar. 2009. “Cautionary Insights on Knockout Mouse Studies: The Gene or Not the Gene?” Brain, Behavior, and Immunity 23 (3): 318–24.

Ermann, Joerg, and Laurie H. Glimcher. 2012. “After GWAS: Mice to the Rescue?” Current Opinion in Immunology 24 (5). Elsevier: 564–70.

Evans, M. J., and M. H. Kaufman. 1981. “Establishment in Culture of Pluripotential Cells from Mouse Embryos.” Nature 292 (5819): 154–56.

Freeman, Helen C., Alison Hugill, Neil T. Dear, Frances M. Ashcroft, and Roger D. Cox. 2006. “Deletion of Nicotinamide Nucleotide Transhydrogenase: A New Quantitive Trait Locus Accounting for Glucose Intolerance in C57BL/6J Mice.” Diabetes 55 (7): 2153–56.

Friedman, James S., Bo Chang, Chitra Kannabiran, Christina Chakarova,

Hardeep P. Singh, Subhadra Jalali, Norman L. Hawes, et al. 2006. “Premature Truncation of a Novel Protein, RD3, Exhibiting Subnuclear Localization Is Associated with Retinal Degeneration.” American Journal of Human Genetics 79 (6): 1059–70.

Gartler, Stanley M. 2015. “Mary Lyon’s X-Inactivation Studies in the Mouse Laid the Foundation for the Field of Mammalian Dosage Compensation.” Journal of Genetics 94 (4). Springer India: 563–65.

Godec, Jernej, Yan Tan, Arthur Liberzon, Pablo Tamayo, Sanchita Bhattacharya, Atul J. Butte, Jill P. Mesirov, and W. Nicholas Haining. 2016. “Compendium of Immune Signatures Identifies Conserved and Species-Specific Biology in Response to Inflammation.” Immunity 44 (1): 194–206.

Gondo, Yoichi. 2008. “Trends in Large-Scale Mouse Mutagenesis: From Genetics to Functional Genomics.” Nature Reviews. Genetics 9 (10): 803–10.

Gordon, J. W., and F. H. Ruddle. 1981. “Integration and Stable Germ Line Transmission of Genes Injected into Mouse Pronuclei.” Science 214 (4526): 1244–46.

Gu, H., J. D. Marth, P. C. Orban, H. Mossmann, and K. Rajewsky. 1994. “Deletion of a DNA Polymerase Beta Gene Segment in T Cells Using Cell Type-Specific Gene Targeting.” Science 265 (5168): 103–6.

Hagmann, M., K. Adlkofer, P. Pfeiffer, R. Bruggmann, O. Georgiev, D. Rungger, and W. Schaffner. 1996. “Dramatic Changes in the Ratio of Homologous Recombination to Nonhomologous DNA-End Joining in Oocytes and Early Embryos of Xenopus Laevis.” Biological Chemistry Hoppe-Seyler 377 (4): 239–50.

Harvey, M., M. J. McArthur, C. A. Montgomery Jr, A. Bradley, and L. A. Donehower. 1993. “Genetic Background Alters the Spectrum of Tumors That Develop in p53-Deficient Mice.” FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology 7 (10): 938–43.

Ivanov, Ivaylo I., Koji Atarashi, Nicolas Manel, Eoin L. Brodie, Tatsuichiro Shima, Ulas Karaoz, Dongguang Wei, et al. 2009. “Induction of Intestinal Th17 Cells by Segmented Filamentous Bacteria.” Cell 139 (3): 485–98.

Jaenisch, R. 1976. “Germ Line Integration and Mendelian Transmission of the Exogenous Moloney Leukemia Virus.” Proceedings of the National Academy of Sciences of the United States of America 73 (4): 1260–64.

Jaenisch, R., and B. Mintz. 1974. “Simian Virus 40 DNA Sequences in DNA of Healthy Adult Mice Derived from Preimplantation Blastocysts Injected with Viral DNA.” Proceedings of the National Academy of Sciences of the United States of America 71 (4): 1250–54.

Jiang, Zhengfan, Philippe Georgel, Xin Du, Louis Shamel, Sosathya Sovath, Suzanne Mudd, Michael Huber, et al. 2005. “CD14 Is Required for MyD88-Independent LPS Signaling.” Nature Immunology 6 (6): 565–70.

Jiang, Zhengfan, Philippe Georgel, Chenglong Li, Jungwoo Choe, Karine Crozat, Sophie Rutschmann, Xin Du, et al. 2006. “Details of Toll-like Receptor:adapter Interaction Revealed by Germ-Line Mutagenesis.” Proceedings of the National Academy of Sciences of the United States of America 103 (29): 10961–66.

Kallapur, S., I. Ormsby, and T. Doetschman. 1999. “Strain Dependency of TGFbeta1 Function during Embryogenesis.” Molecular Reproduction and Development 52 (4): 341–49.

Karp, Christopher L. 2012. “Unstressing Intemperate Models: How Cold Stress Undermines Mouse Modeling.” The Journal of Experimental Medicine 209 (6): 1069–74.

Kawase, E., H. Suemori, N. Takahashi, K. Okazaki, K. Hashimoto, and N. Nakatsuji. 1994. “Strain Difference in Establishment of Mouse Embryonic Stem (ES) Cell Lines.” The International Journal of Developmental Biology 38 (2): 385–90.

King, D. P., Y. Zhao, A. M. Sangoram, L. D. Wilsbacher, M. Tanaka, M. P. Antoch, T. D. Steeves, et al. 1997. “Positional Cloning of the Mouse Circadian Clock Gene.” Cell 89 (4): 641–53.

Klein, Jan. 2001. “George Snell’s First Foray Into the Unexplored Territory of the Major Histocompatibility Complex.” Genetics 159 (2). Genetics: 435–39.

Kumar, Vivek, Kyungin Kim, Chryshanthi Joseph, Saïd Kourrich, Seung-Hee Yoo, Hung Chung Huang, Martha H. Vitaterna, et al. 2013. “C57BL/6N Mutation in Cytoplasmic FMRP Interacting Protein 2 Regulates Cocaine Response.” Science 342 (6165): 1508–12.

Linder, Carol C., and Muriel T. Davisson. 2012. “Chapter 1.2 - Historical Foundations.” In The Laboratory Mouse (Second Edition), edited by Hans J. Hedrich, 21–35. Boston: Academic Press.

Longo-Guess, Chantal M., Leona H. Gagnon, Susan A. Cook, Jian Wu, Qing Y. Zheng, and Kenneth R. Johnson. 2005. “A Missense Mutation in the

Previously Undescribed Gene Tmhs Underlies Deafness in Hurry-Scurry (hscy) Mice." Proceedings of the National Academy of Sciences of the United States of America 102 (22): 7894–99.

Lutz, Cathleen M., Carol C. Linder, and Muriel T. Davisson. 2012. "Chapter 1.3 - Strains, Stocks and Mutant Mice." In The Laboratory Mouse (Second Edition), edited by Hans J. Hedrich, 37–56. Boston: Academic Press.

Lyon, Mary F. 1961. "Gene Action in the X-Chromosome of the Mouse (Mus Musculus L.)." Nature 190 (April). Nature Publishing Group: 372.

Mahajan, Vinay S., Ezana Demissie, Hamid Mattoo, Vinay Viswanadham, Ajit Varki, Robert Morris, and Shiv Pillai. 2016. "Striking Immune Phenotypes in Gene-Targeted Mice Are Driven by a Copy-Number Variant Originating from a Commercially Available C57BL/6 Strain." Cell Reports 15 (9): 1901–9.

Martin, G. R. 1981. "Isolation of a Pluripotent Cell Line from Early Mouse Embryos Cultured in Medium Conditioned by Teratocarcinoma Stem Cells." Proceedings of the National Academy of Sciences of the United States of America 78 (12): 7634–38.

Masopust, David, Christine P. Sivula, and Stephen C. Jameson. 2017. "Of Mice, Dirty Mice, and Men: Using Mice To Understand Human Immunology." Journal of Immunology 199 (2): 383–88.

Mattapallil, Mary J., Eric F. Wawrousek, Chi-Chao Chan, Hui Zhao, Jayeeta Roychoudhury, Thomas A. Ferguson, and Rachel R. Caspi. 2012. "The Rd8 Mutation of the Crb1 Gene Is Present in Vendor Lines of C57BL/6N Mice and Embryonic Stem Cells, and Confounds Ocular Induced Mutant Phenotypes." Investigative Ophthalmology & Visual Science 53 (6): 2921–27.

McCarthy, Derrick P., Maureen H. Richards, and Stephen D. Miller. 2012. "Mouse Models of Multiple Sclerosis: Experimental Autoimmune Encephalomyelitis and Theiler's Virus-Induced Demyelinating Disease." Methods in Molecular Biology 900: 381–401.

Miosge, Lisa A., Julie Blasioli, Mathieu Blery, and Christopher C. Goodnow. 2002. "Analysis of an Ethylnitrosourea-Generated Mouse Mutation Defines a Cell Intrinsic Role of Nuclear Factor kappaB2 in Regulating Circulating B Cell Numbers." The Journal of Experimental Medicine 196 (8): 1113–19.

Mouse Genome Sequencing Consortium, Robert H. Waterston, Kerstin Lindblad-Toh, Ewan Birney, Jane Rogers, Josep F. Abril, Pankaj Agarwal,

et al. 2002. “Initial Sequencing and Comparative Analysis of the Mouse Genome.” Nature 420 (6915): 520–62.

Nelms, K. A., and C. C. Goodnow. 2001. “Genome-Wide ENU Mutagenesis to Reveal Immune Regulators.” Immunity 15 (3): 409–18.

Nolan, P. M., J. Peters, M. Strivens, D. Rogers, J. Hagan, N. Spurr, I. C. Gray, et al. 2000. “A Systematic, Genome-Wide, Phenotype-Driven Mutagenesis Programme for Gene Function Studies in the Mouse.” Nature Genetics 25 (4): 440–43.

Odeh, Hana, Kristina L. Hunker, Inna A. Belyantseva, Hela Azaiez, Matthew R. Avenarius, Lili Zheng, Linda M. Peters, et al. 2010. “Mutations in Grxcr1 Are the Basis for Inner Ear Dysfunction in the Pirouette Mouse.” American Journal of Human Genetics 86 (2): 148–60.

Paigen, Kenneth. 2003. “One Hundred Years of Mouse Genetics: An Intellectual History. I. The Classical Period (1902-1980).” Genetics 163 (1): 1–7.

Pang, Ji-Jing, Bo Chang, Norman L. Hawes, Ronald E. Hurd, Muriel T. Davisson, Jie Li, Syed M. Noorwez, et al. 2005. “Retinal Degeneration 12 (rd12): A New, Spontaneously Arising Mouse Model for Human Leber Congenital Amaurosis (LCA).” Molecular Vision 11 (February): 152–62.

Rivera, Juan, and Lino Tessarollo. 2008. “Genetic Background and the Dilemma of Translating Mouse Studies to Humans.” Immunity 28 (1): 1–4.

Rülicke, Thomas. 2012. “Generation of Mouse Mutants by Genotype-Driven Mutagenesis.” In The Laboratory Mouse, 91–114. Elsevier.

Russell, Liane B. 2013. “The Mouse House: A Brief History of the ORNL Mouse-Genetics Program, 1947-2009.” Mutation Research 753 (2): 69–90.

Russell, W. L., E. M. Kelly, P. R. Hunsicker, J. W. Bangham, S. C. Maddux, and E. L. Phipps. 1979. “Specific-Locus Test Shows Ethylnitrosourea to Be the Most Potent Mutagen in the Mouse.” Proceedings of the National Academy of Sciences of the United States of America 76 (11): 5818–19.

Seok, Junhee, H. Shaw Warren, Alex G. Cuenca, Michael N. Mindrinos, Henry V. Baker, Weihong Xu, Daniel R. Richards, et al. 2013. “Genomic Responses in Mouse Models Poorly Mimic Human Inflammatory Diseases.” Proceedings of the National Academy of Sciences of the United States of America 110 (9): 3507–12.

Seong, Eunju, Thomas L. Saunders, Colin L. Stewart, and Margit Burmeister.

2004. “To Knockout in 129 or in C57BL/6: That Is the Question.” Trends in Genetics: TIG 20 (2): 59–62.

Sibilia, M., and E. F. Wagner. 1995. “Strain-Dependent Epithelial Defects in Mice Lacking the EGF Receptor.” Science 269 (5221): 234–38.

Skarnes, William C., Barry Rosen, Anthony P. West, Manousos Koutsourakis, Wendy Bushell, Vivek Iyer, Alejandro O. Mujica, et al. 2011. “A Conditional Knockout Resource for the Genome-Wide Study of Mouse Gene Function.” Nature 474 (7351): 337–42.

Snell, G. D., and S. Reed. 1993. “William Ernest Castle, Pioneer Mammalian Geneticist.” Genetics 133 (4). Genetics: 751–53.

Song, Hyun Keun, and Dae Youn Hwang. 2017. “Use of C57BL/6N Mice on the Variety of Immunological Researches.” Laboratory Animal Research 33 (2): 119–23.

Steensma, David P., Robert A. Kyle, and Marc A. Shampo. 2010. “Abbie Lathrop, the ‘Mouse Woman of Granby’: Rodent Fancier and Accidental Genetics Pioneer.” Mayo Clinic Proceedings. Mayo Clinic 85 (11): e83.

Stenseth, Nils Chr, Herwig Leirs, Anders Skonhoft, Stephen A. Davis, Roger P. Pech, Harry P. Andreassen, Grant R. Singleton, et al. 2003. “Mice, Rats, and People: The Bio-Economics of Agricultural Rodent Pests.” Frontiers in Ecology and the Environment 1 (7). Ecological Society of America: 367–75.

Su, L. K., K. W. Kinzler, B. Vogelstein, A. C. Preisinger, A. R. Moser, C. Luongo, K. A. Gould, and W. F. Dove. 1992. “Multiple Intestinal Neoplasia Caused by a Mutation in the Murine Homolog of the APC Gene.” Science 256 (5057): 668–70.

Tentler, John J., Aik Choon Tan, Colin D. Weekes, Antonio Jimeno, Stephen Leong, Todd M. Pitts, John J. Arcaroli, Wells A. Messersmith, and S. Gail Eckhardt. 2012. “Patient-Derived Tumour Xenografts as Models for Oncology Drug Development.” Nature Reviews. Clinical Oncology 9 (6): 338–50.

Thomas, K. R., and M. R. Capecchi. 1987. “Site-Directed Mutagenesis by Gene Targeting in Mouse Embryo-Derived Stem Cells.” Cell 51 (3): 503–12.

Threadgill, D. W., A. A. Dlugosz, L. A. Hansen, T. Tennenbaum, U. Lichti, D. Yee, C. LaMantia, T. Mourton, K. Herrup, and R. C. Harris. 1995. “Targeted Disruption of Mouse EGF Receptor: Effect of Genetic

Background on Mutant Phenotype.” Science 269 (5221): 230–34.

Tussiwand, Roxane, Wan-Ling Lee, Theresa L. Murphy, Mona Mashayekhi, Wumesh Kc, Jörn C. Albring, Ansuman T. Satpathy, et al. 2012. “Compensatory Dendritic Cell Development Mediated by BATF-IRF Interactions.” Nature 490 (7421): 502–7.

Uchimura, Arikuni, Mayumi Higuchi, Yohei Minakuchi, Mizuki Ohno, Atsushi Toyoda, Asao Fujiyama, Ikuo Miura, Shigeharu Wakana, Jo Nishino, and Takeshi Yagi. 2015. “Germline Mutation Rates and the Long-Term Phenotypic Effects of Mutation Accumulation in Wild-Type Laboratory Mice and Mutator Mice.” Genome Research 25 (8): 1125–34.

Ulland, Tyler K., Nidhi Jain, Emma E. Hornick, Eric I. Elliott, Gwendolyn M. Clay, Jeffrey J. Sadler, Kathleen A. M. Mills, et al. 2016. “Nlrp12 Mutation Causes C57BL/6J Strain-Specific Defect in Neutrophil Recruitment.” Nature Communications 7 (October): 13180.

Valor, Luis M., and Seth G. N. Grant. 2007. “Clustered Gene Expression Changes Flank Targeted Gene Loci in Knockout Mice.” PloS One 2 (12): e1303.

Vandamme, Thierry F. 2014. “Use of Rodents as Models of Human Diseases.” Journal of Pharmacy & Bioallied Sciences 6 (1): 2–9.

Vandenbergh, John G. 2000. “Use of House Mice in Biomedical Research.” ILAR Journal / National Research Council, Institute of Laboratory Animal Resources 41 (3). Oxford University Press: 133–35.

Vinuesa, Carola G., Matthew C. Cook, Constanza Angelucci, Vicki Athanasopoulos, Lixin Rui, Kim M. Hill, Di Yu, et al. 2005. “A RING-Type Ubiquitin Ligase Family Member Required to Repress Follicular Helper T Cells and Autoimmunity.” Nature 435 (May). Macmillian Magazines Ltd.: 452.

Vitaterna, M. H., D. P. King, A. M. Chang, J. M. Kornhauser, P. L. Lowrey, J. D. McDonald, W. F. Dove, L. H. Pinto, F. W. Turek, and J. S. Takahashi. 1994. “Mutagenesis and Mapping of a Mouse Gene, Clock, Essential for Circadian Behavior.” Science 264 (5159): 719–25.

Weyden, Louise van der, Jacqueline K. White, David J. Adams, and Darren W. Logan. 2011. “The Mouse Genetics Toolkit: Revealing Function and Mechanism.” Genome Biology 12 (6): 224.

Yamanouchi, Jun, Dan Rainbow, Pau Serra, Sarah Howlett, Kara Hunter, Valerie E. S. Garner, Andrea Gonzalez-Munoz, et al. 2007. “Interleukin-2 Gene Variation Impairs Regulatory T Cell Function and Causes

Autoimmunity.” Nature Genetics 39 (February). Nature Publishing Group: 329.

Yang, Hyuna, Jeremy R. Wang, John P. Didion, Ryan J. Buus, Timothy A. Bell, Catherine E. Welsh, François Bonhomme, et al. 2011. “Subspecific Origin and Haplotype Diversity in the Laboratory Mouse.” Nature Genetics 43 (7): 648–55.

Zurita, Esther, Mónica Chagoyen, Marta Cantero, Rosario Alonso, Anna González-Neira, Alejandro López-Jiménez, José Antonio López-Moreno, et al. 2011. “Genetic Polymorphisms among C57BL/6 Mouse Inbred Strains.” Transgenic Research 20 (3). Springer: 481–89.

Chapter 2: Landing on DOCK2

Whole genome sequencing has revealed the presence of a homozygous copy number variant that disrupts the function of Dock2 in a specific commercially available C57BL/6 mouse substrain (C57BL/6NHsd). This *Dock2* variant has been inadvertently introduced into multiple gene-targeted mice in a homozygous state when they were backcrossed into this particular C57BL/6 background. This mutant *Dock2* allele has also been previously reported in some sub-lines of *Irf5* knockout mice and was presumed to have spontaneously arisen in the original colony of knockout mice. However, we now report that this mutant *Dock2* allele has been inadvertently introduced into at least two other independently generated mutant mouse lines, engineered to be deficient in Siae and Cmah that were also backcrossed using this C57BL/6 substrain. Given the widespread use of this C57BL/6 mouse line across the world, this mutation may be responsible for phenotypes assumed to be caused by the deletion of a number of specific targeted genes. This particular strain of C57BL/6 mice from Harlan-Sprague (acquired by Envigo Biosciences) exhibits several striking immune phenotypes that have also been described in the context of Dock2 deficiency including the loss of marginal zone B cells, decreased plasmacytoid dendritic cells and a decrease in invariant NKT cells. Contrary to what has been previously reported, *Siae* and *Cmah* mutant mice that lack the *Dock2* mutation exhibit normal B cell development and do not lack marginal zone B cells. Published studies with hematopoietic

phenotypes where these commercially available C57BL/6 mice have been used as controls, as experimental animals, or for backcrossing into the C57BL/6 background will need to be reinterpreted in light of these findings.

2.1 INTRODUCTION

Over the last three decades, gene targeting technology has become a powerful tool for functional analysis of immune genes in vivo. It has become a common practice to backcross gene-targeted mice for ~10 generations into the C57BL/6 background to facilitate comparisons between gene-targeted mice as well as for adoptive transfer experiments. However, numerous C57BL/6 sublines are in use across the world (Zurita et al. 2011), and the potential effect of variability among these C57BL/6 sublines on immune phenotypes is often not considered.

We had previously described defects in B cell development in mice with altered sialic acid physiology(Cariappa et al. 2009). Mice with a germline loss of either *Siae* (sialic acid acetyl esterase) or *Cmah* (cytidine monophosphate-N-acetylneuraminic acid hydroxylase) were found to lack marginal zone B cells and exhibited hyperactive B cell receptor signaling (Cariappa et al. 2009). Given that these mice generate presumably altered forms of sialic acid that are not recognized by key regulatory Siglecs expressed on B cells (such as CD22/Siglec-2 and Siglec-G), the defects in B cell development observed in these mice were presumed to arise from perturbations in Siglec function (Pillai, Cariappa, and Pirnie 2009; Cariappa et al. 2009). In addition, the observed phenotypes were largely compatible with previous studies of Siglec function (Pillai, Cariappa, and Pirnie 2009; Mahajan and Pillai 2016). Notably, both *Siae*$^{\Delta ex2/\Delta ex2}$ and *Cmah* knockout mice were

originally generated by collaborators at the University of California at San Diego and had been backcrossed into a specific commercially-obtained C57BL/6 background for ten generations (Hedlund et al. 2007; Cariappa et al. 2009). We found that Siae deficient mice unexpectedly lost their aberrant B cell development phenotype upon backcrossing for 13 additional generations into the C57BL/6J (Jackson Laboratories) background. We created an independent knockout line of Siae deficient mice in the C57BL/6N background, and these mice exhibited no defects in B cell development.

Given these discrepant results, we re-examined the genetic basis of aberrant B cell development in *Siae* $^{\Delta ex2/\Delta ex2}$ mice using genetic crosses, SNP arrays and whole genome sequencing. These studies revealed that the defects in B cell development were not linked to *Siae,* which is present on chromosome 9, but instead to a gene encoding a guanine nucleotide exchange factor, *Dock2*, on chromosome 11. Whole genome sequencing revealed a duplication of exons 28 and 29 of *Dock2*. Surprisingly, an identical mutation in *Dock2* had been previously reported in two different colonies of *Irf5*$^{-/-}$ mice (Purtha et al. 2012). Given that the same mutation in *Dock2* was identified in multiple gene-targeted mouse colonies despite different ES cell lines being used to generate these mice, it appeared that it was most likely introduced during backcrossing into the C57BL/6 background. Furthermore, the presence of C57BL/6N SNPs in close linkage with the duplication

suggests that it arose in a C57BL/6N subline. Indeed, we were able to find this *Dock2* variant (*Dock2*^{Hsd}) in the homozygous state in the colony of C57BL/6NHsd mice maintained at Harlan Sprague Laboratories (now Envigo Biosciences). We also identified the *Dock2*^{Hsd} variant in a colony of *Cmah* knockout mice that had been backcrossed into C57BL/6 mice obtained from Harlan Laboratories. Examination of a range of other commercially available C57BL/6J and C57BL/6N mice revealed only wild type *Dock2*. It is therefore likely that the *Dock2*^{Hsd} variant arose within the C57BL/6NHsd mouse colony at Harlan Laboratories. This study raises concerns that many other lines of gene-targeted mice bearing hematopoietic phenotypes may have been inadvertently compromised by backcrosses involving the use of C57BL/6NHsd mice.

2.2 RESULTS

2.2.1 The loss of marginal zone B cells and enhanced CD8^{+} memory T cell phenotype observed in *Siae* $^{\Delta ex2/\Delta ex2}$ mice is not linked to the loss of Siae

We have previously reported that *Siae* $^{\Delta ex2/\Delta ex2}$ mice, which were backcrossed into the C57BL/6 background for 10 generations (henceforth

referred to as N10-*Siae* mice), show a profound loss of marginal zone (MZ) B cells (Cariappa et al. 2009). In subsequent studies, we noted that N10-*Siae* mice also exhibit a marked increase in $CD8^+$ $CD44^+$ $CD122^{hi}$ memory-phenotype (MP) cells in the blood and spleen (Figure 1A). Surprisingly, both the defect in MZ B cell development as well as the enhanced $CD8^+$ MP cell phenotype were completely lost upon further backcrossing of N10-*Siae* mice into the C57BL/6J background for an additional 13 generations (henceforth referred to as N23-*Siae* mice) (Figure 1B). To further determine whether the loss of Siae *per se* was responsible for any of the N10 phenotypes, we generated an independent line of Siae deficient mice ($Siae^{tm1a/tm1a}$) using an ES cell clone bearing a gene-targeted allele of Siae ($Siae^{tm1a(EUCOMM)Wtsi}$; MGI ID: 4842607) in the C57BL/6N background (Figure 2) . These mice also exhibited normal numbers of MZ B cells and no increase in $CD8^+$ MP cells, confirming that the loss of Siae *per se* was not responsible for the phenotypes previously observed in N10-*Siae* mice (Figure 1B).

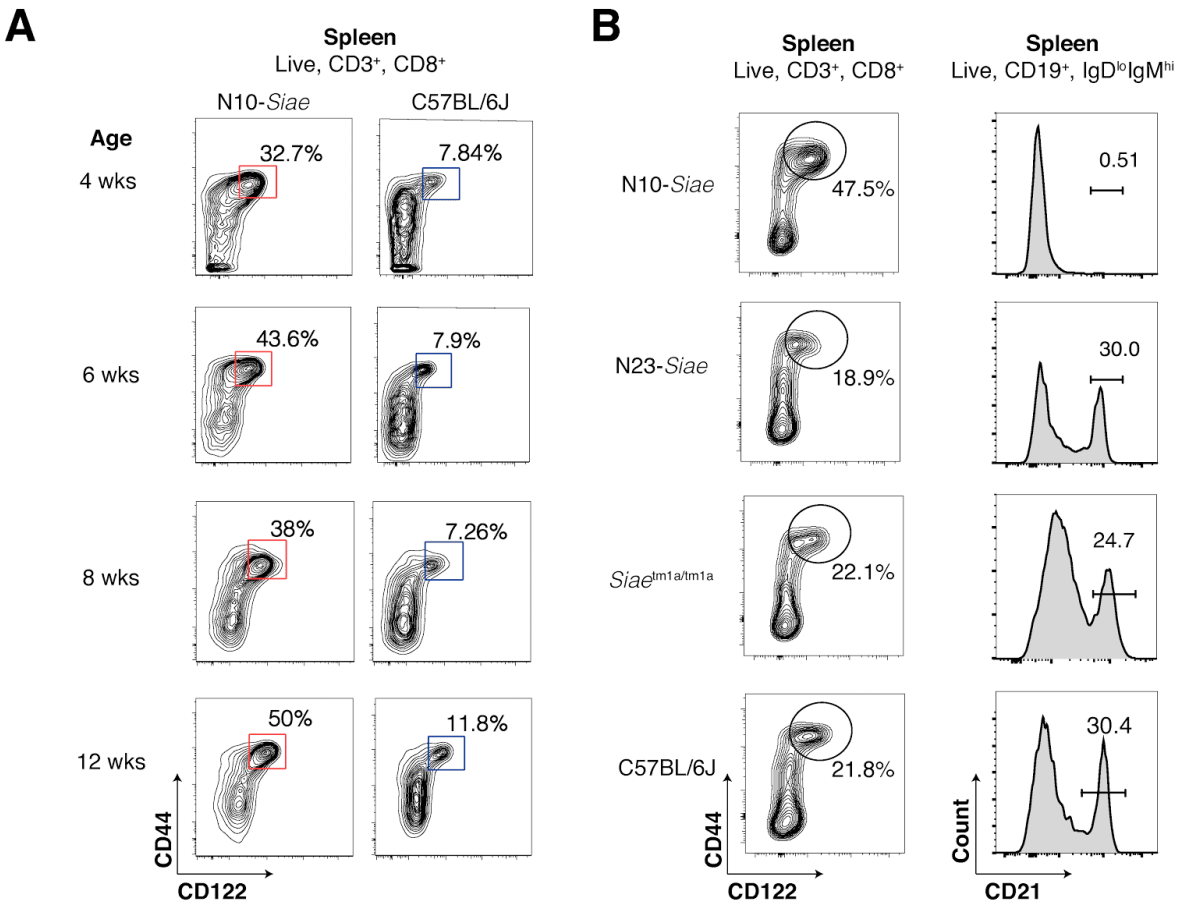


Figure 1: *Siae* deficient mice exhibit a background dependent increase in CD8+ MP cells and a loss of marginal zone B cells. **(A)** The proportion of CD8+ MP cells in the spleens of one to three month old N10*Siae* mice with age matched C57BL/6J controls are shown. **(B)** CD8+ MP and marginal zone B cells in N10*Siae*, N23*Siae*, Siaetm1a/tm1a and C57BL/6J mice.

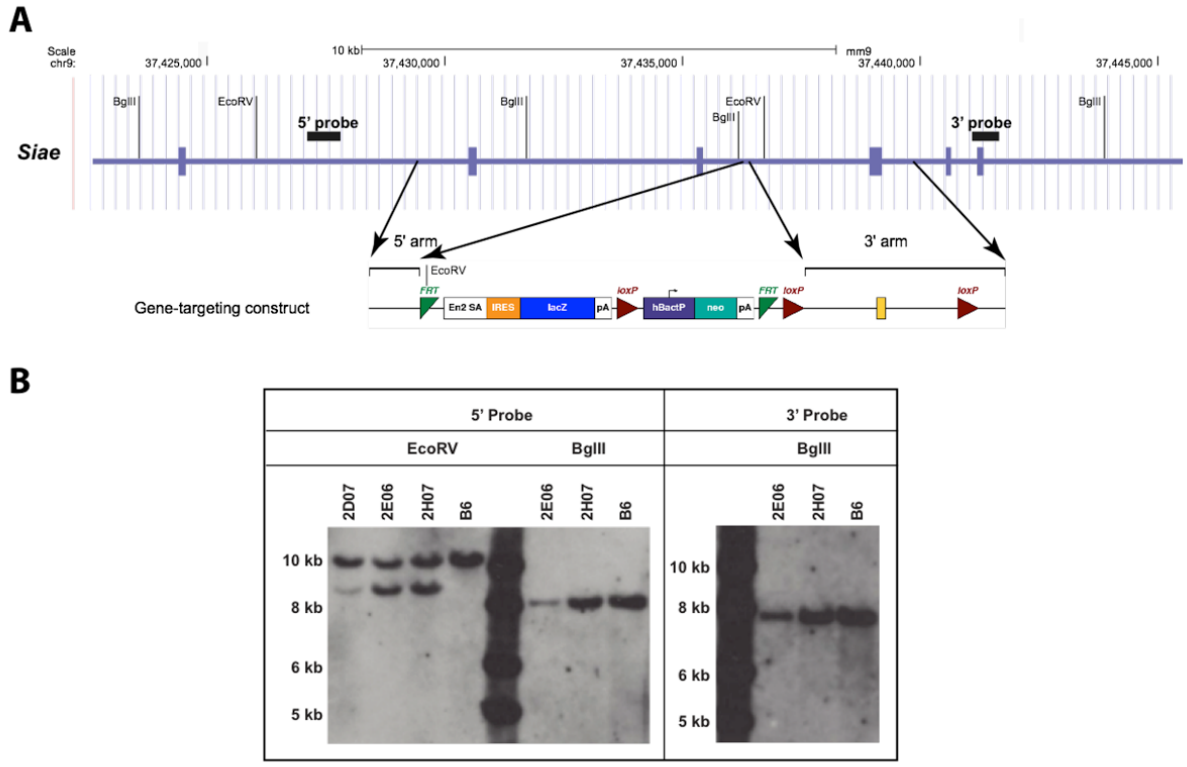


Figure 2: **A schematic of the targeting vector used to generate the EUCOMM genetargeted allele, *Siae*tm1a(EUCOMM)Wtsi, is shown.** Three genetargeted embryoinc stem cell clones (D07, E06, H07) were obtained from EUCOMM and checked for appropriate homologous recombination using a Southern Blot with flanking 5' and 3' probes as shown above. The D07 clone exhibited allelic imbalance suggesting that the D07 ES cell line was subclonal and was excluded from further analysis. The E06 clone was injected into C57BL/6 blastocysts.

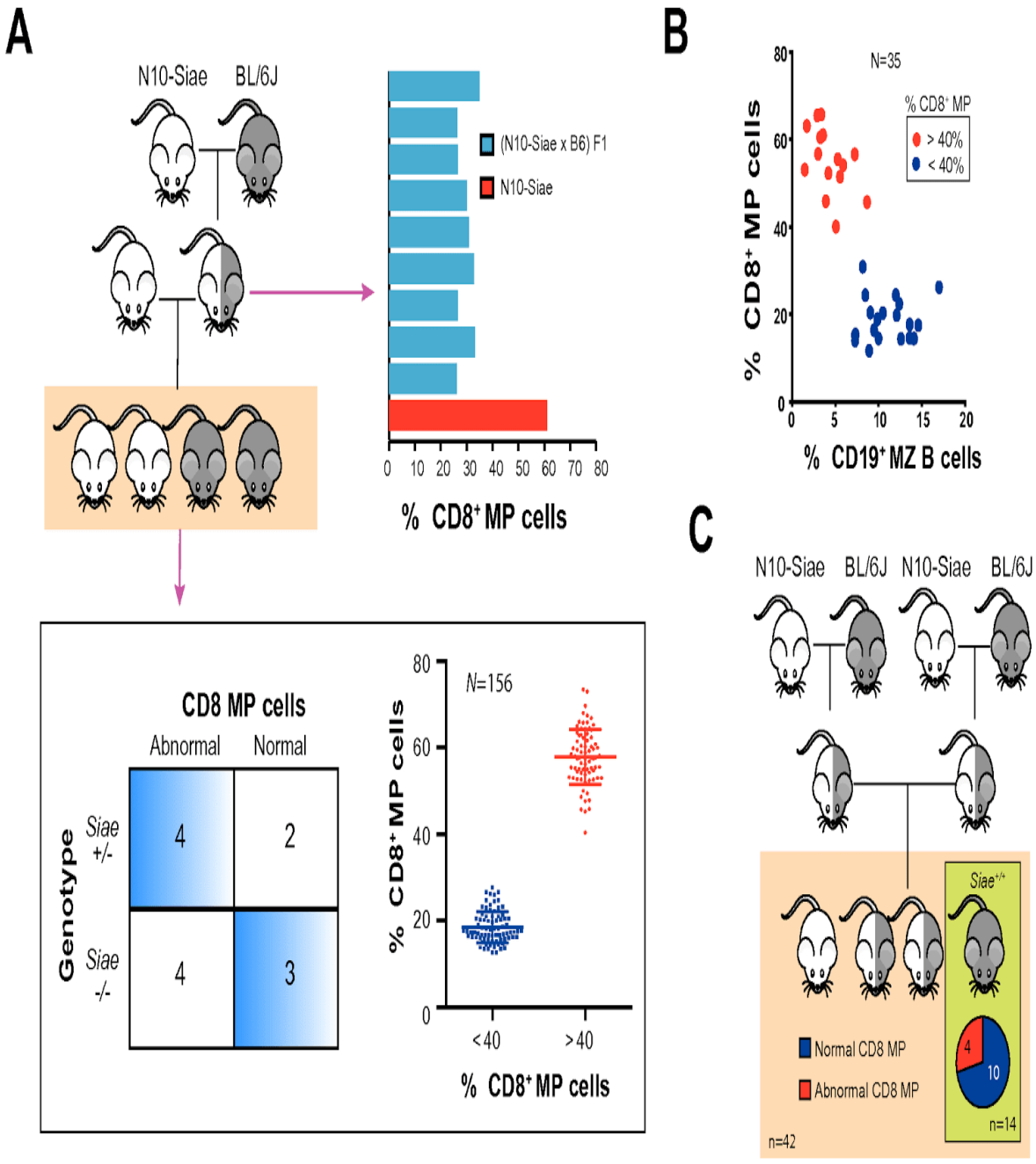
2.2.2 The phenotypes observed in N10-*Siae* mice are inherited in a Mendelian fashion and linked to a locus distinct from *Siae*

To begin to identify the genetic locus responsible for the anomalous phenotypes observed in N10-*Siae* mice, we did a test cross to assess the inheritance pattern. N10-*Siae* x C57BL/6J mice (F1) were generated and backcrossed to N10-*Siae* mice. We found that the increase in $CD8^+$ MP cells was a Mendelian recessive trait and segregated independently of *Siae* (Figure 3A). In addition, analysis of these mice showed a 100% linkage between the loss of MZ B cells and an increase $CD8^+$ MP T cells, suggesting that a single genetic locus was likely responsible for both these phenotypes (Figure 3B). Since the proportion of CD8+ MP cells in blood could be easily measured, we used this trait for mapping the pathogenic locus. We also found that 4 of 18 mice that lacked both copies of the SiaeΔex2 allele had an abnormal CD8+ MP population, suggesting that Siae had no contribution to these observed phenotypes (Figure 3C).

Figure 3: The constellation of hematopoietic phenotypes observed in N10*Siae* mice are Mendelian, recessive & unlinked to *Siae*

(A) (N10Siae × C57BL/6J) F1 animals have a normal CD8+ MP cell compartment. Upon subsequent backcrossing to the N10Siae background, Siae does not segregate with the altered CD8+ MP cells (n=13). Approximately equal number of mice with (n=79) and without (n=77) the phenotype (<40% and >40% CD8 MP cells respectively) are observed. **(B)** A strong correlation between loss of marginal zone B cells and increase in CD8 MP is observed. **(C)** ~25% of Siae+/+ (wildtype) mice obtained from an intercross between heterozygous N10*Siae* mice exhibit normal CD8+ MP cells.

Figure 3 (Continued)

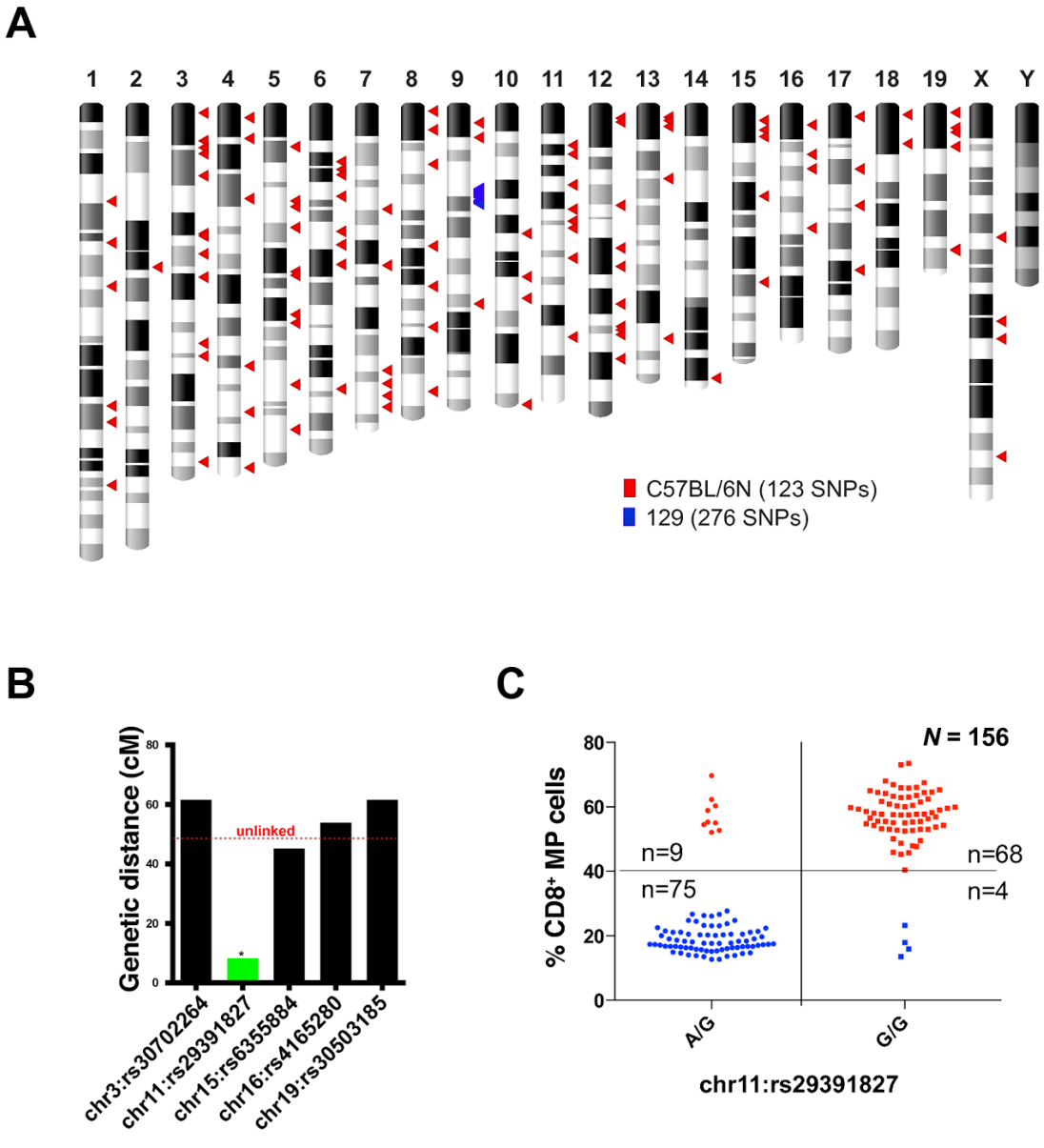


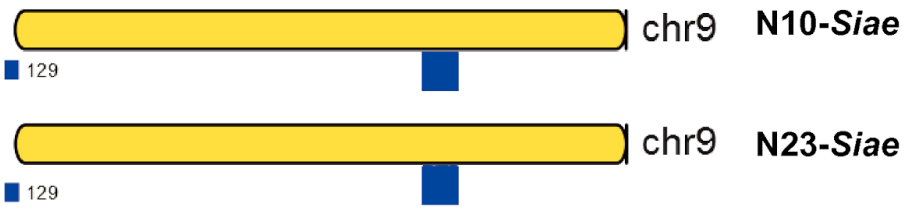
2.2.3 A genetic marker on chromosome 11 segregates with the phenotype

The N10-*Siae* whole genome SNP arrays yielded an average call rate of 99% with 98.7% homozygosity. Publicly available SNP array data from various 129 substrains and C57BL/6 sublines were used for comparison(Didion et al. 2012). 538,667 SNPs showed a 100% call rate across all the arrays. Given that the *Siae*$^{\Delta ex2}$ allele was generated in an R1 ES cell that is derived from the 129X1/SvJ x 129S1/Sv F1-Kitl<+> background(Nagy et al. 1993), it was not surprising that all the SNPs (276 of 276) in a 6 Mbp region surrounding the *Siae* locus (chr9:35022155-41040054) that differed from the C57BL/6 reference genome were of 129 origin (Figure 4A). Surprisingly, all these 129 origin SNPs were retained in the N23-*Siae* mice, suggesting that 13 additional generations of backcrossing into C57BL6/J mice had had no further effect on reducing the 129 contribution on chromosome 9 (Figure 5). Considering that ~17.4% of the 538,667 SNPs calls analyzed in the whole genome SNP array were different between the 129 and C57BL/6 backgrounds and the fact that we did not find any genomic region containing two or more *contiguous* SNPs of 129 origin outside of the chromosome 9 locus containing *Siae*, we recognized that it was extremely unlikely that a mutation in the gene-targeted ES cell clone of 129 origin had been retained in the N10-*Siae* mice.

Figure 4: A SNP on chromosome 11 segregates with increase in CD8 memory phenotype T cells (A) Whole genome SNP array revealed 399 SNPs that differ between N10Siae and C57BL/6J. 276 SNPs (blue triangles) on chromosome 9 in the *Siae* locus are shared by 129 strains. 123 SNPs (red triangles) appear to be of C57BL/6N origin. **(B)** Linkage between candidate SNPs and CD8+ MP expansion in the progeny of test crosses (n=18) described in Figure 2A. A SNP on chromosome 11 (rs29391827) was found to segregate with the CD8 memory phenotype (n=18). **(C)** This observation was validated in a larger cohort (n=156).

Figure 4 (Continued)





SNP_ID	Location	Allele	N10-SIAE	N23-SIAE	C57BL/6J
rs32509547	9:34153897-34153897	T	T/T	T/T	T/T
rs3711756	9:34458844-34458844	C	C/C	C/C	C/C
rs3669204	9:34721457-34721457	T	T/T	T/T	T/T
rs6162143	9:34985485-34985485	A	A/A	A/A	A/A
rs13469717	9:35047277-35047277	A	G/G	G/G	A/A
rs29942698	9:35054520-35054520	G	T/T	T/T	G/G
rs30188025	9:35129933-35129933	G	A/A	A/A	G/G
rs29690638	9:35142525-35142525	A	G/G	G/G	A/A
rs30178192	9:35221038-35221038	G	C/C	C/C	G/G
rs30040613	9:37071278-37071278	A	T/T	T/T	A/A
rs37411281	9:40137298-40137298	G	C/C	C/C	G/G
rs29828932	9:40408637-40408637	C	T/T	T/T	C/C
rs3715411	9:40707199-40707199	T	C/C	C/C	T/T
rs6323544	9:40962391-40962391	C	T/T	T/T	C/C
rs6186137	9:41014146-41014146	T	C/C	C/C	T/T
rs33737439	9:41231561-41231561	G	G/G	G/G	G/G

Figure 5: There is no difference N10 and N23 mice on chromosome 9

SNPs on chromosome 9 of 129 origin in the *Siae* locus (chr9:3502215541040054) in N10*Siae* and N23*Siae* mice. The genotypes of individual SNPs tested are listed.

All (123 of 123) SNPs on the remaining N10-*Siae* chromosomes that differed from the C57BL/6J allele, matched the C57BL/6N consensus allele (shared between C57BL/6NTac, C57BL/6Nci and C57BL/6NCrl), suggesting that the pathogenic locus could have arisen from a C57BL/6N strain used previously for backcrossing before we obtained the mice from the University of California at San Diego. We next tested the genetic linkage to a few randomly chosen candidate C57BL/6N SNPs in the (C57BL/6N X C57BL/6J) test crosses (Figure 3A, B). Fortuitously, a SNP of C57BL/6N origin, rs29391827, on chromosome 11 (chr11:50962622) was linked to the increase in $CD8^+$ MP cells in the blood of 18 mice (Figure 3B). The linkage to rs29391827 was validated in a larger cohort of 156 mice and estimated to be at a distance of about 8.3 cM (Figure 4B).

2.2.4 A region on chr11:30552213-35421130 is tightly associated with the N10-*Siae* phenotype

We next performed whole genome sequencing of an N10-*Siae* mouse that exhibited the phenotypic changes described. Single end sequencing was performed for 85 cycles on a NextSeq500 instrument, yielding a total of ~40Gbp of sequence data. 93.9% of the reads mapped to the reference mouse genome (65.42% of the reads mapped to unique sites on the genome).

A mean coverage depth of 15.33 fold was obtained. The GATK variant calling pipeline was used to identify positions on chromosome 11 that differed from the reference genome (Figure 6) (McKenna et al. 2010). This analysis confirmed the results of the whole genome SNP array (for SNPs with adequate sequencing coverage) and also revealed a few additional SNPs that were not covered by the SNP array that differed between N10-*Siae* and C57BL/6J. All the protein coding variants (3 of 3) that we observed within 25 Mbp of rs29391827 on chromosome 11 were previously reported SNPs present in one or more inbred mouse strains (rs240617401, rs8261521 and rs6268547). As structural variants could not be ruled out on the basis of this analysis, we used a panel of SNPs that differed between N10-*Siae* and C57BL/6J to analyze mice shown in Figure 4 that showed discrepant phenotypes with respect to rs29391827. This allowed us to progressively narrow the pathogenic genetic locus to a 5 Mbp region on chromosome 11 (Ichr11:30552213-35421130) (Figure 7).

Figure 6: Variants in the N10Siae genome (relative to the C57BL/6J reference genome, GRCm38/mm10 assembly) are depicted as a density plot (per million bp).

Figure 6 (Continued)

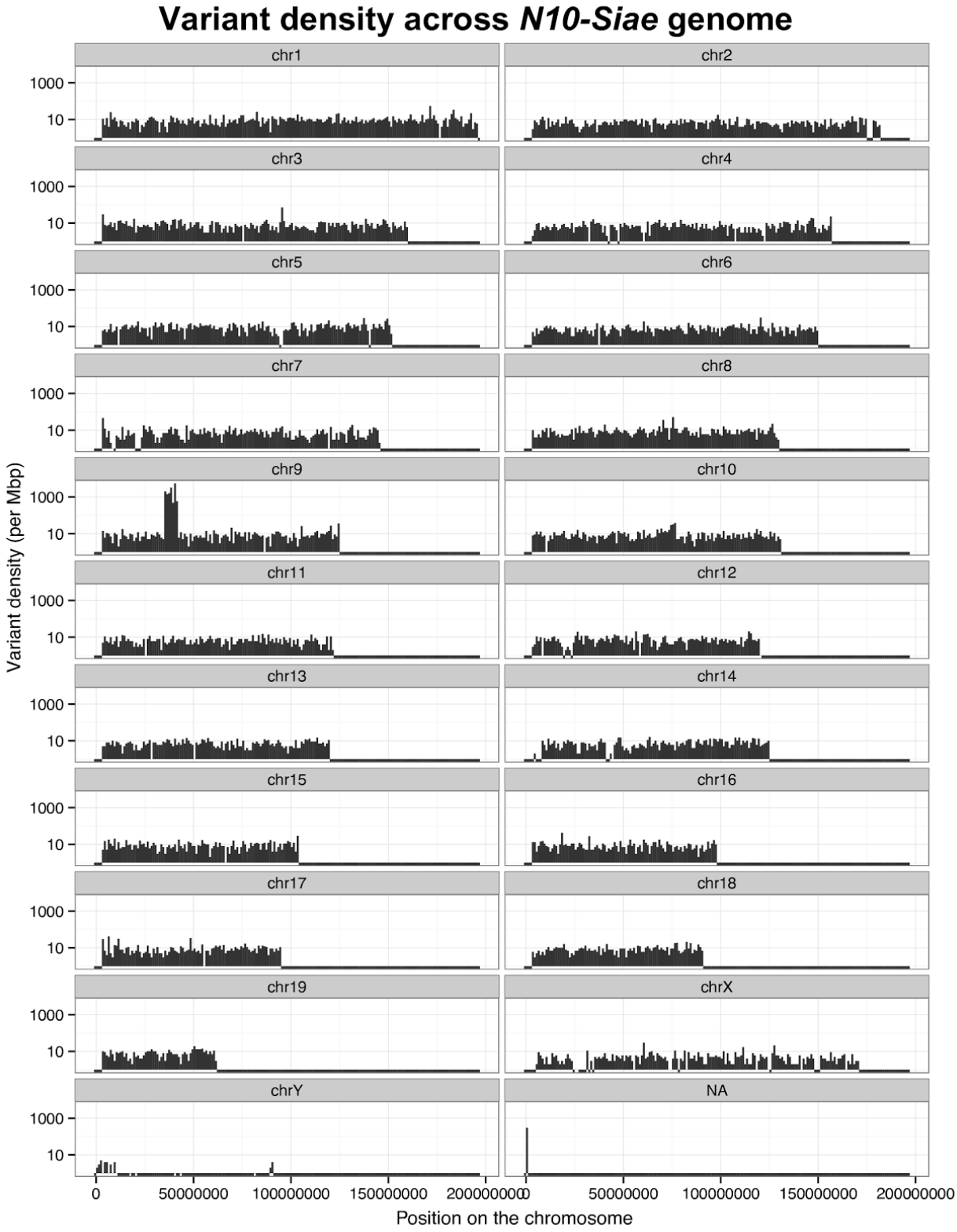
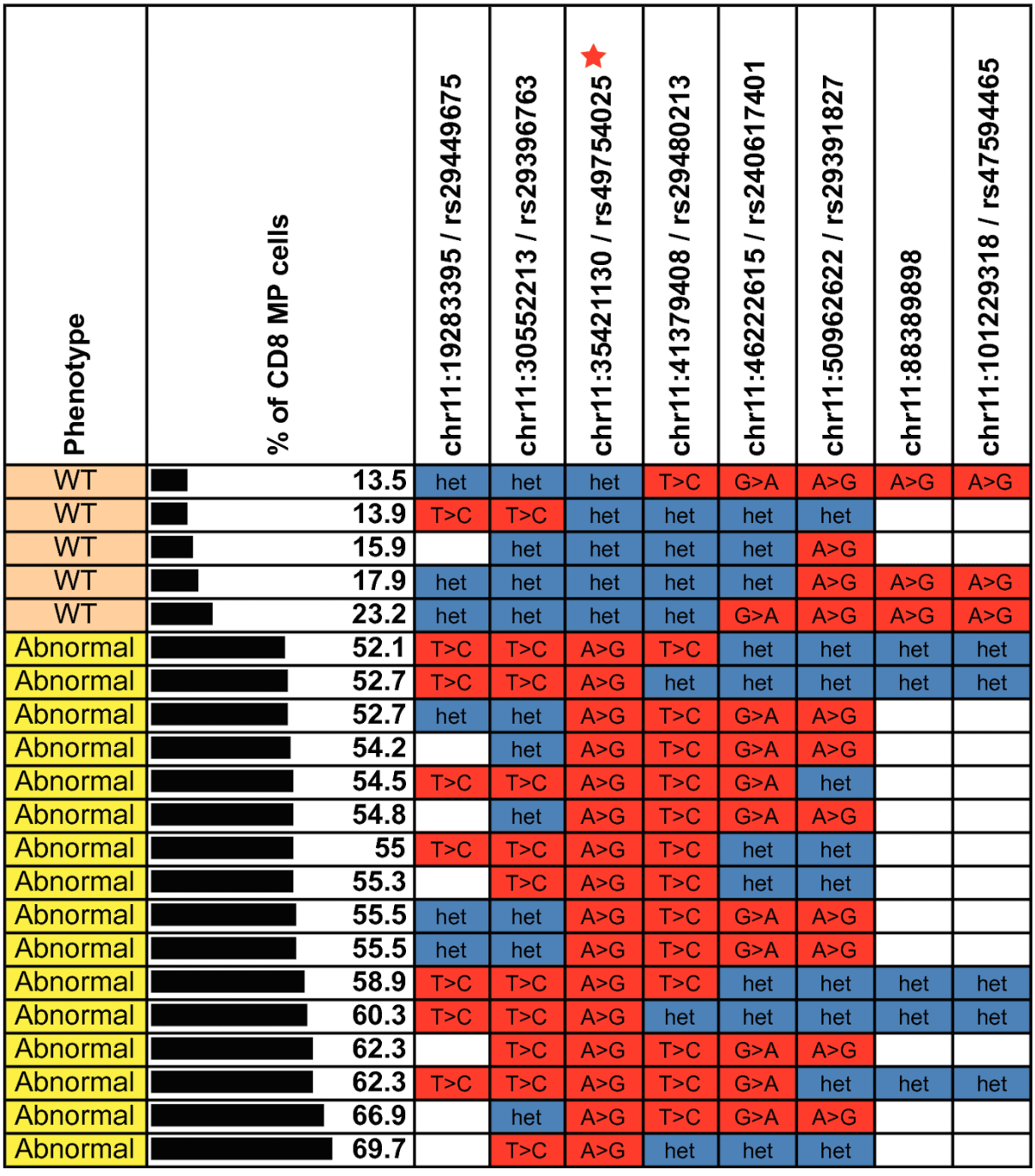


Figure 7: Pathogenic locus narrowed down to chr11:3055221335421130. Mice with suspected informative crossovers were used to progressively narrow the pathogenic locus to chr11:3055221335421130. The rows depict the percentage of MP CD8+ T cells in individual mice with their respective SNP genotypes.

Figure 7 (Continued)



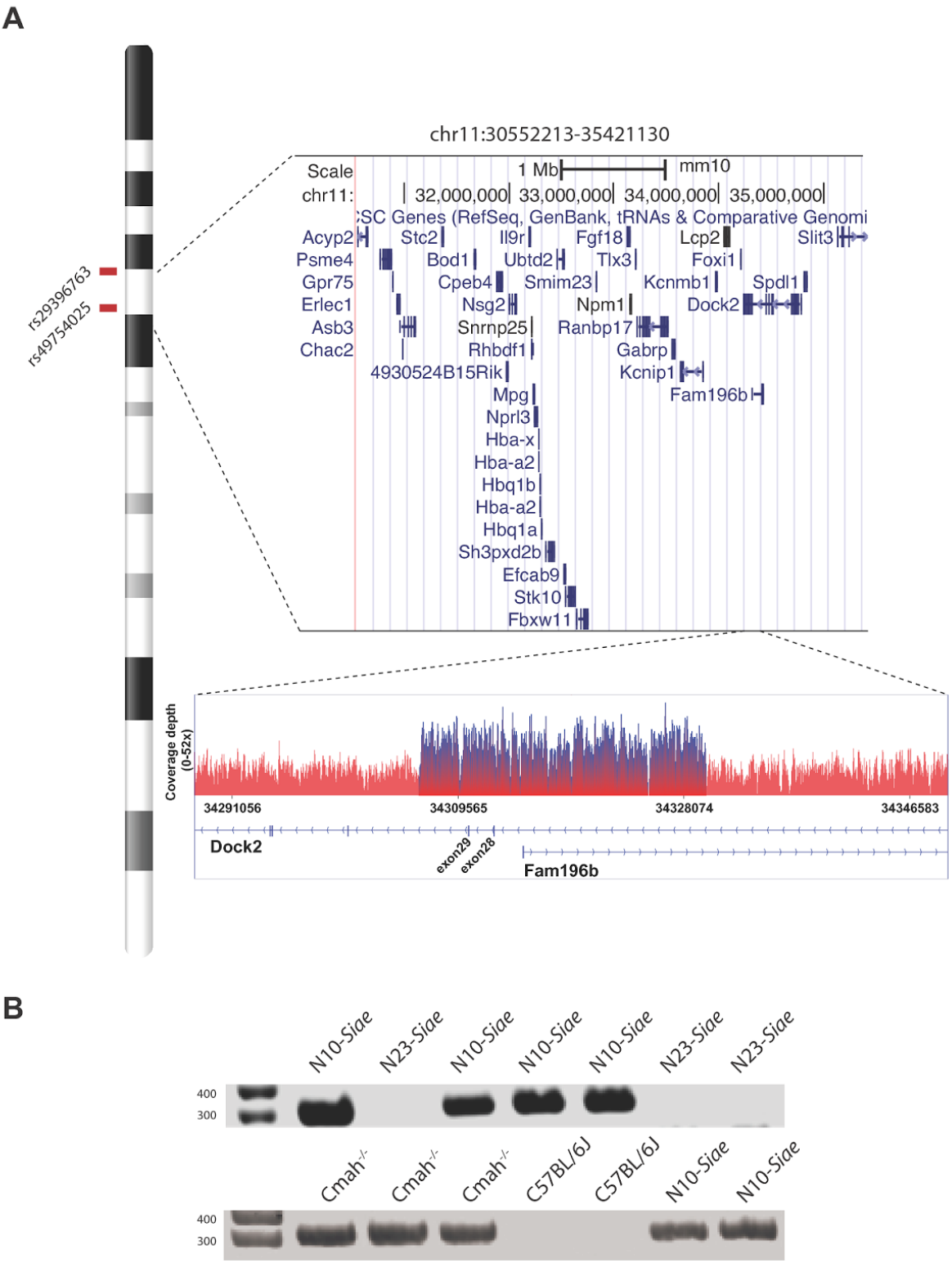
Phenotype	% of CD8 MP cells	chr11:19283395 / rs29449675	chr11:30552213 / rs29396763	chr11:35421130 / rs49754025 ★	chr11:41379408 / rs29480213	chr11:46222615 / rs240617401	chr11:50962622 / rs29391827	chr11:88389898	chr11:101229318 / rs47594465
WT	13.5	het	het	het	T>C	G>A	A>G	A>G	A>G
WT	13.9	T>C	T>C	het	het	het	het		
WT	15.9		het	het	het	het	A>G		
WT	17.9	het	het	het	het	het	A>G	A>G	A>G
WT	23.2	het	het	het	het	G>A	A>G	A>G	A>G
Abnormal	52.1	T>C	T>C	A>G	T>C	het	het	het	het
Abnormal	52.7	T>C	T>C	A>G	het	het	het	het	het
Abnormal	52.7	het	het	A>G	T>C	G>A	A>G		
Abnormal	54.2		het	A>G	T>C	G>A	A>G		
Abnormal	54.5	T>C	T>C	A>G	T>C	G>A	het		
Abnormal	54.8		het	A>G	T>C	G>A	A>G		
Abnormal	55	T>C	T>C	A>G	T>C	het	het		
Abnormal	55.3		T>C	A>G	T>C	het	het		
Abnormal	55.5	het	het	A>G	T>C	G>A	A>G		
Abnormal	55.5	het	het	A>G	T>C	G>A	A>G		
Abnormal	58.9	T>C	T>C	A>G	T>C	het	het	het	het
Abnormal	60.3	T>C	T>C	A>G	het	het	het	het	het
Abnormal	62.3		T>C	A>G	T>C	G>A	A>G		
Abnormal	62.3	T>C	T>C	A>G	T>C	G>A	het	het	het
Abnormal	66.9		het	A>G	T>C	G>A	A>G		
Abnormal	69.7		T>C	A>G	het	het	het		

2.2.5 A *Dock2* duplication on chromosome 11 accounts for the observed N10-*Siae* phenotype

A focused examination of the whole genome sequence in the chr11:30,552,213-35,421,130 region revealed a ~2-fold increase in coverage in a 23.5 Kbp region encompassing exons 28 and 29 of the *Dock2* gene (Figure 8A). This was reminiscent of a previously published report identifying a duplication involving exons 28 and 29 of *Dock2* in *Irf5*$^{-/-}$ mice(Purtha et al. 2012). Indeed, the presence of an identical duplication in *Dock2* was confirmed in N10-*Siae* mice using a previously reported PCR designed to identify this duplication (Figure 8B) (Kei Yasuda et al. 2013). Sequencing of the PCR product revealed that the breakpoint in the *Dock2* duplication (chr11:34329863-34306424; GRCm38/mm10 assembly) was identical in *Irf5*$^{-/-}$ and N10-*Siae* mice (Figure 9). As *Cmah*$^{-/-}$ mice also lack MZ B cells(Cariappa et al. 2009), we looked for the presence of the *Dock2*^{Hsd} allele in two *Cmah*$^{-/-}$ mice. The *Dock2*^{Hsd} allele was present in 3 of 3 *Cmah*$^{-/-}$ mice analyzed (Figure 8B). This duplication has been demonstrated to be a loss-of-function allele as it results in a frameshift mutation and nonsense-mediated decay of *Dock2* mRNA(Purtha et al. 2012).

Figure 8: A genomic segment encoding exons 28 and 29 of Dock2 is duplicated in N10Siae and Cmah/ mice. (A) An abrupt increase in coverage depth within the putative pathogenic locus on chromosome 11 mapped by SNPs ecompasses exons 28 and 29 of Dock2 in the N10Siae genome. (B) PCR detection of Dock2 duplication in N10Siae, Cmah/, N23Siae and C57BL6/J mice.

Figure 8 (Continued)



A

Duplicated region in *Dock2* (chr11:34329863-34306424)

```
> IRF5-/-; Dock2 duplication; Yasuda et al.
GCCCCTCAAGACGACCTTATGAGGTGGAACCACAACCATTTGCATTTTACAAATGACACAGCTAACTGACTTACTCATCATTGCAATGCC
AGTAAGGAAGAGAAACACCATTCAACCTGAGCCAACTTCTGAGCCCCGGCTTCATCCTGTGAGGGAGGGAGGGAGGGAGGGAGAGAGAGA
GAGAGAGAAAATCTCCCCTACCCCACACAGAAGCTAAAAGATCATTAGACATGGTTGCTCTGGAGAAAGCAATATTCCTGTTTCTTAAAC
CTAAACCCAATGAACACAGAAGTGGAGCTGTAGGGAATCTTTGGATCCCCAGTCCCCAACCGCAGCCTCTTGAAAATAGCTC


> N10-SIAE; Dock2 duplication; amplicon sequencing
GCTAACTGACTTACTCATCATTGCAATGCCAGTAAGGAAGAGAAACACCATTCAACCTGAGCCAACTTCTGAGCCCCGGCTTCATCCTGT
GAGGGAGGGAGGGAGGGAGGGAGAGAGAGAGAGAGAGAAAATCTCCCCTACCCCACACAGAAGCTAAAAGATCATTAGACATGGTTGCTC
TGGAGAAAGCAATATTCCTGTTTCTTAAACCTAAACCCAATGAACACAGAAGTGGAGCTGTAGGGAATCTTTGGATCCCCAGTCCCCAAC
CGCAGCCTCTTGAAAATAGC
```

B

```
Pairwise BLAST Comparison:

N10-SIAE        1    GCTAACTGACTTACTCATCATTGCAATGCCAGTAAGGAAGAGAAACACCATTCAACCTGA  60
IRF5-/-         61   GCTAACTGACTTACTCATCATTGCAATGCCAGTAAGGAAGAGAAACACCATTCAACCTGA  120

N10-SIAE        61   GCCAACTTCTGAGCCCCGGCTTCATCCTGTgagggagggagggagggagggagagagaga  120
IRF5-/-         121  GCCAACTTCTGAGCCCCGGCTTCATCCTGTGAGGGAGGGAGGGAGGGAGGGAGAGAGAGA  180

N10-SIAE        121  gagagagaAAATCTCCCCTACCCCACACAGAAGCTAAAAGATCATTAGACATGGTTGCTC  180
IRF5-/-         181  GAGAGAGAAAATCTCCCCTACCCCACACAGAAGCTAAAAGATCATTAGACATGGTTGCTC  240

N10-SIAE        181  TGGAGAAAGCAATATTCCTGTTTCTTAAACCTAAACCCAATGAACACAGAAGTGGAGCTG  240
IRF5-/-         241  TGGAGAAAGCAATATTCCTGTTTCTTAAACCTAAACCCAATGAACACAGAAGTGGAGCTG  300

N10-SIAE        241  TAGGGAATCTTTGGATCCCCAGTCCCCAACCGCAGCCTCTTGAAAATAGC  290
IRF5-/-         301  TAGGGAATCTTTGGATCCCCAGTCCCCAACCGCAGCCTCTTGAAAATAGC  350
```

Figure 9: Identical duplication of Dock2 is present in N10Siae and IRF5/ mice. (A)Sequence of the Dock2 duplication from an N10Siae mouse obtained with Sanger sequencing of the PCR amplicon, and the previously published Dock2 duplication sequence from IRF5/ (Yasuda et al. Int. Immunol. 25, 295–306, 2013). (B) Pairwise alignment of the Dock2 duplication from N10Siae and IRF5/ mice. The breakpoint is indicated by an arrow. A low complexity sequence at the site of the breakpoint is shown in lower case.

2.2.6 The mutant Dock2 allele was introduced by backcrossing into C57BL/6NHsd mice

Given that the same duplication was seen in three independently gene-targeted mice, it confirmed our suspicion that the mutant *Dock2* allele was inadvertently introduced during backcrossing. As previously noted, the SNPs flanking the duplication were of C57BL6/N origin (rs29473246 at chr11:33548367, rs29414108 at chr11:43358462), suggesting that this duplication arose in a C57BL/6N substrain. Indeed, a survey of several commercially available C57BL6/N substrains showed the presence of a duplication in exons 28 and 29 of Dock2 with an identical breakpoint in 44 of 44 C57BL/6NHsd mice (Harlan Sprague Dawley, acquired by Envigo Biosciences in 2015) analyzed (Figure 10A and Figure 11). We henceforth refer to this allele as *Dock2*^{Hsd}. The *Dock2*^{Hsd} allele was not detected in any of the other C57BL/6N strains analyzed viz. C57BL/6NTac, C57BL/6NCrl, C57BL/6NJ (Figure 11). This Dock2 duplication was also not observed in the C57/BL6NJ genome published by the Mouse Genomes Project (Wellcome Trust Sanger Institute)(Yalcin et al. 2012). Thus, the *Dock2*^{Hsd} allele appears to be a copy number variant that has been fixed in the C57BL/6NHsd strain and may have been introduced into numerous other gene-targeted mouse lines during backcrossing.

Figure 10: The Dock2 duplication is present in the C57BL/6NHsd strain, which exhibits multiple characteristics of Dock2 deficiency. (A) PCR detection of Dock2 duplication in C57BL/6NHsd mice and other C57BL6/N sublines (C57BL6/NTac and C57BL6/NJ). (B) Sequences of the Dock2 duplication breakpoint (arrow) from C57BL6/NHsd, N10Siae and IRF5/ mice. A low complexity sequence at the site of the breakpoint is shown in lower case. (C) Proportion of CD8+ MP cells, marginal zone B cells, invariant NK T cells and plamsmacytoid dendritic cells C57BL/6NHsd, C57BL6/NTac and C57BL6/J mice.

Figure 10 (Continued)

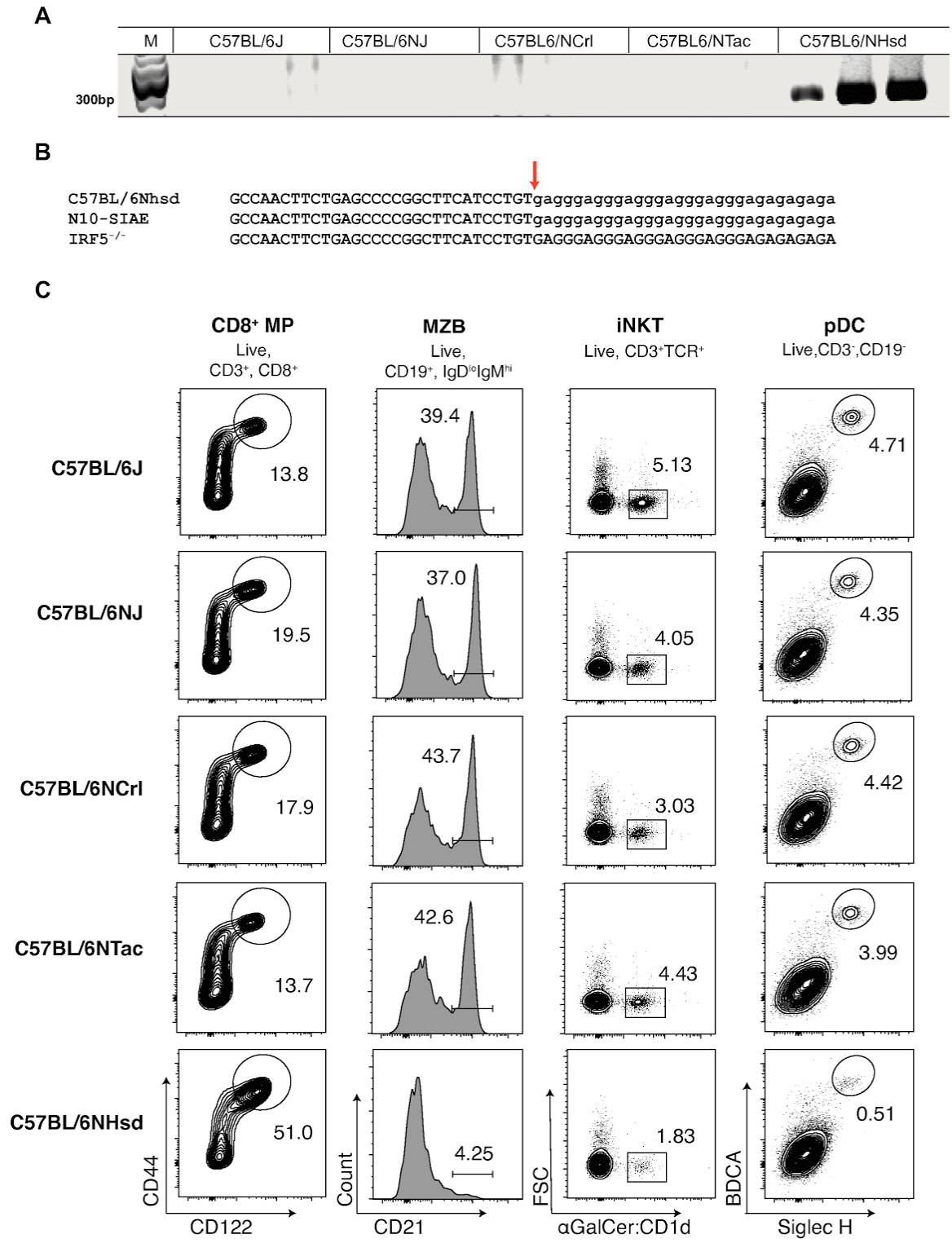
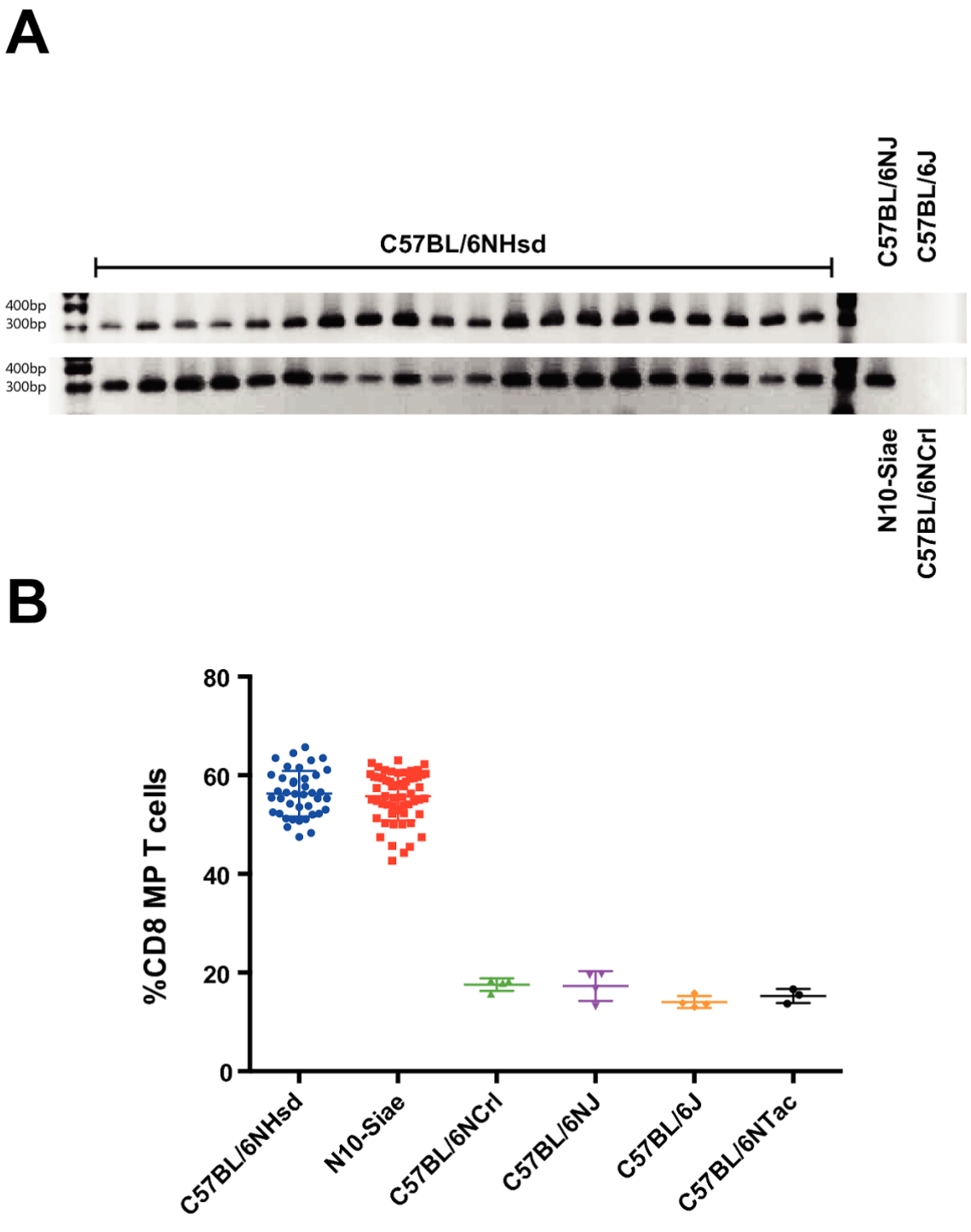


Figure 11 Dock2 duplication is found at a very high allele frequency in the C57BL/6NHsd strain. (A) 40 C57BL/6NHsd and other C57BL/6N sublines were genotyped for the Dock2 duplication using a previously described PCR (Yasuda et al. Int. Immunol. 25, 295–306, 2013) and (B) phenotyped for the expansion of CD8+ MP T cells in the blood.

Figure 11 (Continued)



Consistent with this, both *Siae* $^{\Delta ex2/\Delta ex2}$ and *Cmah*$^{-/-}$ mice were generated from ES cells of 129 origin and backcrossed into the C57BL/6 background at UCSD using mice obtained from Harlan Laboratories (personal communication from Ajit Varki). The *Irf5*$^{-/-}$ mice were generated from a E14K ES cell (129OlaHsd background) and backcrossed into the C57BL/6 background before being distributed to various investigators by laboratory of Tadatsugu Taniguchi (Takaoka et al. 2005). However, the Dock2 duplication was found only in two of three colonies derived from the original *Irf5*$^{-/-}$ colony, which led the investigators to believe that the Dock2 duplication arose spontaneously within a subline of *Irf5*$^{-/-}$ mice (Purtha et al. 2012). Like the N10-*Siae* mice, C57BL/6NHsd mice exhibit an expansion of $CD8^+$ MP T cells (Figure 10C and Figure 11). C57BL/6NHsd mice also exhibit other phenotypes that have been previously observed in the *Dock2*$^{-/-}$ mice (Fukui et al. 2001). Both *Dock2*$^{-/-}$ and C57BL/6NHsd mice exhibit splenic lymphopenia as well as a profound loss of marginal zone B cells, invariant NKT cells and plasmacytoid dendritic cells (Figure 10C) (Fukui et al. 2001; Kunisaki et al. 2006; Gotoh et al. 2008). The similarities between *Dock2*$^{-/-}$ and C57BL/6NHsd mice support the notion that the *Dock2*^{Hsd} allele inactivates the function of *Dock2* by creating a premature stop codon that contributes to nonsense mediated decay of the resultant mRNA.

2.3 DISCUSSION

Striking phenotypes have been discovered in a number of gene-targeted mice that were previously assumed to be the result of the homozygous loss of a specific gene. To clarify the loss of phenotype in *Siae* $^{\Delta ex2/\Delta ex2}$ mice upon additional backcrossing into a C57BL/6J background, we generated an additional mutant allele of *Siae* and also failed to observe any of the previously reported phenotypes. We therefore undertook a series of studies involving conventional genetics, whole genome genotyping and whole genome sequencing to identify a causal mutation on chromosome 11. We then also found this same mutation, a homozygous duplication of two exons of *Dock2* that creates a non-functional product, in *Cmah* mutant mice and noted that an identical mutation had been previously found in *Irf5*$^{-/-}$ mice, originally generated in Japan, and assumed to have been a spontaneous mutation (Purtha et al. 2012; Takaoka et al. 2005). The most parsimonious explanation for a specific mutation being present in three different gene-targeted mice, all derived from different ES cells, two generated originally in California and a third in Japan, is that they had all been acquired by backcrossing into an unidentified mutant C57BL/6 strain.

Because our SNP arrays showed the presence of specific C57BL/6N SNPs we had a high degree of suspicion that this mutation had been acquired from a commercial C57BL/6N strain. We surveyed commercial C57BL/6N and C57BL/6J strains and discovered a specific *Dock2* gene duplication that was present only in C57BL/6N mice obtained from Harlan Laboratories (C57BL/6NHsd). Intriguingly, the set of 123 SNPs outside the *Siae* locus that differed between N10-*Siae* and C57BL/6J also comprised 97% of the SNP alleles (n = 128) from a pool of 538,667 SNPs analyzed in the whole genome arrays that differ between C57BL/6J and C57BL/6N sublines but are identical to the consensus 129 alleles. Given that C57BL/6J and C57BL/6N strains were derived from a common stock of C57BL/6 mice, the differences between C57BL/6J and C57BL/6N substrains have been so far attributed to genetic drift. In contrast, our analysis indicates that C57BL/6N mice may have derived some genetic contribution from the 129 strain at a remote time during their history, prior to the divergence of various C57BL/6N sublines.

We have verified that the *Siae* $^{\Delta ex2/\Delta ex2}$ and *Cmah* deficient mice were backcrossed into a C57BL/6 background using mice obtained by our collaborators at UCSD from Harlan Laboratories Inc. between 2005-2008. We have also been informed that C57BL/6 mice from Harlan Laboratories were being routinely used for backcrossing and some other lines generated at UCSD have also have been compromised (personal communication from Ajit Varki). Initial reports of *Irf5*$^{-/-}$ mice backcrossed into C57BL/6 background that

bear phenotypes (e.g. lack of type I interferon) attributable to Dock2 deficiency were published in 2007 (Gotoh et al. 2010, 2008; K. Yasuda et al. 2007). Thus, we believe that the *Dock2* copy number variant has been present in the C57BL/6NHsd colony at Harlan Laboratories for at least a decade. The C57BL/6NHsd mice were derived by Harlan Laboratories from the breeding nucleus of the C57BL/6N colony maintained at the NIH in 1988. Harlan Laboratories Inc was acquired by Envigo Biosciences in 2015. The C57BL/6N subline originated from the C57BL/6J colony at Jackson Laboratory in 1951 (Mekada et al. 2009). Given that other sublines derived from C57BL/6N and C57BL/6J strains lack the *Dock2*^{Hsd} variant, it is formally possible that a spontaneous mutation present in the C57BL/6N founders was expanded at Harlan Laboratories. However, it is more likely that the *Dock2* duplication arose later in the C57BL/6NHsd colony after it diverged from the NIH breeding stock and was fixed by the breeding strategies based on continued brother-sister mating that are designed to minimize genetic drift in mouse sublines. Analysis of archived tissues from C57BL/6NHsd mice will help estimate when the *Dock2*^{Hsd} variant arose, and help assess how many studies have been affected. Despite the relatively large size of the Dock2 gene, coding variants of *Dock2* have not been reported in other inbred mouse strains (Sherry et al. 2001).

The *Dock2*^{Hsd} variant is particularly significant as it causes wide-ranging hematopoietic phenotypes in a strain of mice that is widely used for immunological studies. Since 2003, the Jackson Laboratory has implemented a Genetic Stability Program in selected mouse strains to limit cumulative genetic drift, including that caused by copy number variation, by regularly rebuilding foundation stocks from cryopreserved, pedigreed embryos every few generations (Taft, Davisson, and Wiles 2006). Commercial breeders of inbred mouse strains as well as individual research groups maintaining knockout lines over long periods should consider the implications of the choice of their breeding strategies on genetic drift within the colony.

The recent availability of ES cells from a C57BL/6 background obviates the need for backcrossing but ES cells derived from the C57BL/6NHsd background should be tested for the presence of the *Dock2*^{Hsd} allele (Skarnes et al. 2011). Given that Dock2 is expressed primarily in the hematopoietic lineage, immune phenotypes that have been studied in the context of B6 mice from Harlan Laboratories should be reviewed carefully. The presence of the Dock2^{Hsd} allele also affects studies when mutant strains that lack the Dock2^{Hsd} allele were compared with C57BL/6 mice obtained from Harlan Sprague as controls.

Loss of function variants of other genes such as *Nnt*, *Mnrn1*, *Rd8* and *Cyfip2* have been described in one or more sublines of C57BL/6 and have been linked to glucose intolerance, impaired platelet function, retinal degeneration and altered cocaine response respectively (Mattapallil et al. 2012; Kumar et al. 2013; Reheman et al. 2010; Freeman et al. 2006). Loss of function variants of *Scna* are seen in some sublines of C57BL/6N as well as in ES cells of C57BL/6N origin but result in no obvious phenotype (Specht and Schoepfer 2001). A copy number variant that alters the expression of insulin-degrading enzyme (*Ide*) and fibroblast growth factor binding protein 3 (*Fgfbp3*) genes has been reported to have achieved a high allele frequency in C57BL6/J mice (Watkins-Chow and Pavan 2008). While it is difficult to estimate the precise degree of genetic drift that results in deleterious or loss-of-function mutations, subline-specific variants are not uncommon and should be considered as a possible cause of phenotypic discrepancies in mouse sublines. Furthermore, the specific substrain of mice used for experiments or for backcrossing should be clearly documented.

2.4 MATERIALS AND METHODS

Generation of *Siae* knockout mice (*Siae*$^{tm1a/tm1a}$).

An ES cell clone (EPD0679_2_E06) bearing a targeted disruption of the Siae locus (*Siae*$^{tm1a(EUCOMM)Wtsi}$; MGI ID: 4842607) was obtained from the European Conditional Mouse Mutagenesis Program (EUCOMM) (Collins, Rossant, and Wurst 2007). The *Siae*$^{tm1a(EUCOMM)Wtsi}$ allele (*Siae*^{tm1a}) contains a beta-galactosidase reporter and a polyadenylation signal in intron 3 of *Siae*. Gene targeting had been initially confirmed using long-range PCR by EUCOMM, and additionally validated by us with a Southern Blot using the non-isotopic BrightStar Biodetect kit (Ambion Inc.) (Figure 2). The ES cells were then injected into C57BL/6 blastocysts at the Transgenic Core Facility at Brigham and Women's Hospital. The resulting chimeric mice were crossed into C57BL6/NTac. An *Siae*^{tm1a} founder was backcrossed for one generation into C57BL6/NTac and intercrossed to obtain *Siae*$^{tm1a/tm1a}$ homozygous mice.

Genetic mapping

Whole genome SNP arrays were performed at the Microarray Core Facility at the Dana Farber Cancer Institute using the Affymetrix Mouse Diversity Array platform. Analysis was performed using the Affymetrix Genotyping Console version 4.2.0.26. Publicly available SNP array data from various 129 substrains and C57BL/6 sublines were used for comparison (Didion et al. 2012). We focused on 538,667 SNPs that were concordant among all the C57BL/6N substrains, all the 129 strains and both the N10-Siae mice analyzed respectively. These were intended to represent the ancestral alleles of the 129 and C57BL6/N strains. Candidate SNPs were evaluated using PCR and Sanger sequencing. Sequencing chromatograms were analyzed using Mutation Surveyor v3.24 (Softgenetics).

Whole genome sequencing

A whole genome library was constructed using the Kapa HyperPlus kit following manufacturer's recommendations. Briefly, 1 µg of genomic DNA was enzymatically fragmented, end repaired, A-tailed and ligated to Illumina Truseq adapters. No library amplification was performed to avoid introducing any coverage bias. The library was sequenced on the Illumina NextSeq 500

for 85 cycles using the NextSeq 500/550 High Output v2 kit. A total of 567,104,840 single-end reads were aligned to the *Mus musculus* genome (GRCm38/mm10, December 2011 build) using Bowtie2 (Langmead and Salzberg 2012). For a total 93.96% of reads mapping to the reference genome, 65.42% of reads mapped to unique sites on the genome, and 28.54% of reads mapped to multiple regions. The GATK pipeline was used to identify variants (McKenna et al. 2010). Variants with quality scores < 30 or allele frequency <100% were excluded, yielding 35,340 variants. Variant and read densities were calculated using BEDtools on Galaxy and visualized using the R statistical programming language (Quinlan and Hall 2010).

Flow cytometry

Mouse splenocytes or peripheral blood were hemolyzed with ACK buffer, stained with fluorescently conjugated antibodies analyzed on a BD LSR II flow cytometer. The following antibody clones were used in this study: anti-CD122 (TMB1), anti-CD44 (IM7), anti-CD3 (17A2), anti-CD8 (53-6.7) anti-CD19 (6D5) anti-B220 (RA3-6B2), anti-CD21(7E9), anti-IgM (RMM-1), anti-CD1d (1B1), and anti-IgD (11-26c.2a) from Biolegend.

Genotyping strategy

Multiple mice should be tested to check if a mouse colony has been compromised with the Dock2Hsd allele.

A PCR screening approach is outlined below.

1. PCR screening for the Dock2Hsd variant from genomic DNA

1) Extract genomic DNA from tail clips or splenocytes using KAPA Mouse Genotyping kit (KapaBiosystems, #KK7352) or Qiagen DNeasy Blood & Tissue Kit (Qiagen, #69504).

2) Perform PCR using Kapa2G HotStart ReadyMix (Kapa Biosystems, #KK5701) with the following primers as described by

GAC CTT ATG AGG TGG AAC CAC AAC C

GAT CCA AAG ATT CCC TAC AGC TCC AC

→ Cycling conditions: 95°C 3min; (95°C 10sec, 62.5°C 10 sec, 72°C 15 sec) x 35 cycles; 72°C 1 min.

A 305 bp product indicates the presence of the Dock2Hsd copy number variant. This PCR does not distinguish between homozygous or heterozygous variants. In order to distinguish whether the Dock2Hsd variant is homozygous or heterozygous, one could look for the presence of the WT Dock2 mRNA in

the hematopoietic tissues of Dock2Hsd PCR-positive animals by RT-PCR. Flow cytometric quantitation of marginal zone B cells in the spleen or of CD8+ memory T cells in the blood as described in this paper can also serve as a surrogate assay for homozygosity and may be used in combination with DNA genotyping.

This chapter was published in cell reports (Cell Reports 15 (9):1901–9*) under the title:*

Striking immune phenotypes in diverse gene-targeted mice are driven by a confounding copy number variant originating from a commercially available C57BLB6/N strain

Authors: Vinay S Mahajan, Ezana Demissie*, Hamid Mattoo, Vinay Viswanadham, Robert Morris, Shiv Pillai*

**These authors contributed equally*

References

Cariappa, Annaiah, Hiromu Takematsu, Haoyuan Liu, Sandra Diaz, Khaleda Haider, Cristian Boboila, Geetika Kalloo, et al. 2009. “B Cell Antigen Receptor Signal Strength and Peripheral B Cell Development Are Regulated by a 9-O-Acetyl Sialic Acid Esterase.” The Journal of Experimental Medicine 206 (1):125–38.

Collins, Francis S., Janet Rossant, and Wolfgang Wurst. 2007. “A Mouse for All Reasons.” Cell 128 (1):9–13.

Didion, John P., Hyuna Yang, Keith Sheppard, Chen-Ping Fu, Leonard McMillan, Fernando Pardo-Manuel de Villena, and Gary A. Churchill. 2012. “Discovery of Novel Variants in Genotyping Arrays Improves Genotype Retention and Reduces Ascertainment Bias.” BMC Genomics 13 (January):34.

Freeman, Helen C., Alison Hugill, Neil T. Dear, Frances M. Ashcroft, and Roger D. Cox. 2006. “Deletion of Nicotinamide Nucleotide Transhydrogenase: A New Quantitive Trait Locus Accounting for Glucose Intolerance in C57BL/6J Mice.” Diabetes 55 (7):2153–56.

Fukui, Y., O. Hashimoto, T. Sanui, T. Oono, H. Koga, M. Abe, A. Inayoshi, et al. 2001. “Haematopoietic Cell-Specific CDM Family Protein DOCK2 Is Essential for Lymphocyte Migration.” Nature. https://doi.org/10.1038/35090591.

Gotoh, Kazuhito, Yoshihiko Tanaka, Akihiko Nishikimi, Ayumi Inayoshi, Munechika Enjoji, Ryoichi Takayanagi, Takehiko Sasazuki, and Yoshinori Fukui. 2008. “Differential Requirement for DOCK2 in Migration of Plasmacytoid Dendritic Cells versus Myeloid Dendritic Cells.” Blood 111 (6):2973–76.

Gotoh, Kazuhito, Yoshihiko Tanaka, Akihiko Nishikimi, Risa Nakamura, Hisakata Yamada, Naoyoshi Maeda, Takahiro Ishikawa, et al. 2010. “Selective Control of Type I IFN Induction by the Rac Activator DOCK2 during TLR-Mediated Plasmacytoid Dendritic Cell Activation.” The Journal of Experimental Medicine. https://doi.org/10.1084/jem.20091776.

Hedlund, Maria, Pam Tangvoranuntakul, Hiromu Takematsu, Jeffrey M. Long, Gary D. Housley, Yasunori Kozutsumi, Akemi Suzuki, et al. 2007. “N-Glycolylneuraminic Acid Deficiency in Mice: Implications for Human Biology and Evolution.” Molecular and Cellular Biology 27 (12). Am Soc Microbiol:4340–46.

Kumar, Vivek, Kyungin Kim, Chryshanthi Joseph, Saïd Kourrich, Seung-Hee Yoo, Hung Chung Huang, Martha H. Vitaterna, et al. 2013. “C57BL/6N

Mutation in Cytoplasmic FMRP Interacting Protein 2 Regulates Cocaine Response." Science 342 (6165):1508–12.

Kunisaki, Yuya, Yoshihiko Tanaka, Terukazu Sanui, Ayumi Inayoshi, Mayuko Noda, Toshinori Nakayama, Michishige Harada, Masaru Taniguchi, Takehiko Sasazuki, and Yoshinori Fukui. 2006. "DOCK2 Is Required in T Cell Precursors for Development of Valpha14 NK T Cells." Journal of Immunology 176 (8):4640–45.

Langmead, Ben, and Steven L. Salzberg. 2012. "Fast Gapped-Read Alignment with Bowtie 2." Nature Methods 9 (4):357–59.

Mahajan, Vinay S., and Shiv Pillai. 2016. "Sialic Acids and Autoimmune Disease." Immunological Reviews 269 (1):145–61.

Mattapallil, Mary J., Eric F. Wawrousek, Chi-Chao Chan, Hui Zhao, Jayeeta Roychoudhury, Thomas A. Ferguson, and Rachel R. Caspi. 2012. "The Rd8 Mutation of the Crb1 Gene Is Present in Vendor Lines of C57BL/6N Mice and Embryonic Stem Cells, and Confounds Ocular Induced Mutant Phenotypes." Investigative Opthalmology & Visual Science 53 (6):2921.

McKenna, Aaron, Matthew Hanna, Eric Banks, Andrey Sivachenko, Kristian Cibulskis, Andrew Kernytsky, Kiran Garimella, et al. 2010. "The Genome Analysis Toolkit: A MapReduce Framework for Analyzing next-Generation DNA Sequencing Data." Genome Research 20 (9):1297–1303.

Mekada, Kazuyuki, Kuniya Abe, Ayumi Murakami, Satoe Nakamura, Hatsumi Nakata, Kazuo Moriwaki, Yuichi Obata, and Atsushi Yoshiki. 2009. "Genetic Differences among C57BL/6 Substrains." Experimental Animals / Japanese Association for Laboratory Animal Science 58 (2):141–49.

Nagy, A., J. Rossant, R. Nagy, W. Abramow-Newerly, and J. C. Roder. 1993. "Derivation of Completely Cell Culture-Derived Mice from Early-Passage Embryonic Stem Cells." Proceedings of the National Academy of Sciences of the United States of America 90 (18):8424–28.

Pillai, Shiv, Annaiah Cariappa, and Stephan P. Pirnie. 2009. "Esterases and Autoimmunity: The Sialic Acid Acetylesterase Pathway and the Regulation of Peripheral B Cell Tolerance." Trends in Immunology 30 (10):488–93.

Purtha, Whitney E., Melissa Swiecki, Marco Colonna, Michael S. Diamond, and Deepta Bhattacharya. 2012. "Spontaneous Mutation of the Dock2 Gene in Irf5−/− Mice Complicates Interpretation of Type I Interferon Production and Antibody Responses." Proceedings of National Academy of Sciences 109 (15). https://doi.org/10.1073/pnas.1118155109.

Quinlan, Aaron R., and Ira M. Hall. 2010. "BEDTools: A Flexible Suite of

Utilities for Comparing Genomic Features." Bioinformatics 26 (6):841–42.

Reheman, Adili, Subia Tasneem, Heyu Ni, and Catherine P. M. Hayward. 2010. "Mice with Deleted Multimerin 1 and Alpha-Synuclein Genes Have Impaired Platelet Adhesion and Impaired Thrombus Formation That Is Corrected by Multimerin 1." Thrombosis Research 125 (5):e177–83.

Sherry, S. T., M. H. Ward, M. Kholodov, J. Baker, L. Phan, E. M. Smigielski, and K. Sirotkin. 2001. "dbSNP: The NCBI Database of Genetic Variation." Nucleic Acids Research 29 (1):308–11.

Skarnes, William C., Barry Rosen, Anthony P. West, Manousos Koutsourakis, Wendy Bushell, Vivek Iyer, Alejandro O. Mujica, et al. 2011. "A Conditional Knockout Resource for the Genome-Wide Study of Mouse Gene Function." Nature 474 (7351):337–42.

Specht, C. G., and R. Schoepfer. 2001. "Deletion of the Alpha-Synuclein Locus in a Subpopulation of C57BL/6J Inbred Mice." BMC Neuroscience 2 (August):11.

Taft, Robert A., Muriel Davisson, and Michael V. Wiles. 2006. "Know Thy Mouse." Trends in Genetics: TIG 22 (12):649–53.

Takaoka, Akinori, Hideyuki Yanai, Seiji Kondo, Gordon Duncan, Hideo Negishi, Tatsuaki Mizutani, Shin-Ichi Kano, et al. 2005. "Integral Role of IRF-5 in the Gene Induction Programme Activated by Toll-like Receptors." Nature 434 (7030):243–49.

Watkins-Chow, Dawn E., and William J. Pavan. 2008. "Genomic Copy Number and Expression Variation within the C57BL/6J Inbred Mouse Strain." Genome Research 18 (1):60–66.

Yalcin, Binnaz, David J. Adams, Jonathan Flint, and Thomas M. Keane. 2012. "Next-Generation Sequencing of Experimental Mouse Strains." Mammalian Genome: Official Journal of the International Mammalian Genome Society 23 (9-10):490–98.

Yasuda, Kei, Kerstin Nündel, Amanda A. Watkins, Tania Dhawan, Ramon G. Bonegio, Jessalyn M. Ubellacker, Ann Marshak-Rothstein, and Ian R. Rifkin. 2013. "Phenotype and Function of B Cells and Dendritic Cells from Interferon Regulatory Factor 5-Deficient Mice with and without a Mutation in DOCK2." International Immunology 25 (5). Jpn Soc Immunol:295–306.

Yasuda, K., C. Richez, J. W. Maciaszek, N. Agrawal, S. Akira, A. Marshak-Rothstein, and I. R. Rifkin. 2007. "Murine Dendritic Cell Type I IFN Production Induced by Human IgG-RNA Immune Complexes Is IFN Regulatory Factor (IRF)5 and IRF7 Dependent and Is Required for IL-6

Production.” Journal of Immunology 178 (11):6876–85.

Zurita, Esther, Mónica Chagoyen, Marta Cantero, Rosario Alonso, Anna González-Neira, Alejandro López-Jiménez, José Antonio López-Moreno, et al. 2011. “Genetic Polymorphisms among C57BL/6 Mouse Inbred Strains.” Transgenic Research 20 (3):481–89.

Chapter 3: TCR and DOCK2

We previously demonstrated that a loss function *Dock2* variant was introduced into several knockout mouse strains during backcrossing. In this study, show that loss of DOCK2 results in a 3-5 fold expansion of virtual memory cells that correlates with increased resistance to intracellular infection. Bone marrow chimera and adoptive transfer studies indicate that these cells arise in a cell intrinsic manner following thymic egress. In addition, close inspection of their transcriptional profile, TCR repertoire and cell surface marker expression shows that *Dock2* deficient naive $CD8^+$ T cells directly convert into virtual memory cells bypassing the effector T cell stage. This phenomenon also occurs in the context of a restricted TCR repertoire as DOCK2 deficient T cells expressing the OT1 transgene show a similarly dramatic expansion of their memory compartment. This direct conversion to memory is associated with TCR hypersensitivity to ex vivo weak agonist peptide and in vivo self-peptide triggering. Collectively, these findings suggest that DOCK2 sets the threshold for entry into the “virtual” memory compartment by negatively regulating tonic TCR triggering.

3.1 Introduction

$CD8^+$ T cells are critical effectors of cellular immunity. A cardinal feature of T cell responses is the establishment of long lived memory that rapidly responds during subsequent challenge with the same antigen. The widely held view of memory formation posits that cognate peptide activation drives naive T cell clonal expansion and subsequent contraction leaving behind memory cells that persist following the clearance of an acute infection (Kaech et al. 2003; Yu et al. 2017; Akondy et al. 2017).

CD8 T cell memory cell formation that is not accompanied by effector cell differentiation but sharing most conventional memory transcriptional and functional traits has also been extensively documented (Jameson 2002; Ananda W. Goldrath et al. 2004; Hamilton and Jameson 2008; Cheung, Yang, and Goldrath 2009). Cognate antigen independent memory cells are also found in the sterile environment of human fetal spleen and cord blood (Zhang et al. 2014; Jacomet et al. 2015; Van Kaer 2015) suggesting that this is an evolutionary conserved bonafide naive T cell differentiation pathway. This phenomenon, initially uncovered by adoptively transferring naive T cells into lymphopenic settings such as neonatal and irradiated mice (A. W. Goldrath and Bevan 1999; Ernst et al. 1999; Min et al. 2003), requires a combination of common γ-chain cytokine (mainly by IL-7) and self peptide TCR triggering (J.

T. Tan et al. 2001). These and other studies suggested that a transient increase in “survival” signals drive such memory cell differentiation (Sprent and Surh 2011). Consistent with this idea, there is a direct correlation between homeostatic conversion to memory and the strength of tonic self peptide signaling that a T cell is receiving (Kieper, Burghardt, and Surh 2004).

Recently, several studies have described cognate antigen-independent $CD8^+$ T cell memory that arises without any apparent lymphopenia (White, Cross, and Kedl 2017; Haluszczak et al. 2009). In mice, this type of memory can be divided into two categories: “innate” and “virtual” memory. Innate memory cells arise in the thymus and are dependent on IL-4 production from $PLZF^+$ NKT cells (Lee, Jameson, and Hogquist 2011). Virtual memory cells arise in the periphery requiring common γ-chain cytokines (mainly IL-15 and IL-4) (Sosinowski et al. 2013; Akue, Lee, and Jameson 2012; Fiege, Burbach, and Shimizu 2015) and self-peptide TCR triggering (White et al. 2016), in a manner reminiscent of lymphopenia induced memory cells. These cells are also seen in germ free mice consistent with their dependence on homeostatic signals for generation and survival. Bystander virtual memory cells also provide innate-like protection in the of context intracellular infection (White et al. 2016) similar to conventional memory T cells in mice (R. E. Berg et al. 2003; Chu et al. 2013) and humans (Kim et al. 2017). Interestingly, very little is

known about negative regulators of the $CD8^+$ T cell virtual memory compartment (Fiege, Burbach, and Shimizu 2015).

DOCK2, a hematopoietic restricted guanine exchange factor (GEF), activates the Rho GTPase Rac by catalyzing the transition from the inactive GDP-bound state to the active GTP-bound state. DOCK2 ensures spatially controlled activation of Rac by localizing to cell membrane via DHR1 domain mediated interactions with PIP3 and polybasic amino acid cluster based interactions with phosphatidic acid (Fukui et al. 2001; Nishikimi et al. 2009, 2013). GTP bound RAC1 subsequently drives actin polymerization enabling cytoskeletal rearrangements required for lymphocyte chemotaxis (Fukui et al. 2001; Terasawa et al. 2012; Gotoh et al. 2008), T cell interstitial motility (Nombela-Arrieta et al. 2007), plasmacytoid dendritic cell cytokine secretion (Gotoh et al. 2010), and TCR activation (Sanui et al. 2003; Le Floc'h et al. 2013). In $CD8^+$ T cells, these studies used relatively strong stimuli, i.e. anti-CD3 or cognate antigen, to identify DOCK2 as a negative regulator of TCR clustering and immune synapse formation. In addition, Sanui et al. used a weak agonist against the MHC-II restricted 2B4 TCR transgene to show a defect in the proliferative response of stimulated *Dock2*$^{-/-}$ $CD4^+$ T-cells (Sanui et al. 2003). However, it is likely that DOCK2's role in regulating lymphocyte proliferation independent of the TCR complicates interpretation of this experiment (Wang et al. 2010).

In an earlier study, we used the inheritance pattern of expanded memory T cells in the blood to identify a loss function *Dock2* allele ($Dock2^{hsd/hsd}$) in several independently generated knockout mice. We further went on to show that this allele was introduced during backcrossing into the commercially available C57BL/6NHsd substrain (Mahajan et al. 2016). In this study, we show that the memory phenotype cells expanded in the absence of *Dock2* are polyclonal virtual memory cells. Interestingly, the differentiation of these cells is not dependent on the presence of a polyclonal repertoire as they are also seen in OT1 transgenic $Dock2^{hsd/hsd}$ mice. In addition, bone marrow chimera and adoptive transfer studies suggest that these cells arise in a cell intrinsic manner following thymic egress. Importantly, a close inspection of their surface phenotype, transcriptional profile and TCR repertoire suggests that $Dock2^{hsd/hsd}$ naive T cells are directly converting into memory cells in the absence of cognate antigen stimulation. Importantly, such direct conversion to memory is associated with hyperresponsiveness to self peptide-MHC in lymphoreplete and lymphopenic environments as well as ex vivo TCR triggering with weak peptide agonists. Collectively, these findings suggest that the absence of DOCK2 lowers the threshold of self-peptide triggering required to enter the virtual memory cell compartment

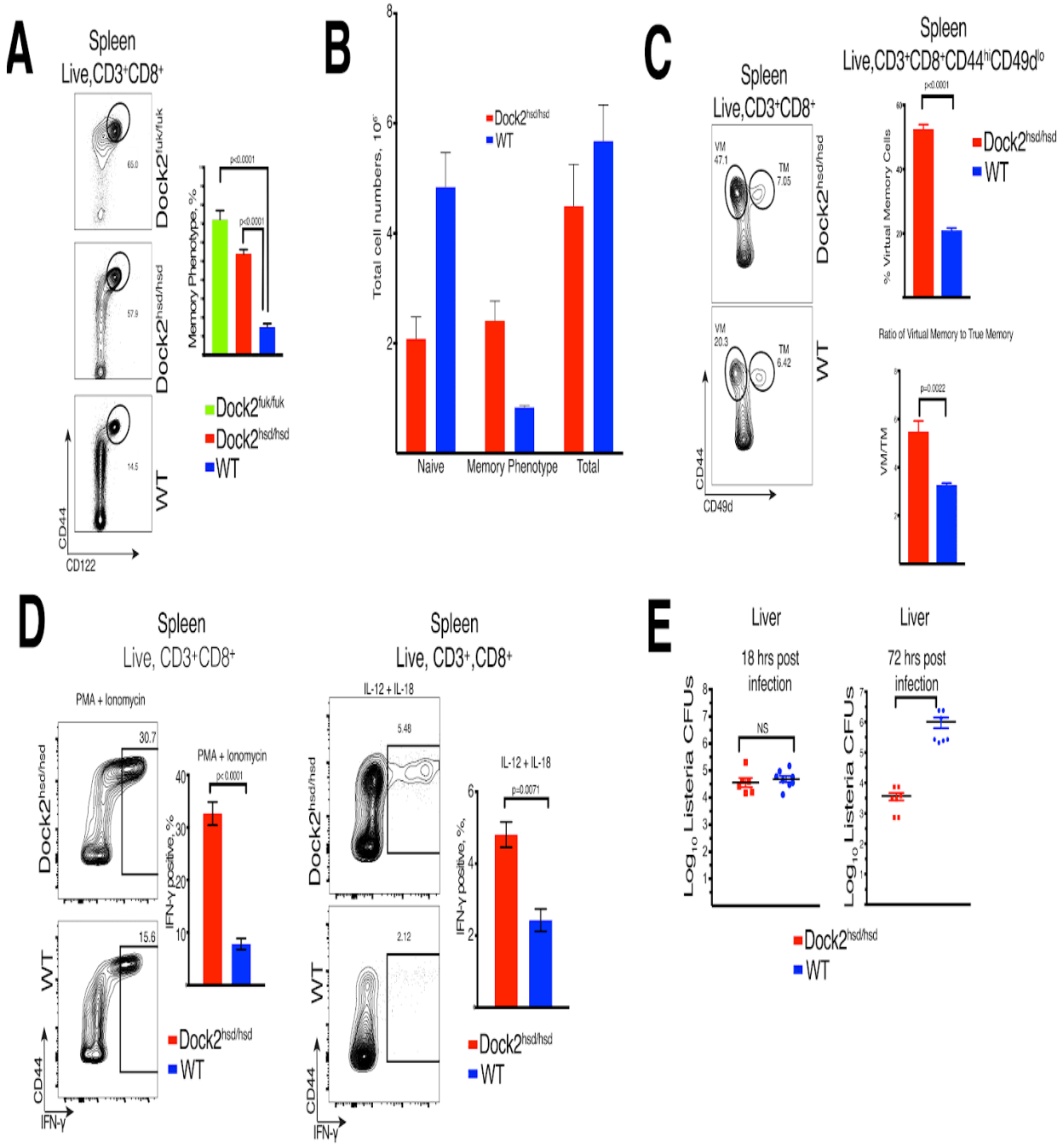
3.2 Results

3.2.1 Absence of Dock2 results in expansion of virtual memory T cells

In an earlier study, we identified an approximately 3 fold expansion of memory phenotype cells in mice carrying spontaneous loss of function mutation in the guanine exchange factor DOCK2 (*Dock2*$^{hsd/hsd}$) (Mahajan et al. 2016). This expansion isn't specific to this particular allele of *Dock2* as it is also present in gene targeted Dock2 deficient mice (Fukui et al. 2001) (Figure 12A). The increase in *Dock2*$^{hsd/hsd}$ memory phenotype cell percentage is also mirrored by a similar increase in total numbers of such cells (Figure 12B). It is important to note that these cells lack surface expression of NK1.1 but express both CD8α and CD8β (Data not shown).

Figure 12: Dockhsd/hsd mice have an expanded virtual T cell compartment and are resistant (A) The proportion of CD8+ memory phenotype cells (MP) in the spleens of Dock2hsd/hsd, DockFukui/Fukui and WT mice as measured by flow cytometry. (B) Number of splenic naive (CD44loCD122), memory phenotype, MP (CD44hiCD122+) and total T cells in indicated mice. (C) Percentage and ratio of virtual memory (VM) (CD44hiCD49dlo) T cells relative to true memory (TM) (CD44hiCD49dhi) T cells among total splenic T cells from indicated mice as assessed by flow cytometry. (D) CD8 T cells stimulated for four hours with PMAIonomycin or IL12+IL18 were analyzed for interferonγ production by intracellular staining. (E) Listeria (LM1043S) infected mice were sacked at indicated time points followed by measurement of bacterial CFUs in the liver. Representative or pooled data from experiments done twice with groups of 3 or 4 mice is shown. Data was analyzed for statistical significance by unpaired twotailed Student's t test using Prism (GraphPad Software).

Figure 12 (Continued)



Recent studies have identified cognate antigen independent memory cells that arise in unmanipulated lymporeplete mice. (White, Cross, and Kedl 2017; Haluszczak et al. 2009; Akue, Lee, and Jameson 2012). Such “virtual” memory cells can be distinguished from conventional “true” memory cells by their low expression of CD49d (Haluszczak et al. 2009). Thus, we used surface staining of Dock2$^{hsd/hsd}$ splenic T cells to show that the majority of expanded memory phenotype cells phenotypically resemble virtual memory cells (Figure 12C). Although there is a slight increase in true memory cell percentages, the ratio of virtual to true memory is significantly increased in the absence of DOCK2 (Figure 12C).

3.2.2 *Dock2*$^{hsd/hsd}$ virtual memory cells are functional and their presence correlates with protection from intracellular infection

A key feature of true memory cells is rapid cognate antigen independent production of interferon (IFN)γ that is protective following intracellular bacterial infection (R. E. Berg, Cordes, and Forman 2002; R. E. Berg et al. 2003; Soudja et al. 2012; Chu et al. 2013). Virtual memory T cells can similarly undergo bystander activation and produce protective IFNγ in Listeria infected mice (White et al. 2016). Consistent with these studies, a higher proportion of *Dock-Hsd* T cells rapidly respond to in vitro polyclonal stimuli (PMA+Ionomycin or IL-12+IL-18) by secreting IFNγ (Figure 12D).

Virtual memory cells extensively traffic to the liver. (White et al. 2016). This led us to hypothesize that increased IFNγ production by such cells provides innate like protection in the context of listeria infection, a bacterium that preferentially replicates in the liver. Indeed, the virtual memory expansion in the absence of DOCK2 correlated with increased resistance to intracellular infection as *Dock2*$^{hsd/hsd}$ mice show significantly lower bacterial burden 3 days after i.v. infection with LM1043S (Figure 12E). Importantly, there are no differences between wild type and *Dock2*$^{hsd/hsd}$ liver bacterial CFUs 18 hours post-infection, consistent with published studies showing that the protective effect of CD8 derived IFNγ doesn't manifest itself until 3 days after infection (R. E. Berg et al. 2003).

3.2.3 *Dock2*$^{hsd/hsd}$ memory phenotype T cells arise a hematopoietic intrinsic manner following thymic egress

Several studies have identified key roles for radioresistant stromal cells in maintaining peripheral T cell homeostasis (Roozendaal and Mebius 2011). With this in mind, we attempted to rule out any non-hematopoietic role in behind the expansion of virtual cells by transferring bone marrow into irradiated RAG deficient recipients. This experiments revealed that only recipients of Dock2$^{hsd/hsd}$ bone marrow have a robust expansion of memory

cells while mice that receive wild type bone marrow have a much smaller virtual compartment (Figure 13A). Mixed chimeras, with wild type and *Dock2*$^{hsd/hsd}$ mice bone marrow injected into same recipient, are not possible due to a severe defect in competitive lymphoid reconstitution seen with *Dock2* deficient precursors (Kikuchi et al. 2008).

Dock2$^{hsd/hsd}$ virtual memory cells arise in the periphery as there are no mature CD8^{+} single positive CD44^{hi} cells in the thymus (Figure 13B), in contrast to recently described innate CD8 T cells which are initially generated prior to thymic egress (L. J. Berg 2007; Lee, Jameson, and Hogquist 2011). Conventional true memory cells are generated following clonal activation, effector expansion and contraction to form long lived memory (Kaech and Cui 2012). Consistent with the lack of expression of CD49d and markers of early effector activation (CD69 and CD25), there is no significant increase in the percentage of cycling naive or memory T cells in DOCK2 deficient mice (Figure 13C & 13D).

3.2.4 *Dock2*$^{hsd/hsd}$ naive T cells directly convert into memory phenotype cells

In order to get some insight into the mechanism underlying the enhanced generation of *Dock2*$^{hsd/hsd}$ virtual memory cells, we compared the

transcriptional profile of DOCK2 deficient and wild type naive T cells. Interestingly, we found that genes upregulated in *Dock2*$^{hsd/hsd}$ naive T cells are significantly enriched for several gene sets. associated with memory T cell differentiation (Figure 14A). A more focussed examination of individual genes upregulated in the absence of DOCK2 also revealed that transcription factors associated with memory such as *Tcf7* and *IL-7R* had higher expression in *Dock2*$^{hsd/hsd}$ naive T cells (data not shown). Consistent with direct conversion of such cells to memory T cells bypassing effector T cell clonal expansion, we also didn't observe any gene sets associated with early T cell activation.

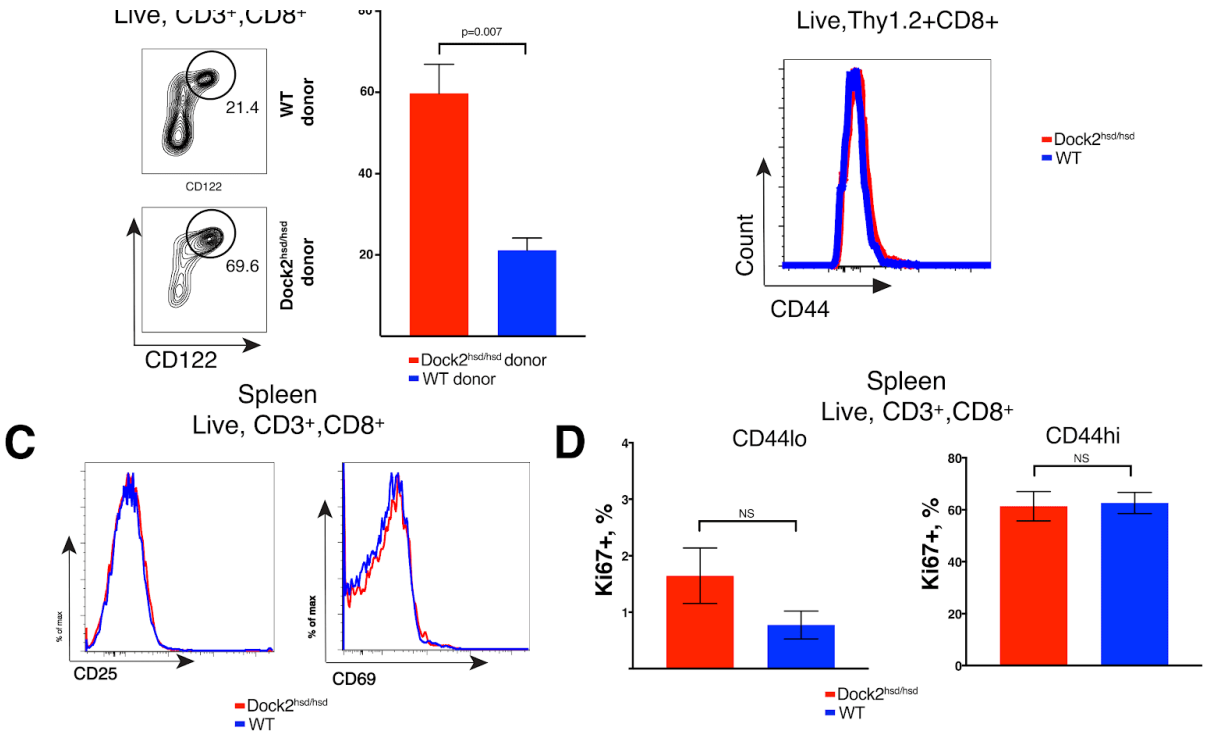


Figure 13: Dockhsd/hsd virtual memory T cells arise in a hematopoietic intrinsic manner in the periphery (A) Bone marrow from either Dock2hsd/hsd or WT mice was injected into irradiated RAG1/ mice. Recipient mice were examined for spontaneous memory formation 12 weeks after injection. (B) Thymocytes from indicated mice were examined for the presence of CD8 single positive CD44hi memory cells. (C) Splenic T cells were assessed for markers of activation and (D) cell cycle status. Representative or pooled data from experiments done twice with groups of 3 or 4 mice is shown. Data was analyzed for statistical significance by unpaired twotailed Student's t test using Prism (GraphPad Software).

Consistent with a lack of effector activation, TCR sequencing of *Dock2*$^{hsd/hsd}$ CD8 T cells revealed a very diverse repertoire with no dominant clonal expansions (Figure 14B). One measure of subtle peripheral self-antigen selection is convergence, the number of unique CDR3 sequences coding for the same amino acid sequence (Shugay et al. 2015). Analysis of the *Dock2*$^{hsd/hsd}$ memory T cell repertoire revealed a significantly lower number of unique CDR3 nucleotide sequences that code for the same amino acid sequence compared with wild type memory cells (Figure 14C). This decrease in convergence suggests that the self-antigen affinity threshold for entering the virtual memory compartment is lowered in the absence of DOCK2 allowing more “naive” TCRs to enter this compartment. Further supporting this conclusion, we found that TCR Vβ expression profile of *Dock2*$^{hsd/hsd}$ virtual memory cells bears remarkable similarity to naive T cells and is quite distinct from wild type virtual memory cells (Figure 14D).

In order to determine whether Dock2$^{hsd/hsd}$ naive T cells have an increased cell intrinsic propensity to convert to virtual memory, we investigated how adoptively co-transferred Dock2$^{hsd/hsd}$ and WT naive T cells responded to homeostatic signals *in vivo*. As can be seen in figure 14, *Dock2*$^{hsd/hsd}$ T cells proliferate at a strikingly faster rate, completely diluting CFSE, than co-transferred WT T cells with concomitant CD44 upregulation when injected

Figure 14: Dock2hsd/hsd T cells directly convert to virtual memory (A) GSEA plots of gene sets enriched in a list of genes upregulated in Dock2hsd/hsd naive (CD44loCD122) T cells when compared to WT naive T cells. (B) Clone frequency of individual TCRs of naive (CD44loCD122) and MP (CD44hiCD122+) T cells FACS sorted from Dock2hsd/hsd and WT mice. (C) Comparison of TCR repertoire convergence (number of unique CDR3 sequences that code for the same amino acid sequence) between Dock2hsd/hsd and WT mice naive (CD44loCD122) and MP (CD44hiCD122+) T cells. (D) Hierarchical clustering of naive (CD44loCD122) and MP (CD44hiCD122+) T cells from Dockhsd/hsd and WT mice based on TCR Vβ usage. (E) Naive T cells fromCD45.2+ Thy1.2+ Dock2hsd/hsd and CD45.2+ Thy1.1+ WT mice were cotransferred into irradiated CD45.1+ Thy1.2+ lymphopenic mice and assessed for CFSE dilution and CD44 upregulation after 1 week. (F) Naive T cells from CD45.2+ Thy1.2+ Dock2hsd/hsd and CD45.2+ Thy1.1+ WT mice were cotransferred into unmanipulated lymphoreplete CD45.1+ Thy1.2+ mice and assessed for upregulation of memory markers after 3 weeks. (G) Surface CD5 and TCR levels of T cells from Dock2hsd/hsd and WT T cells as assessed by flow cytometry. Representative or pooled data from experiments done twice with groups of 3 or 4 mice is shown. Data was analyzed for statistical significance by unpaired twotailed Student's t test using Prism (GraphPad Software).

Figure 14 (Continued)

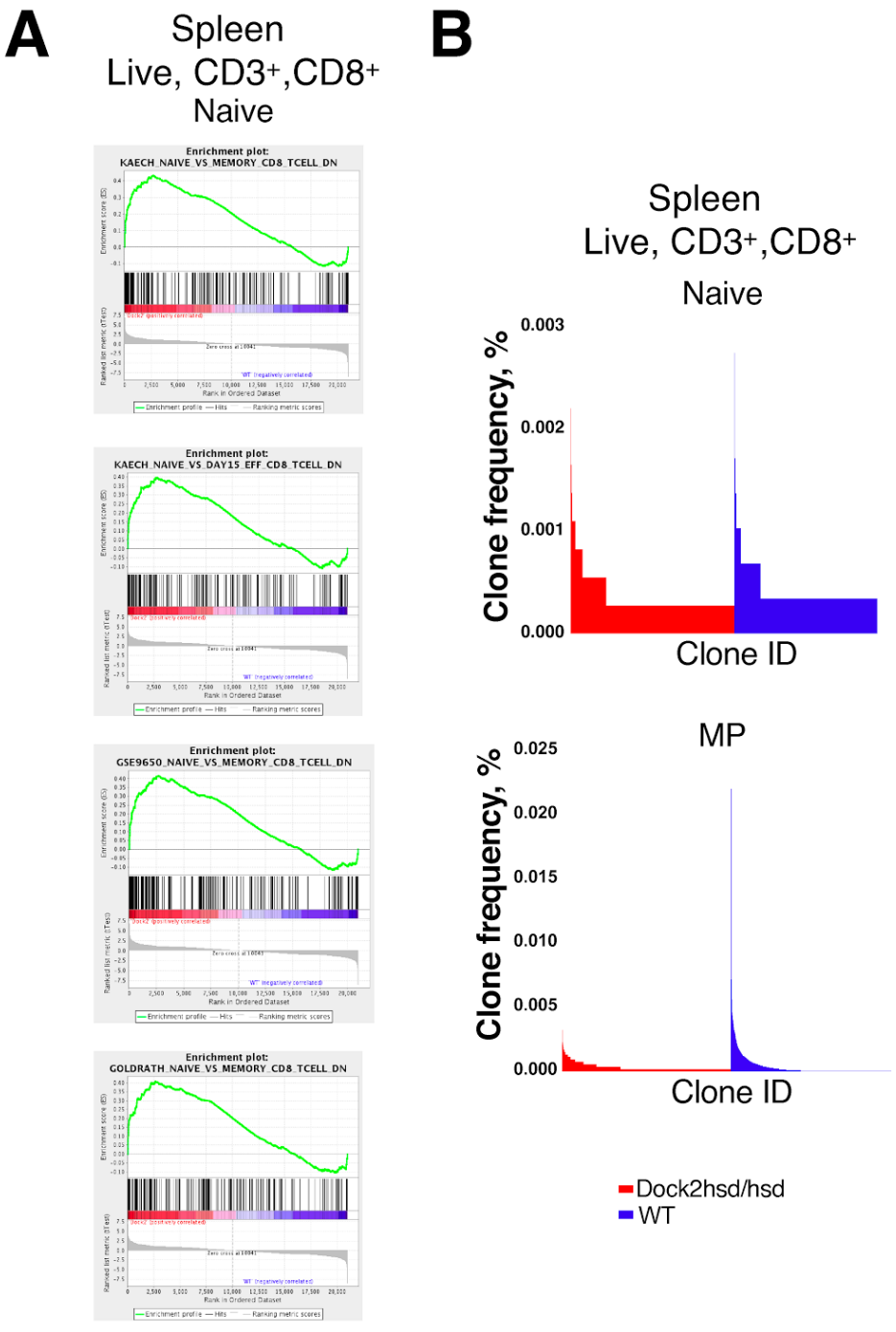
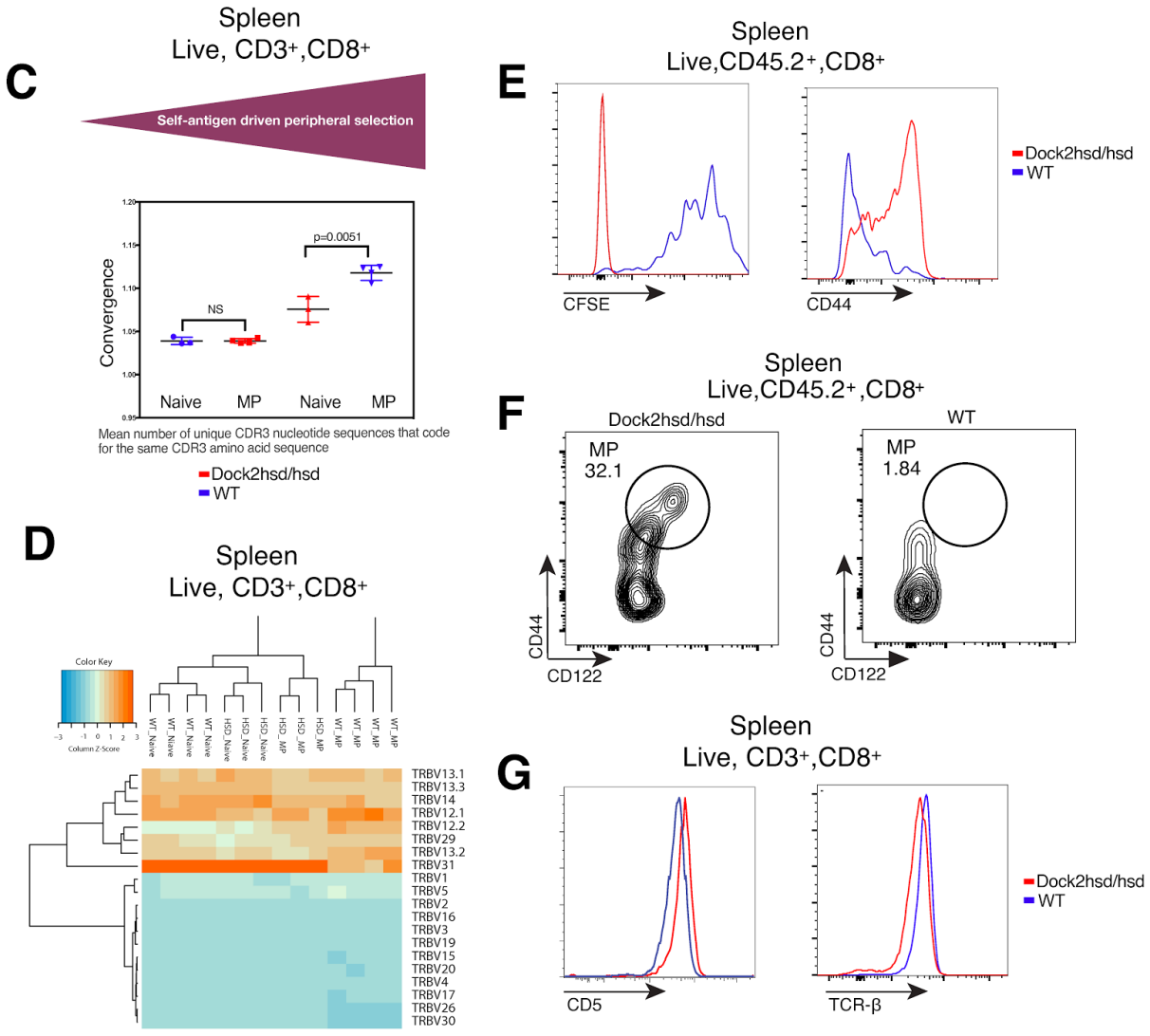


Figure 14 (Continued)



into irradiated lymphopenic mice (Figure 14E). Importantly, this increased rate of conversion to memory is also seen when naive *Dock2*$^{hsd/hsd}$ T cells are transferred to a lymphoreplete mouse without any irradiation (Figure 14F).

Recent studies have suggested that IL-15 and IL-4 derived from dendritic cells and iNKT cells respectively control the size of the peripheral memory like T cell compartment (Tripathi et al. 2016; Kurzweil, LaRoche, and Oliver 2014; Akue, Lee, and Jameson 2012; Sosinowski et al. 2013). The expansion of virtual memory cells in Dock2$^{hsd/hsd}$ mice, however, is unlikely to be secondary to changes in cytokine producing cells as the number of iNKT cells are reduced (Mahajan et al. 2016) while there are no obvious changes in CD8α dendritic cell homeostasis in these mice (Data not shown).

Surface CD5 levels, a proxy for tonic TCR signalling (Azzam et al. 1998, 2001), were significantly higher on naive T cells from Dock2$^{hsd/hsd}$ mice (Figure 14G). In addition, there is lower surface TCR expression in Dock2$^{hsd/hsd}$ T cells (Figure 14G). These findings suggested that an increase in self peptide-MHC triggering is associated with the expansion of memory phenotype cells. Interestingly, increased CD5 levels are also seen in *Dock2*$^{hsd/hsd}$ CD8 single positive thymocytes suggesting that the such cells are receiving stronger positively selecting signals. However, this isn't

accompanied with dramatic changes in proportions of thymic precursors (Figure 15).

3.2.5 *Dock2*$^{hsd/hsd}$ OT1 transgenic mice have an expanded memory phenotype compartment and increased ex vivo responses to weak TCR agonists

As our aforementioned experiments implicated tonic TCR signalling in driving the conversion to virtual memory, we generated DOCK2 deficient OT1-TCR transgenic mice (OT1-*Dock2*$^{hsd/hsd}$) to dissect this phenomena in context of a restricted TCR repertoire. OT1-*Dock2*$^{hsd/hsd}$ mice were born at mendelian ratios and exhibited no obvious developmental defects. A close examination of T cells in the blood of these mice, however, revealed a striking accumulation of virtual memory cells compared to OT1 mice with normal expression of DOCK2 (Figure 16A). Thus, the DOCK2 dependent expansion of CD8+ T cells is independent of a polyclonal repertoire. Similar our findings in mice that have a diverse repertoire, there is also lower surface TCR (Vα2) and no increased CD69 expression in OT1-*Dock2*$^{hsd/hsd}$ mice (Data not shown).

These findings led us to explore the intriguing possibility that DOCK2 directly regulates ex vivo TCR triggering in CD8 T cells. We designed

experiments to test T cell responses to stimuli across a wide range of affinities by taking advantage of altered peptide ligands for the OT1 TCR transgene. The peptides we chose, in increasing order of affinity, are SIIGFEKL (G4), SIIQFEKL (Q4) SIITFEKL (T4) and SIINFEKL (N4). (Salmond et al. 2014). As a proxy for early TCR triggering, we decided to measure TCR downregulation (Itoh et al. 1999) and CD69 upregulation (Cibrián and Sánchez-Madrid 2017). Consistent with our in vivo findings, ex vivo peptide stimulation of DOCK2 deficient T cells resulted in significantly more TCR downregulation when compared to cells from *Docks*$^{+/hsd}$ littermates (Figure 16B). Further, the per cell expression of CD69 is much higher in DOCK2 deficient T cells (Figure 16B and Supplementary Figure 17). Importantly, stimulation of OT1-*Dock2*$^{hsd/hsd}$ with a very weak ligand, SIIGFEKL (G4), resulted in pronounced increase in the percentage of CD69 positive cells compared with OT1 T cells (Figure 16B).

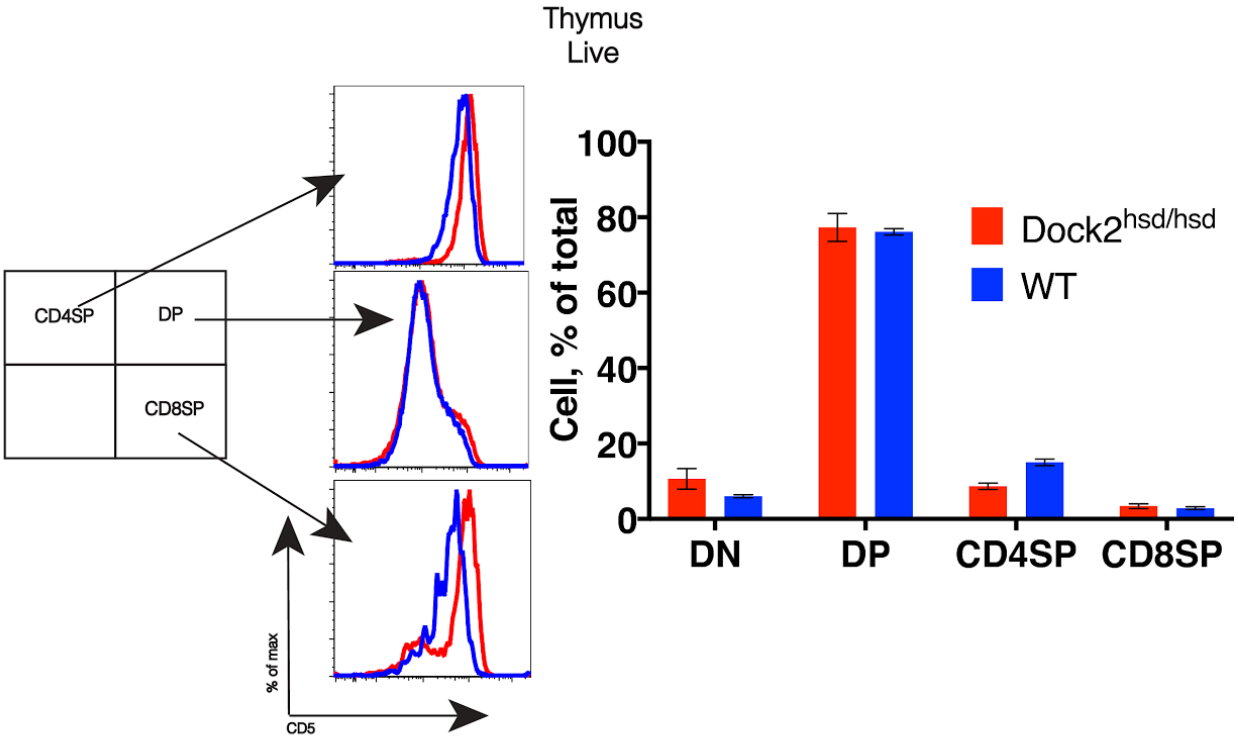
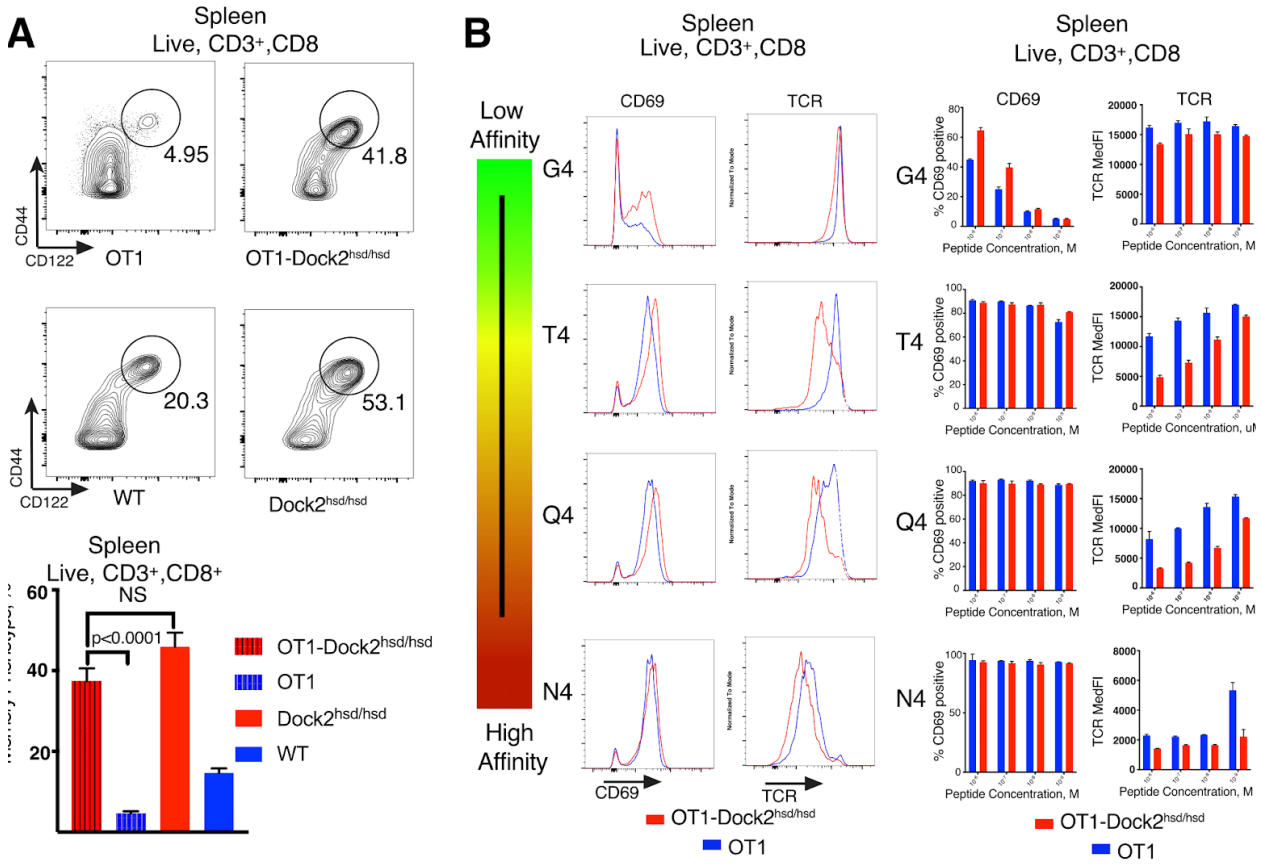


Figure 15: CD5 expression and proportion of T cell precursors in the thymus.

Figure 16 : OT1 Dock2hsd/hsd mice have an expansion of memory phenotype (MP) T cells and show increased ex vivo responses to TCR stimulation (A) The proportion of CD8+ memory phenotype cells, MP (CD44hiCD122+) in the spleens of indicated mice. (B) TCR responses to ex vivo stimulation as measured by CD69 upregulation and TCR downregulation following incubation for four hours with the indicated peptides. FACS plots show representative responses to stimulation with 1μM of peptide while bar graphs show responses across a range of concentrations. Peptides are arranged, top to bottom, in increasing order of affinity to the OT1 TCR. Representative or pooled data from experiments done twice with groups of 3 or 4 mice is shown.

Figure 16 (Continued)



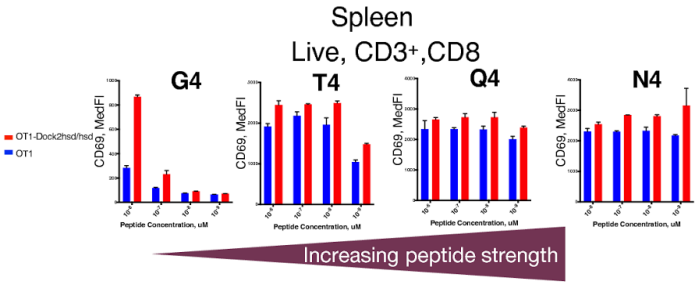


Figure 17: TCR responses to ex vivo stimulation as measured by CD69 median fluorescence intensity. Peptides are arranged, right to left, in increasing order of affinity to the OT1 TCR.

3.3 Discussion

We took advantage of a prominent expansion of $CD8^+$ blood memory phenotype cells to identify a loss function *Dock2* variant that was inadvertently introduced into several knockout mice during backcrossing into the commercially available C57BL6/NHsd substrain. In this study, we show that $CD8^+$ T cell memory expansion is also found *Dock2* knockout gene targeted mice. In contrast to antigen experienced memory cells, the majority of spontaneously arisen $Dock2^{hsd/hsd}$ memory T cells have low surface expression of CD49d similar to virtual memory cells. This, combined with their peripheral origin, lack of activation makers and absence of dominant clones, leads us to conclude that memory phenotype cells seen in $Dock2^{hsd/hsd}$ mice are virtual memory cells.

Virtual memory cells rapidly respond to innate cytokine and adaptive TCR triggering conferring substantial protection. This is especially valuable in immunodeficient hosts particularly when the T cell compartment is compromised. For instance, absence of *Dock2* in mice and humans leads to immunodeficiency with increased susceptibility to certain infections (Liu et al. 2016; Dobbs et al. 2015). However, we find that $Dock2^{hsd/hsd}$ mice are more resistant to Listeria infection consistent with increased ex vivo innate like IFNγ

production by T cells from these mice. Interestingly, resistance from infection manifests itself 3 days post infection ruling out most innate mechanisms of protection while being consistent with kinetics of innate protection by memory T cells (R. E. Berg et al. 2003). However, T cell depletion and adoptive transfer studies are currently underway to rule out contributions from other cells. Regardless, it is tempting to speculate that the overlap between the replicative niche of this bacterium and extensive trafficking of virtual memory cells to the liver (White et al. 2016) is responsible for this phenotype. It is possible that DOCK2 loss results in increased susceptibility to pathogens replicating at anatomical sites where virtual memory cells are found at lower frequency prior to infection.

Virtual memory cell generation and maintenance is heavily dependent on TCR and common γ-chain cytokine triggering. However, there is limited understanding of negative regulators of the the virtual memory compartment. In this study, we provide evidence that DOCK2 sets threshold for direct entry into the virtual memory compartment. We show that there is (i) reduced convergence and naive like TCR Vβ usage in the *Dock2*$^{hsd/hsd}$ virtual memory TCR repertoire suggesting that more “naive” TCRs able to enter this compartment; (ii) a striking enrichment gene sets associated with memory differentiation in genes upregulated in *Dock2*$^{hsd/hsd}$ naive T cells relative to WT naive T cells . Notably, there is no enrichment of gene sets associated

with effector differentiation; and (iii) an increased propensity of adoptively co-transferred *Dock2*$^{hsd/hsd}$ naive T cells to convert to memory phenotype in response homeostatic signals. In contrast to Dock2 gene targeted mice,which show a substantial T cell lymphopenia (Fukui et al. 2001), total T cell numbers are unchanged in *Dock*$^{hsd/hsd}$ mice. As described in our earlier study, we also find a substantial number of T cells in the blood of *Dock*$^{hsd/hsd}$ mice, again contrasting with gene targeted mice. One possible explanation for this discordance is that *Dock*$^{hsd/hsd}$ is a hypomorphic allele retaining some residual function.

The TCR is one source of homeostatic signals promoting virtual memory formation. Profiling of *Dock2*$^{hsd/hsd}$ naive T cell for reporters of tonic signalling, CD5 upregulation (Azzam et al. 2001) and TCR downregulation (Itoh et al. 1999), suggested that conversion to memory is associated with increased TCR responsiveness. We further probed this in the context of a restricted repertoire by crossing *Dock2*$^{hsd/hsd}$ mice to OT1 transgenic mice. Interestingly, we find the expansion of virtual memory cells is independent of a polyclonal repertoire as there is a robust expansion of cells with memory markers in *Dock*$^{hsd/hsd}$-OT1 mice whch accompanied with TCR downregulation. To directly assess the consequence of DOCK2 loss on TCR triggering, we tested several altered peptide ligands with varying levels of affinity to the OT-1 TCR. These studies revealed a pronounced decrease in

surface TCR levels and a simultaneous increase per cell levels of surface CD69. Importantly, stimulation with a very weak agonist peptide (SIIGFEKL) results in a pronounced increase in the percentage of cells upregulating CD69, consistent with our in vivo findings of increased self peptide-MHC triggering. These finding leads us to conclude that increased tonic TCR triggering in the absence of DOCK2 likely drives the direct conversion of naive T cells in to virtual memory cells. As levels of self peptide triggering are associated with increases T cell sensitivity to common γ-chain cytokines, future studies will need to address the effect DOCK2 loss has on responses to such stimuli.

It is currently unclear how loss of DOCK2 sets the threshold for weak agonist TCR stimulation. One possibility is that the decreased interstitial motility (Nombela-Arrieta et al. 2007) observed by *Dock2*$^{-/-}$ T cells results in increased TCR-MHC contact duration. It is possible that increases in antigen presenting cell “residency” time results in increased amounts of TCR triggering and concomitant conversion of to memory with weak peptides. A recent elegant study has shown that a cortical actin network restricts the initiation of TCR triggering in the absence of CD28 dependent costimulation (Y. X. Tan et al. 2014). Given its role in actin polymerization via activation of RAC, it is possible that loss of DOCK2 disrupts this cortical network lowering the threshold for activation. Intriguingly, loss of cortical actin integrity might

also promote the formation of activation promoting TCR and LAT microclusters (Lillemeier et al. 2010). With this in mind, it will be interesting to investigate plasma membrane architecture *Dock2*$^{hsd/hsd}$ mice. Studies on the role of RAC in TCR signaling and virtual memory differentiation following egress into the periphery should also be informative as current studies of this GTPase have been, with the exception of homing, largely focussed on thymic development (Guo et al. 2008; Dumont et al. 2009).

In conclusion, we show a novel role for DOCK2 in restricting the size of the virtual memory T cell compartment most likely by setting the threshold for responses against weak agonists. These findings suggest that efforts to dampen immune responses using a small molecule inhibitor of DOCK2 should tempered with an understanding that this protein has pleiotropic effects on peripheral T cell homeostasis (Nishikimi et al. 2012). However, they also point to the possibility of modulating the size of the virtual memory compartment by targeting DOCK2 using such an approach.

3.4 Materials and Methods

Mice

Dockhsd/hsd mice were purchased from Harlan laboratories and maintained as a separate colony in specific pathogen free environments. C57BL/6J, OT1 and RAG1 knockout mice were purchased from Jackson labs.

Bone Marrow Chimeras

Bone marrow was isolated from the indicated mice. Pooled bone marrow was depleted of mature T cells using biotinylated antibody mediated magnetic enrichment. Cells were then resuspended in PBS and 1x106 cells were transferred to each mouse.

Cell Transfers

For lymphopenia induced proliferation experiments, mice were irradiated at a dose of 600 cGy. 6 hours later, a total of 1X106 CFSE labeled naive T cells from $Dock^{hsd/hsd}$ and WT mice were co-injected into irradiated congenic hosts. 1 week later, transferred cells were recovered and assessed for CFSE dilution and CD44 upregulation.

For experiments in lymphoreplete mice, a total of 2x106 naive T cells from $Dock^{hsd/hsd}$ and WT mice were co-injected into unmanipulated congenic hosts. 3 weeks later, transferred cells were assessed for CD44 and CD122 upregulation.

RNA-Sequencing

RNA was isolated with QIAGEN RNA isolation kits according to the manufacturer instructions. RNA-Seq libraries were then prepared using a modified version of a previously published Smart-Seq2 protocol (Picelli et al. 2013). Libraries were sequence on NextSeq sequencer using V2 kits (number of cycles). Reads were aligned using RSEM package and differentially expressed genes identified by EBSEQ with PPDE cutoff of 0.95. TCR repertoire libraries were prepared using commercially available irepertoire kits. VD and J segment assignment and clonotypes identification was performed using previously published MiXCR pipeline (Bolotin et al. 2015). VDJtools was used for post analysis determination of convergence and hierarchical clustering of samples based on TCR Vβ usage.

T cell stimulation

Ex vivo polyclonal stimulation was performed by incubating T cells in cell stimulation cocktail from ebioscience (cat: 00-4975-03) for 4 hours. Cytokines were purchased from peprotech and used at a concentration of 10 ng/ml for 18 hours. Brfelding A (cat: 00-4506-51) was added for the last 4 hours of cytokine culture according to manufacturer's instruction. For OT1-TCR

stimulation altered peptide ligands were purchased from anaspec and used at the indicated concentrations to stimulate cells for 4 hours.

Listeria Infection.

Mice were intravenoulsy infected with 2x104 CFUs of LM1043S. At the indicated times, livers were homogenized with 0.05% Triton X in PBS followed by plating of serial dilutions on BH-Agar plates.

Flow cytometry

The following antibodies were used for surface staining CD8 (53-6.7), CD44 (IM7), CD49d (R1-2), CD122 (TM-β1), IFN-g (XMG1.2), CD69 (H1.2F3), CD25 (, Ki67 (16A8), TCR, Vβ5 (MR9-4), Vα2 (B20.1), CD45.1 (A20), CD45.2 (104), CD90.1 (OX-7), CD90.2 (30-H12)

Cells were permeabilized for intracellular staining using foxp3 permeabilization kit (cat:00-5523-00).

References

Akondy, Rama S., Mark Fitch, Srilatha Edupuganti, Shu Yang, Haydn T. Kissick, Kelvin W. Li, Ben A. Youngblood, et al. 2017. “Origin and Differentiation of Human Memory CD8 T Cells after Vaccination.” Nature, December. https://doi.org/10.1038/nature24633.

Akue, Adovi D., June-Yong Lee, and Stephen C. Jameson. 2012. “Derivation and Maintenance of Virtual Memory CD8 T Cells.” Journal of Immunology 188 (6):2516–23.

Azzam, H. S., J. B. DeJarnette, K. Huang, R. Emmons, C. S. Park, C. L. Sommers, D. El-Khoury, E. W. Shores, and P. E. Love. 2001. “Fine Tuning of TCR Signaling by CD5.” Journal of Immunology 166 (9):5464–72.

Azzam, H. S., A. Grinberg, K. Lui, H. Shen, E. W. Shores, and P. E. Love. 1998. “CD5 Expression Is Developmentally Regulated by T Cell Receptor (TCR) Signals and TCR Avidity.” The Journal of Experimental Medicine 188 (12):2301–11.

Berg, Leslie J. 2007. “Signalling through TEC Kinases Regulates Conventional versus Innate CD8(+) T-Cell Development.” Nature Reviews. Immunology 7 (6):479–85.

Berg, Rance E., Christoph J. Cordes, and James Forman. 2002. “Contribution of CD8+ T Cells to Innate Immunity: IFN-γ Secretion Induced by IL-12 and IL-18.” European Journal of Immunology 32 (10). WILEY-VCH Verlag:2807–16.

Berg, Rance E., Emily Crossley, Sean Murray, and James Forman. 2003. “Memory CD8+ T Cells Provide Innate Immune Protection against Listeria Monocytogenes in the Absence of Cognate Antigen.” The Journal of Experimental Medicine 198 (10):1583–93.

Cheung, Kitty P., Edward Yang, and Ananda W. Goldrath. 2009. “Memory-like CD8+ T Cells Generated during Homeostatic Proliferation Defer to Antigen-Experienced Memory Cells.” Journal of Immunology 183 (5):3364–72.

Chu, Talyn, Aaron J. Tyznik, Sarah Roepke, Amy M. Berkley, Amanda Woodward-Davis, Laura Pattacini, Michael J. Bevan, Dietmar Zehn, and Martin Prlic. 2013. “Bystander-Activated Memory CD8 T Cells Control Early Pathogen Load in an Innate-Like, NKG2D-Dependent Manner.” Cell Reports 3 (3). Elsevier:701–8.

Cibrián, Danay, and Francisco Sánchez-Madrid. 2017. “CD69: From Activation

Marker to Metabolic Gatekeeper." European Journal of Immunology 47 (6):946–53.

Dobbs, Kerry, Cecilia Domínguez Conde, Shen-Ying Zhang, Silvia Parolini, Magali Audry, Janet Chou, Emma Haapaniemi, et al. 2015. "Inherited DOCK2 Deficiency in Patients with Early-Onset Invasive Infections." The New England Journal of Medicine 372 (25):2409–22.

Dumont, Celine, Agnieszka Corsoni-Tadrzak, Sandra Ruf, Jasper de Boer, Adam Williams, Martin Turner, Dimitris Kioussis, and Victor L. J. Tybulewicz. 2009. "Rac GTPases Play Critical Roles in Early T-Cell Development." Blood 113 (17):3990–98.

Ernst, B., D. S. Lee, J. M. Chang, J. Sprent, and C. D. Surh. 1999. "The Peptide Ligands Mediating Positive Selection in the Thymus Control T Cell Survival and Homeostatic Proliferation in the Periphery." Immunity 11 (2):173–81.

Fiege, Jessica K., Brandon J. Burbach, and Yoji Shimizu. 2015. "Negative Regulation of Memory Phenotype CD8 T Cell Conversion by Adhesion and Degranulation-Promoting Adapter Protein." Journal of Immunology 195 (7):3119–28.

Fukui, Yoshinori, Osamu Hashimoto, Terukazu Sanui, Takamasa Oono, Hironori Koga, Masaaki Abe, Ayumi Inayoshi, et al. 2001. "Haematopoietic Cell-Specific CDM Family Protein DOCK2 Is Essential for Lymphocyte Migration." Nature 412 (6849). Nature Publishing Group:826–31.

Goldrath, Ananda W., C. John Luckey, Richard Park, Christophe Benoist, and Diane Mathis. 2004. "The Molecular Program Induced in T Cells Undergoing Homeostatic Proliferation." Proceedings of the National Academy of Sciences of the United States of America 101 (48):16885–90.

Goldrath, A. W., and M. J. Bevan. 1999. "Low-Affinity Ligands for the TCR Drive Proliferation of Mature CD8+ T Cells in Lymphopenic Hosts." Immunity 11 (2):183–90.

Gotoh, Kazuhito, Yoshihiko Tanaka, Akihiko Nishikimi, Ayumi Inayoshi, Munechika Enjoji, Ryoichi Takayanagi, Takehiko Sasazuki, and Yoshinori Fukui. 2008. "Differential Requirement for DOCK2 in Migration of Plasmacytoid Dendritic Cells versus Myeloid Dendritic Cells." Blood 111 (6):2973–76.

Gotoh, Kazuhito, Yoshihiko Tanaka, Akihiko Nishikimi, Risa Nakamura, Hisakata Yamada, Naoyoshi Maeda, Takahiro Ishikawa, et al. 2010. "Selective Control of Type I IFN Induction by the Rac Activator DOCK2

during TLR-Mediated Plasmacytoid Dendritic Cell Activation." The Journal of Experimental Medicine 207 (4):721–30.

Guo, Fukun, Jose A. Cancelas, David Hildeman, David A. Williams, and Yi Zheng. 2008. "Rac GTPase Isoforms Rac1 and Rac2 Play a Redundant and Crucial Role in T-Cell Development." Blood 112 (5):1767–75.

Haluszczak, Catherine, Adovi D. Akue, Sara E. Hamilton, Lisa D. S. Johnson, Lindsey Pujanauski, Lenka Teodorovic, Stephen C. Jameson, and Ross M. Kedl. 2009. "The Antigen-Specific CD8+ T Cell Repertoire in Unimmunized Mice Includes Memory Phenotype Cells Bearing Markers of Homeostatic Expansion." The Journal of Experimental Medicine 206 (2):435–48.

Hamilton, Sara E., and Stephen C. Jameson. 2008. "The Nature of the Lymphopenic Environment Dictates Protective Function of Homeostatic-Memory CD8+ T Cells." Proceedings of the National Academy of Sciences 105 (47):18484–89.

Itoh, Y., B. Hemmer, R. Martin, and R. N. Germain. 1999. "Serial TCR Engagement and down-Modulation by peptide:MHC Molecule Ligands: Relationship to the Quality of Individual TCR Signaling Events." Journal of Immunology 162 (4):2073–80.

Jacomet, Florence, Emilie Cayssials, Sara Basbous, Anaïs Levescot, Nathalie Piccirilli, Deborah Desmier, Aurélie Robin, et al. 2015. "Evidence for Eomesodermin-Expressing Innate-like CD8(+) KIR/NKG2A(+) T Cells in Human Adults and Cord Blood Samples." European Journal of Immunology 45 (7):1926–33.

Jameson, Stephen C. 2002. "Maintaining the Norm: T-Cell Homeostasis." Nature Reviews. Immunology 2 (8):547–56.

Kaech, Susan M., and Weiguo Cui. 2012. "Transcriptional Control of Effector and Memory CD8+ T Cell Differentiation." Nature Reviews. Immunology 12 (11):749–61.

Kaech, Susan M., Joyce T. Tan, E. John Wherry, Bogumila T. Konieczny, Charles D. Surh, and Rafi Ahmed. 2003. "Selective Expression of the Interleukin 7 Receptor Identifies Effector CD8 T Cells That Give Rise to Long-Lived Memory Cells." Nature Immunology 4 (12):1191–98.

Kieper, William C., J. Theodore Burghardt, and Charles D. Surh. 2004. "A Role for TCR Affinity in Regulating Naive T Cell Homeostasis." Journal of Immunology 172 (1):40–44.

Kikuchi, Tomoko, Shiro Kubonishi, Misako Shibakura, Noriko Namba,

Toshimitsu Matsui, Yoshinori Fukui, Mitsune Tanimoto, and Yoshio Katayama. 2008. “Dock2 Participates in Bone Marrow Lympho-Hematopoiesis.” Biochemical and Biophysical Research Communications 367 (1):90–96.

Kim, Jihye, Dong-Yeop Chang, Hyun Woong Lee, Hoyoung Lee, Jong Hoon Kim, Pil Soo Sung, Kyung Hwan Kim, et al. 2017. “Innate-like Cytotoxic Function of Bystander-Activated CD8+ T Cells Is Associated with Liver Injury in Acute Hepatitis A.” Immunity, December. https://doi.org/10.1016/j.immuni.2017.11.025.

Kurzweil, Vanessa, Ami LaRoche, and Paula M. Oliver. 2014. “Increased Peripheral IL-4 Leads to an Expanded Virtual Memory CD8+ Population.” Journal of Immunology 192 (12):5643–51.

Lee, You Jeong, Stephen C. Jameson, and Kristin A. Hogquist. 2011. “Alternative Memory in the CD8 T Cell Lineage.” Trends in Immunology 32 (2):50–56.

Le Floc’h, Audrey, Yoshihiko Tanaka, Niels S. Bantilan, Guillaume Voisinne, Grégoire Altan-Bonnet, Yoshinori Fukui, and Morgan Huse. 2013. “Annular PIP3 Accumulation Controls Actin Architecture and Modulates Cytotoxicity at the Immunological Synapse.” The Journal of Experimental Medicine 210 (12). Rockefeller University Press:2721–37.

Lillemeier, Björn F., Manuel A. Mörtelmaier, Martin B. Forstner, Johannes B. Huppa, Jay T. Groves, and Mark M. Davis. 2010. “TCR and Lat Are Expressed on Separate Protein Islands on T Cell Membranes and Concatenate during Activation.” Nature Immunology 11 (1):90–96.

Liu, Zhiping, Si Ming Man, Qifan Zhu, Peter Vogel, Sharon Frase, Yoshinori Fukui, and Thirumala-Devi Kanneganti. 2016. “DOCK2 Confers Immunity and Intestinal Colonization Resistance to Citrobacter Rodentium Infection.” Scientific Reports 6 (June):27814.

Mahajan, Vinay S., Ezana Demissie, Hamid Mattoo, Vinay Viswanadham, Ajit Varki, Robert Morris, and Shiv Pillai. 2016. “Striking Immune Phenotypes in Gene-Targeted Mice Are Driven by a Copy-Number Variant Originating from a Commercially Available C57BL/6 Strain.” Cell Reports 15 (9):1901–9.

Min, Booki, Rebecca McHugh, Gregory D. Sempowski, Crystal Mackall, Gilles Foucras, and William E. Paul. 2003. “Neonates Support Lymphopenia-Induced Proliferation.” Immunity 18 (1):131–40.

Nishikimi, Akihiko, Hideo Fukuhara, Wenjuan Su, Tsunaki Hongu, Shunsuke

Takasuga, Hisashi Mihara, Qinhong Cao, et al. 2009. “Sequential Regulation of DOCK2 Dynamics by Two Phospholipids during Neutrophil Chemotaxis.” Science 324 (5925):384–87.

Nishikimi, Akihiko, Mutsuko Kukimoto-Niino, Shigeyuki Yokoyama, and Yoshinori Fukui. 2013. “Immune Regulatory Functions of DOCK Family Proteins in Health and Disease.” Experimental Cell Research 319 (15):2343–49.

Nishikimi, Akihiko, Takehito Uruno, Xuefeng Duan, Qinhong Cao, Yuji Okamura, Takashi Saitoh, Nae Saito, et al. 2012. “Blockade of Inflammatory Responses by a Small-Molecule Inhibitor of the Rac Activator DOCK2.” Chemistry & Biology 19 (4):488–97.

Nombela-Arrieta, César, Thorsten R. Mempel, Silvia F. Soriano, Irina Mazo, Matthias P. Wymann, Emilio Hirsch, Carlos Martínez-A, Yoshinori Fukui, Ulrich H. von Andrian, and Jens V. Stein. 2007. “A Central Role for DOCK2 during Interstitial Lymphocyte Motility and Sphingosine-1-Phosphate-Mediated Egress.” The Journal of Experimental Medicine 204 (3):497–510.

Picelli, Simone, Åsa K. Björklund, Omid R. Faridani, Sven Sagasser, Gösta Winberg, and Rickard Sandberg. 2013. “Smart-seq2 for Sensitive Full-Length Transcriptome Profiling in Single Cells.” Nature Methods 10 (11):1096–98.

Roozendaal, Ramon, and Reina E. Mebius. 2011. “Stromal Cell-Immune Cell Interactions.” Annual Review of Immunology 29:23–43.

Salmond, Robert J., Rebecca J. Brownlie, Vicky L. Morrison, and Rose Zamoyska. 2014. “The Tyrosine Phosphatase PTPN22 Discriminates Weak Self Peptides from Strong Agonist TCR Signals.” Nature Immunology 15 (9):875–83.

Sanui, Terukazu, Ayumi Inayoshi, Mayuko Noda, Eiko Iwata, Masahiro Oike, Takehiko Sasazuki, and Yoshinori Fukui. 2003. “DOCK2 Is Essential for Antigen-Induced Translocation of TCR and Lipid Rafts, but Not PKC-Theta and LFA-1, in T Cells.” Immunity 19 (1):119–29.

Shugay, Mikhail, Dmitriy V. Bagaev, Maria A. Turchaninova, Dmitriy A. Bolotin, Olga V. Britanova, Ekaterina V. Putintseva, Mikhail V. Pogorelyy, et al. 2015. “VDJtools: Unifying Post-Analysis of T Cell Receptor Repertoires.” PLoS Computational Biology 11 (11):e1004503.

Sosinowski, Tomasz, Jason T. White, Eric W. Cross, Catherine Haluszczak, Philippa Marrack, Laurent Gapin, and Ross M. Kedl. 2013. “CD8α+

Dendritic Cell Trans Presentation of IL-15 to Naive CD8+ T Cells Produces Antigen-Inexperienced T Cells in the Periphery with Memory Phenotype and Function." Journal of Immunology 190 (5):1936–47.

Soudja, Saïdi M'homa, Anne L. Ruiz, Julien C. Marie, and Grégoire Lauvau. 2012. "Inflammatory Monocytes Activate Memory CD8(+) T and Innate NK Lymphocytes Independent of Cognate Antigen during Microbial Pathogen Invasion." Immunity 37 (3):549–62.

Sprent, Jonathan, and Charles D. Surh. 2011. "Normal T Cell Homeostasis: The Conversion of Naive Cells into Memory-Phenotype Cells." Nature Immunology 12 (6):478–84.

Tan, J. T., E. Dudl, E. LeRoy, R. Murray, J. Sprent, K. I. Weinberg, and C. D. Surh. 2001. "IL-7 Is Critical for Homeostatic Proliferation and Survival of Naive T Cells." Proceedings of the National Academy of Sciences of the United States of America 98 (15):8732–37.

Tan, Ying Xim, Boryana N. Manz, Tanya S. Freedman, Chao Zhang, Kevan M. Shokat, and Arthur Weiss. 2014. "Inhibition of the Kinase Csk in Thymocytes Reveals a Requirement for Actin Remodeling in the Initiation of Full TCR Signaling." Nature Immunology 15 (2):186–94.

Terasawa, Masao, Takehito Uruno, Sayako Mori, Mutsuko Kukimoto-Niino, Akihiko Nishikimi, Fumiyuki Sanematsu, Yoshihiko Tanaka, Shigeyuki Yokoyama, and Yoshinori Fukui. 2012. "Dimerization of DOCK2 Is Essential for DOCK2-Mediated Rac Activation and Lymphocyte Migration." PloS One 7 (9):e46277.

Tripathi, Pulak, Suzanne C. Morris, Charles Perkins, Allyson Sholl, Fred D. Finkelman, and David A. Hildeman. 2016. "IL-4 and IL-15 Promotion of Virtual Memory CD8+ T Cells Is Determined by Genetic Background." European Journal of Immunology 46 (10):2333–39.

Van Kaer, Luc. 2015. "Innate and Virtual Memory T Cells in Man." European Journal of Immunology 45 (7):1916–20.

Wang, Lei, Hiroshi Nishihara, Taichi Kimura, Yasutaka Kato, Mishie Tanino, Mitsufumi Nishio, Masato Obara, Tomoyuki Endo, Takao Koike, and Shinya Tanaka. 2010. "DOCK2 Regulates Cell Proliferation through Rac and ERK Activation in B Cell Lymphoma." Biochemical and Biophysical Research Communications 395 (1):111–15.

White, Jason T., Eric W. Cross, Matthew A. Burchill, Thomas Danhorn, Martin D. McCarter, Hugo R. Rosen, Brian O'Connor, and Ross M. Kedl. 2016. "Virtual Memory T Cells Develop and Mediate Bystander Protective

Immunity in an IL-15-Dependent Manner." Nature Communications 7 (April):11291.

White, Jason T., Eric W. Cross, and Ross M. Kedl. 2017. "Antigen-Inexperienced Memory CD8(+) T Cells: Where They Come from and Why We Need Them." Nature Reviews. Immunology, May. https://doi.org/10.1038/nri.2017.34.

Yu, Bingfei, Kai Zhang, J. Justin Milner, Clara Toma, Runqiang Chen, James P. Scott-Browne, Renata M. Pereira, et al. 2017. "Epigenetic Landscapes Reveal Transcription Factors That Regulate CD8+ T Cell Differentiation." Nature Immunology 18 (5). Nature Research:573–82.

Zhang, Xiaoming, Brian Mozeleski, Sebastien Lemoine, Edith Dériaud, Annick Lim, Dania Zhivaki, Elie Azria, et al. 2014. "CD4 T Cells with Effector Memory Phenotype and Function Develop in the Sterile Environment of the Fetus." Science Translational Medicine 6 (238):238ra72.

General Discussion

4.1 Dock2$^{hsd/hsd}$ allele: Implications for knockout mouse studies.

In chapter 2, we described how a *Dock2* copy number variant was is responsible for immune phenotypes in two independently generated knockout strains (*Siae*$^{-/-}$ and Cmah$^{-/-}$) (Mahajan et al. 2016). These findings, combined with an earlier publications identifying the same allele in *Irf5*$^{-/-}$ mice, led us to hypothesize that the duplication came from a common source. We ruled out the possibility that the duplication arose in ES cells used for gene targeting as different ES lines were used to generate these mice. Interestingly, SNP arrays and whole genome sequencing identified several C57BL/6N SNPs closely linked to the duplication in Siae$^{-/-}$ mice. This suggested that the mutation was introduced during backcrossing into this substrain. Indeed, a survey of commercially available C57BL/6N strains identified the presence of an identical duplication around exons 28 and 29 of *Dock2* in C57BL/6NHsd mice.

Envigo, previously known as Harlan, is the commercial vendor of C57BL/6NHsd mice. The company recently conducted an investigation identifying this mutation at multiple facilities across the world. The results of this investigation summarized in the table below:

Table 2: Envigo facilities carrying *Dock2* copy number variant. *(Adapted from Envigo press release, June 2016)*

Location	**Dock2$^{hsd/hsd}$ allele**
Italy	Y
Mexico	Y
Netherlands	Y
USA, Virginia	Y
USA, Maryland	Y
USA, Michigan	Y
Germany	Pending
South Korea	Pending
USA, California	N
France	N
Israel	N
Indiana	N
Texas	N

The presence of this allele in facilities across the world, for at least a 13 years, suggests that numerous studies are compromised by using C57BL6/Nhsd mice as controls, experimental animals or for backcrossing. Unfortunately, compiling a comprehensive list of suspect studies is not feasible as very few publications clearly document what specific substrain was used during backcrossing or for experiments. Thus, we leave it to individual researchers to test any banked tissue samples and current colonies using a previously published PCR screening approach (Yasuda et al. 2013). Following amplification with the primers* shown below, a 305 bp product indicates the presence of the duplication. This strategy doesn't distinguish between homozygous and heterozygous mice. We recommend using RT-PCR for *Dock2* expression or flow cytometric quantitation of marginal zone B cells in the spleen or of CD8+ memory T cells in the blood in combination with the outlined PCR strategy to identify heterozygous mice. (See section 2.4 for details)

***Primers:** GAC CTT ATG AGG TGG AAC CAC AAC C

GAT CCA AAG ATT CCC TAC AGC TCC AC

4.2 Insights from DOCK2 deficiency about memory development and differentiation

In this study, we provide evidence that DOCK2 sets threshold for direct entry into the virtual memory compartment by increasing affinity to weak agonist TCR stimulation. We suspect that differences in the underlying actin cytoskeleton of DOCK2 deficient T cells drives this conversion. This could work affecting TCR triggering through three ways (i) Decreasing interstitial motility (Nombela-Arrieta et al. 2007) and increasing TCR-MHC contact duration which results in increased amounts of TCR triggering. This could promote concomitant conversion of to memory with weak peptides. (ii) By disrupting a cortical actin network and lowering the threshold for activation. (iii) Disrupting cortical actin integrity which might also promote the formation of activation promoting TCR and LAT microclusters (Lillemeier et al. 2010). With this in mind, it will be interesting to investigate plasma membrane architecture *Dock2*$^{hsd/hsd}$ mice. Importantly, these findings suggest that it might be possible to increase the size of the virtual memory compartment by targeting DOCK2 using small molecule inhibitors (Nishikimi et al. 2012).

4.3 Modern genetic manipulation of mice

Two major advances, availability of ES cells from the C57BL/6 background and nucleases for gene editing, have had a transformative effect bringing down costs and time required to generate knockout mice.

4.3.1 ES cells derived from C57BL/6 mice.

The first attempts to culture pluripotent murine ES cells revealed that there cell lines from C57BL/6 that these differentiated prematurely in culture (Rülicke 2012). As a result, the vast majority of knockout mice were made using the more permissible 129 substrain. Knockout mice generated in this way were subsequently backcrossed to the C57BL/6 ground as that is the most widely used strain in biomedical laboratories.This approach was not only time consuming, (backcrossing can take years) it can also introduce mutations that confound interpretations of future studies (see chapter 2).

Luckily advances in ES cell culture have obviated the need for onerous backcrossing (Seong et al. 2004). For example, the inhibition of differentiation by targeting ERK, FGFR and GSK3 (3-inhibitor culture system) was instrumental in generating germline-competent ES cells from C57BL/6N mice (Kiyonari et al. 2010). As a result, C57BL/6 ES cells are have largely

displaced cells of 129 origin in projects involving mouse genetic engineering (Skarnes et al. 2011). However, care should be taken so that the use of C57BL/6 ES cells research does not reinforce an over reliance C57BL/6 mice as a single strain can not adequately model a diverse and outbred human population (Rivera and Tessarollo 2008).

4.3.2 Nucleases for gene editing

Genome editing using nucleases has allowed cost effective and rapid generation of mice with designed gene modification. These molecular tools allow high efficiency gene modification in single cell embryos without the need for cumbersome ES cell culture. This means that researchers can generate mutant mice in one step, in relatively painless process that takes as a little time as a few weeks (Aida 2015). One common theme is shared with genome editing tools: double stranded breaks at precisely targeted loci which will be repaired by one of two of endogenous DNA repair pathways. Double stranded break repair via the error prone non-homologous end joining (NHEJ) can potentially introduce frameshift mutations that result in a gene knockout allele. On the other hand, homology directed repair (HDR) can be used to introduce pinpoint mutations by providing a repair template with homologous arms (Gupta and Musunuru 2014).

Editing mammalian genomes with nucleases got it start with the introduction of Zinc finger nucleases (ZFNs). ZFNs are fusion proteins that are targeted to a particular locus through specific DNA binding domains of zinc finger containing transcription factors fused to endonuclease domain of the bacterial restriction enzyme, Fok1. This approach was first used to make a knockout rat (Geurts et al. 2009) and subsequently in mice (Carbery et al. 2010). However, assembling zinc finger domains can be challenge as site selection in the genome is limited (Gupta and Musunuru 2014). This system is also plagued by relatively high off target cutting (Pattanayak et al. 2011).

The next stage in mammalian genome editing came with the use of Transcription activator like nucleases (TALENs) for gene targeting. Originally discovered in plant pathogens, TALENs are arrays of 10 to 30 aa that recognize DNA through two adjacent amino acids called repeat-variable di-residue (RVD) (Aida 2015). Shortly after the first ZFN based knockout, TALENs were used to target two genes with more than 50% of the resultant mice being knockouts. Importantly, off target cutting of TALENs is about 10% of levels seen with ZFNs (Sung et al. 2013).

Although ZFNs and TALENs were very valuable, mammalian genome editing has been truly revolutionized with the development of CRISPR/Cas9 based approaches (Liu et al. 2017). CRISPR/Cas9 mediated gene is very

efficient with over 80% of gene targeted progeny carrying homozygous mutations. This system allows multiplexed gene targeting by using several 20-bp protospacer gRNA sequence to efficiently target it to multiple locations (Wang et al. 2013). In what is sure to be a critical contribution of biomedical research, Jaenisch and colleagues have also used this system to make mice carrying conditional and reporter alleles (Yang et al. 2013).

In conclusion, changes in the way gene targeted mice are engineered makes unlikely that a confounding mutation will be introduced from a common source. However, this shouldn't lull researchers into a fall sense of security. Spontaneous mutations arise every generation. Thus, we should all be vigilant in ensuring that the appropriate breeding schemes are used, inbred strains are regularly checked for contamination and littermates used as controls whenever possible.

References

Aida, Tomomi. 2015. “Genome Editing in Mice Using TALENs.” In Targeted Genome Editing Using Site-Specific Nucleases, 167–82. Springer, Tokyo.

Carbery, Iara D., Diana Ji, Anne Harrington, Victoria Brown, Edward J. Weinstein, Lucy Liaw, and Xiaoxia Cui. 2010. “Targeted Genome Modification in Mice Using Zinc-Finger Nucleases.” Genetics 186 (2): 451–59.

Geurts, Aron M., Gregory J. Cost, Yevgeniy Freyvert, Bryan Zeitler, Jeffrey C. Miller, Vivian M. Choi, Shirin S. Jenkins, et al. 2009. “Knockout Rats via Embryo Microinjection of Zinc-Finger Nucleases.” Science 325 (5939): 433.

Gupta, Rajat M., and Kiran Musunuru. 2014. “Expanding the Genetic Editing Tool Kit: ZFNs, TALENs, and CRISPR-Cas9.” The Journal of Clinical Investigation 124 (10): 4154–61.

Kiyonari, Hiroshi, Mari Kaneko, Shin-Ichi Abe, and Shinichi Aizawa. 2010. “Three Inhibitors of FGF Receptor, ERK, and GSK3 Establishes Germline-Competent Embryonic Stem Cells of C57BL/6N Mouse Strain with High Efficiency and Stability.” Genesis 48 (5): 317–27.

Lillemeier, Björn F., Manuel A. Mörtelmaier, Martin B. Forstner, Johannes B. Huppa, Jay T. Groves, and Mark M. Davis. 2010. “TCR and Lat Are Expressed on Separate Protein Islands on T Cell Membranes and Concatenate during Activation.” Nature Immunology 11 (1): 90–96.

Liu, Edison T., Ewelina Bolcun-Filas, David S. Grass, Cathleen Lutz, Steve Murray, Lenny Shultz, and Nadia Rosenthal. 2017. “Of Mice and CRISPR.” EMBO Reports 18 (2): 187–93.

Mahajan, Vinay S., Ezana Demissie, Hamid Mattoo, Vinay Viswanadham, Ajit Varki, Robert Morris, and Shiv Pillai. 2016. “Striking Immune Phenotypes in Gene-Targeted Mice Are Driven by a Copy-Number Variant Originating from a Commercially Available C57BL/6 Strain.” Cell Reports 15 (9): 1901–9.

Nishikimi, Akihiko, Takehito Uruno, Xuefeng Duan, Qinhong Cao, Yuji Okamura, Takashi Saitoh, Nae Saito, et al. 2012. “Blockade of Inflammatory Responses by a Small-Molecule Inhibitor of the Rac Activator DOCK2.” Chemistry & Biology 19 (4): 488–97.

Pattanayak, Vikram, Cherie L. Ramirez, J. Keith Joung, and David R. Liu. 2011. “Revealing off-Target Cleavage Specificities of Zinc-Finger

Nucleases by in Vitro Selection.” Nature Methods 8 (9): 765–70.

Rivera, Juan, and Lino Tessarollo. 2008. “Genetic Background and the Dilemma of Translating Mouse Studies to Humans.” Immunity 28 (1): 1–4.

Rülicke, Thomas. 2012. “Generation of Mouse Mutants by Genotype-Driven Mutagenesis.” In The Laboratory Mouse, 91–114. Elsevier.

Seong, Eunju, Thomas L. Saunders, Colin L. Stewart, and Margit Burmeister. 2004. “To Knockout in 129 or in C57BL/6: That Is the Question.” Trends in Genetics: TIG 20 (2): 59–62.

Skarnes, William C., Barry Rosen, Anthony P. West, Manousos Koutsourakis, Wendy Bushell, Vivek Iyer, Alejandro O. Mujica, et al. 2011. “A Conditional Knockout Resource for the Genome-Wide Study of Mouse Gene Function.” Nature 474 (7351): 337–42.

Sung, Young Hoon, In-Jeoung Baek, Duk Hyoung Kim, Jisun Jeon, Jaehoon Lee, Kyunghee Lee, Daewon Jeong, Jin-Soo Kim, and Han-Woong Lee. 2013. “Knockout Mice Created by TALEN-Mediated Gene Targeting.” Nature Biotechnology 31 (January). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 23.

Wang, Haoyi, Hui Yang, Chikdu S. Shivalila, Meelad M. Dawlaty, Albert W. Cheng, Feng Zhang, and Rudolf Jaenisch. 2013. “One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering.” Cell 153 (4): 910–18.

Yang, Hui, Haoyi Wang, Chikdu S. Shivalila, Albert W. Cheng, Linyu Shi, and Rudolf Jaenisch. 2013. “One-Step Generation of Mice Carrying Reporter and Conditional Alleles by CRISPR/Cas-Mediated Genome Engineering.” Cell 154 (6): 1370–79.

Yasuda, Kei, Kerstin Nündel, Amanda A. Watkins, Tania Dhawan, Ramon G. Bonegio, Jessalyn M. Ubellacker, Ann Marshak-Rothstein, and Ian R. Rifkin. 2013. “Phenotype and Function of B Cells and Dendritic Cells from Interferon Regulatory Factor 5-Deficient Mice with and without a Mutation in DOCK2.” International Immunology 25 (5): 295–306.