מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

עבודת גמר (תזה) לתואר
דוקטור לפילוסופיה

מוגשת למועצה המדעית של
מכון ויצמן למדע
רחובות, ישראל

מאת
סרקאלם איינאו

Thesis for the degree
Doctor of Philosophy

Submitted to the Scientific Council of the
Weizmann Institute of Science
Rehovot, Israel

By
Serkalem Ayanaw

חקירת הספציפיות של מעצמי STAT3 בתאים שונים ותפקידם
בהתפתחות מחלות מעיים דלקתיות (IBD)

Defining cell-type specific Stat3 enhancers and their role in IBD development

מנחה:

פרופ' סטפן יונג

Advisor:

Prof. Steffen Jung

תשרי תשפ"ג

September 2022

Acknowledgments

I would like to express my deepest appreciation to my advisor, Prof. Steffen Jung, for his guidance, help, and support in the past four and a half years. Steffen inspires me with his love for science and exploring the unknown. I would also like to express my deepest gratitude to Sigalit Boura Halfon for helping me throughout my Ph.D studies for her support, advice, and productive discussion. I have learned so much from her. Additionally, I would like to thank all the current and previous members of the Jung laboratory. Specifically, Jung-Seok and Biana for helping me throughout my first year in the lab. Many thanks to Yaara for her support and for going above and beyond her role. Special thanks to my advising committee, Prof. Rivka Dikstein, and Prof. Nahum Shpigel, for guidance throughout my studies. I want to acknowledge our collaborator Dr. Simon Fishilevich from Lancet laboratory, WIS. In addition to Dr. Shifra Ben-Dor from Bioinformatics and Biological Computing Unit and Dr. Rebecca Haffner-Krausz from Veterinary Resources for helping to generate the Stat3 enhancer mice. Finally, and most importantly, I would like to thank my family, especially my parents, who have always encouraged me that everything is possible.

Table of Contents

Abstract

Ulcerative colitis and Crohn's disease are inflammatory bowel disorders (IBD) resulting from a breakdown of gut homoeostasis. A critical module required to maintain gut health is the IL-10/ IL-10 receptor axis and its associated signaling pathways. Thus, children lacking the IL10 receptor (IL-10R) develop severe colitis. Also, mice lacking IL-10 or the IL1-0R specifically in macrophages develop fulminant gut inflammation. We recently showed that the disease was caused by mutant macrophage-derived IL23 that induced TH17 cells to produce IL-22, that in turn activated epithelial cells to recruit detrimental neutrophils.

IL-10-, IL-23- and IL-22 receptors all require the transcription factor Stat3 for signaling, although the specific receptors are expressed on distinct cells. Interestingly, patients carrying Stat3 loss-of-function mutations do not display gut pathology, but rather suffer from an autosomal dominant Hyper-IgE (AD-HIES) or 'Job's' syndrome. We hypothesized that this might be due to the fact that STAT3 open reading frame variants will cause deleterious IL-10R impairment, but concurrently neutralize the critical pro-inflammatory executor elements of this IBD cascade, i.e. IL-23R and IL-22R. Conversely, we predicted that reported *STAT3* loci-associated single nucleotide polymorphisms (SNPs) identified by GWAS as IBD risk factors might affect cell type-specific regulatory regions and hence impair cell type-specific STAT3 expression in selected cell types. To test our hypothesis, we decided to define the enhancer elements that drive Stat3 expression in macrophages, T cells and non-immune cells, respectively. Specifically, we used CRISPR/Cas9 technology to generate the respective mutant mice and tested the animals for Stat3 expression in macrophages, as well as colitis development.

תקציר

קרוהן והקוליטיס הינם שני סוגי מחלות דלקתיות במעי (Inflammatory bowel disorders (IBD, אשר נגרמות כתוצאה מפגיעה בהומאוסטאזיס של המעי. על מנת לשמור על בריאות המעי נדרש תהליך מעברי אותות חיוני, המתווך על ידי אינטרלוקין-10 והקולטן אינטרלוקין-10 (IL-10/IL-10R). בילדים אשר חסר להם את הגן האחראי לביטוי של הרצפטור אינטרלוקין-10 (IL-10R), מתפתח קוליטיס חמור. ממצא דומה נצפה במודל עכברי, המכיל מוטציה גנטית שפוגעת בביטוי של IL-10 כך גם במודל נוסף לדלקת מעי, המכיל מוטציה ספציפית הפוגעת בביטוי של הקולטן IL-10 (IL-10R) בתאי מקרופאג'ים בלבד. לאחרונה הראינו שהמחלה נגרמת על ידי אינטרלוקין 23 (IL-23), שמקורם מתאי ממקרופאג'ם מוטנטיים שגרמו לתאי TH17 להפריש ציטוקין נוסף IL-22, הגורם לתאי אפיתל של המעי להזעיק את הנויטרופילים המחוללים את הדלקת.

בתהליך מעברי האותות, גורם השעתוק STAT3 הינו הכרחי לקולטנים IL10, IL-22, IL-23 , למרות מעורבותם בתאים שונים. מעניין לציין, שמטופלים בעלי מוטציות בגן STAT3 הפוגעת בתפקוד החלבון אינם מראים פתולוגיות במעי, אולם הם סובלים מתסמונת יתר - IgE, המכונה גם תסמונת ג'וב (איוב) (Autosomal dominant Hyper-IgE (AD-HIES. ההיפותזה שלנו היא שתופעה זו עשויה לנבוע מכך שוואריאנטים במסגרת קריאה פתוחה, (open reading frame (ORF של הגן Stat3 יכולים לגרום לפגיעה בתפקוד של IL-10R. אך במקביל ינטרלו את המרכיבים הפרו-דלקתיים הקריטיים לתהליך מעבר האותות הגורמות למחלות מעי דלקתיות (IBD) כגון, IL-23R ו-IL-22R. לחילופין, אנו משערים כי הסניפים single nucleotide polymorphisms SNPs) בגן STAT3 שצוינו על ידי GWAS כגורמי סיכון למחלות דלקתיות במעי (IBD), מצבעים על אזורים בקרתיים שהינם ספציפיים לתא. בהתאם לכך, הם פוגעים בביטוי של הגן STAT3 בתאים אלו. כדי לבחון את השערת המחקר שלנו, החלטנו לאפיין את המעצימים של הגן STAT3 המבוטאים במקרופג'ים, בתאי T ובתאים שאינם מהמערכת החיסונית. ובאופן ספציפי, השתמשנו בטכנולוגיית CRISPR/Cas9 כדי ליצור עכברים מוטנטים בהתאמה למעצימים שאפיינו. במוטנטיים אלו, אנו נבדוק ביטוי של STAT3 במקרופג'ים, וכן התפתחות של קוליטיס.

Introduction

Inflammatory bowel disease

Inflammatory bowel diseases (IBD) are a group of inflammatory conditions of the colon and small intestine [1, 2]. IBD incidence is on the rise in the western world. Furthermore, in the 21st century the disease emerged in newly industrialized countries with similar phenotypic manifestation seen in the western world [3]. Experts predict that over 1 million residents in the USA and 2.5 million in Europe will develop IBD by 2025 [3]. IBD is a complex and multifactorial disease, affected by genetics, environment and diet [2]. Another prominent characteristic of IBD is the reduced diversity of the microbiota, and dysbiosis [4]. Crohn's disease (CD) and ulcerative colitis (UC) are the main types of IBD with chronic and relapsing pathologies of the intestine. UC and CD differ in the location of the pathology and type of gut inflammation [5]. UC primarily affects the colon, including mucosal and submucosal layers of the intestinal wall, as well the rectum, and is characterized by continuous inflammation. CD involves the entire gastrointestinal tract, although most commonly the ileum and colon. In CD, the inflammation is frequently trans-mural and typically discontinuous and patchy with granulomas. Despite these differences, UC and CD share common symptoms, i.e. ulceration, abdominal pain, bleeding, diarrhea, and malnutrition. IBD often displays additional associated diseases, such as primary sclerosing cholangitis, ankylosing spondylitis and psoriasis [5].

The pathways and the mechanisms that lead to IBD pathogenesis remain incompletely understood and are under intense investigation. Recent studies reported for instance a link between IBD and noncoding RNAs (ncRNAs), that include microRNA , long noncoding RNAs and circular RNAs [6]. These new elements may serve as potential biomarkers for diagnosis and perhaps as novel therapeutic targets [6]. Genome Wide Association Study (GWAS) have reported more than 240 risk loci associated with IBD pathogenesis [7, 8], most of which are located outside of protein coding regions. These variants may affect regulatory element potentially impairing mRNA and ncRNA expression [6-8]. In most cases, the underlying mechanism of why the variants confer IBD risk remains unknown, but recurrent candidate loci include STAT3, IL23R and JAK2 [9-11], pointing at a critical role of immune cells and cytokine circuits in the pathology, and specifically intestinal macrophages.

Gastrointestinal microbiome and inflammatory bowel diseases (IBD)

The gastrointestinal microbiome has an essential role in different functions in the gut, both in steady state and in disease. That includes defense as a biological barrier, in digestion processes, and in the training of the immune system [12]. The bacteria density in the gut increases along the intestine, and differences depend on different organs and the area reaching from 10^9 bacteria per ml in the oral cavity over 10^1-10^2 per ml in the stomach to 10^4-10^7 per ml in the ileum and 10^{10}-10^{12} per ml in cecum and

colon [12, 13]. Although the commensal microbiome has an important positive role for its host, they also can be drivers of pathologies, including IBD, obesity, and diabetes [12].

CD patients display reduced diversity of their commensal bacteria [14]. Moreover, some bacteria have particular associations with gut inflammation, such as *Mucispirillum schaedleri* and *Helicobacter* species. However, not only bacteria have been linked to IBD but also fungi and increased fungal diversity has been reported for CD patients. For example, this includes expansion of *Candida albicans*, *Aspergillus clavatus* and *Cryptococcus neoformans,* and a decrease in *Saccharomyces cerevisiae* [15].

The mucus layer plays a critical role in maintaining gut homeostasis. In UC, the mucus layer is thin because of a reduction in the production of MUC2. In contrast, in CD, the layers of the mucus are maintained, but the composition of MUC2 is different with respect to O-linked glycosylation. As a result, permeability of bacteria in the colon is increased [16].

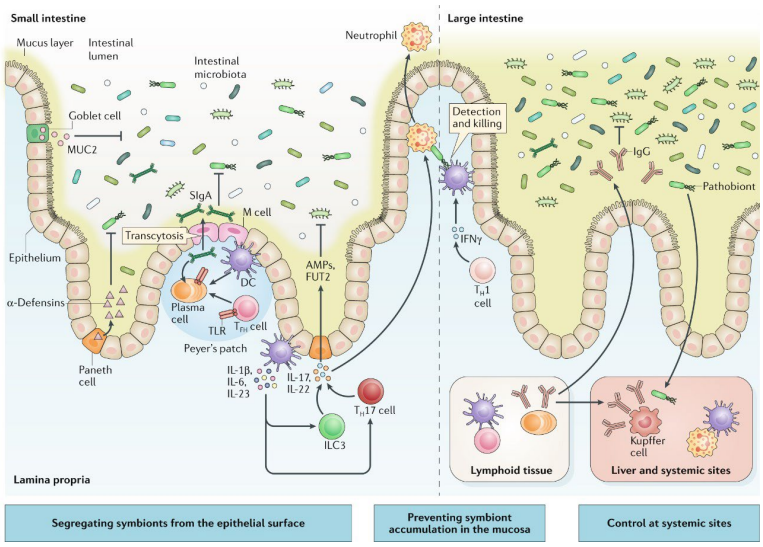
In summary, IBD is a complex disorder, involving microbiome changes, alterations of different layers of the gut, immune cells, and genetics, that still need to be studied.

The gastrointestinal tract (GIT)

The gastrointestinal (GI) tract, which starts from the mouth and ends in the anus, is a significant extended interface between the body and the environment. This system is highly complex and needs to protect the organism from toxins and pathogens. The digestive system includes, among other organs that take part in the digestive process, the esophagus, stomach, and intestines. The intestines have two parts, the small intestine, and the large intestine, which differ in structure. The small intestine includes the *duodenum, jejunum,* and *ileum*, and the large intestine comprises the *cecum* and *colon.* In general, digestion and absorption occur in the small intestine. Absorption of water and nutrients continues in the large intestine[17, 18].

The small and large intestines differ with respect to their physiological roles and anatomy (**Fig 1A**). The small intestine is characterized by finger-like structures, which are absent from cecum and colon. Villi reach into the gut lumen and increase the surface area of the intestinal epithelial cell (IEC) layer. Also, the IEC composition of small and large intestines differ. For instance only the small intestine harbors Paneth cells which express multiple genes that have been linked to Crohn's disease, including the ATG16-like 1 protein, transcription factor 4 (TCF4), NOD2, and immunity-related GTPase family M protein 1 (IRGM1). Unlike Paneth cells, goblet cells, which secret mucus, are found more abundant in the colon and cecum [17].

A



B

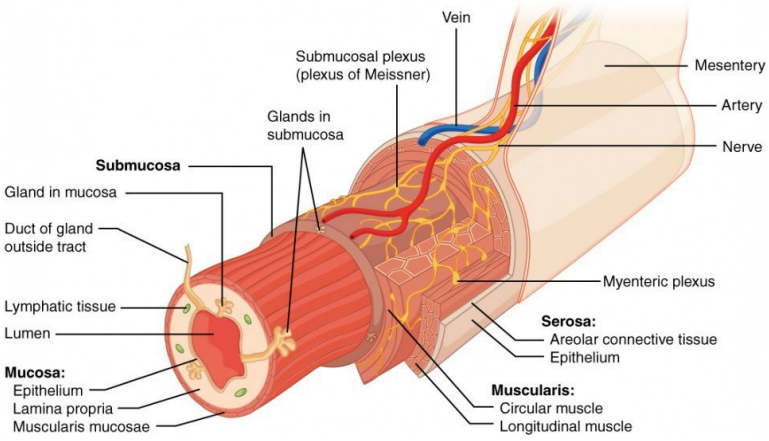


Figure 1. The gastrointestinal (GI) tract (A) Graphical representation of the cellular composition of the intestine. Adapted from [15]. (B) The four layers of the GI tract adapted from: https://courses.lumenlearning.com/suny-ap2/chapter/overview-of-the-digestive-system/

Histology and function of the four layers of the gastrointestinal tract

The GI tract is composed of four layers which give it its tubular shape. The largest layer which faces the lumen is the *mucosa,* followed by the *submucosa,* the *muscular layer,* and lastly the *serosa* or *adventitia* [18] (**Fig 1B**).

Mucosa - this layer is the most extensive layer that includes a simple epithelium and glandular tissue (like goblet cells), essential for secretory and absorptive functions of the intestine. In addition, the mucosa comprises connective tissue the *lamina propria* (LP), which is the location of most immune cells, including macrophages, dendritic cells (DC) and lymphocytes. The mucosa is also highly vascularized and innervated. Barrier function of the mucosa which is critical to keep gut homeostasis

includes physical, chemical, immunological, and biological mechanisms. The immunological LP barrier propria includes T cells, B cells, DC and macrophages [19]. Macrophages in LP are vital in protecting the gut from infection and display in this tissue a mild pro-inflammatory profile. However, they also maintain tissue tolerance and balance between inflammation to protect the body and calm the immune system. The best known tolerogenic cytokine is IL-10, which is essential for balancing T cell immunity in the intestine. The lack of IL-10 has been shown to cause colitis in humans and mice [20]. Most intestinal macrophages are derived from monocytes and show a continuous turnover. More recently though, a new resident long-lived intestinal macrophage population was detected in the LP close to the submucosal enteric nervous system (ENS), blood vessels and Peyer's patches (PP). In the small intestine, long-lived macrophages are located in proximity to Paneth cells, and it was proposed that these cells contribute to Paneth cell differentiation [20, 21].

Submucosa - a connective tissue layer, including lymphatics, blood vessels, and nerves, as well as mucous secreting glands [17]. The submucosa includes the submucosal plexus that constitutes part of the ENS, that plays role in secretion. Furthermore, long-lived macrophages detected in the submucosa are close to the blood vessels and neurons and depletion of these cells was shown to lead to vascular leakage and loss of submucosal neurons [20, 21].

Muscular layer – This smooth muscle layer maintains the rhythmic waves of contraction or peristalsis, that is critical to move food down through the gut. The layer includes the myenteric plexus, a part of the ENS. Muscular macrophages are located in the *muscularis* externa and show differential gene expression as compared to macrophages in LP, including expression of *Retnla, Mrc1* (CD206), and CD163. Muscularis macrophages (MM) and enteric neurons closely interact. MM are thought to regulate peristalsis of the smooth muscle. Moreover, secretion of bone morphogenic protein 2 (BMP2) by these cells which engages BMPR on enteric neurons to induce secretion of colony stimulatory factor 1 (CSF1), which is essential for normal development [22].

Serosa or adventitia- this tissue which faces the peritoneal cavity consists of loose connective tissue, vascularized, innervated and populated by another population of macrophages.

Taken together, The GIT is comprised of multiple layers which contain distinct populations of tissue macrophages. The complex and unique structure of the gut allows the GI tract to perform its role of protecting and digesting food.

Intestinal macrophages

Resident intestinal macrophages are essential for gut homeostasis. They constantly sense the unique and dynamic environment in the gut and maintain a delicate balance between tolerance and response to the commensal microbiome, food, and toxins [23, 24]. Tissue resident macrophage and DC in the gut are located in connective tissue or *lamina propria* underlying a single layer of epithelial cells (EC). Intestinal macrophages are characterized by expression of the integrins CD11b and CD11c, the Fc-γ receptor 1 (Fc-γRI; CD64), the chemokine receptor CX_3CR1 and F4/80 (EGF-like module containing

mucin-like hormone receptor-like 1-EMR1). The majority of gut macrophages is, unlike other tissue-resident macrophages, constantly replenished by bone marrow (BM)-derived monocytes [14]. The predicted half-life of these macrophages is 4-6 weeks [15]. The cascade of differentiation from classical monocytes into gut macrophages has been termed "monocyte waterfall" [16, 17]. Specifically, $Ly6C^{hi}$ $CD64^{-}$ $CX3CR1^{int}$ $MHCII^{-}$ monocytes (P1) differentiate through an intermediate into $Ly6C^{lo}$ $CD64^{+}$ $CX3CR1^{hi}$ $MHCII^{hi}$ (P3-P4) [14]. The "monocyte waterfall" has also been reported for humans, comprising a differentiation cascade from $CD14^{high}$ CCR2+ CD11c ^{high} classical monocytes into $CD14^{lo}$ $CCR2^{-}$ $CD11c^{high}$ gut macrophages, [20]. Most tissue macrophages, such as microglia, alveolar macrophages, Langerhans cells, and Kupffer cells, are derived from the yolk sac and fetal liver precursors. In contrast as noted above, resident intestinal macrophages are considered to originate form continuous monocyte replenishment. However, the notion that all gut macrophages are short-lived was recently challenged. Using a fate mapping and flow cytometry approach [14, 15], Grainger and colleagues showed that the gut also hosts long-lived macrophages, which self-maintain and are defined by the surface markers Tim4 and CD4.

The role of IL-10 and STAT3 in gut homeostasis

Studies of others and our lab have established the critical role of the IL-10 / IL-10 receptor (IL-10R) axis and its associated pathways in the maintenance of gut homeostasis [25-27]. IL-10, provided by T regulatory cells [28], is a critical factor ensuring colon homeostasis and preventing the emergence of pro-inflammatory monocyte-derived cells. Specifically, colonic macrophages unable to sense IL10 due to an IL10 receptor deficiency fail to be restrained in patients, and children bearing this mutation develop severe early onset colitis [26]. Conversely, our laboratory has shown that mice harboring Il10Ra-deficient macrophages develop severe gut inflammation [29]. These animals hence provide a valuable model for mechanistic studies of the human disorder caused by the Il-10ra loss-of-function mutation. In extension of our original work we, for instance, showed that the mutant gut macrophages trigger T cells to produce deleterious IL-22 that induces epithelial chemokine expression and leads to a detrimental neutrophil recruitment [25]. Notably, the cytokine receptors involved in this intercellular communication circuits, i.e. IL-10R, IL-23R and IL-22R, all engage the signal transducer and activator of transcription 3 (STAT3), although in distinct cell types. Stat3 signaling is hence required for both, the inducer and the effector module of the pathology associated with IL10R pathology (**Fig 2**).

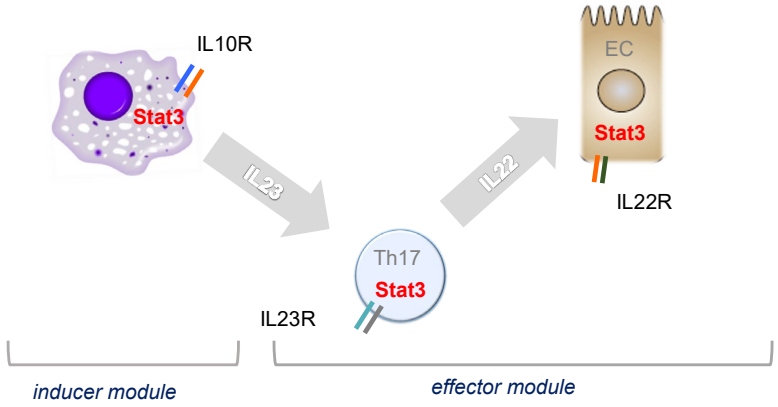


Figure 2. Impairment of the IL0/ IL10R axis results in macrophages producing IL23 (*inducer module*) that triggers T cells to produce IL22 that chemokine production by EC (*effector module*). Stat3 is a signal transducer associated with IL10R, IL23R and IL22R [26].

STAT3 (Signal transducer and activator of transcription 3)

Signal transducer and activator of transcription 3 (STAT3) is a member of the STAT family of transcription factors (TF) with pleiotropic functions in steady state and under pathology, including the regulation of proliferation, survival, cell growth, development, apoptosis, inflammation and immunity [14, 15]. Stat3 functions have been divided into two categories, i.e. canonical/ transcriptional and non-canonical/ non-transcriptional activities. Canonical STAT3 functions involve STAT3 tyrosine and serine phosphorylation by JAK and Src kinases [16], which is stimulated by a large battery of cytokines and growth factors (e.g. IL10, IL-6 family members, leptin, IL-12, IL-2, IFNs, G-CSF, EGF, HGF, LIF). STAT3 phosphorylation results in its translocation to the nucleus, DNA binding and the stimulation of target gene transcription [14, 15]. Non-canonical STAT3 activities, mostly refer to functions of un-phosphorylated STAT3, which is thought to have inhibitory effects, and has for instance been reported to be involved in mitochondria activation, metabolism, oxidative phosphorylation (ROS) and autophagy [14, 17]. The importance of STAT3 in development is underlined by the fact that a STAT3 germ line mutation in mice results in embryonic lethality [14, 18]. STAT3 loss-of-function (LOF) mutations in humans cause Autosomal Dominant Hyper-IgE (AD-HIES), which commonly is a result of STAT3 LOF by missense and point mutations in the DNA-binding/ SRC homology 2 (SH2) /transactivation domain [15, 19-21]. While AD-HIES covers a wide spectrum of pathologies, it somewhat surprisingly does not include gut inflammation or colitis [22].

Cell type specific Stat3 expression and its impact in Human and mice

Stat3 is expressed in all cell types [30]. While Stat3 LOF mutations are not associated with gut inflammation or colitis [7], genetic Stat3 variants have been associated with IBD risk [31]. These SNPs could affect regulatory elements, such as enhancers, and thereby cell type-specific Stat3 expression. Accordingly, a hematopoietic deficiency of Stat3 in $Tie2^{Cre}$:$Stat3^{fl}$ mice was shown to cause severe pathology, including abnormal myeloid development, CD-like gut inflammation and lethality [32]. Likewise, $LysM^{Cre}$:$Stat3^{fl}$ animals that harbor a Stat3 deficiency in macrophages and neutrophils develop colitis [33], likely because the anti-inflammatory effect of IL-10 on macrophages is impaired in these

mice. Interestingly, also a Stat3 deficiency in epithelial cells, as generated in *Villin*^{Cre}*:Stat3*^{fl} mice, was reported to result in a breakdown of gut homeostasis, potentially because the cells can no longer respond to IL-22 [34]. In contrast, mice that harbor a T cell specific *Stat3* deletion are protected from colitis [9]. It was proposed that this is due to a role of Stat3 in differentiation of TH17 cells and regulation of proliferation and survival of T cells.

Enhancers / regulatory element

Enhancers are *cis*-regulatory DNA elements that positively regulate the transcription of target genes in a cell type-specific manner. Enhancers are affected by the environment and respond to external stimuli [35] by recruiting DNA-binding TF. Enhancers regulate expression of coding genes but also ncRNAs, such as microRNAs. Enhancer mutations can cause a variety of diseases [36]. The activation state of enhancers can be inferred from DNA hypersensitivity, i.e. the 'openness' of the chromatin, as well as histone modifications, such as H3K27ac, H3K4me1 and H3K4me3[37] (**Fig 3**). Recent studies have highlighted that enhancers and promotors share several histone modifications and other properties suggesting that they may also share function [35] . Moreover, enhancers can initiate chromatin loops in order to interact with the promoter of their target genes [31, 38]. As a consequence, the distance of an enhancer to its target promoter can range from 100bp to megabases [39].

In this thesis, I focused on dissecting cell type-specific regulatory elements and putative enhancers of the murine and human STAT3 locus that might explain distinct disease associations, including IBD. Specifically, we used CRISPR/Cas9 technology to generate mice harboring Stat3 enhancer mutations.

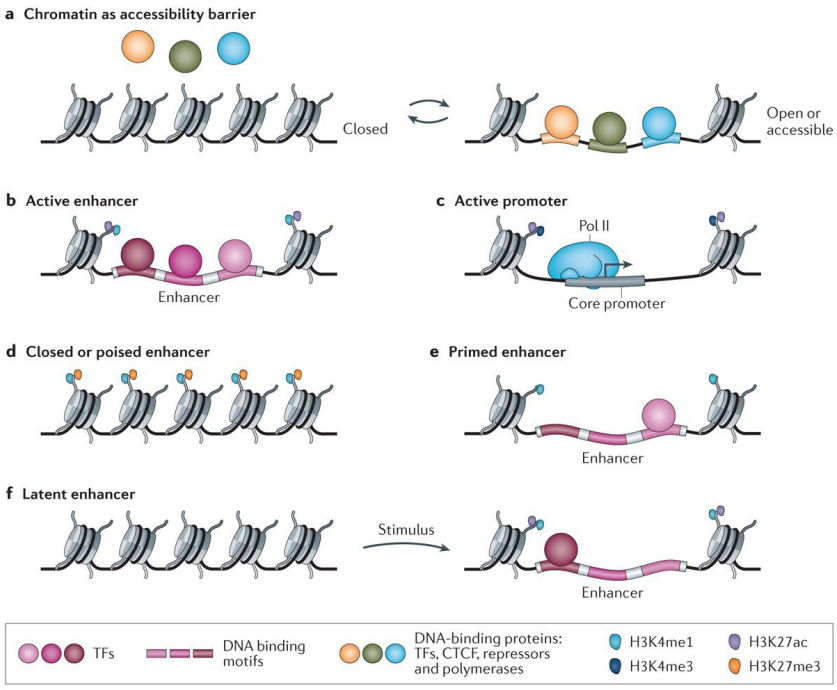


Figure 3. Enhancer Scheme. Adapted from [37]

Results

Putative regulatory elements / Enhancers controlling Stat3 expression

Enhancers are *cis*-regulatory DNA elements that positively regulate the transcription of target genes in a cell type-specific manner. GWAS have discovered more than 240 genetic variants that are associated with IBD risk [7, 8]. In most cases, the underlying mechanism for the increased disease susceptibility remains unknown, but there are recurrent candidate loci that have drawn particular attention [9-11]. In the STAT3 gene **(Fig 4)**, GWAS data and GeneCard analysis identified nine single nucleotide polymorphisms (SNPs) associated with diseases; four of these variants represent risk factors for IBD, including UC and CD [40]. These four SNPs are located in intron #1 and intron #2 of the human STAT3 gene. Notably, three of the IBD risk SNPs (rs6503695, rs9891119, rs744166) are located within elements that are recognized to be associated with regulatory activity, or enhancers, by the GeneHancer software [31] **(Fig 4B).** Two of these IBD risk-associated SNPs in the STAT3 loci (GH17J042346, GH17J042360) are predicted to exhibit activity in macrophages and monocytes (**Fig 4B**) [31]. This might imply that these IBD risk variants are associated with cell-type specific STAT3 expression defects in the myeloid lineage and in particular macrophages, rather than T cells. As would be expected, these regulatory elements are highly conserved among human and mice [31], although in the mouse genome element can be subdivided, as seen in enhancer GH17J042346. (**Fig 4B, C**). The whole region in the mouse is therefore larger, but is spans like in the human most of the second intron of the *Stat3* gene.

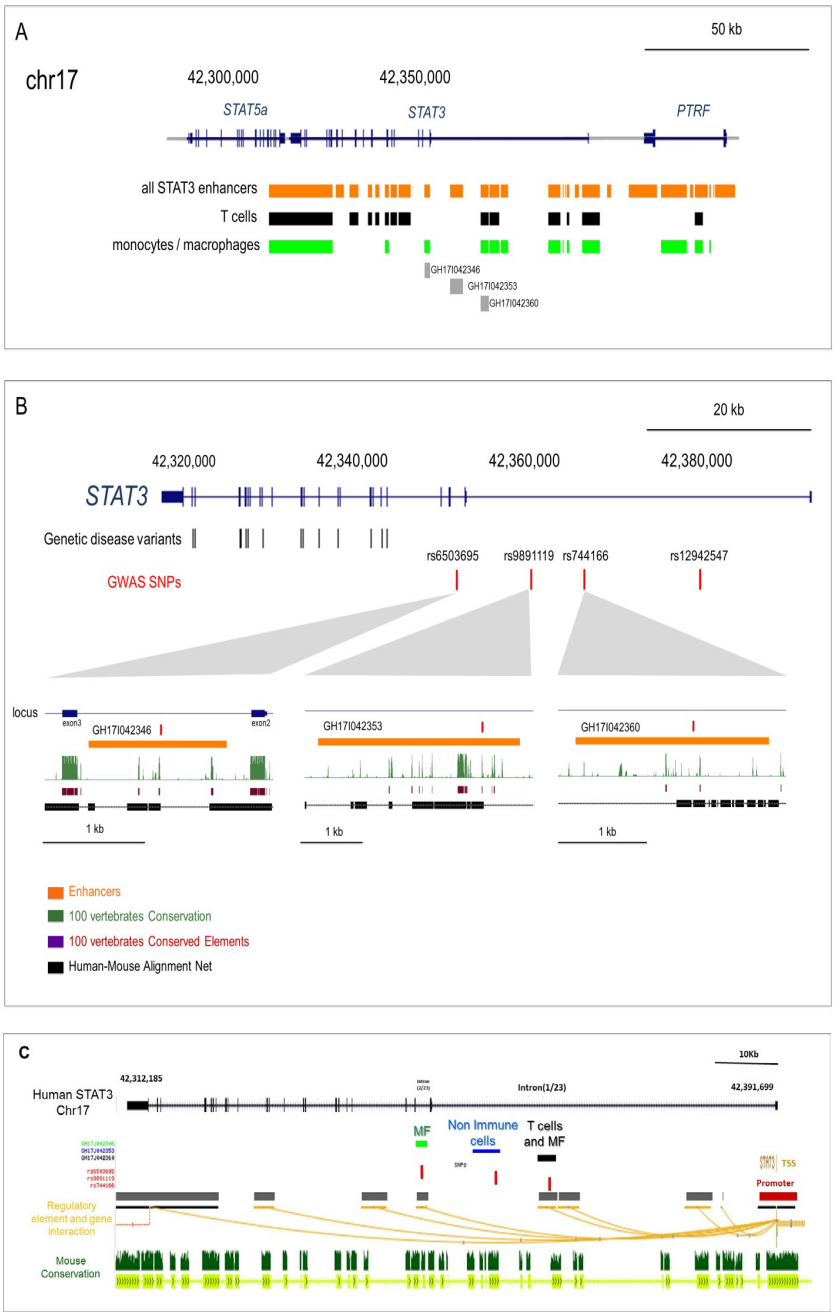


Figure 4 : Human STAT3 locus (Chromosome 17 q21.2, GRCh38/hg38).

(A) The STAT3 locus harbors 42 candidate regulatory elements according to GeneHancer predictions (13 distal to the shown locus). Enhancer activity in myeloid cells (green) and T cells (black).

(B) The STAT3 locus contains 42 genetic disease variants with "pathogenic" clinical significance status (black).

(C) Three SNPs are located near or in predicted enhancer elements and linked to the STAT3 promoter**.** 22 variants have a disease association, parted to two diseases: "Hyperimmunoglobulin E syndrome" and "Autoimmune disease, multisystem, infantile-onset, 1". The STAT3 locus is also associated with 4 GWAS SNPs associated with Crohn's disease, Ulcerative colitis or IBD (red). Three of those intronic SNPs are located within enhancer candidates, namely GH17I042346, GH17I042353 and GH17I042360. The proximal surrounding of each enhancer shows conservation and human-mouse Alignment Net (GRCm38/mm10).

STAT3 enhancer Transcription factor binding sites (TFBS) with potential relevance for Stat3 expression in monocyte and macrophages

The focus of our study is on the definition of cell-type specific activities of putative enhancer elements in murine and human Stat3 loci. To probe the function of a predicted macrophage-specific enhancer we generated mice harboring genomic deletions. In parallel, we performed a detailed bioinformatic analysis for putative TF binding sites (TFBS). This part of the project was a collaboration with Dr. Simon Fishilevich (Lancet laboratory, Dept of Molecular Genetics, Weizmann Institute) who helped us also with the initial identification of the enhancer elements. We focused our analysis on two critical lineage determining TF (LDTF) that have been shown to act as pioneering factors in monocyte/macrophage differentiation.

- **Spi-1/PU.1 (ETS family) binding sites in putative Stat3 enhancers (Fig 5A).**

Spi-1/PU.1 is an ETS family master TF that plays a key role as pioneering factor in myeloid differentiation [41, 42]. Moreover, ETS family members regulate many different processes in steady-state and in diseases, including immune cell functions [3]. ETS family members can function as activators or as repressors [3, 4]. The Spi1 motif spans 15-20 bases with a core sequence of 4-6 bp that is recognized [3] **(Fig 5B)**. The Genomatix software [43] identifies 6 potential ETS sites in the predicted human macrophage (MF) enhancer (GH17J042346), including sites for PU.1/Spi1 binding. In the respective putative mouse MF enhancer, 8 potential ETS family sites were identified, three of which are Spi-1/PU.1 sites. Predicted human T cell and MF enhancer (GH17J042360) harbour only two ETS sites. In the corresponding mouse enhancer, three ETS family sites were recognized with only one expected Spi-1/PU.1 binding site. Lastly, a predicted human non-immune (NI) enhancer (GH17J042353) includes 5 ETS family sites, whereas the mouse equivalent contains 12 expected ETS family sites, with only one defined as Spi-1/PU.1 site. Our lab showed that expression of Spi-1/PU.1 TF is down-regulated in generic intestinal macrophages as compared to monocytes, their precursors [23].

- **C/EBP (CEBP family) binding sites in putative Stat3 enhancers (Fig 5A)**. TF of the CCAAT/enhancer binding protein (C/EBP) family were shown to control differentiation of a range of cell types and have key roles in regulating cellular proliferation [44]. The motif of C/EPB spans between10-12 bases with a core sequence of 4-5 bp that are recognized [45] (**Fig 5B)**. The Genomatix software predicts four C/EBP sites in the putative mouse MF enhancer (GH17J042346); however none were found for the human MF enhancer. The human T cell and MF enhancer (GH17J042360) harbors 5 sites expected to bind C/EBP, in contrast the mouse equivalent lacked C/EBP binding sites. The human NI enhancer (GH17J042353) contains four expected C/EBP sites, whereas in the mouse counterpart only one C/EBP binding site was observed.

- **STAT family TF binding sites in the putative MF enhancer (Fig 5A).**

STAT family members are classical TF that bind to different DNA regulatory elements in various genes [6] under different conditions and cellular signaling. The Genomatix and JASPR software [3, 4]

identifies 12 potential STAT sites in the putative mouse MF enhancer located in intron #2 in the *Stat3* locus. Interestingly, two of these show a Stat3 motif spanning 10 bases with 4-8 bases that are conserved (**Fig 5B)**, suggesting the existence of an autocrine loop. Indeed, STAT family TF are well known to be involved in autoregulatory circuits, inducing epigenetic changes [46]. Two other STAT sites are predicted to bind dimers of Stat1and Stat2, five binding sites only Stat2, one Stat5b, and two Stat6. Moreover, expression of the Stat3 TF is down-regulated in generic intestinal macrophages compared to their precursors, the blood monocytes [23].

Generation of mice harboring STAT3 enhancer mutations using CRISPR/Cas9 genome editing

To functionally define critical enhancers that control cell-type specific Stat3 expression, impairment of which might cause gut pathology, we generated mutant mice using the CRISPR/Cas9 approach [33]. Specifically, we established 4 mutant mouse strains (**Fig 6, 7**). The predicted MF enhancer in humans (GH17J042346 (MF) is split into two regions in mice. We therefore generated two mutant strains to probe the activity of the mouse equivalent of this predicted human MF enhancer: ΔE MF1 harboring a 475bp deletion and ΔE MF2 harboring a 2051bp deletion. In addition we generated one mouse strain with a deletion in the putative T/MF cell enhancer (ΔE T/MF, 673 bp deletion), and as a control, one strain with a deletion in the putative non-immune cell enhancer (ΔE NI, 2145 bp deletion).

With expert help of Dr. Shifra Ben-Dor of the WIS Core facilities we designed two guide RNAs (gRNAs) for each enhancer and used an electroporation-based strategy [33] to deliver the Cas9/sgRNA complexes into mouse zygotes for *in vivo* genome editing. All mice were on C57BL/6 background. A summary of the yield of offspring and the result of the genotyping is shown in **Table 1**. Notably, we directly obtained both heterozygote and homozygote mutant animals. The deletions were verified by genomic PCR analysis and sequencing of the respective PCR products **(Fig 7**). Of note, CRISPR/Cas9 mice have been reported to have fewer off-targets than cell lines [34]. Moreover, potential off-target events that are not close to the *Stat3* loci are removed by the subsequent crossing to WT animals. Heterozygous F1 mice were intercrossed to generate the respective homozygote mutant mouse strains termed ΔE MF1,ΔE MF2, ΔE T/MF and ΔE NI.

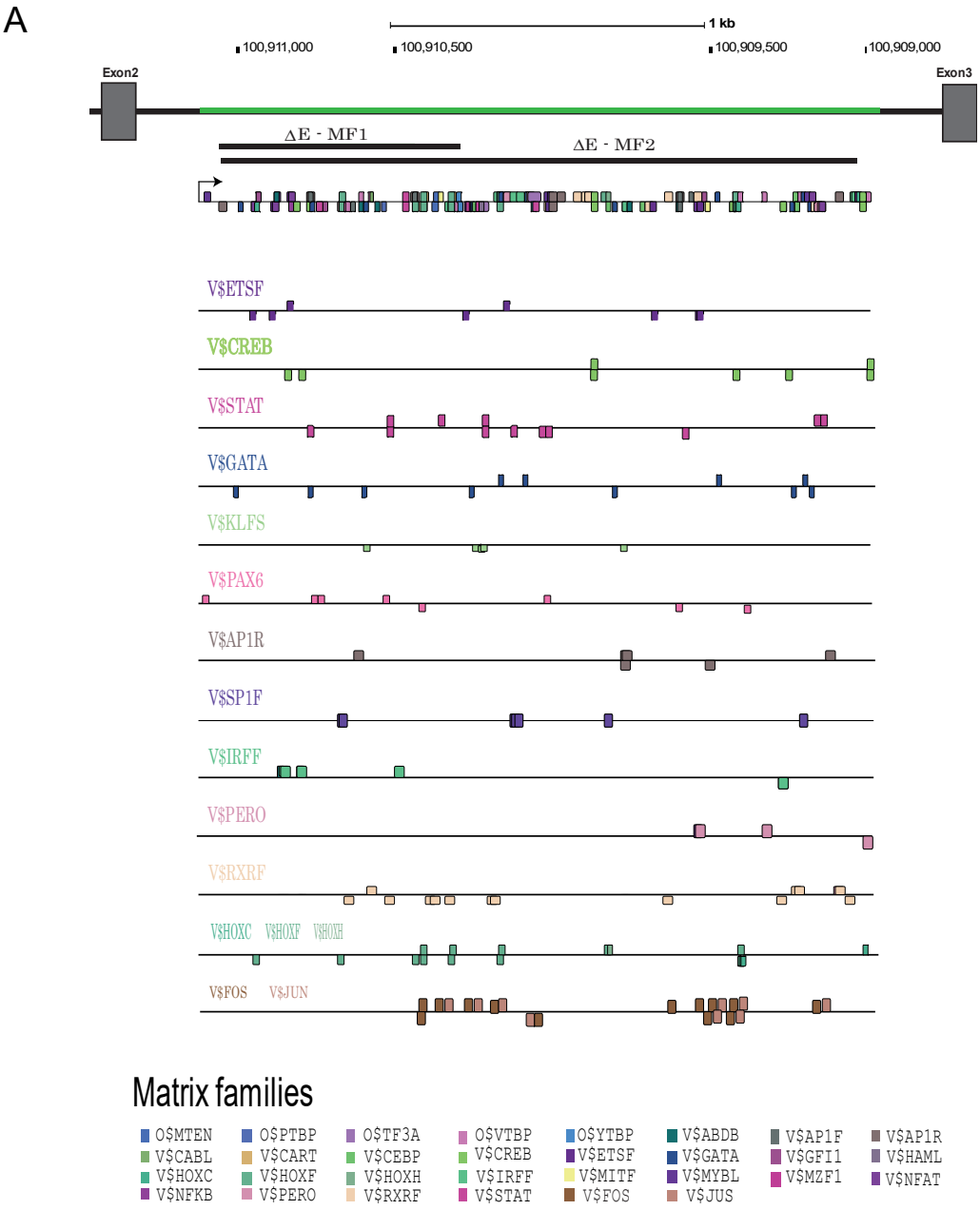


Figure 5: Intron 2 of the murine *Stat3* locus. (A) TF binding sites in the MF putative enhancer on myeloid cells. This area is conserved with human genome: activity in myeloid cells GH17J042346, MF (green) ΔMF-1 and ΔMF-2 (black). (Genomatix , JASPAR database) , p-value 10^{-5}. (B) TF motif for SPI1, C/EBP and STAT3 (Genomatix and JASPAR database)

Stat3 enhancer Deeletion	Date	Total mutants	Homozygote	Compound heterozygotes	chimeras	% of mutant
ΔE MF1	6/8/2019	19	1	3	15	30.2
ΔE MF2	31/12/2019	7	2	1	4	14.8
ΔE T/MF	14/08/2019	16	1	2	13	37.2
ΔE NI	11/8/2019	8	3	0	5	18.6

Table 1. Summary of the generation of Stat3 enhancer mutant mice

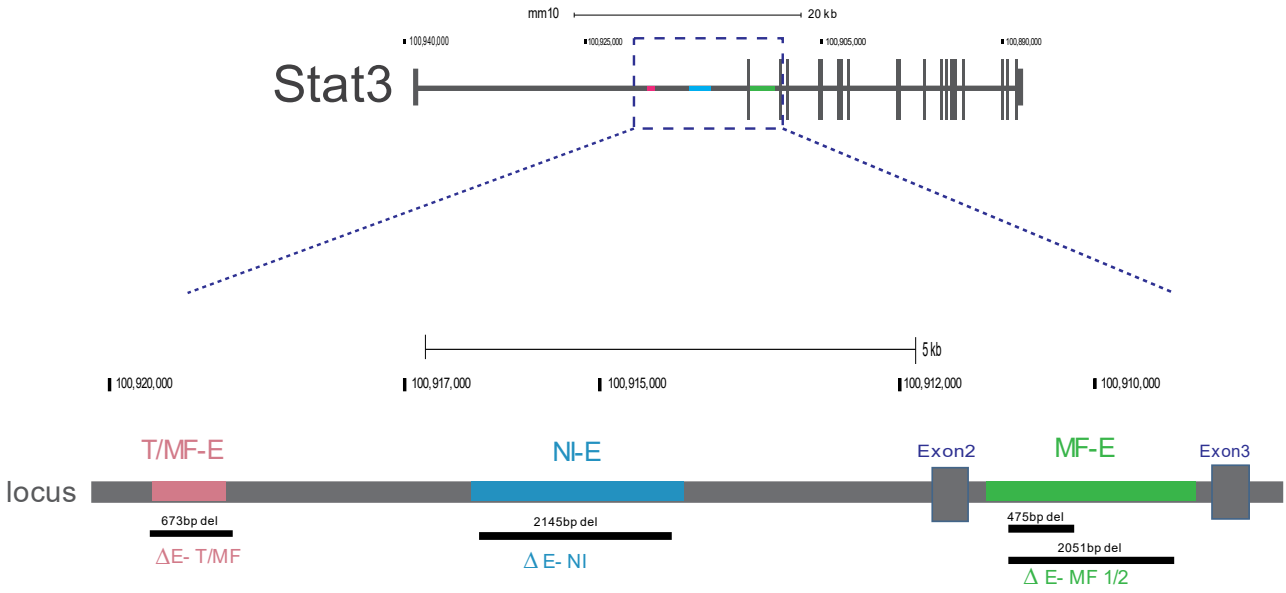


Figure 6: The murine Mouse *Stat3* locus. Mouse conserved sequences to human enhancers and activity in myeloid cells GH17J042346, (MF-E1, MF-E2 (green)) and T cells GH17J042360 (T/MF-E) (pink) and no immune cells-enhancer GH17J042353 (NI-E) (blue).The human mouse conservation enhancer located in introns: intron#1 two enhancers, in intron #2 one enhancer (GH17J042346).

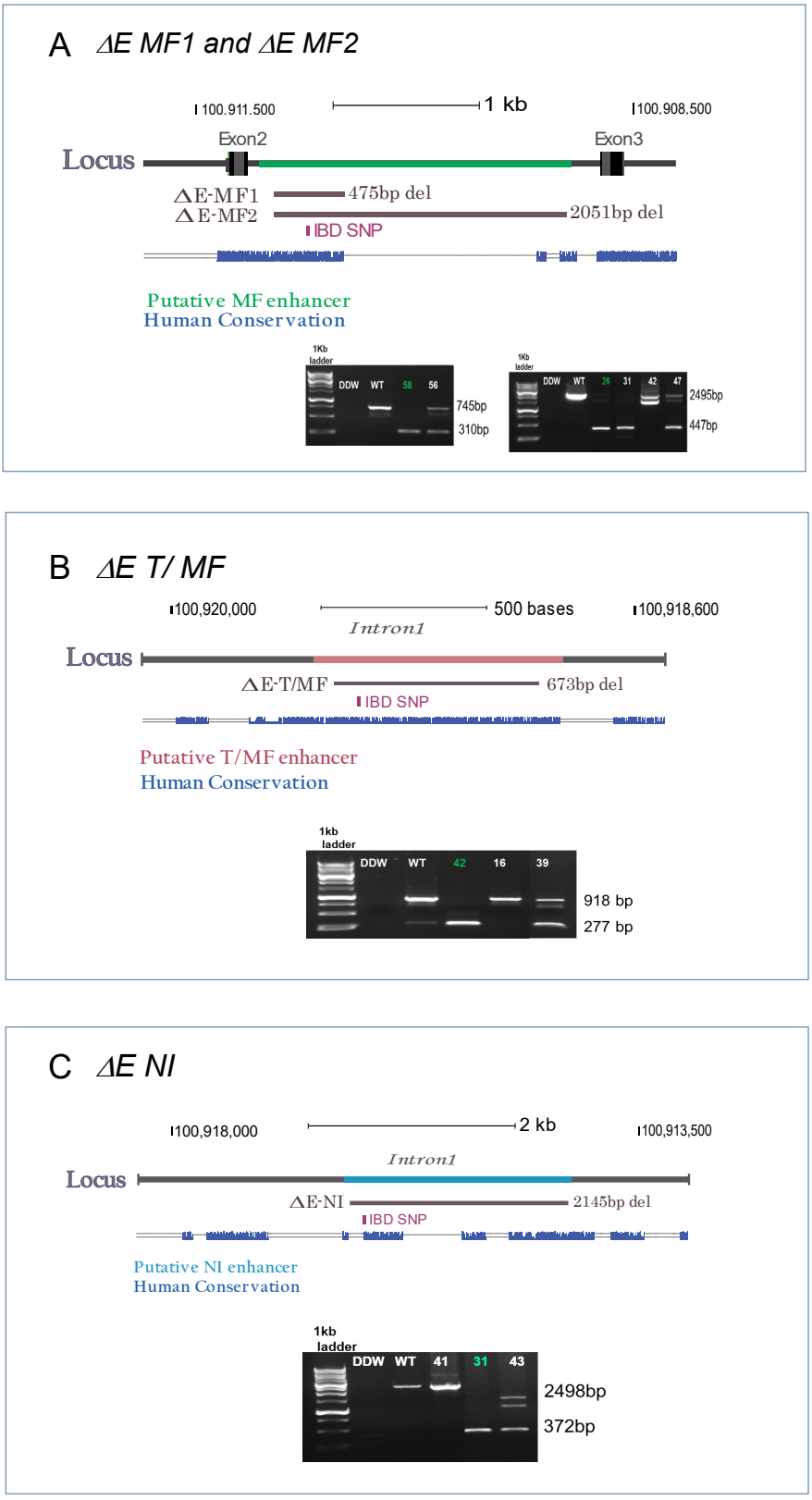


Figure 7. Schematic of CRISPR/Cas9-mediated deletions of regulatory regions in *Stat3* loci and genotyping results of established mutant mouse strains. Mouse STAT3 locus Chromosome 11 qD (GRCm38/mm10). The

human mouse conservation enhancer located in intron #2 and #1. Bottom panel- green- shows conservation details and the human-mouse Alignment Net (GRCh38/hg38).

(A) Putative MF / myeloid cell enhancer (GH17J042346). Representative genomic PCR analysis of mice generated by CRISPR/Cas9 approach. For MF-E1 we obtained 61 pups. We identified by genotyping and sequencing one homozygote mutant strain, mouse #58. The size of the PCR product of WT alleles is 745 bp and of the mutant alleles is 310 bp. For MF-E2 we obtained 47 pups. We identified by genotyping and sequencing one homozygote mutant strain, mouse #26. The size of the PCR product of WT alleles is 2495 bp and of the mutant alleles is 447 bp.

(B) Putative enhancer active in T cells and MF (GH17J042360). Representative genomic PCR analysis of mice generated by CRISPR/Cas9 approach. For T/MF we obtained 43 pups (using electroporation), among them one mouse was identified as being homozygote mutant (mouse #42). The size of the PCR product of WT alleles is 918 bp and of the mutant alleles is 277 bp.

(C) Putative enhancer active in non-immune cells (NI-E) (GH17J042360). Representative genomic PCR analysis of mice generated by CRISPR/Cas9 approach. For NI-E we obtained 43 pups. We identified by genotyping and sequencing one homozygote mutant strain, mouse #31. The size of the PCR product of WT alleles is ~2498 bp and of the mutant alleles is 372 bp.

Analysis of *cis*-impact of macrophage enhancer deletions on allele-specific Stat3 transcription

Enhancers are by definition *cis*-regulatory elements that control the transcription of nearby target genes [24]. Therefore, enhancer effects should in principle be detectable in heterozygote cells, provided that the gene products of the two alleles can be discriminated. If the deletions affect *cis*-regulatory enhancer elements, we reasoned that we might detect their effect on the same allele as a reduced amount of *Stat3* transcription. To this end, we isolated bone marrow (BM) monocytes of the heterozygote mutant mice (ΔE MF1, ΔE MF2, ΔE T/MF) and prepared macrophages (MoMF). Notably, also most gut macrophages are monocyte-derived [47, 48], and although the cultured cells differ from tissue macrophages, MoMF can hence serve as a reasonable proxy. Following culture, RNA was extracted from MoMF and cDNA was synthesized. In general, we designed 5 pairs of primers to detect transcripts emanating from the WT and mutant *Stat3* genes (**Fig 8**). Two of them were located in the area of the putative intronic enhancer, i.e. specific for unspliced RNA: one pair identifying only WT allele product and the second pair of primers detecting only products of the deleted allele. Three other primer pairs were designed to detect transcripts arising from both WT and mutant alleles, including pre-mRNA and spliced mRNA.

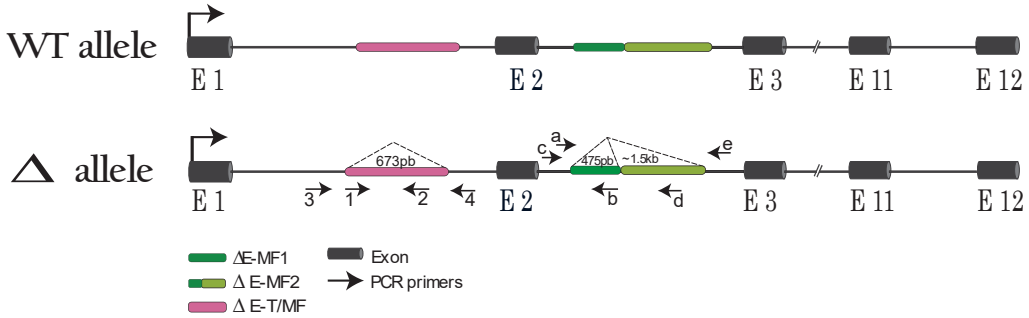


Figure 8. Schematic of strategy to analyze cis impact of Stat3 enhancer deletions in heterozygote mutant cells. For the putative MF / myeloid cell enhancer (GH17J042346)- ΔE MF1 (475bp) and ΔE MF2 (2051bp), located in intron #2 :Primers b+c (product size 123bp) detected only WT alleles since one primer is located inside of the deleted sequence. Primers a+d (product size, mut=118bp, WT=593bp) detected only deleted alleles for ΔE MF1 and primers a+e detected only deleted alleles for ΔE MF2 . Putative enhancer active in T cells and MF (GH17J042360)-ΔE T/MF1(670bp) : located in intron #1 :Primers 1+2 (product size 70bp) detected only WT alleles since one primer is located inside of the deleted sequence. Primers 3+4 (product size, mut=245bp, WT=918bp). Internal control located on intron #11 (product size 92bp).

The ΔE MF1 deletion is located in intron 2 of the *Stat3* gene. Primers b+c detected only WT alleles since one primer is located inside of the deleted sequence. Primers a+d detected only deleted alleles, since the WT PCR product was with 475bp too big for real-time PCR (**Fig 9A**). Analysis of WT and $Stat3^{\Delta E\ MF1/WT}$ MoMF by real-time PCR revealed that Stat3 transcription in WT cells was higher than in heterozygote mutant cells. This is due to the fact that in $Stat3^{WT/WT}$ cells primers b+c detect products from both alleles, but in $Stat3^{\Delta E\ MF1/WT}$ cells, the primers detected only transcripts from the WT allele (**Fig 8, 9A**). The real-time PCR with Primers a+d validated the Δ*E MF1* deletion in $Stat3^{\Delta E\ MF1/WT}$ MoMF. In addition, we performed qPCR with primers specific for *Stat3* intron #11, and pair of primers to detect pre-mRNA, and mRNA on exon 8-9 (**Fig 9A**). For the small deletion (ΔE MF1) transcriptome levels of Stat3 were not found to be affected, i.e. there was no difference in the amount of *Stat3* transcripts detected in $Stat3^{WT/WT}$ and $Stat3^{\Delta E\ MF1/WT}$ cells. Importantly, we also performed an RT-PCR for ΔE MF1 with same primers, as used for the genotyping. The PCR products confirmed the $Stat3^{\Delta E\ MF1}$ and $Stat3^{WT}$ allele status, 375bp and 745bp, respectively. Absence of a product in the RNA samples before cDNA preparation confirmed that the products we got with the primers for pre-RNA are not due to DNA contamination (**Fig 9E**).

Also, the ΔE MF2 deletion is located in intron 2 of the *Stat3* gene. Primers b+c detected only WT alleles since one primer is located inside of the deleted sequence (the same primers like ΔE MF1). Primers a+e detect only deleted alleles, since the WT PCR product was with ~2kb too big for real-time PCR (**Fig 8, 9B**). For the large deletion (ΔE MF2) transcriptome levels of Stat3 were not found affected,

i.e. there was no difference in the amount of *Stat3* transcripts detected in $Stat3^{WT/WT}$ and $Stat3^{\Delta EMF2/WT}$ cells.

Collectively, this analysis suggests that at least in the in vitro assays the deleted enhancer elements are not critical for Stat3 expression. However, the function of the putative enhancer elements might only be revealed upon stimulation of the cells. We therefore repeated the analysis with cells that had been exposed to 3-deoxy-D-*manno*-octulosonic acid $(Kdo)_2$-Lipid A (Kdo_2-Lipid A , KLA) for 6hrs [49]. KLA is a synthetic nearly homogeneous lipopolysaccharide (LPS) substructure with endotoxin activity equal to that of native LPS. However, KLA is more potent and absolutely TLR4 specific [50]. Six hours post stimulation *Stat3* transcription levels were not significantly affected in $Stat3^{\Delta E\ MF2/WT}$ MoMF as compared to WT cells (**Fig 9C**). These data show that, at least in the in vitro system and MoMF, Stat3 transcripts can be efficiently induced in MF that lack the complete putative MF enhancer in intron 2 of the *Stat3* locus.
The ΔE T/MF deletion is located in intron 1 of the *Stat3* gene. We designed Primers 1+2 that detected only WT alleles and Primers 3+4 that detected only deleted alleles, since the PCR product of the WT locus would be too big for real time PCR (~670bp) (**Fig 8**). Primers 3+4 validated the deletion in heterozygote mutant $Stat3^{\Delta E\ T/MF/WT}$ MoMF (**Fig 9D**). The PCRs for transcripts arising from both the WT and the mutant loci showed that *Stat3* transcription levels were not significantly affected, as compared to WT cells.

In conclusion, the in vitro analysis of MoMF harboring an ΔE MF1/ 2 deletion did not provide evidence that that the predicted MF enhancer are critical *cis*-regulatory element required for Stat3 expression in steady state or following activation.

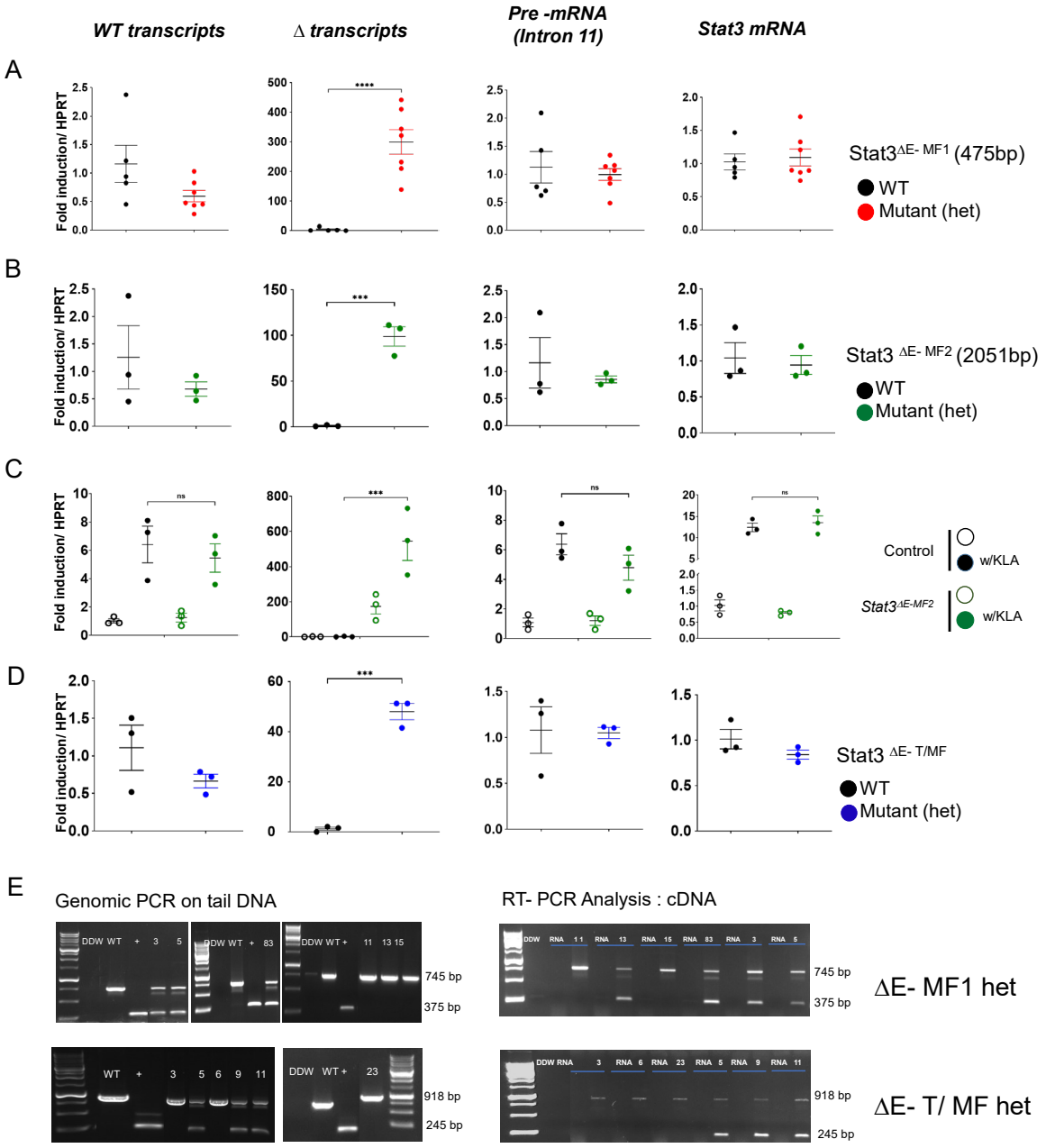


Figure 9. Analysis of *cis*-impact of enhancer deletions on allele-specific Stat3 transcription: qPCR analysis for Stat3 transcription in heterozygote and WT macrophages (MoMF) . For each mutant, we designed primers that detected only WT, mutant alleles, and three primers as controls for all the mutant strains, two of them detecting pre- mRNA, and one for mRNA.

A) ΔE MF1 heterozygoet (small deletion) -Primers b+c detected only WT, for both ΔE MF1 , Stat3 transcription in WT cells was higher than in heterozygote mutants (p-value < 0.05), a+d primer detected for Δ E MF1, only mutant alleles.

B) ΔE MF2 heterozygoete (big deletion) - Primers b+c detected only WT, for both ΔE MF1 , Stat3 transcription in WT cells was higher than in heterozygote mutants (p-value < 0.05), a+d primer detected for Δ E MF1, only mutant alleles.

C) Kdo2-Lipid A (KLA) treatment for ΔE MF2 heterozygotes

D) ΔE T/MF /WT - for this mutant we showed the same pattern as ΔE MF1/wt.

E) RT-PCR and genomic PCR on tail DNA - we used the same primers for both PCRs . ΔE MF1 /WT =n=6 (2WT mice and 4 mutant mice), ΔE T/ MF / WT n=3-6, ΔE MF1 /WT =n=4 (3 femalee 1 male , 3 technical repeat for WT) Unpaired t-test Two-tailed analysis. KLA= 0.1ug/ml, cell number = $1x10^6$

Probing spontaneous colitis development in Stat3 enhancer mutant animals and challenge of colitis- animals with piroxicam

Like children that harbor IL10 receptor (IL10R) deficiencies, also *CX₃CR1^{Cre}:IL10ra$^{fl/fl}$* mice whose gut macrophages fail to respond to IL10 develop spontaneous colitis [11, 13]. Likewise, animals that harbor a Stat3 deficiency in macrophages develop gut inflammation [25, 26], although this signal transducer is involved with multiple cytokine receptors. If the Stat3$^{\Delta E\ MF1}$ and Stat3$^{\Delta E\ MF2}$ mutations affect macrophage Stat3 expression, the respective homozygote mice should display deficient IL10R signaling and might hence be expected to show signs of colitis, either spontaneous or following challenge.

To probe for signs of pathology, homozygote mutant Stat3$^{\Delta E\ MF1}$, Stat3$^{\Delta E\ T/MF}$ and Stat3$^{\Delta E\ NI}$ animals that did not show any overt phenotypes were subjected to weight analysis and Fecal lipocalin-2 (Lnc2) measurement which revealed no significant differences between the groups (**Fig 10A, B)**.

Fecal lipocalin-2 (Lnc2) is indicative of colitis or IBD [51] and can serve as highly sensitive biomarker for gut inflammation [52]. We therefore decided to measure Lnc2 levels in feces of *Stat3$^{\Delta E\ MF1}$, Stat3$^{\Delta E\ T/MF}$* and *Stat3$^{\Delta E\ NI}$* animals and controls. As positive control we used feces of colitic *CX₃CR1^{Cre}:IL10ra$^{fl/fl}$* mice. Cre-negative *IL10ra$^{fl/fl}$* ***and*** *IL10ra$^{fl/wt}$* littermates served as negative controls. Feces of sick *CX₃CR1^{Cre}:IL10ra$^{fl/fl}$* mice [12, 16] as expected showed high Lnc2 levels in the ELISA assay (**Fig 10 C**). Although there was some variation between the different groups, analysis of feces of homozygous mutant unchallenged *Stat3$^{\Delta E\ MF1}$, Stat3$^{\Delta E\ T/MF}$* and *Stat3$^{\Delta E\ NI}$* animals showed however no consistently elevated Lnc2 levels, as compared to negative controls (**Fig 10B)**. In conclusion, both weight measurements and fecal Lnc2 analysis of *Stat3$^{\Delta E\ MF1}$, Stat3$^{\Delta E\ T/MF}$* and *Stat3$^{\Delta E\ NI}$* animals suggest that these animals do not display signs for spontaneous gut inflammation.

To test whether the mutant animals might be more sensitive to challenge, we treated the animals with Piroxicam, a non-steroidal anti-inflammatory drug (NSAID) [39], which inhibits cyclooxygenase (COX) enzymes, causing reduction of prostaglandin E2 levels [38]. Colitis-prone animals, such as IL10 deficient mice, develop Piroxicam accelerated colitis (PAC), as compared to WT controls [26, 27]. The treatment is thought to impair mucosal integrity and cause penetration of luminal bacteria. To test the resistance of *Stat3$^{\Delta E\ MF1}$, Stat3$^{\Delta E\ T/MF}$* and *Stat3$^{\Delta E\ NI}$* animals to Piroxicam challenge we exposed them to the drug in the diet .For this experiment we used 10-12 weeks old homozygote mutant males. Homozygote mutant ΔE MF2 mice were not yet available when the experiment was performed. Colitis-prone *Cx₃cr1^{Cre} :IL-10R$^{fl/fl}$* mice [12, 16] served as positive controls. Weight and fecal Lpcn2 ELISA

measurement for one week, starting from the day of addition of the drug to the diet (day 0) (**Fig 10 C, E**), did not show a consistent significant difference between the tested animals. Specifically, piroxicam challenge induced weight loss of *Cx₃cr1*^{Cre} *:IL-10R*$^{fl/fl}$ mice , but not in any of the other groups, except for day 10 measurement of *Stat3*$^{\Delta E\ MF1}$ mice. The fecal Lipocalin- 2 found to be significantly elevated in *Stat3*$^{\Delta E\ MF1}$ mice following the Piroxicam treatment (**Fig 10D, F**), potentially indicates moderate inflammation.

In conclusion, neither *Stat3*$^{\Delta E\ MF1}$, *Stat3*$^{\Delta E\ T/MF}$ and *Stat3*$^{\Delta E\ NI}$ animals responded to the Piroxicam challenge with consistent robust development of colitis. ΔE MF1 mice displayed signs of mild inflammation in the PAC paradigm.

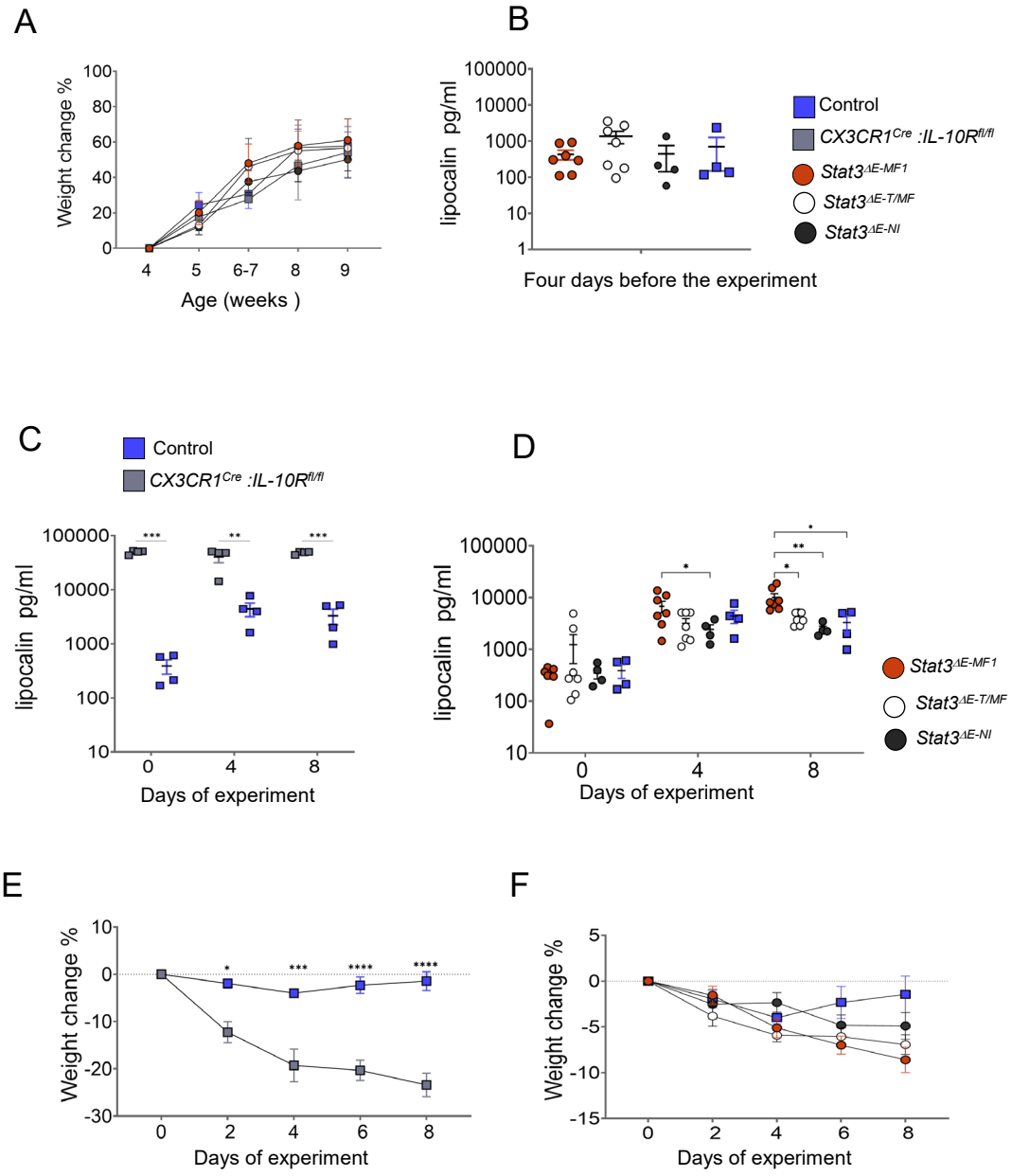


Figure 10. Piroxicam-accelerated colitis (PAC) and Lipocalin -2 (Lnc-2) measurement

A) Weight measurement of indicated mice harboring homozygous deletions of Stat3 enhancers before challenge including $Cx3cr1^{Cre}$:IL-$10R^{fl/fl}$ mice and controls (IL10r $^{fl/fl}$ and IL10$r^{fl/wt}$). $Stat3^{\Delta E\text{-}MF1}$:n= 6, $Stat3^{\Delta E\text{-}T/MF}$ n= 6, $Stat3^{\Delta E\text{-}NI}$: n= 6, $Cx3cr1^{Cre}$:IL-$10R^{fl/fl}$ n=2 , controls (IL-$10R^{fl/fl}$:IL-$10R^{fl/wt}$) n=3

B) Lipocalin measurement of indicated mice harboring homozygous deletions of Stat3 enhancers before challenge including controls (IL10r $^{fl/fl}$ and IL10$r^{fl/wt}$).

C) Lipocalin measurement after challenge of $Cx3cr1^{Cre}$:IL-$10R^{fl/fl}$ mice and controls

D) Lipocalin measurement of indicated mice harboring homozygous deletions of Stat3 enhancers after challenge including controls (IL10r $^{fl/fl}$ and IL10$r^{fl/wt}$).

E) Weight measurement after challenge of $Cx3cr1^{Cre}$:IL-$10R^{fl/fl}$ mice and controls (IL10r $^{fl/fl}$ and IL10$r^{fl/wt}$).

F) Weight measurement of indicated mice harboring homozygous deletions of Stat3 enhancers after challenge, including controls (IL10r $^{fl/fl}$ and IL10$r^{fl/wt}$). animals were exposed to Piroxicam food for 10 days. Homozygous mice : $Stat3^{\Delta E\text{-}MF1}$:n= 7, $Stat3^{\Delta E\text{-}T/MF}$ n= 7, $Sat3^{\Delta E\text{-}NI}$: n= 4, $Cx3cr1^{Cre}$:IL-$10R^{fl/fl}$ n=5 , controls (IL-$10R^{fl/fl}$:IL-$10R^{fl/wt}$) n=4. one-way ANOVA analysis, and two-way ANOVA analysis was preformed.

Analysis of BM chimeras for spontaneous colitis development and piroxicam-accelerated colitis (PAC)

Previously, we had established that irradiated WT recipient animals that are transplanted with *$Cx3cr1^{Cre}$:Il-$10r^{fl/fl}$* BM develop colitis within 6-7 weeks following engraftment [12]. We therefore decided to use this sensitive approach to independently test whether immune cells harboring the ΔE MF1 or ΔE MF2 enhancer mutation might cause gut inflammation. [*$Cx3cr1^{Cre}$:IL-10Rfl/fl* > WT] chimeras that we generated as positive controls showed clear signs of pathology as indicated by the impaired weight gain (**Fig 11A)**. In contrast, [*$Stat3^{\Delta E\ MF1}$* > WT] and [*$Stat3^{\Delta E\ MF2}$* > WT] chimeras gained similar weight during the time of analysis. To investigate whether the chimeras generated with the mutant BM might be more sensitive to challenge, we treated them with Piroxicam. At week eight post-transplantation, we exposed [*$Stat3^{\Delta E\ MF1}$* > WT], [*$Stat3^{\Delta E\ MF2}$* > WT] and [WT > WT] chimeras to the drug in the diet. Weight measurement did not show a significant difference between the animals (**Fig 11 B**). In order to study the effect of the mutations on the generation and differentiation the donor cells from hematopoietic stem cells (HCSs) [53], we performed a peripheral blood analysis Flow cytometric analysis of the blood of ΔE MF1 and ΔE MF2 chimeras for different myeloid cells, including classical and non-classical monocytes, as well as neutrophils did not reveal alterations as compared to controls (**Fig 11 C**). Notably, [*$Cx3cr1^{Cre}$:IL-10Rfl/fl* > WT] chimeras showed slightly elevated classical monocytes numbers indicating inflammation. B cell served as control.

We can conclude that no sign of gut pathology or leukocyte alterations were observed for the mutant chimeric animals, neither spontaneous inflammation nor after the challenge.

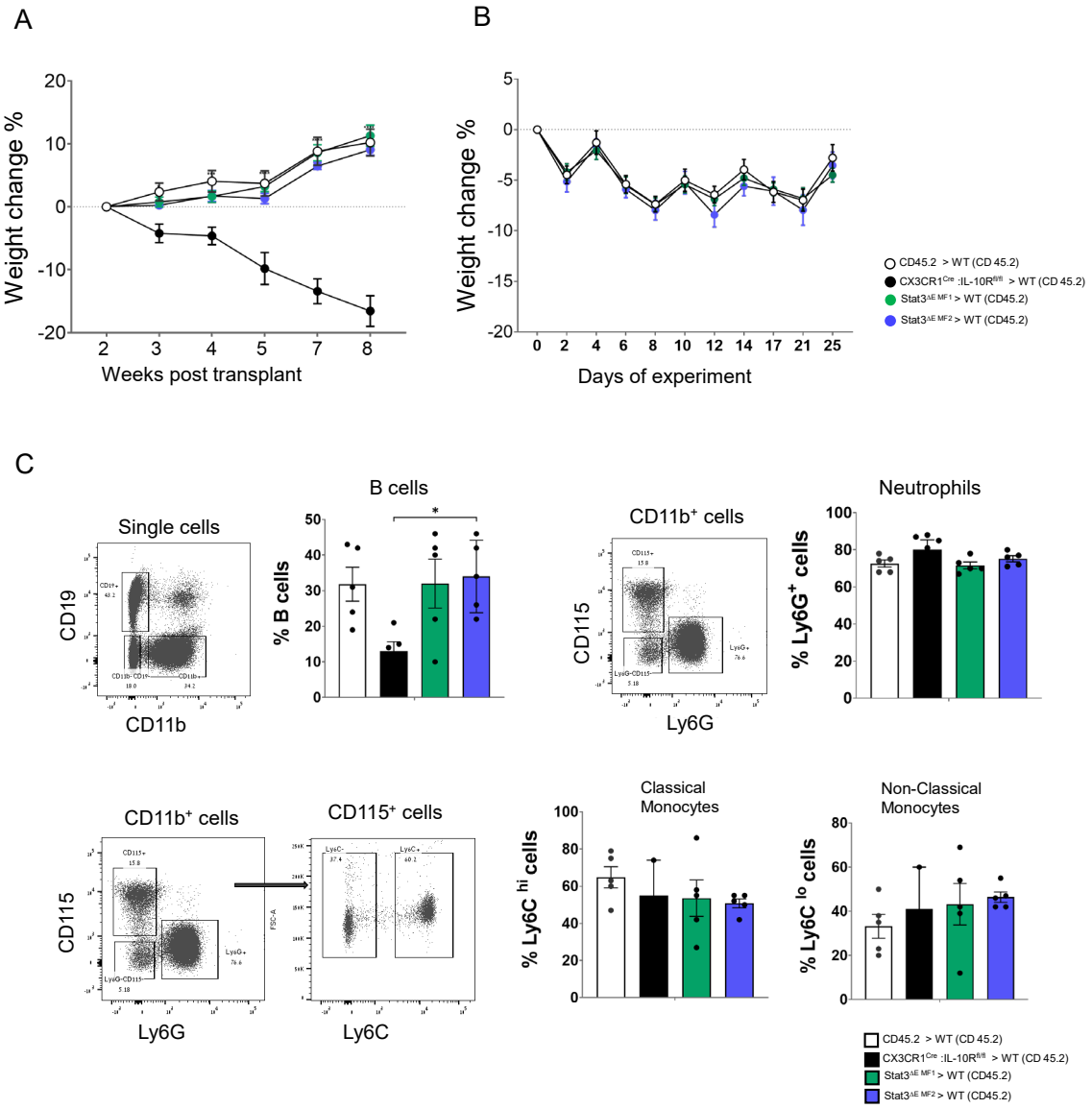


Figure 11. Analysis of BM chimeras.

(A) Weight analysis of [Stat3$^{\Delta E\text{-}MF2}$ > WT (CD45.2)] , [Stat3$^{\Delta E\text{-}MF1}$ > WT (CD45.2)] [*Cx₃cr1^{Cre}:IL-10R$^{fl/fl}$* > CD45.2 (WT)] and CD45.2 >WT (CD45.2) chimeras .

(B) Weight measurements after piroxicam treatment

(C) Flow cytometric blood analysis of BM chimeras.

n=8-10 in each and number of mice for blood analysis n= 4-5, one-way ANOVA analysis

Probing for macrophage intrinsic defects of the ΔE MF1 and ΔE MF2 enhancer mutations in mixed BM chimeras

Mixed BM chimeras generated with WT and mutant BM allow for a very sensitive comparative analysis and can reveal early cell-intrinsic events. We therefore next generated mixed BM chimeras focusing first on the *Stat3$^{\Delta E\ MF2}$* animals that harbor the extended deletion (**Fig 12 A**). Specifically, irradiated WT recipient mice (CD45.2) were engrafted with a mixture of BM isolated from homozygote mutant *Stat3$^{\Delta E\ MF2}$* and CD45.1^{+} WT mice, respectively. As controls, we generated mixed chimeras with *Cx₃cr1^{Cre} :Il-*

10r$^{fl/fl}$ and WT BM. As already mentioned, [*Cx₃cr1*^{Cre} *: Il-10r*$^{fl/fl}$ > WT] chimeras develop signs of colitis [25], but mixed chimeras with *Cx₃cr1*^{Cre} *: Il-10r*$^{fl/fl}$ and WT BM had not been analyzed before. [WT / WT > WT] and [*Stat3*$^{\Delta E\ MF2}$ / WT > WT] chimeras displayed similar weight gain following the irradiation and engraftment (**Fig 112B**). In line with potential mild colitis development, weight gain of the *Cx₃cr1*^{Cre} *:IL-10R*$^{fl/fl}$ BM recipients was attenuated (**Fig 12B**). [*Stat3*$^{\Delta E\ MF2}$ / WT > WT] and [WT / WT > WT] chimeras however displayed similar weight gain following
the irradiation and engraftment (**Fig 12B**). Blood analysis (**Fig 12C**) of the mixed [*Stat3*$^{\Delta E\ MF2}$ / WT (CD45.1) > WT (CD45.2)] chimeras and controls showed an equal representation of CD45.1 mutant and CD45.2 WT HSC-derived cells. Interestingly though, these animals showed a strong bias of non-classical monocytes towards the CD45.1 haplotype. This might be related to additional genetic differences between the strains beyond the CD45 allele, whose impact in BM chimeras has been previously reported [54].

Mixed BM chimeras generated from *Stat3*$^{\Delta E\ MF1}$ mice harboring the small deletion with positive and negative controls (**Fig 13A, B**) gave similar results as the ones obtained with *Stat3*$^{\Delta E\ MF2}$ BM, Chimeras harboring mutant BM gained weight following the irradiation and engraftment, as compared to the controls harboring colitogenic *Cx₃cr1*^{Cre} *:IL-10R*$^{fl/fl}$ cells. Blood analysis (**Fig 13C**) of the mixed [*Stat3*$^{\Delta E\ MF1}$/ WT (CD45.1) > WT (CD45.2)] chimeras and controls showed an equal representation of CD45.1 mutant and CD45.2 like mixed *Stat3*$^{\Delta E\ MF2}$ chimeras.

To investigate monocyte differentiation into gut macrophages in more detail we performed a flow cytometric analysis of the characteristic P1-P3 'waterfall' [17] (**Fig 14**). Alterations in the frequency of the three phenotypically defined differentiation stages, and specifically differences between CD45.1^{+} (WT) and CD45.2^{+} (*Cx₃cr1*^{Cre} *:IL-10R*$^{fl/fl}$ or *Stat3*$^{\Delta E\ MF1}$) cells, might indicate intrinsic impairments due to deficient signaling of IL10R or other Stat3-associated surface receptors. Notably, sick animals were reported to display a higher percent of Ly6C$^{hi\text{-}int}$ MHC2 ^{hi} P2 cells [17], which we could confirm in our study (**Fig 14A**). Gut analysis of [*Stat3*$^{\Delta E\ MF1}$ / WT > WT], [*Cx₃cr1*^{Cre} *:IL-10R*$^{fl/fl}$ / WT> WT and [WT / WT > WT] showed that the ratios of CD45. 1 and CD45.2 cells were is similar between the groups, for both colon and cecum (**Fig 14B)**. Likewise, the dynamics of the P1-P3 waterfall were similar between the CD45.1 and CD45. 2 cells. In the [*Cx₃cr1*^{Cre} *:IL-10R*$^{fl/fl}$ / WT> WT] chimeras the P2 frequency was higher for CD45.1 (WT) cells in colon and cecum. Finally, we also analyzed the abundance of neutrophils in the chimeras, as a proxy for signs of inflammation (**Fig. 14C**). In [*Cx₃cr1*^{Cre} *:IL-10R*$^{fl/fl}$ / WT > WT] chimeras neutrophils were found elevated in line with the fact that these animals develop pathology. The other chimeras did not show evidence for neutrophil accumulation.
Collectively, the detailed analysis of the mixed BM chimeras did not reveal evidence for early cell-intrinsic defects of the *Stat3*$^{\Delta E\ MF1}$ and *Stat3*$^{\Delta E\ MF2}$ mutant cells, and also corroborated the earlier notion that these animals do not show signs of inflammation.

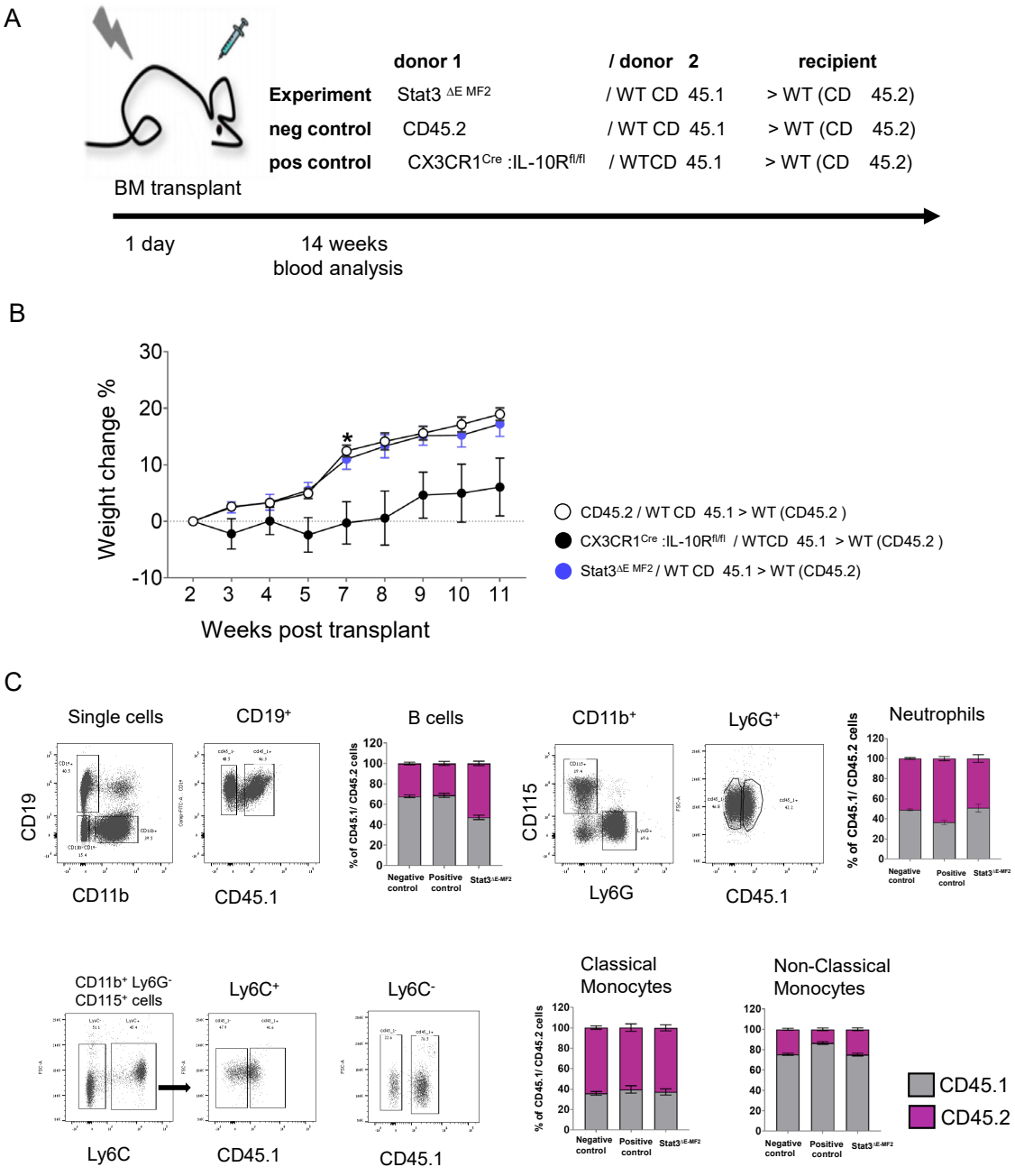


Figure 12 . Analysis of mixed BM chimeras - ΔE MF2.

(A) Schematic of BM chimera generation and analysis

(B) Weight analysis of *[Stat3$^{\Delta E\text{-}MF2}$/* CD45.1 > CD45.2 (WT)] ,[*Cx$_3$cr1^{Cre}:IL-10R$^{fl/fl}$* / CD45.1 > CD45.2 (WT)] and CD45.2 / WT CD45.1 > WT (CD45.2) chimeras .

(C) Flow cytometric blood analysis of mix BM chimeras 4 weeks post-transplant. n=10 and # of mice for blood n= 4-5 , one-way ANOVA

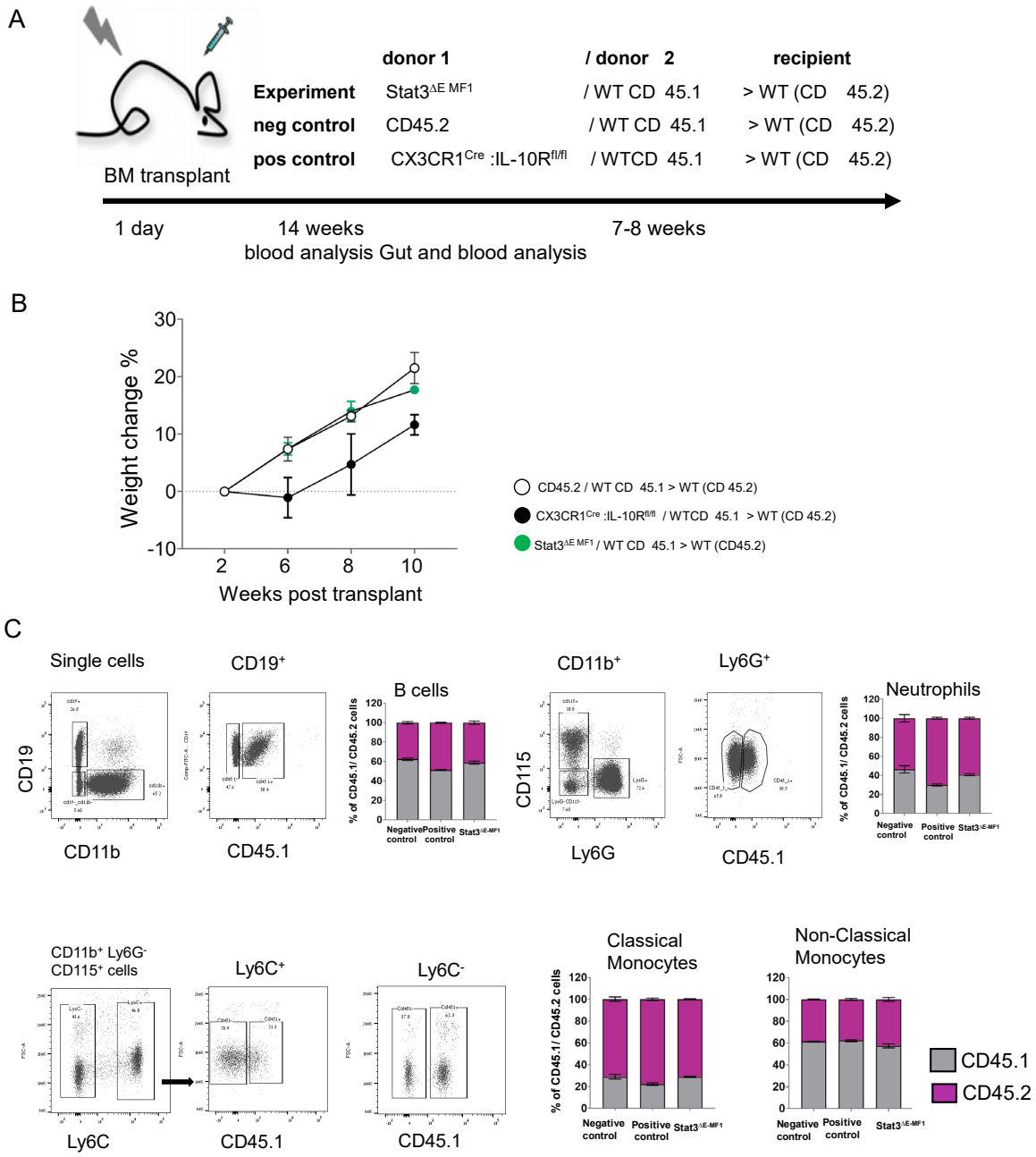


Figure 13 . Analysis of mixed BM chimeras - ΔE MF1.

(A) Schematic description of BM chimera experiment

(B) Weight analysis of [Stat3$^{\Delta E\text{-}MF1}$/ CD45.1 > CD45.2 (WT)] ,[*Cx₃cr1*^{Cre}*:IL-10R*$^{fl/fl}$ / CD45.1 > CD45.2 (WT)] and CD45.2 / WT CD45.1 > WT (CD45.2) chimeras . n= 5-7

(C) Flow cytometric blood analysis of mix BM chimeras, 2 weeks post transplant. n= 3-4 one-way ANOVA

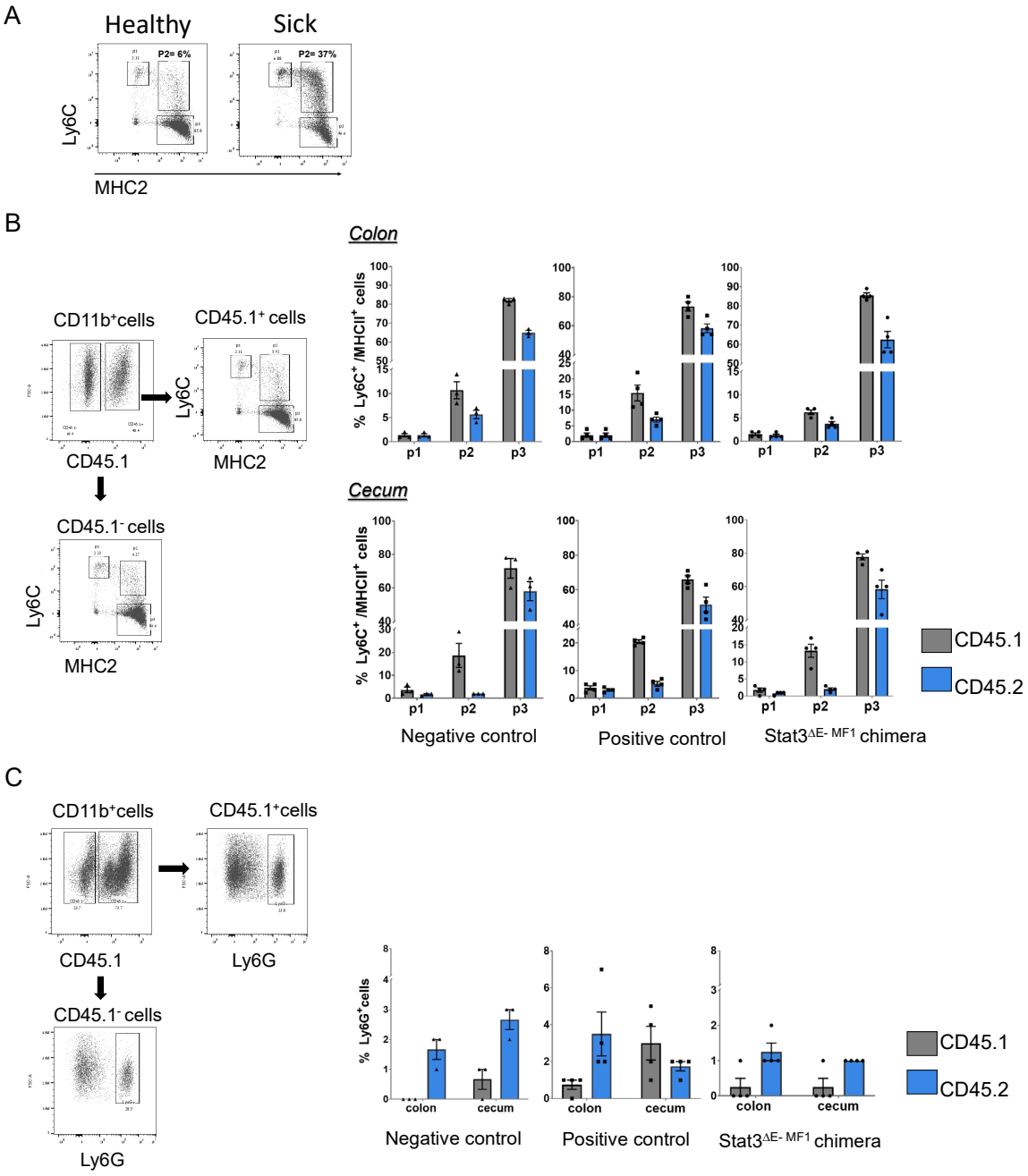


Figure 14 . Gut analysis of mixed BM chimeras - ΔE MF1.

Gut (colon and cecum) analysis of [*Stat3*$^{\Delta E\text{-}MF1}$/ CD45.1 > CD45.2 (WT)] and [*Cx*$_3$*cr1*^{Cre}*:IL-10R*$^{fl/fl}$ / CD45.1 > CD45.2 (WT)] chimeras. Representative picture of flow cytometry analysis of gut monocytes, macrophages (P1-P3, "waterfall") ,neutrophils.

(A) Flow cytometry analysis of healthy and sick mice to show monocyte differentiation path (P1-P3, "waterfall").

(B) Flow cytometry analysis of % CD45.1+ and CD45.2+ cells for each chimera, then analysis of (p1-p3, "waterfall"). colon, n=5 , cecum n=4,

(C) Abundance of neutrophils . n=3 two-way ANOVA analysis

Analysis of abundance of long-lived Tim4^{+} intestinal macrophages

Intestinal resident macrophages include two types of macrophages: short-lived macrophages, which rely on continuous replenishment from monocytes , via the 'monocyte waterfall' –and long-lived macrophages. The discovery of the long-lived macrophages by Grainger and colleagues challenged the original concept that all gut macrophages are short-lived and rely on continuous replenishment by monocytes [14, 15]. Long-lived macrophages, which could potentially like other tissue macrophages also self-maintain by low level proliferation are currently defined by the surface markers Tim4 and CD4. To probe for a potential impact of the *Stat3*$^{\Delta E\text{-}MF1}$ and *Stat3*$^{\Delta E\text{-}MF2}$ mutations on intestinal macrophage differentiation and turnover, we performed a comprehensive flow cytometry analysis. Specifically, we analyzed Stat3$^{\Delta E\text{-}MF1}$, *Stat3*$^{\Delta E\text{-}MF2}$, and WT mice of different ages. We isolated macrophages from the cecum and colon, for both tissues, the mutant mice did not show alterations in the abundance of monocytes, intermediate and mature macrophages as identified as P1, P2, and P3 population in the waterfall scheme (**Fig 15A**). However, mice harboring deletions in the *Stat3* intron showed a significant reduction of the population of long-lived CD4^{+} Tim4^{+} macrophages compering to WT from Harlan and not from our facility. (**Fig 15, B C**). When comparing *Plxna4*$^{fl/fl}$ mice from our facility to Stat3MF mutant, we found that both CD4+ Tim4+ macrophages decreased. This result may be related to the commensal microbiome. However it might be worth to repeat this experiment with appropriate controls, such as WT mice from our facility and of the same age to address this point.

To test if the effect was cell-intrinsic or if the mutant macrophages cause tissue inflammation that impedes the establishment of long-lived CD4^{+} Tim4^{+} macrophages. we generated mixed BM chimeras with wt C57Bl/6 BM (CD45.1) and *Stat3*$^{\Delta E\text{-}MF2}$ BM. Eight months after their generation, the resulting chimeras were analyzed by flow cytometry to distribute WT and mutant cells among the P1-P3 ,'waterfall' populations and the long-lived CD4+ Tim4+ macrophages. As seen in **Fig 16** no significant difference between the ratio of CD45.1 (WT) and the CD45.2 (mutant) for the P1 and P2 waterfall populations in BM chimeras *Stat3*$^{\Delta E\text{-}MF2}$ BM and wt C57Bl/6 BM (CD45.1). The CD45.1/2 ratio of the P3 (Tim4+CD4+, Tim4-CD4+, and Tim4-CD4-) sub-population was decreased in the *Stat3*$^{\Delta E\text{-}MF2}$ BM compared to WT mix chimeras in the cecum, although not reaching significance. The same subpopulation in the colon showed an increase of the CD45.1/ 2 ratio of, indicating a reduction of mutant cells (CD45.2) in the *Stat3*$^{\Delta E\text{-}MF2}$ BM.

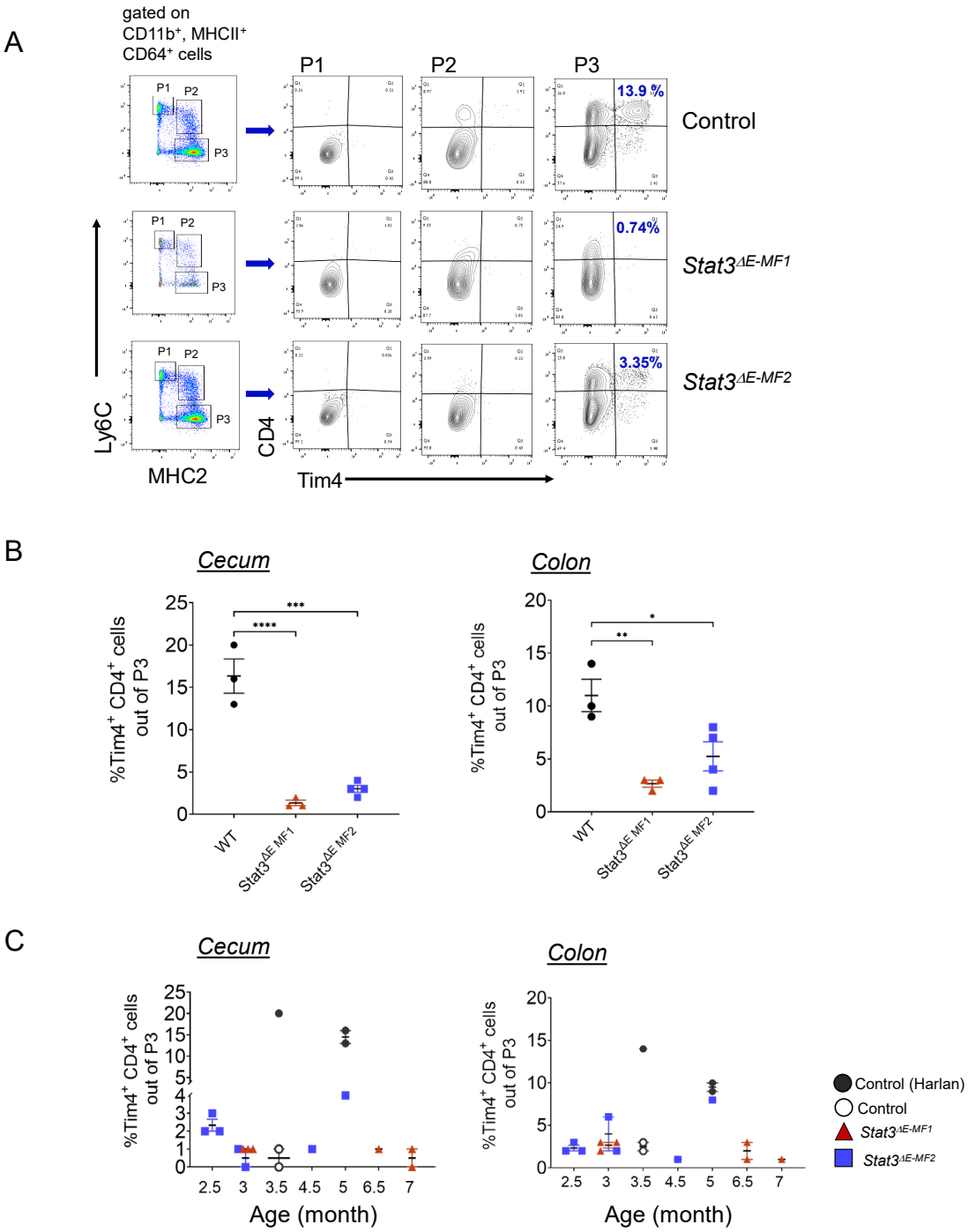


Figure 15. Gut analysis long-lived Tim4⁺ intestinal macrophages

(A) Flow cytometry analysis schematic

(B) The percent of long live macrophage Tm4⁺ CD4⁺ cells from P3 pupletion, for cecum and colon .

(C) Timeline of ages for the mice. *Stat3*$^{\Delta E\text{-}MF1}$, *Stat3*$^{\Delta E\text{-}MF2}$, WT (BL6 mice from Harlan) *Plxna4*$^{fl/fl}$ control from our facility. All of them are males , n=2-7 ,one-way ANOVA analysis

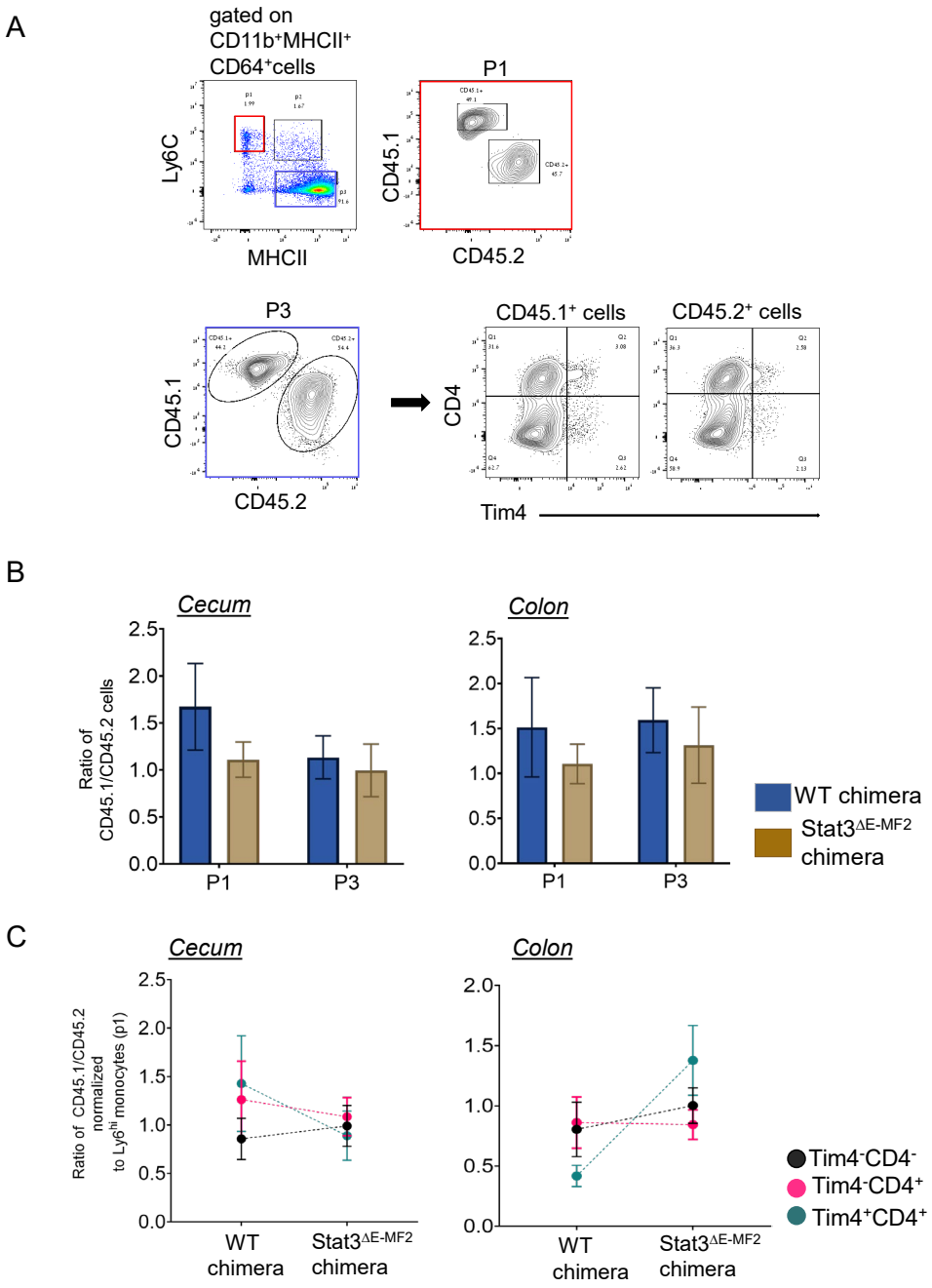


Figure 16 : ΔE-MF2 mixed BM chimeras mice, gut analysis long-lived Tim4^{+} intestinal macrophages.

(A) Flow cytometry analysis schematic

(B) The ration of CD45.1/CD45.2 for p1 and p3 population [*Stat3*$^{\Delta E\ MF2}$/ CD45.1 > CD45.2 (WT)] and CD45.2 / WT CD45.1 > WT (CD45.2) chimeras .

(C) The ration of CD45.1/CD45.2 for Tim+ CD4+ population (flow cytometry analysis) 8 month post trnaplentation . n= 4-5 one-way ANOVA

In vitro analysis of MF harboring homozygote ΔMF-1 and ΔMF-2 mutations with and without challenge

Following the Piroxicam challenge mice harboring the small ΔE MF1 deletion showed mild signs of inflammation **(Fig 10D)**. This might suggest that the enhancer mutation impairs binding of signal-dependent TF (SDTFs) and that activity of the enhancer element requires prior cell activation. In order to investigate the expression of Stat3, selected target genes and cytokines post stimulation, we prepared BMDM of homozygote mutant *Stat3*$^{\Delta E\ MF1}$ mice, as well as WT controls.

Interestingly, following 6hr stimulation with stimulation we used KLA [49], homozygote mutant *Stat3*$^{\Delta E\ MF1}$ BMDM showed elevated *Stat3* pre-mRNA and mRNA levels, as compared to WT BMDM although not reaching significance in the former case **(Fig 17A, 18 A)**. Likewise, transcripts of pro-inflammatory cytokines (IL1β, TNFα, and IL-6) were found increased in the mutant cells following stimulation though not reaching significance in all cases (**Fig 17A, 18B)**. Furthermore, we also observed significant elevated expression of anti-inflammatory IL-10. Western blot analysis of Stat3 protein revealed comparable expression in WT and ΔE MF1 mutant BMDM (**Fig 17C).** Also phospho-Stat3 after KLA stimulation was induced to a similar in WT and ΔE MF1 BMDM.

Taken together, although the mRNA data suggested that mutant cells might express more Stat3, this was not reflected in the Stat3 protein levels.

To investigate if the elevated Stat3 transcription results in altered expression of Stat3 target genes, we compared expression of *Socs3*, *Bcl3*, and *Stat1* in the WT and mutant BMDM (**Fig 19A)**. This analysis revealed significantly elevated target gene expression. In contrast, non-direct targets of Stat3, such as *Marc1*, *Arg1*[55], *Nod2, iNOS*[56] and *Irf1*[57] were not significantly affected in the mutant cells (**Fig 19B)**.

To substantiate this finding, we performed the same experiment with *Stat3*$^{\Delta E\ MF2}$ BMDM harboring the larger enhancer deletion. We obtained similar results to the ones obtained with *Stat3*$^{\Delta E\ MF1}$ BMDM corroborating the notion that the mutation alters Stat3 expression following activation. *Stat3*$^{\Delta E\ MF2}$ BMDM showed elevated Stat3 pre-mRNA and mRNA levels **(Fig 20A)** and transcripts of the Stat3 targets genes were found increased in the mutant cells (**Fig 20C)**. However, also here we did not observe a change in the Stat3 protein expression **(Fig 20D)** and likewise, the p-STAT3 induction was comparable after KLA stimulation for both mutant and WT BMDM. The alterations are hence likely below our level of detection.

In conclusion, these data establish that the lack of the MF enhancer impairs binding of SDTF and suggests that activity of the enhancer element requires prior cell activation

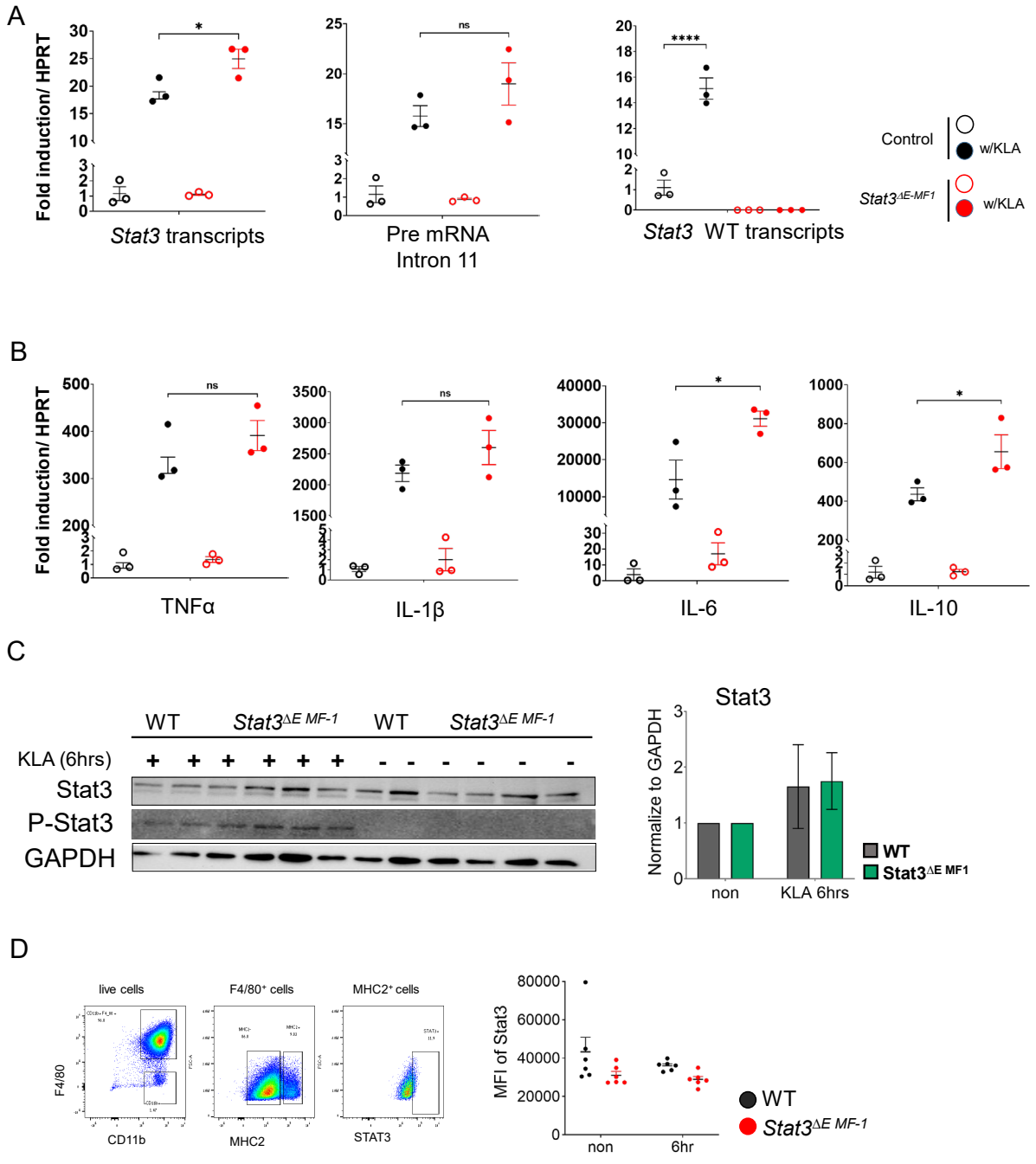


Figure 17 Analysis of Kdo2-Lipid A (KLA) challenged in vitro cultured macrophages - ΔE MF1

(A) *Stat3*$^{\Delta E\ MF1}$ homozygote and WT After stimulation with KLA, the mRNA expression of Stat3, pre- mRNA (primers detected only WT alleles and intron 11).

(B) TNFα and IL-1β , IL-6, and IL-10 mRNA . (p-value >0.05).

(C) western blot and contification for *Stat3*$^{\Delta E\ MF1}$ BMDM. The blot is for Stat3 protein and pospo-Stat3 *Stat3*$^{\Delta E\ MF1}$ n=6 (homozygote males) ,WT (C57BL 6) n=6 (5 meals and 1 female) , (p-value >0.05). KLA= 0.1ug/ml, cell number = 1x10^{6}

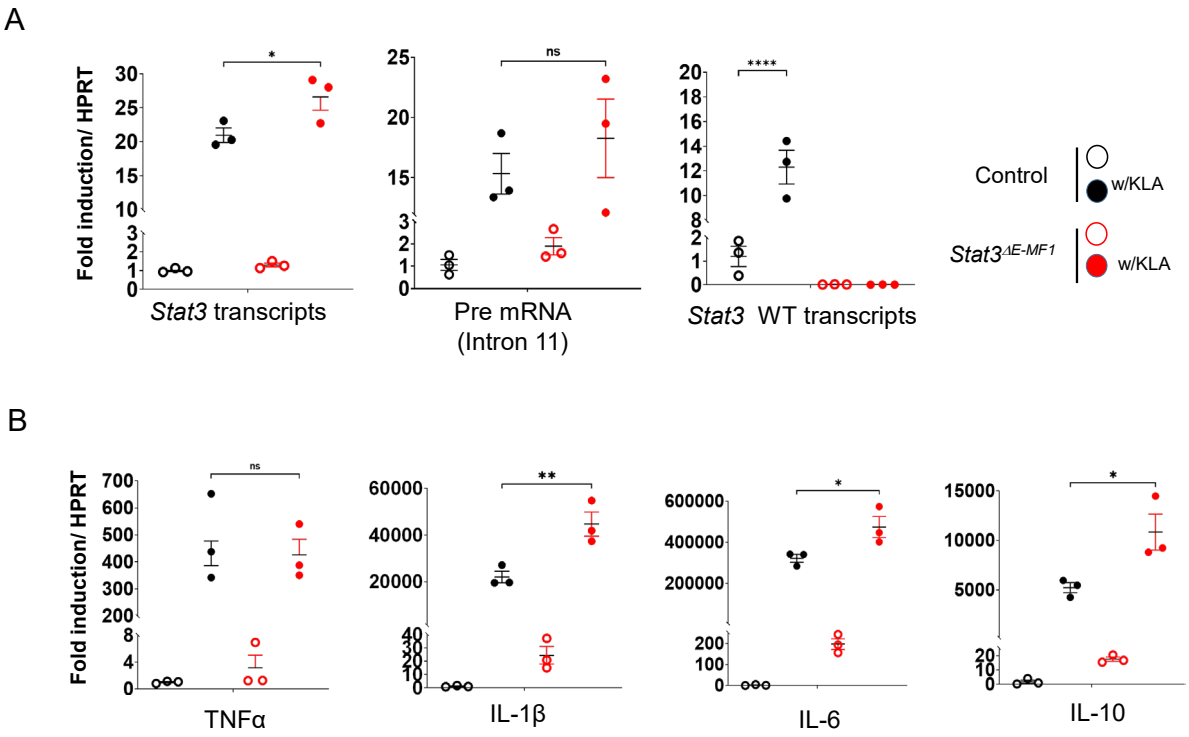


Figure 18 Analysis of Kdo2-Lipid A (KLA) challenged in vitro cultured macrophages - ΔE MF1

(A) *Stat3*$^{\Delta E\ MF1}$ homozygote and WT After stimulation with KLA, the mRNA expression of Stat3, pre- mRNA (primers detected only WT alleles and intron 11).

(B) TNFα and IL-1β , IL-6, and IL-10 mRNA . (p-value >0.05).

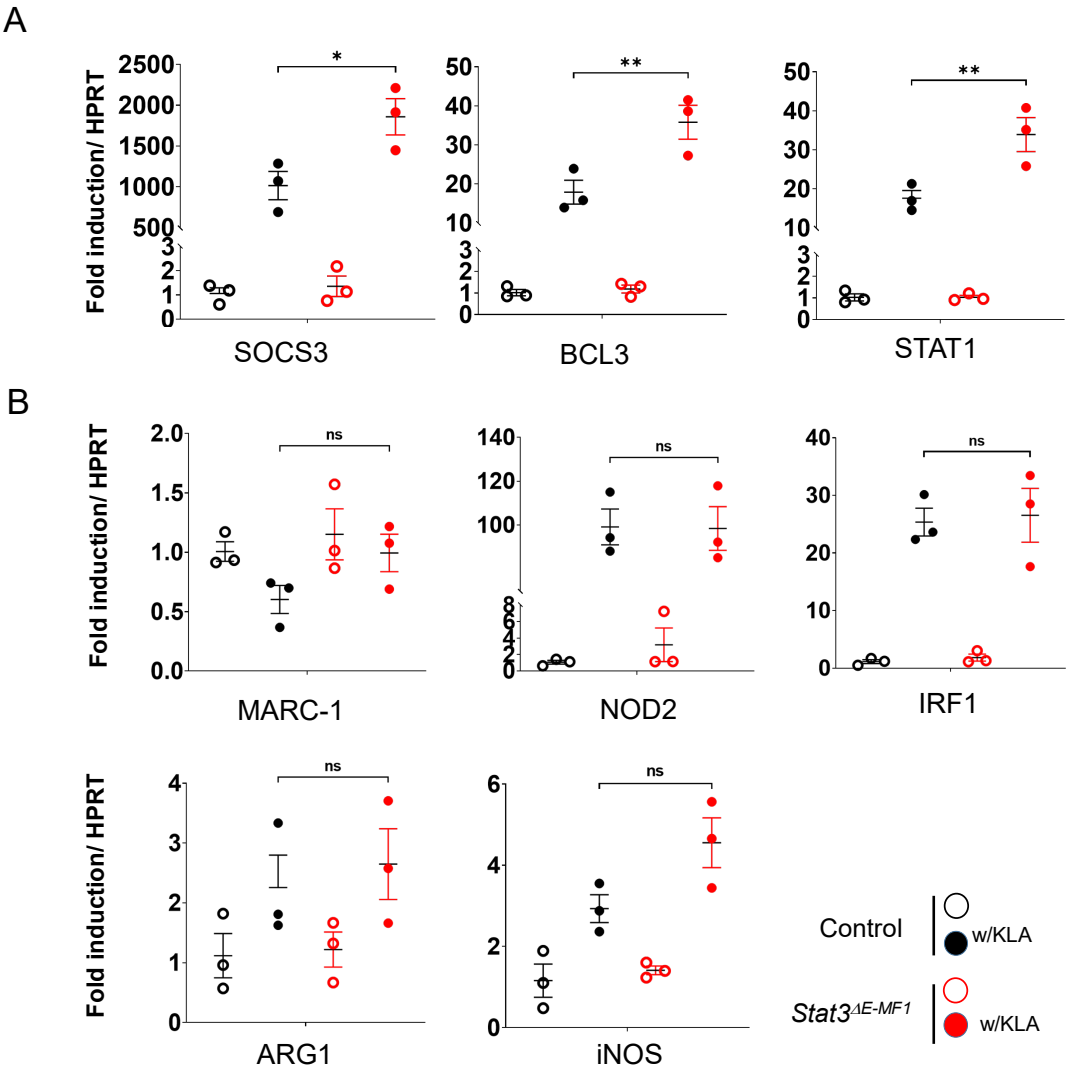


Figure 19. Analysis of Kdo2-Lipid A (KLA) challenged in vitro cultured macrophages - Stat3 target gene for ΔE MF1

(A)target genes (SOCS3, BCL3 and Stat1) (p-value >0.05)

(B) indirect Stat3 target genes (MAR-1, ARG1, NOD2, iNOS and IRF1) (p-value >0.05)

(C)*Stat3*$^{\Delta E\ MF1}$ n=3 (homozygote males) ,WT (C57BL 6) n=3 meals , KLA= 0.1ug/ml, cell number = 1x10^6

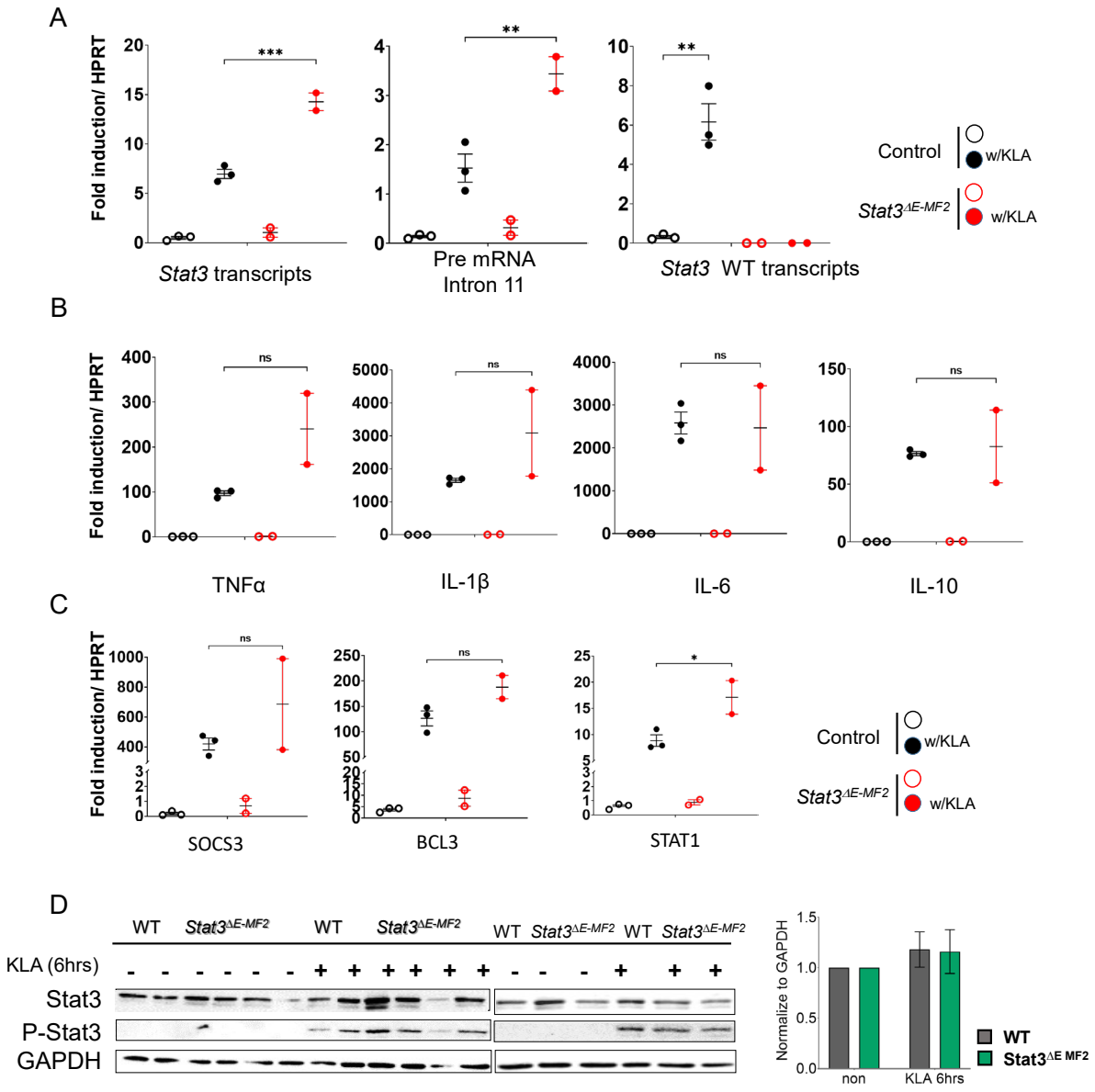


Figure 20 . Analysis of Kdo2-Lipid A (KLA) challenged in vitro cultured macrophages - ΔE MF2

(A) *Stat3*$^{\Delta E\ MF2}$ homozygote and WT After stimulation with KLA, the mRNA expression of Stat3, pre- mRNA (primers detected only WT alleles and intron 11).

(B) TNFα and IL-1β , IL-6, and IL-10 mRNA . (p-value >0.05).

(C) stat3 target genes (SOCS3, BCL3 and Stat1) (p-value >0.05) .

(D) western blot and contification for ΔE MF2 BMDM. The blot is for Stat3 protein and pospo-Stat3

Stat3$^{\Delta E\ MF2}$n=3 (2 females +1 males) ,WT (C57BL 6) n= 3 technical reputes (females)

(p-value >0.05). KLA= 0.1ug/ml, cell number = 1x10^6

Discussion

This study aimed to identify the cell type-specific regulatory elements of the STAT3 locus that might explain distinct disease associations of genetic STAT3 variants, including colitis and AD-HIES. Specifically, we hypothesized that STAT3 open reading frame (ORF) variants will cause deleterious IL-10R impairment, but concurrently neutralize the critical pro-inflammatory executor elements of an IBD cascade, i.e. IL-23R and IL-22R. Conversely, we predicted that reported *STAT3* loci-associated single nucleotide polymorphisms (SNPs) identified by GWAS as IBD risk factors might affect cell type-specific regulatory regions and hence impair cell type-specific STAT3 expression in selected cell types. Using the GeneHancer database [31] we found that reported IBD risk-associated STAT3 SNPs are close to putative Stat3 enhancer regions that are highly conserved among humans and mice. To test our hypothesis we decided to define the enhancer elements that drive Stat3 expression in macrophages, T cells and non-immune cells, respectively. We generated mutant mice lacking defined conserved putative regulatory elements and subjected the animals to a number of assays, with a particular focus on the monocytes and macrophages.

First, we defined the enhancer elements of Stat3, depending on bioinformatics prediction from GeneHancer database and GWAS from human loci. The criteria for determination of an enhancer by the GeneHancer software relies on database sources established 37 from tissues and cell lines [35]. Meta-analyses of different assays, such as eQTLs, CHi-C, enhancer RNA (eRNA) co-expression, TF co-expression and Nearest neighbor analysis aid to define the score of individual regulatory elements and determine the probability of a given DNA sequence to be an enhancer or promoter[35]. GeneHancer database includes meta-analyses only for human loci but not for mouse loci. We thus depended in our study for the three putative enhancers on data of the human Stat3 gene. Notably though, these loci are highly conserved between humans and mice, and we used the conservation to delineate the putative murine Stat3 enhancers.

Second, we analyzed the cis-impact of macrophage enhancer deletions on allele-specific Stat3 transcription. However, this in vitro analysis did not indicate that the putative MF enhancer located in Stat3 intron 2 is a critical *cis*-regulatory element required for Stat3 expression in steady state or following activation, at least in the BMDM studied. Enhancers are by definition *cis* regulatory elements that control the transcription of nearby target genes [24]. Therefore, enhancer effects should in principle be detectable in heterozygote cells, provided that, as in our case, the gene products of the two alleles can be discriminated.

To reveal potential enhancer activity of the respective regulatory elements in the in vivo context, we analyzed the mutant mice with or without challenge with piroxicam. The mutant *Stat3*$^{\Delta E\ MF1}$ mice, that harbor the shorter deletion, displayed signs of mild inflammation in the PAC paradigm, as compared to controls. However, this phenotype was not observed in Stat3$^{\Delta E\ MF2}$ mice and also could not be

recapitulated in BM chimeras generated with BM of Stat3$^{\Delta E\ MF1}$ or Stat3$^{\Delta E\ MF2}$ mice challenged with piroxicam.

Surprisingly, and contradicting our prediction that absence of the regulatory element would impair Stat3 transcription, in the in vitro assay, BMDM generated from *Stat3$^{\Delta E\ MF1}$* and *Stat3$^{\Delta E\ MF2}$* mice showed following stimulation with KLA increased levels of Stat3 transcript as compared to controls. Moreover, that Stat3 target gene increased accordingly. The results of the in vitro experiments obtained with the BMDM after KLA stimulation suggest that the *Stat3$^{\Delta E\ MF1}$* and *Stat3$^{\Delta E\ MF2}$* enhancer deletions might affect Stat3 expression, but this could be more complicated, as these cultures do not only contain macrophages and also include complex inter-cellular interactions via secreted molecules. In our data, in mutant BMDM *Stat3$^{\Delta E\ MF}$*, the level of Stat3 mRNA increased post-stimulation with KLA in BMDM. This finding suggested that the deletion may impair suppressor or silencer binding. Previous studies support this assumption. Primarily Stat3 is a TF that is activated through cytokine receptors such as IL-10 and IL-6, the subunit gp130, such as the IL-6 receptor, and for IL-10 receptor is a docking motif (YxxQ) [58]. In our data, the summation of KLA, the orthologue of LPS, initiated the signal transduction of Stat3 and it might be that this cascade of Stat3 signaling starts from IL-10 on activated macrophages, as shown before in human macrophages (hMΦ) post-infection *M. tuberculosis* to reduce the production of the pro-inflammatory cytokines. It was shown that removing IL-10 blocked the activation of Stat3 in human macrophages (hMΦ), but the blocking of IL-6 did not affect Stat3 activation post *M. tuberculosis* exposure. Additionally, in bacterial infection model, Stat3 initially repressed the production of pro-inflammatory factors like IL-6 but induced anti-inflammatory IL-10. [58] . Moreover, a previous study showed that in conditional Stat3 KO mice that harbor a myeloid specific ablation in macrophages and neutrophils (LysM^{cre}/Stat3$^{flox/-}$) after treatment with LPS there is an increase of pro-inflammatory cytokines and reduction of IL-10 production by macrophages. Likewise, BMDM from *LysM^{cre}/Stat3$^{flox/-}$* mice treated with IL-10 and then with LPS did not reduce expression of pro-inflammatory cytokines like IL-6 and TNFa [33, 59] . Additionally, BMDM from IL-10 KO mice did not display phosphorylation of Stat3 post -infection *M. tuberculosis* [58]. In our data, the BMDM of the mutant mice displayed a similar pattern of upregulation of inflammatory cytokines due to the KLA stimulation. In contrast to a previous study on LysM^{cre}/Stat3$^{flox/-}$ BMDM after LPS treatment mutant and human macrophages (hMΦ) after infection, we observed an increase in IL-10 levels.

KLA exposure triggers aTLR4 specific pathway in macrophages. We see upregulation of SOCS3 and BCL3 in BMDM of the mutant Stat3 starts transcription by binding to the intrinsic enhancer HS4 (BCL3 enhancers), which induces the BCL3, which has a negative feedback loop by himself [60, 61]. SOCS3 is regulated by IL-6, IL-10, and LPS in macrophages. It is downstream of Stat3 activation and has different dynamics depending on the inducing signal, i.e. IL-10 or IL-6.[60]. Furthermore, SOCS3 specifically blocks activation of Stat3 signaling that starts from IL-6. In BMDM Stat3^{MF}, we see an

increased level of SOCS3 and BCL3 due to the elevation of Stat3 in mutant BMDM after stimulation by KLA.

The Stat3 locus intron 2 deletion may remover a silencer element, rather than an enhancer, as we predicted. There are two types of silencers; one motif that binds TF, repressor TF, together with (pre-initiation complex) PIC, in the silencer element. The second silencer is a negative regulatory elements (NREs), which prevent TFs from binding their respective cis-regulatory motifs. The second type of silencer can be within introns and exons, downstream of the transcription start site (TSS) or upstream of the TSS, in a passive way [62]. Studies indicate that silencers may be enhancers, and the silencing and activation are the two sides of the same cis-regulatory element. Namely, this sequence can have a bifunctional role depending on cell type and cellular context [63]. However, for this regulatory element, commonly histone modification that knowing for enhancers may not sufficient to identify them. For example histone H3 trimethyllysine 27 (H3K27me3), histone H3 monomethyllysine 4 (H3K4me1) and histone H3 acetyllysine 27 (H3K27me3). *Cis*-regulatory elements (CREs), like enhancers and silencers, have TFs that bind to them dependent on cell type, cellular state, and environment. Doni Jayavelu N. et al. reported TF motifs enriched in candidate silencers, like TFAP2C, NF1, BATF, BACH2, FRA1, ATF3, FOSL2, ZFX, EBF1, NFE2, and RFX family. These TFs described that they have repressor activity. In addition, the authors compared active enhancers and silencers and found distinct motifs of TFs that enrich silencers such as REST, ZFX, PITX1, ZNF family and CUX1, FOSL1, GFI1B, and GATA family. In our analysis, described in Fig 5, we found the Stat3 locus enriched with TFBS, like FOS (FOSL2/1), JUN ,GATA, and ZNF (GFI1B) family according to Genomatix database and UCSC browser (JASPAR database) for our locus. Moreover, they also have STAT and CREB, ETSF-PU.1 (SFPI1) and IRFF (IRF3/4) families enriched in candidate silencers that we also fined in our Stat3 locus. Like enhancers, silencer elements can display disease association variants (SNPs) from GWAS in different cell type or tissue including colitis [64].

It is possible that in our case, the deletion we made on Stat3 for putatively macrophage enhancer may be silencers, like in the *Drosophila* study[65]. It depends on the cell contexts and the environment, the regulatory elements can be enhancers or silencers. For example, in this study, the authors deleted the hkb_0.6RIRV element, which was reported as an enhancer of the blastoderm embryo's gene *Huckebein* (*hkb)*. However, they showed that it is a silencer on mesodermal cells in *Drosophila* [65]. Namely, the mRNA level of *hkb* was upregulated compared to the WT. Similar results by another study, showed that deletion of complex array of cis-regulatory elements (CREs) regions in K562 cell lines increases target gene expression [64]. We observed a similar result in $Stat3^{\Delta E\ MF}$ BMDM, where the deletion of putative enhancer led to upregulation of Stat3 mRNA levels after stimulation with KLA. It might be that this regulatory element of Stat3 is significant and activated only after the challenge particularly *in vitro.*

Follow up experiment could be to sort the macrophages from $Stat3^{\Delta E\ MF}$ from the gut and different organs and test for Stat3, IL-10, and IL-6 mRNA levels with and without challenge. Additionally, we

can perform motif enrichment analysis for our Stat3 locus on intron 2 to identify TF motifs enriched in silencer elements. This experiment can be performed on WT-sorted macrophages from the gut, brain, spleen, and more tissues. And it can also be performed in vitro, for example, BMDM.
To summarize this part, our data for BMDM suggest that the deletion we introduced in the murine Stat3 locus may span a suppressor or silencer. The detailed study of cis-regulatory elements, enhancers, silencers, promotors, and insulators is challenging. Moreover, cis-regulatory elements can have different roles depending on the developmental stage, cell type and environment, as well as combinatorial TF binding. Enhancers can function as a silencer in different cells. No specific histone modification is associated with enhancers or silencers. We need better tools to characterize the regulatory elements together with what exists and combine more bioinformatics prediction with experiments in animals.

Intestinal macrophages have a variety of functions, including roles in the maintenance of homeostasis and tissue development. Monocytes constantly replace the majority of resident macrophages in the intestine in a 'monocyte waterfall' [54]. However, recent studies identified a macrophage population that not replaced, but long-lived and self-maintaining cells in *lamina propria* (LP) and *muscularis externa* (ME), essential for supporting Enteric Neurons and submucosal vasculature, specifically support permeability in the LP. Ablation of these cells did not lead to overt ME inflammation [55]. In addition, these cells occupy different niches of the gut. We observed no phenotype indicating gut inflammation in the mutant mice, even not following challenge of piroxicam. Surprisingly, Stat3 mutant mice however displayed decreased levels of Tim4+ CD4+ long-lived macrophages among the P3 macrophage population in cecum and colon. However, we also observed this phenotype in *plxna4*$^{fl/fl}$ control mice obtained from our mammalian facility. In contrast, controls that we purchased from a vendor (Harlan) display a normal population for Tim4+ CD4+ cells. The observed phenotype is hence not related to the Stat3 mutation, but rather results from the different microbiomes of animal facilities. Establishment of the macrophage network in the gut requires live microbiota, including long-lived and short-lived macrophages [47]. Shaw et al. performed an experiment to understand the contribution of the commensal microbiome on Tim4+ CD4+ macrophage population. Comparison of animals held under germ-free (GF) and specific pathogen-free (SPF) conditions, revealed no difference in the frequency of Tim4-CD4- and Tim4- CD4+, (monocyte derived) and short and long lived macrophage subpopulations. However, the total number of all macrophages was reduced under GF condition [47].
To expand our finding, one would have to first repeat this experiment with appropriate controls. Once could compare the abundance of the Tim4+ CD4+ population in different facilities at the Weizmann institute and other institutions in Israel. If a differential abundance is confirmed one could analyze the microbiome of the animals. Moreover, one could test if the Tim4+ CD4+ population interacts with the specific commensal bacteria or members of the microbiome, such *Saccharomyces cerevisiae*. Tim4+ CD4+ cells can be found both in the LP and the ME, and one should test which population is

preferentially affected. Presence of the macrophages might also depend on the presence of T cells, such as TH17 and Treg cells, or innate lymphoid cells (ILC).

In the ME, the Tim4+ CD4+ population is located close to blood vessel and supporting enteric neurons, and communicating with enteric glia cells.

Taken together our results did not confirm our hypothesis, that the predicted and conserved enhancer elements are critical for Stat3 expression in macrophages. This project relied on bioinformatics prediction from GeneHancer database and GWAS from human loci. GeneHancer database includes meta analyses only for human loci but not for mouse loci. We thus depended in our study for the three putative enhancers on data of the human Stat3 gene. Notably though, these loci are highly conserved between humans and mice and we used the conservation to delineate the putative murine Stat3 enhancers. However, not all enhancers are conserved, and not all conserved sequences are enhancers [48]. This could have been a shortcoming of our study. Building on our previous work on the IL10R mutants [12, 13], we focused our analysis on the context of gut inflammation and colitis. Specifically, we investigated and IBD cascade that associates Stat3 with Il10r signaling. We therefore chose to work with gut macrophages, monocytes and BMDM, but we did not test the functionality of other tissue macrophages. We hence cannot rule out that the enhancer elements we targeted play a role in other cells, such as Kupffer cells or microglia. Notably, these latter populations are established in the embryo and therefore are distinct from the populations we chose to study which are all adult hematopoiesis derived. In vivo enhancer functions are depended on the cell state, the environmental context, external stimulation, cell type-specific TFs, the chromatin state, the 3D chromatin structure and genetic variants [48]. Stat3 is a well-regulated gene involved in numerous essential processes, and it has been shown that there is dosage compensation for Stat3 gene [49]. Moreover, studies have shown that mutants often do not reveal clear phenotypes, due to a phenomenon called genetic robustness, the capability of an organism to keep fitness in the presence of harmful mutations [50]. This can be achieved by protein feedback loops or by transcriptional adaptation, the molecular mechanisms for which are unknown. In addition, genetic robustness can be explain by functional redundancy of genetic networks. There are diverse genetic networks like transcription-regulatory networks, signaling, metabolic, and protein-protein interaction. The sum of these interactions altogether determines the molecular behavior of the cell [51]; and transcriptional adaptation is the ability of certain gene mutations to affect other gene expression or even his mRNA. This phenomenon can upregulate gene that compensate the mutation gene, thus masking the predictable phenotype [50, 52]. This may also be why we did not see a phenotype in our mutants, especially in the in vivo studies of gut macrophages or monocytes. It also may explain the upregulation of Stat3 mRNA and Stat1 mRNA after stimulation with KLA for BMDM for MF mutant. It is less clear how this phenomenon could affect our in vitro studies on the cis effect. Additionally previous studies showed redundancy of enhancers in development. Namely, deletions of one enhancer did not affect the tissue- specific expression of gene in context of development. This may also can be the case in our study of the enhancers in Stat3 gene [48, 53].

To conclude this thesis, we started our story with the identification of human SNPs associated with IBD. We hypothesized that these sequences are putative cell type-specific enhancers in macrophages, macrophages and T cells, and non-immune cells, respectively. According to the human database, we predicted murine Stat3 enhancers and tested the prediction in mouse models. Two independent mouse lines harboring *Stat3*$^{\Delta EMF}$ mutations did not display inflammation in the gut, neither spontaneous, nor following challenge. Results of in vitro studies with primary BMDM of the animals suggest that the targeted elements comprise a silencer that might be required upon macrophage activation to curb Stat3 expression. Further experimentation will be required to establish the function of this element, confirm its presence in the human genome and address its relevance for the IBD association of Stat3.

Materials and Methods

STAT3 regulatory element analysis. Regulatory elements were obtained from GeneHancer (version 4.7,[31]). GWAS SNPs from the GWAS Catalog [66] were retrieved from GeneCards (version 4.7[40]). ClinVar genetic disease variants [67] were mined from MalaCards (version 4.7,[68]). Figures 3, 4, 5 ware generated with the UCSC genome browser [69], using the GRCh38/hg38 human / GRCm38/mm10 Mouse genome assembly with the following tracks: (1) NCBI RefSeq; (2) Vertebrate (100 /60 Species) Conservation (phastCons) and Conserved Elements ;(3) Mouse (GRCm38/mm10) / (GRCh38/hg38) Alignment Net. TFBS for STAT3 regulatory element analysis was performed by Geomatics software.

Generation of STAT3 enhancer deletion mice . Animal protocols were approved by the Weizmann Institute IACUC and were carried out in accordance with their approved guidelines. *Stat3 enhancer* mutant embryos were generated by CRISPR/Cas9 gene targeting technology [70, 71]. In vitro transcribed Cas9 (Alt-R® S.p. Cas9 Nuclease 3NLS, IDT) and sgRNA RNA were electroporated into mouse zygotes and the embryos were transferred into the oviducts of pseudo-pregnant ICR females as described [70]

PCR genotyping. Mouse tails were incubated with sodium dodecyl sulfate (SDS) and Protease K solution at 55 ℃ overnight, and DNA from mouse tail was precipitated by Isopropanol. Precipitated DNA was dissolved into Tris-EDTA buffer. Taq DNA polymerase from Amplicon (Cat#A111103) was used for genotyping PCR. See Table 2 for primers used for genotyping.

Table 2. Primers used in genotyping and qPCR.

Gene / Stat3 locus	***Forward (5'->3')***	***Reverse (5'->3')***
ΔE MF1 (intron 2)	AGTTCCTGGCACCTTGGATT	TCTATACAAACGTGCACACATACT
ΔE MF2 (intron 2)	AGTTCCTGGCACCTTGGATT	TCTATACAAACGTGCACACATACT TGGCATGTGACTCTTTGCTG
ΔE T/MF (intron 1)	CCTTAGAATGAAGGAGTTGAAGC	ATGGAATTCTAAACTGGGTGGT
ΔE NI (intron 1)	GAAGTGAAGAGGAAGTGGTGAA	GACACTGTCGTTCTGGGC
WT transcript for	TTGATCAGACATGTGAGGGCAG	GCAAGCTGTTTCAAAGTGCA

ΔE MF (intron 2)		
Δ transcript for ΔE MF1 (intron 2)	GAAGGATTGGGAGAACAGAGGT	TCTATACAAACGTGCACACATACT
Δ transcript for ΔE MF2 (intron 2)	GAAGGATTGGGAGAACAGAGGT	TGGAGAGCGTTTTGTGAGAGC
Stat3 mRNA (exon 8 and 9)	TGGACCGTCTGGAAAACTGG	ATAGGGTCGCCCTTGTAGGA
Stat3 pre-mRNA (intron 11)	GGAAATGCAGTTGGGGCAAG	CACGGTAAGGATGGCATGGA
WT transcript for ΔE T/MF(intron 1)	GCATAGGTCCCCTGATACTGA	TTGGCTACTCTACTCCTCC
Δ transcript for ΔE T/MF(intron 1)	CCTTAGAATGAAGGAGTTGAAGC	ATGGAATTCTAAACTGGGTGGT
HPRT	TCCAAATCCTCGGCATAATG	CGTCGTGATTAGCGATGATG
TNFa	CTGAACTTCGGGGTGATCGG	GGCTTGTCACTCGAATTTTGAGA
IL-1b	GAAGGTCCACGGGAAAGACAC	TTCAGGCAGGCAGTATCACTC
IL-6	TAGTCCTTCCTACCCCAATTTCC	TTGGTCCTTAGCCACTCCTTC
IL-10	TGAATTCCCTGGGTGAGAAGC	AGACACCTTGGTCTTGGAGCTTATT
SOCS3[i]	GGACCAAGAACCTACGCATCCA	CACCAGCTTGAGTACACAGTCG
BCL3	AGCAGTCGTCTCAGCTCCAATG	AGGCAGGTGTAGATGTTGTGGG
STAT1	TGGTGAAATTGCAAGAGCTG	CAGACTTCCGTTGGTGGATT
Arg1	TTTTTCCAGCAGACCAGCTT	AGAGATTATCGGAGCGCCTT
iNOS	CCCTTCCGAAGTTTCTGGCAGCAGC	CCAAAGCCACGAGGCTCTGACAGCC
MARC-1	CAGGTGTGGGCTCAGGTAGT	TGGCATGTCCTGGAATGAT
IRF-1	CAGAGGAAAGAGAGAAAGTCC	CACACGGTGACAGTGCTG
SOD2	TAACGCGCAGATCATGCAGCTG	AGGCTGAAGAGCGACCTGAGTT

Piroxicam treatment. *Stat3*$^{\Delta E\ MF1}$ mice and *Cx₃cr1*^{Cre} *:IL-10R*$^{fl/fl}$ 10–12 week old male mice were used in the experiments. The mice had free access to water and Teklad 2918SC Rodent Chow with piroxicam (Sigma Aldrich, Broendby, Denmark) 200 ppm homogenized in the diet (research diet, INC). Piroxicam chow was administrated orally to STAT3 enhancers and *Cx₃cr1*^{Cre} *:IL-10R*$^{fl/fl}$ mice [25] from day 0 until day 10 of the experiment after which the treatment was terminated. Weight change were examined 2 times weekly and used as measures of the clinical status of the mice. Mice were sacrificed by cervical dislocation when severe disease was observed. Severe disease was defined as a weight loss exceeding 20% of the initial weight, rectal bleeding for more than two succeeding observations or a morbid appearance

ELISA. Quantification of Fecal Lipocalin-2/NGAL was measured using the DuoSet ELISA kit (R&D Systems) according to manufacturer's instructions.

BM chimeras. The following mice strains all on C57BL6 background were used: *Stat3*$^{\Delta E\ MF1}$, *Stat3*$^{\Delta E\ MF2}$,*Cx₃cr1*^{Cre} *:IL-10R*$^{fl/fl}$ and WT (C57BL6). BM chimeras were generated by engraftment of 7–8 week old recipient mice that were irradiated the day before with a single dose of 950 cGy using a XRAD 320 machine (Precision X-Ray (PXI). Femurs and tibiae of donor mice were removed and BM was flushed with cold PBS. BM was washed with cold PBS twice. BM cells were suspended in PBS and 2.5 10^6 cells were injected by IV into irradiated recipient. Mice were handled and experiments were performed under protocols approved by the Weizmann Institute Animal Care Committee (IACUC) in accordance with international guidelines.

Preparation of BM monocyte-derived macrophages (MoMF). MACS extraction was used to isolate of CD115 positive monocytes form BM (Miltenyi kit). Monocytes were seeded on 6 well tissue culture plates, 2.5-5x10^5 cells per well in RPMI medium (Biological industries) with 10% heat-inactivated fetal bovine serum (FBS) (Life technologies), 1% penicillin streptomycin amphotericin b solution (Biological industries), 1% L-Glutamine (Biological industries) and 10 ng/ml M-CSF (Sigma-Aldrich). 3 days later, half the medium was renewed. Cells were analyzed on day seven. The BMDM cells were stimulated with 3-deoxy-D-*manno*-octulosonic acid $(Kdo)_2$-Lipid A (KLA, Sigma-Aldrich) 10 ng /ml for 6 hours on day seven.

Preparation of BM-derived macrophages (BMDM). BM cells were seeded on 6-well tissue culture plates, 1 x10^6 cells per well in RPMI medium (Biological industries) with 10% heat-inactivated fetal bovine serum (FBS) (Life technologies), 1% penicillin streptomycin amphotericin solution (Biological industries), 1% L-Glutamine (Biological industries) and 10 ng/ml M-CSF (Sigma-Aldrich). 3 days later, half the medium was renewed. The BMDM cells were stimulated with 3-deoxy-D-*manno*-octulosonic acid $(Kdo)_2$-Lipid A (Kdo_2-Lipid A) (KLA), (Sigma-Aldrich) 10 ng /ml for 6 hours on day seven.

RNA extraction and qRT-PCR. Total RNA was extracted from MoMF cells with Dynabeads kit (Dynabeads™ mRNA DIRECT™ Purification Kit); PCR was performed with SYBR Green PCR Master Mix kit (Applied Biosystems). See Table 2 for primers used for qPCR.

Western blot. Proteins were extracted from BMDM using RIPA buffer with protease inhibitor. Supernatants of cell extracts resolved by SDS-PAGE and Western blotted with the indicated antibodies. The antibodies used were anti-STAT3 (Cell Signaling no. 9139), anti-pSTAT3 (Cell Signaling no. 4113), and anti-GPDH. Quantification of the proteins was performed using the ImageJ software.

Isolation of intestinal lamina propria cells. Intestines were removed and fecal content flushed out with PBS; tubes were opened longitudinally and cut into 0.5 cm sections. Pieces were placed in 10ml/sample with 3% heat-inactivated FCS/FBS, 5 mM EDTA and 2 mM DL-Dithiothreitol ((DTT), Sigma-Aldrich Cat# D9779) and 20mM HEPES (Sigma-Aldrich) and incubated on a 37°C shaker at 250 rpm for 20 min to remove mucus and epithelial cells. Following incubation, samples were washing with PBS for and filtered through a crude cell strainer. Pieces that did not pass the strainer were collected and transferred to 2 ml/sample of PBS, 20mM HEPES (Sigma-Aldrich) and 0.1 mg/ml DNase I (Roche Cat# 10104159001) and 0.1mg/ml Liberase. Tissue was incubated in a 37°C at 20 (colon) or 15 min (cecum) in the solution. After incubation, samples were dissociated with a syringe and then filtered through a crude cell strainer. The strainer was washed with PBS and centrifuged in 4ºC, 375G for 5 min. Cells were stained and subjected to FACS analysis or sorting.

Isolation of murine blood monocytes. Blood collected from the mouse, immediately placed in 40ul/sample Heparin. Tubes were centrifuged 1400rpm in 4c degree with PBS for 5 min. Then adding 1X of ACK centrifuged 1400rp in 4c degree with PBS. Additionally, the cells were collected and washed with cold PBS. Cells were analysis according to the following parameters: $CD45^+$ $CD11b^+$ $CD115^+$ $Ly6C^{+/-}$ $Ly6G^+$

Flow cytometry analysis. Samples were suspended and incubated in staining medium [PBS without calcium and magnesium with 2% heat-inactivated Fetal Calf/Bovine Serum (FCS/FBS) and 1 mM EDTA] containing fluorescent antibodies. Following incubation, cells were washed with staining buffer only or staining buffer with DAPI, centrifuged, filtered through 70 μm filter and read. For FACS analysis, Cytek® Aurora and LSRFortessa (BD Biosciences) was used. Results were analyzed in FlowJo software. Staining antibodies (clones indicated within brackets): anti-CD45 (30-F11), CD11b (M1/70),), Ly-6C (HK1.4), CD64/FcγRI (X54-5/7.1), anti-I-Ab (MHCII) (MS/114.15.2), DAPI. Ly6G (1A8) , CD45.1 (A20) , Tim4 (RMT4-54) and CD4 (GK1.5) from Biolegend.

Quantification and statistical analysis. Results are presented as mean ± SEM. For comparisons between two groups, we used Student's two-tailed, unpaired *t*-test. One-way analysis of variance, and two way analysis of variance followed by Newman-Keuls or Tukey's post-test analysis, was used for comparisons of more than two groups. The numbers of samples/group (*n*) or the numbers of experiments are specified in the figure legends. Statistical significance is accepted at $P < 0.05$.

References

1. Gearry, R.B., *IBD and Environment: Are There Differences between East and West.* Dig Dis, 2016. **34**(1-2): p. 84-9.
2. Ananthakrishnan, A.N., et al., *Environmental triggers in IBD: a review of progress and evidence.* Nat Rev Gastroenterol Hepatol, 2018. **15**(1): p. 39-49.
3. Kaplan, G.G., *The global burden of IBD: from 2015 to 2025.* Nat Rev Gastroenterol Hepatol, 2015. **12**(12): p. 720-7.
4. Sartor, R.B. and G.D. Wu, *Roles for Intestinal Bacteria, Viruses, and Fungi in Pathogenesis of Inflammatory Bowel Diseases and Therapeutic Approaches.* Gastroenterology, 2017. **152**(2): p. 327-339 e4.
5. Cho, J.H., *The genetics and immunopathogenesis of inflammatory bowel disease.* Nat Rev Immunol, 2008. **8**(6): p. 458-66.
6. Lin, L., et al., *Which long noncoding RNAs and circular RNAs contribute to inflammatory bowel disease?* Cell Death Dis, 2020. **11**(6): p. 456.
7. Duerr, R.H., et al., *A genome-wide association study identifies IL23R as an inflammatory bowel disease gene.* Science, 2006. **314**(5804): p. 1461-3.
8. Huang, H., et al., *Fine-mapping inflammatory bowel disease loci to single-variant resolution.* Nature, 2017. **547**(7662): p. 173-178.
9. Durant, L., et al., *Diverse targets of the transcription factor STAT3 contribute to T cell pathogenicity and homeostasis.* Immunity, 2010. **32**(5): p. 605-15.
10. McGovern, D.P., et al., *Genome-wide association identifies multiple ulcerative colitis susceptibility loci.* Nat Genet, 2010. **42**(4): p. 332-7.
11. Barrett, J.C., et al., *Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease.* Nat Genet, 2008. **40**(8): p. 955-62.
12. Alarcon, T., L. Llorca, and G. Perez-Perez, *Impact of the Microbiota and Gastric Disease Development by Helicobacter pylori.* Curr Top Microbiol Immunol, 2017. **400**: p. 253-275.
13. Mowat, A.M. and W.W. Agace, *Regional specialization within the intestinal immune system.* Nat Rev Immunol, 2014. **14**(10): p. 667-85.
14. Manichanh, C., et al., *The gut microbiota in IBD.* Nat Rev Gastroenterol Hepatol, 2012. **9**(10): p. 599-608.
15. Caruso, R., B.C. Lo, and G. Nunez, *Host-microbiota interactions in inflammatory bowel disease.* Nat Rev Immunol, 2020. **20**(7): p. 411-426.
16. Paone, P. and P.D. Cani, *Mucus barrier, mucins and gut microbiota: the expected slimy partners?* Gut, 2020. **69**(12): p. 2232-2243.
17. Juge, N., *Microbial adhesins to gastrointestinal mucus.* Trends Microbiol, 2012. **20**(1): p. 30-9.
18. Fernandes, T., et al., *Bowel wall thickening at CT: simplifying the diagnosis.* Insights Imaging, 2014. **5**(2): p. 195-208.
19. Bao, X. and J. Wu, *Impact of food-derived bioactive peptides on gut function and health.* Food Res Int, 2021. **147**: p. 110485.
20. Muller, P.A., F. Matheis, and D. Mucida, *Gut macrophages: key players in intestinal immunity and tissue physiology.* Curr Opin Immunol, 2020. **62**: p. 54-61.
21. De Schepper, S., et al., *Self-Maintaining Gut Macrophages Are Essential for Intestinal Homeostasis.* Cell, 2018. **175**(2): p. 400-415 e13.
22. Muller, P.A., et al., *Crosstalk between muscularis macrophages and enteric neurons regulates gastrointestinal motility.* Cell, 2014. **158**(2): p. 300-313.
23. Gross-Vered, M., et al., *Defining murine monocyte differentiation into colonic and ileal macrophages.* Elife, 2020. **9**.
24. Bain, C.C., et al., *Constant replenishment from circulating monocytes maintains the macrophage pool in the intestine of adult mice.* Nat Immunol, 2014. **15**(10): p. 929-937.

25. Bernshtein, B., et al., *IL-23-producing IL-10Ralpha-deficient gut macrophages elicit an IL-22-driven proinflammatory epithelial cell response.* Sci Immunol, 2019. **4**(36).
26. Zigmond, E., et al., *Macrophage-restricted interleukin-10 receptor deficiency, but not IL-10 deficiency, causes severe spontaneous colitis.* Immunity, 2014. **40**(5): p. 720-33.
27. Shouval, D.S., et al., *Interleukin-10 receptor signaling in innate immune cells regulates mucosal immune tolerance and anti-inflammatory macrophage function.* Immunity, 2014. **40**(5): p. 706-19.
28. Rubtsov, Y.P., et al., *Regulatory T cell-derived interleukin-10 limits inflammation at environmental interfaces.* Immunity, 2008. **28**(4): p. 546-58.
29. Glocker, E.O., et al., *Inflammatory bowel disease and mutations affecting the interleukin-10 receptor.* N Engl J Med, 2009. **361**(21): p. 2033-45.
30. Maritano, D., et al., *The STAT3 isoforms alpha and beta have unique and specific functions.* Nat Immunol, 2004. **5**(4): p. 401-9.
31. Fishilevich, S., et al., *GeneHancer: genome-wide integration of enhancers and target genes in GeneCards.* Database (Oxford), 2017. **2017**.
32. Welte, T., et al., *STAT3 deletion during hematopoiesis causes Crohn's disease-like pathogenesis and lethality: a critical role of STAT3 in innate immunity.* Proc Natl Acad Sci U S A, 2003. **100**(4): p. 1879-84.
33. Takeda, K., et al., *Enhanced Th1 activity and development of chronic enterocolitis in mice devoid of Stat3 in macrophages and neutrophils.* Immunity, 1999. **10**(1): p. 39-49.
34. Pickert, G., et al., *STAT3 links IL-22 signaling in intestinal epithelial cells to mucosal wound healing.* J Exp Med, 2009. **206**(7): p. 1465-72.
35. Andersson, R. and A. Sandelin, *Determinants of enhancer and promoter activities of regulatory elements.* Nat Rev Genet, 2020. **21**(2): p. 71-87.
36. Kang, R., et al., *EnhancerDB: a resource of transcriptional regulation in the context of enhancers.* Database (Oxford), 2019. **2019**.
37. Shlyueva, D., G. Stampfel, and A. Stark, *Transcriptional enhancers: from properties to genome-wide predictions.* Nat Rev Genet, 2014. **15**(4): p. 272-86.
38. Doyle, B., et al., *Chromatin loops as allosteric modulators of enhancer-promoter interactions.* PLoS Comput Biol, 2014. **10**(10): p. e1003867.
39. Mora, A., et al., *In the loop: promoter-enhancer interactions and bioinformatics.* Brief Bioinform, 2016. **17**(6): p. 980-995.
40. Stelzer, G., et al., *VarElect: the phenotype-based variation prioritizer of the GeneCards Suite.* BMC Genomics, 2016. **17 Suppl 2**: p. 444.
41. Hallier, M., et al., *The transcription factor Spi-1/PU.1 interacts with the potential splicing factor TLS.* J Biol Chem, 1998. **273**(9): p. 4838-42.
42. Rosa, A., et al., *The interplay between the master transcription factor PU.1 and miR-424 regulates human monocyte/macrophage differentiation.* Proc Natl Acad Sci U S A, 2007. **104**(50): p. 19849-54.
43. Frisch, M., et al., *LitInspector: literature and signal transduction pathway mining in PubMed abstracts.* Nucleic Acids Res, 2009. **37**(Web Server issue): p. W135-40.
44. Nerlov, C., *The C/EBP family of transcription factors: a paradigm for interaction between gene expression and proliferation control.* Trends Cell Biol, 2007. **17**(7): p. 318-24.
45. Castro-Mondragon, J.A., et al., *JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles.* Nucleic Acids Res, 2022. **50**(D1): p. D165-D173.
46. Verhoeven, Y., et al., *The potential and controversy of targeting STAT family members in cancer.* Semin Cancer Biol, 2020. **60**: p. 41-56.
47. Shaw, T.N., et al., *Tissue-resident macrophages in the intestine are long lived and defined by Tim-4 and CD4 expression.* J Exp Med, 2018. **215**(6): p. 1507-1518.
48. Lefere, S. and F. Tacke, *Macrophages in obesity and non-alcoholic fatty liver disease: Crosstalk with metabolism.* JHEP Rep, 2019. **1**(1): p. 30-43.

49. Raetz, C.R., et al., *Kdo2-Lipid A of Escherichia coli, a defined endotoxin that activates macrophages via TLR-4.* J Lipid Res, 2006. **47**(5): p. 1097-111.
50. Saito, O., et al., *Spinal glial TLR4-mediated nociception and production of prostaglandin E(2) and TNF.* Br J Pharmacol, 2010. **160**(7): p. 1754-64.
51. Playford, R.J., et al., *Effects of mouse and human lipocalin homologues 24p3/lcn2 and neutrophil gelatinase-associated lipocalin on gastrointestinal mucosal integrity and repair.* Gastroenterology, 2006. **131**(3): p. 809-17.
52. Chassaing, B., et al., *Fecal lipocalin 2, a sensitive and broadly dynamic non-invasive biomarker for intestinal inflammation.* PLoS One, 2012. **7**(9): p. e44328.
53. Challen, G.A., et al., *Mouse hematopoietic stem cell identification and analysis.* Cytometry A, 2009. **75**(1): p. 14-24.
54. Mercier, F.E., D.B. Sykes, and D.T. Scadden, *Single Targeted Exon Mutation Creates a True Congenic Mouse for Competitive Hematopoietic Stem Cell Transplantation: The C57BL/6-CD45.1(STEM) Mouse.* Stem Cell Reports, 2016. **6**(6): p. 985-992.
55. Yu, T., et al., *Modulation of M2 macrophage polarization by the crosstalk between Stat6 and Trim24.* Nat Commun, 2019. **10**(1): p. 4353.
56. Gu, W., et al., *ICAM-1 regulates macrophage polarization by suppressing MCP-1 expression via miR-124 upregulation.* Oncotarget, 2017. **8**(67): p. 111882-111901.
57. Liu, J., X. Guan, and X. Ma, *Interferon regulatory factor 1 is an essential and direct transcriptional activator for interferon {gamma}-induced RANTES/CCl5 expression in macrophages.* J Biol Chem, 2005. **280**(26): p. 24347-55.
58. Queval, C.J., et al., *STAT3 Represses Nitric Oxide Synthesis in Human Macrophages upon Mycobacterium tuberculosis Infection.* Sci Rep, 2016. **6**: p. 29297.
59. Akira, S., *Functional roles of STAT family proteins: lessons from knockout mice.* Stem Cells, 1999. **17**(3): p. 138-46.
60. Braun, D.A., M. Fribourg, and S.C. Sealfon, *Cytokine response is determined by duration of receptor and signal transducers and activators of transcription 3 (STAT3) activation.* J Biol Chem, 2013. **288**(5): p. 2986-93.
61. Brocke-Heidrich, K., et al., *BCL3 is induced by IL-6 via Stat3 binding to intronic enhancer HS4 and represses its own transcription.* Oncogene, 2006. **25**(55): p. 7297-304.
62. Riethoven, J.J., *Regulatory regions in DNA: promoters, enhancers, silencers, and insulators.* Methods Mol Biol, 2010. **674**: p. 33-42.
63. Halfon, M.S., *Silencers, Enhancers, and the Multifunctional Regulatory Genome.* Trends Genet, 2020. **36**(3): p. 149-151.
64. Doni Jayavelu, N., et al., *Candidate silencer elements for the human and mouse genomes.* Nat Commun, 2020. **11**(1): p. 1061.
65. Gisselbrecht, S.S., et al., *Transcriptional Silencers in Drosophila Serve a Dual Role as Transcriptional Enhancers in Alternate Cellular Contexts.* Mol Cell, 2020. **77**(2): p. 324-337 e8.
66. MacArthur, J., et al., *The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog).* Nucleic Acids Res, 2017. **45**(D1): p. D896-D901.
67. Landrum, M.J., et al., *ClinVar: public archive of interpretations of clinically relevant variants.* Nucleic Acids Res, 2016. **44**(D1): p. D862-8.
68. Rappaport, N., et al., *MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search.* Nucleic Acids Res, 2017. **45**(D1): p. D877-D887.
69. Casper, J., et al., *The UCSC Genome Browser database: 2018 update.* Nucleic Acids Res, 2018. **46**(D1): p. D762-D769.
70. Wang, H., et al., *One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering.* Cell, 2013. **153**(4): p. 910-8.
71. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-2308.

Appendices

Summary of STAT3 enhancer mutations, Sanger sequence result
Deletion (red), gRNAs (yellow, **bold**), CRISPR /Cas9 cutting site (blue), PCR primers (purple).

MF ΔE1 Mouse #58

>chr11: 100910514-100911328 (plus strand)

atgacacttgaggttgtcttctggcctacacgcatacacgaatgcacaca 100910514
cgcctgaatctatacaaacgtgcacacatactggtaaatgttttgttttt 100910564
GTTTTTTTAAAGGACAGataaggttgtttgccttaaattcttggccaacc 100910614
ct**gtattcagtcagactagtgt**ctataagtgcttcttagcttaatgagca 100910664
atgccaggctgaggtgtgatccacacctctttggcccatcaccaaatcag 100910714
gtatgtcttctggaaaaagtaggtcacacgcagatagggctgacttaggc 100910764
cagaactagcacacgggcctgaatacactaaagtgaggtaaacagctggc 100910814
atttcctctatgcctctagttccctcctcatcccatgcctcctctttacc 100910864
aggggaagccacgtgtgtggtagaccccaatttctacctgacatagaatt 100910914
cttagatttaaaaaggaaaaacatctgtatctcactgttctagactcaaa 100910964
atgtgcagagatctcaccaagtaagtgagactcaaacctgactccaaaac 100911014
atctaccagcaagctgtttcaaagtgcaaagaggccg**ctgtc**A**TATGCAC** 100911064
TGTGAACAGATGCGTTAAGAGACTAAGTTACCTCTGTTCTCCCAATCCTT 100911114
CCCCCGCACCTGCCCTCACATGTCTGATCAAAGGGTTAGTAAGGTGTGAA 100911164
CCTAGAACAGGTTGTGTCCTGACTCAATGCTAAACCCCCCAAGTCAGGGA 100911214
GTTTTCAAGAAGGActcaccagtcttgactctcaatccaaggtgccagga 100911264
actgccgcagctccatggggaagctgtcgctgtacagctggtgcagctgc 100911314
tccaggtagcgtgt

MF ΔE2 Mouse #26

>chr11: 100911266-100911509 (reverse complement)

agcagctgcaccagctgtacagcgacagcttccccatggagctgcggcag 100911266
ttcctggcaccttggattgagagtcaagactggtgagtccttcttgaaaa 100911216
CTCCCTGACTTGGGGGGTTTAGCATTGAGTCAGGACACAACCTGTTCTAG 100911166
GTTCACACCTTACTAACCCTTTGATCAGACATGTGAGGGCAGGTGCGGGG 100911116
GAAGGATTGGGAGAACAGAGGTAACTTAGTCTCTTAACGCATCT**GTTCAC** 100911066
AGTGCATAT**T**gacagcggcctctttgcactttgaaacagcttgctggtaga 100911016
tgttttggagtcaggtttgagtctcacttacttggtgagatctctgcaca 100910966
ttttgagtctagaacagtgagatacagatgtttttccttttaaatctaa 100910916

gaattctatgtcaggtagaaattggggtctaccacacacgtggcttcccc 100910866
tggtaaagaggaggcatgggatgaggagggaactagaggcatagaggaaa 100910816
…… ggtgtgctacaccacctattgatggtactgttgatctgag 100909066
aatgagaaatggaaaatgatggatttgttatgacatcagatgttataaca 100909016
acccttgatGGTGGATTCTGGGGGTATACGTTTTAGAGTCCTATTATTA 100908966
CTGCCTTTGAGAGAATGTCATCCAAACTGGCTTGGAACTCTCTAGGTAGG 100908916
GAAAGCTGGCCTTGAACTCTTGATCCTCCTGTCTCCACCACCCCGTATAC 100908866
AGATGTGTGCTCTCACAAAACGCTCTCCAAATTACTCATGCAGCTTACCT 100908816
CCTCTTCCTAGGGCATATGCAGCCAGCAAAGAGTCACATGCCAcgttggt 100908766
gtttcataatctcttgggtgaaattgaccagcaatatagccgattcctgc 100908716
aagagtccaatgtcctctatcagcacaaccttcgaagaatcaa

T/MF ΔE Mouse #42

>chr11: 100919598 -100918753 (reverse complement)

tggacactgtgctggatgctggagaggggggcacattagtgagccgaagc 100919598
caaggttcttcctgtcatcatggaccttagaatgaaggagttgaagcaag 100919548
GCATAGGTCCCCTGATACTGAAGGCAAGCACACATGTGAAGGCCCGTGAA 100919498
cggaggagtagagtagccaagtgagtgtaggagtgtagctagcaggcaag 100919448
tggcagccagacagacacagggctatcttgtcatgcatgtttgctctgtc 100919398
ctgctagctttgtgtgctttaagtccggtctgacactgaccctgtgaagt 100919348
gagtacagttgtcttctatttttctgcataaggaaggctagacacctggc 100919298
acgcatctcggaggggagtcagggtccacaggctgttgctttcatagttg 100919248
atgatttttagtctcagagcaaaaagtggtgagcctttatggacgtggat 100919198
ggcagctcttcagaaagcctggagggatgcaggcagatgtgcagatacag 100919148
gtaaaagctcctagctgtgggtttgtagaagcctgggagtgttttagaaa 100919098
cagctagagtacagactctggacatctgacccttttggttaaatgtcaac 100919048
ctgctttatctctgcgcatgtttattattgggatgcagcggtctcagtgt 100918998
cccaagtgagcggagctgaaatggcttcagaggccgtgcagccccttgcc 100918948
ctcctggagtctgcattatggctgtgatacctcaaggaagagcatctcca 100918898
tgtaagtgtcagatgcaacaggaacgaggcagctgcccgtacataactc 100918848
tcttcatgccattgcccagtgaaGTTATATCTGAGAGGTCTGTTTTTTGG 100918798
AGACGGCTTCTCAGCACTGACTGTcctggatctcactctgtagaccaggc 100918748
tggccttgaactcacagagatcctccatgcctctgcctcccaagtactgg 100918698
gattaaaagcatttgacaccaccaCCCAGTTTAGAATTCCATTTAAATTGTCATC

NI ΔE Mouse #31

>chr11: 100913937 -100916531 (plus strand)

gagacccccatctaaaaaacaaaaacaaaacaaacaaaaccttaagaaaa 100913937

gaaaatgtagtatgacatatgtgacccgctttctatttatgaaaatcaaa 100913987

GACACTGTCGTTCTGGGCCACACTGTAAAAAAAATGCCATGATCCTAATT 100914037

AAATCATCTGCTTGAAATAGTAAATTCCAATGCTATAAGTTACAAAACCA 100914087

GATGATTTCATCATTTATCTTGGACACAAATGCAGAGCCAGTCCCAAGAT 100914137

AAAAACCATAAAACCTCACCCACTCTACCTACTCCATCACAGATGACCAc 100914187

tccagtcggggaacattttcaatctaacataagaaaaaagctaagttac 100914237

ccatagccgacttatttctgtctgtaactcaagaggcttgtttcagtgtt 100914287

gatgtttggtgtatttcgctggtagaatggctttttcttaaccagagacc 100914337

agcttgtgccttttttccaattaactggaccatcaagaacgactaattttg 100914387

tgttttgttttgatctaaccaagataggtaatactgccaaactattctac 100914437

……acttcctaaactctctaagacatgtatgaacataaatgaaccag 100916087

taaaatggaggatgaagagagggacaaaatgcaaatgcaaataaaaatcc 100916137

acaacagaaagtgatacagagacatttctgaagcgagtgtaatgggactc 100916187

acccttaatcccagcattcggaaggaagaaacagcaggctcaggagttta 100916237

aagtctgcctcagctacataataagttcagtgcccacctgagctaagtgg 100916287

gactctatctttaaaagaggacaattcctgttcaatttcaCAGTGTGTCA 100916337

AAAAGCCTTCTTTGGGGCCATATATGTAGCTCGGGCTGACCTGAAATTGG 100916387

ATATGTAGAGCAGACTGGCCTCAACATCAAGAGATCTGCCTGCCtggcct 100916437

cctgggtactaggattaaaggtgtgattcaccacttcctcttcacttcca 100916487

gcatgaaagtcaggaaaaaggggggtaggggaaggagacagaga

PubMed ID	Chromosome	Position	SNP	P-VALUE	Phenotype	Strongest SNP risk allele	Study accession
27863252	17	42314074	rs1053004	7.00E-10	mean corpuscular volume	rs1053004-A	GCST004602
28135244	17	42314074	rs1053004	4.00E-07	pulse pressure measurement	rs1053004-A	GCST004278
22190364	17	42329511	rs2293152	4.00E-08	multiple sclerosis	rs2293152-C	GCST001341
26482879	17	42333221	rs17881320	1.00E-07	atopic eczema	rs17881320-T	GCST003184
26482879	17	42333221	rs17881320	3.00E-06	atopic eczema	rs17881320-T	GCST003184
26192919	17	42347515	rs6503695	1.00E-14	Crohn's disease	rs6503695-?	GCST003044
23266558	17	42355962	rs9891119	2.00E-15	Crohn's disease	rs9891119-A	GCST001785
21833088	17	42355962	rs9891119	2.00E-10	multiple sclerosis	rs9891119-C	GCST001198
18587394	17	42362183	rs744166	7.00E-12	Crohn's disease	rs744166-A	GCST000207
20159113	17	42362183	rs744166	3.00E-10	multiple sclerosis	rs744166-G	GCST000593
26192919	17	42362183	rs744166	1.00E-22	inflammatory bowel disease	rs744166-?	GCST003043
28199695	17	42371707	rs9897389	1.00E-07	mosquito bite reaction itch intensity measurement, mosquito bite reaction size measurement	rs9897389-T	GCST004862
23128233	17	42375526	rs12942547	6.00E-22	inflammatory bowel disease	rs12942547-A	GCST001725
26192919	17	42375526	rs12942547	7.00E-11	ulcerative colitis	rs12942547-A	GCST003045
26192919	17	42375526	rs12942547	1.00E-16	Crohn's disease	rs12942547-G	GCST003044
27569725	17	42375526	rs12942547	2.00E-09	inflammatory bowel disease	rs12942547-G	GCST003602
28067908	17	42375526	rs12942547	1.00E-10	ulcerative colitis	rs12942547-?	GCST004133
28067908	17	42375526	rs12942547	2.00E-17	inflammatory bowel disease	rs12942547-?	GCST004131
28067908	17	42375526	rs12942547	2.00E-11	Crohn's disease	rs12942547-?	GCST004132
28135244	17	42384786	rs62075782	6.00E-08	pulse pressure measurement	rs62075782-C	GCST004278

***Table 1*. GWAS data identifies SNPs associated with diseases**

GWAS SNP	GeneHancer ID	Chrom #	start	end	Summary Myeloid cells	Summary T cells	CD14+ monocyte	CD14+ monocytes	CD14+ monocytes (PB)	CD14+CD16- monocytes (CB)	CD14+ CD16- monocytes (VB)	M0 macrophage (CB)	M0 macrophage (VB)	M1 macrophage (CB)	M1 macrophage (VB)	M2 macrophage (CB)	M2 macrophage (VB)	macrophage	monocyte	CD14+ monocytes	CD4+ab T cells (VB)	Central memory CD4+ab T cells (VB)	T cells (PB)	T cell	CD4+ CD25- CD45RA+ naive primary cells	CD4+ CD25- CD45RO+ memory primary cells	CD4+ CD25- IL17+ PMA Th17 primary cells	CD4+ CD25- IL17- PMA Th primary cells	CD4+ CD25int CD127+ Tmem primary cells
	GH17I041674	17	41674903	41680514	1	1	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	-	+	+	+	+	+	-	+
	GH17I042222	17	42222243	42222472	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
	GH17I042245	17	42245834	42256824	1	1	+	-	-	+	+	-	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	-
	GH17I042266	17	42266056	42278823	1	1	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+
	GH17I042308	17	42308128	42323984	1	1	+	-	-	+	+	-	+	+	+	+	-	-	-	-	+	-	+	-	-	-	-	+	-
	GH17I042324	17	42324683	42326649	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042328	17	42328075	42330438	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
	GH17I042332	17	42332762	42333784	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
	GH17I042334	17	42334638	42335529	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
	GH17I042336	17	42336966	42338004	1	1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
	GH17I042338	17	42338419	42339897	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
	GH17I042340	17	42340479	42343407	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
rs6503695	GH17I042346	17	42346823	42348152	1	0	-	-	-	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs9891119	GH17I042353	17	42353360	42356542	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs744166	GH17I042360	17	42360898	42363013	1	1	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-
	GH17I042363	17	42363201	42365623	1	1	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-
	GH17I042366	17	42366002	42367800	1	0	+	-	-	+	+	-	+	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042377	17	42377861	42380858	1	1	-	-	-	+	+	-	+	+	+	+	+	-	-	-	+	-	+	-	-	-	+	-	-
	GH17I042381	17	42381401	42381600	1	0	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042382	17	42382003	42382152	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042383	17	42382401	42383000	1	1	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-
	GH17I042384	17	42384416	42385372	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042386	17	42386291	42390604	1	1	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	-	+	-	-	-	+	-	-
	GH17I042392	17	42392558	42393426	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042397	17	42397982	42403783	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042403	17	42403827	42404888	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042405	17	42405902	42412371	1	0	-	-	-	-	+	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042413	17	42413267	42413890	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042414	17	42414466	42416398	1	1	-	-	-	+	+	-	+	-	-	-	-	-	-	-	+	-	+	-	-	-	-	-	-
	GH17I042416	17	42416401	42417498	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042417	17	42417975	42418400	1	0	-	+	-	-	+	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
	GH17I042420	17	42419001	42419200	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042419	17	42419366	42424521	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042596	17	42596925	42597208	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042672	17	42672991	42684404	1	0	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042698	17	42698348	42699485	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	GH17I042772	17	42772434	42776394	1	1	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	-	-	-
	GH17I042796	17	42796547	42799332	1	1	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	-	-	-	-	-	-	-	-
	GH17I042832	17	42832071	42835211	1	1	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	-	-	-
	GH17I042885	17	42885967	42887269	1	0	+	+	-	-	+	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-
	GH17I042963	17	42963446	42965776	1	1	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	-	-	-	-	-	-
	GH17I043366	17	43366706	43370722	1	1	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	-	+	-	+	+	+	+	+

Table 2. **GeneHancer database identified a 42 regulatory element that might affect human STAT3 expression**

[i] Primers for SOCS3 , BCL3 and SOD2 Adapted from: https://www.origene.com/catalog/gene-expression/qpcr-primer-pairs/mp216398/socs3-mouse-qpcr-primer-pair-nm_007707, https://www.origene.com/catalog/gene-expression/qpcr-primer-pairs/mp201260/bcl3-mouse-qpcr-primer-pair-nm_033601 https://www.origene.com/catalog/gene-expression/qpcr-primer-pairs/mp216386/sod2-mouse-qpcr-primer-pair-nm_013671