

Article

Not peer-reviewed version

Host Genetic Background Effect on Body Weight Changes Influenced by *Smad4* Knockout Using Collaborative Cross Mouse Population

Nayrouz Qahaz , [Iqbal M LONE](#) , Aya Ghnaim , Aya Khadeja , Osayd Zohud , [Nadav Bennun](#) , [Aysar Nashef](#) , [Imad Abu El-Naaj](#) , [FUAD IRAQI](#) *

Posted Date: 29 September 2023

doi: 10.20944/preprints202309.2084.v1

Keywords: Collaborative cross (CC); Smad4 knock out; body weight gain; machine learning (ML); genotyping



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Host Genetic Background Effect on Body Weight Changes Influenced by *Smad4* Knockout Using Collaborative Cross Mouse Population

Nayrouz Qahaz ¹, Iqbal M. Lone ¹, Aya Khadija ¹, Aya Ghnaim ¹, Osayd Zohud ¹, Nadav Ben Nun ¹, Aysar Nashef ², Imad Abu El-Naaj ² and Fuad A. Iraqi ^{1,*}

¹ Department of Clinical Microbiology and Immunology, Sackler Faculty of Medicine, Tel-Aviv University, Israel. nayrouz.qahaz@gmail.com (N.Q.); iqbalzoo84@gmail.com (I.M.L.); aya.khadeja91@gmail.com (A.K.); ayaghnaime89@gmail.com (A.G.); osaydzohud@gmail.com (O.Z.); nadavbennun1@mail.tau.ac.il (N.B.N.)

² Department of Oral and Maxillofacial Surgery, Baruch Padeh Medical Center, Poriya, Israel. dr.aysarn@gmail.com (A.N.); iabu@poria.health.gov.il (I.A.E.)

* Correspondence: Author: Tel-Aviv 69978; Phone: +972-3-6409321; Fax: +972-3-6409160; Mobile: +972-50-4808036; email: fuadi@tauex.tau.ac.il

Abstract: Background: Obesity and its attendant conditions have become major health problems worldwide and is currently ranked as the fifth most common leading cause of death globally. Complex environmental and genetic factors are the cause of the current obesity epidemic. Diet, lifestyle, chemical exposures, and other confounding factors are difficult to manage in humans. The mice model is helpful in researching genetics BW gain because genetic and environmental risk factors can be controlled in mice. Studies in mouse strains with various genetic backgrounds and established genetic structures provide unparalleled opportunities to find and analyze trait-related genomic loci. In this study, we used the Collaborative Cross (CC), a large panel of recombinant inbred mouse strains, to present a predictive study using CC mice *SMAD4* knockout profiles to understand and effectively identify predisposition to body weight gain. **Materials and Methods:** Male C57Bl/6J *Smad4*^{+/−} mice were mated with female mice from 10 different CC lines to create F1 mice (*Smad4*^{+/−} × CC). Body weight (BW) was measured weekly until week 16, then monthly until the end (week 48). The heritability H² of the assessed traits was estimated and presented. Comparative analysis of various machine learning algorithms for predicting the BW changes and genotype of mice was conducted. **Results and conclusions:** Our data showed that the body weight records of F1 mice with different CC lines differed between wildtype and mutant *Smad4* mice during the experiment. Genetic background affects weight gain and some lines gained more weight in the presence of *Smad4* knockout, while others gained less, but in general, the mutation caused overweight mice, except for a few lines. In both control and mutant groups, female %BW had a higher heritability H² value than males. Additionally, both sexes with wildtype genotypes showed higher heritability values than the mutant group. Logistic Regression provides the most accurate mouse genotype predictions using machine learning. We plan to validate the proposed method on more CC lines and mice per line to expand machine learning for BW prediction.

Keywords: Collaborative cross (CC); *Smad4* knock out; body weight gain; machine learning (ML); genotyping

1. Introduction

Obesity and its attendant conditions have become major health problems worldwide and is currently ranked as the fifth most common leading cause of death globally [1]. Over the last few years, the prevalence of obesity has steadily increased and almost tripled since 1975. Complex environmental and genetic factors are the cause of the current obesity epidemic. The unparalleled availability of inexpensive, appealing, calorie-dense foods, coupled with a decline in occupational physical activity, has created a setting conducive to obesity. Despite this, a substantial proportion of the world's population remains lean, showing the genetic factors that determine an individual's tendency to develop obesity. This genetic predisposition to obesity is determined by the interplay of numerous risk genes, each of which has a negligible influence on its own [2,3].

Numerous chronic disorders, including hypertension, dyslipidemia, metabolic syndrome, type 2 diabetes (T2D), cardiovascular disease (CVD), non-alcoholic fatty liver disease, Alzheimer's disease, and cancer are associated with obesity [4,5]. Being overweight/obese are substantial contributors to the increased incidence and prevalence of cancer and may surpass smoking as the most important avoidable cause of cancer [6]. The cessation of tobacco uses and the reduction of overweight and obesity may be the most influential lifestyle changes on human health and cancer in particular.

Obesity is associated with a greatly higher risk of several malignant carcinomas [7–9] and contributes to a poor prognosis and survival rate. The required processes and mechanisms for clarifying this relationship are poorly understood. Reducing the prevalence of obesity should be the ultimate goal. Therefore, population-wide education regarding a healthy lifestyle must be intensified. However, education alone will not suffice to optimize the weight of the entire population. To minimize the risk of obesity-related disorders, including metabolic syndrome and cancer susceptibility, it is also necessary to investigate druggable targets for weight reduction. Hopefully, a better knowledge of the genetic function in regulating body weight will result in the creation of novel diagnostic and treatment methods.

GWAS research related several loci and genes to illnesses and phenotypic features, including body weight, obesity, and neurological disorders [10–12]. Life history and genetic/environmental combinations often determine disease risk [13–15]. Diet, lifestyle, chemical exposures, and other confounding factors are difficult to manage in humans. The mice model is useful for researching genetics of BW gain, because genetic and environmental risk factors can be controlled in mice. Studies in mouse strains with various genetic backgrounds and established genetic structures provide unparalleled opportunities to find and analyze trait-related genomic loci.

Smad4 was first identified as a tumor suppressor in pancreatic cancer and later as a TGF- β mediator [16]. Smad4 plays a key role in TGF- β /BMP signaling by forming complexes with receptor-activated Smads, Smad2 and 3 or Smad1 and 5. Somatic inactivation of Smad4 is a frequent event in multiple tumor types. Moreover, Smad4 deletion in murine tissues, in combination with other genetic alterations that cause tumor initiation, resulted in cancer lesions of colon [17,18], pancreas [19,20], stomach [21], and liver [22]. Additionally, Smad4 correlated with depth of invasion and pathologic stage as well as regional metastases and decreased survival [23]. In animal studies, global deletion of Smad2 or Smad4 is lethal, whereas Smad3-deficient mice are viable [24]. Consequently, TGF- β /Smad3 signaling has received a great deal of attention. Mice lacking Smad3 are resistant to high fat diet (HFD)-induced obesity, insulin resistance, and diabetes in numerous models [25,26]. By contrast, little is known about Smad4 in obesity or diabetes [27].

In this study, we used a mouse panel formed by crossing Smad4 $^{+/-}$ mice with Collaborative Cross (CC) mice [28], a highly mouse genetically diverse recombinant inbred lines (RILs), allows time and cost-efficient mapping of target regions as quantitative trait loci (QTL) that are responsible for the genetic variance of a specific complex trait [29,30]. The strains are derived from eight genetically diverse founder strains (A/J, C57BL/6J, 129S1/SvImJ, NOD/LtJ, NZO/HiLtJ, CAST/EiJ, PWK/PhJ, and WSB/EiJ). Three founders of the CC strain (CAST/EiJ, PWK/PhJ, and WSB/EiJ) are wild-derived and 5 of which are of inbred lines [31]. The CC mouse population was shown to be the most powerful genetically reference population (GRP) for studying and dissecting complex traits compared to any reported approaches [32,33]. The key features of the CC in relation to gene mapping by providing a very large number of genetic variants segregate in the population (there are over 36 million SNPs) [34], and relatively high level of recombination present compared to other mouse RI sets. The CC mouse GRP will provide a unique platform and resource for studying BW complexity associated with obesity and BW gain influenced by Smad4 knock out gene.

Finally, we used machine learning (ML) methods in our recorded data to classify the different genetic backgrounds under the assessed traits and finally to provide us a tool to predict the outcome, based on earlier data. ML is a field of study that uses computational algorithms to turn empirical data into models for the purpose of analysis. It allows for the strengthening of statistical and computational approaches that auto generate the given data [35]. Recently, ML methods have been implemented in several disease-recognition and other health-related applications [1]. However, there

is a dearth of literature exploring the links between ML methods and weight gain. Furthermore, it is critical to identify the potential relationship between obesity genetics background, however this has often been overlooked in prior studies, particularly in the mouse model.

2. Methods and Material

2.1. Ethical aspects of the project

It is important to note that all animal experiments in this study were compliant with national standards for the care and use of laboratory animals, and experiment was reviewed and approved by Tel Aviv University's Institutional Animal Care and Use Committee (IACUC), with an approved number (01-19-044). Mice were monitored daily for their overall health status. Mice which showed loss of around 10% of their BW between two measure points, or 20% overall of their initial body weight, or which were observed to be suffering (less movement and activity) and based on the consultation with the Veterinarian at the small animal unit, were terminated.

2.2. Study cohort

F1 (Smad4^{+/-} x CC) and the control F1 (Smad4^{+/+} x CC) mice were generated by crossing male mice of C57Bl/6J carrying Smad4^{+/-} with 10 different CC females that it was expected to show, significant differences in their susceptibility to spontaneous cancer development as described previously in perspective review [36]. The CC mice were supplied by the animal facility of Tel Aviv University and raised in the facility. They were maintained under the ethical standards of humidity and temperature (21–23 °C).

2.3. Study Design

F1 Mice were weaned at 4 weeks of age and maintained until 48 weeks of age on a standard rodent chow diet (TD.2018SC, 18% Kcal from fat, 24% from protein, and 58% from carbs; Teklad Global, Harlan Inc., Madison, WI, USA) with a maximum of five mice per cage. The experiment begins when the mice are four weeks old. Throughout the duration of the trial, BW was determined biweekly until week 16, and then monthly until week 48. Throughout the duration, mice had free access to food and water.

2.4. Genotype

At 4 weeks old after weaning, 0.5 cm tail biopsies were collected from all F1 crossed mic of (Smad4^{+/-} x CC), and DNA extracted by NaOH boiling protocol [37]. Using the protocol provided by the Jackson Laboratory, the vendor providing the C57BL/6J- Smad4^{+/-}, mice were genotyped by Polymerase chain reaction (PCR) protocol for distinguishing the Smad4⁻ mutant allele from the wild type (Smad4⁺), using the primers: ID# 30403 (TGT AGT TCT GTC TTT CCT TCC TG), ID# 30404 (ACT GAC CTT TAT ATA CGC GCT TG), ID# oIMR2088 (AGA CTG CCT TGG GAA AAG CG), which were designed by Jackson Laboratory. For Smad4 wild type alleles, we used the primers 30403 and 30404, while for the mutant allele we used 30404 and oIMR2088 primers. For later identification, each mouse was labelled by ear cuts.

2.5. Heritability and genetic coefficient variation

Heritability quantifies the proportion of phenotype variation attributable to genetic variation. Here, we calculated the broad-sense heritability using the ANOVA data using the formula below:

$$H^2 = Vg / (Vg + Ve) \quad (1)$$

where H² is the heritability, V_g is the genetic variance between the CC lines and V_e is the environment variance. Considering the heritability results, we calculated the genetic coefficient of variation.

2.6. Computational methods:

2.6.1. Classification models

Classification is a data mining (machine learning) method that used to predict group connectivity for data instances. Decision tree (DT), Random Forest (RF), Naïve Bayes, K Neighbors, Support Vector Machine Classifier (SVC) and Logistic Regression (LR) classifiers were accomplished for further analysis. They were all applied using the K-fold cross-validation (K=4) implementation of the Python package [Scikit-Learn](#). All models were assessed by the average AUC of the ROC curve, using a 4-fold cross validation 3 times for each algorithm.

2.6.2. Decision Tree

Decision tree algorithms are most used algorithms in classification models. DT contribute a convenient modeling technique and simplifies the classification process. The decision tree is transparent mechanism it assists users to follow a tree structure easily to see how the decision is made. In this study, we used scikit-learn's decision tree default implementation, with maximum depth of 10.

2.6.3. K-Neighbors

In K-nearest neighbor (KNN) technique, nearest neighbor is measured with respect to value of k, that define how many nearest neighbors need to be examined to describe class of a sample data point. In this study, we used scikit-learn's default implementation for K-Neighbors classifier (5 neighbors).

2.6.4. Random Forest

RF is a classification that uses many decision trees. RF is a multifunctional machine learning technique. It can perform the tasks of both prediction and regression. Additionally, RF is based on bagging, and it plays an important role in ensemble ML [38]. RF thus has been implemented in biomedicine research vastly [39]. In this study, we used scikit-learn's default implementation for random forest, with 100 trees.

2.6.5. Naïve Bayes

A Bayesian Network (BN) refers graphical model for probability associations between a set of variables. In this study, we used scikit-learn's default implementation for Naïve Bayes (Gaussian).

2.6.6. Support Vector Machine Classifier

SVM is a supervised learning method that looks at data and sorts it into one of two categories. An SVM outputs a map of the sorted data with the margins between the two as far apart as possible. SVMs are used in text categorization, image classification, handwriting recognition and in the sciences

2.6.7. Logistic Regression (LR)

Logistic regression is a statistical analysis method to predict a binary outcome, such as yes or no, based on prior observations of a data set. A logistic regression model predicts a dependent data variable by analyzing the relationship between one or more existing independent variables.

2.6.8. Model Details

We used a Chained Linear Regression algorithm in this research. This algorithm's basic idea is to predict a set of ordered features using the basic features and previously predicted features. We used the following algorithm to predict the mouse's body weight from week 8 to week 48 in our study:

w_1 = last week we let the algorithm know its data

1. Predict w_2 using Sex, Genotype, BW6, ..., w_1
2. For every week, w_k , from $w_2 \rightarrow 48$ (if data exists):
Predict w_k using Sex, Genotype, BW6, ..., w_{k-1}

Finally, return the R^2 of the predicted features and the actual featu

The model was assessed by a 5-fold cross validation.

3. Results

3.1. Generation of F1 (Smad4 x CC) mice

In this study, we have assessed 391 F1 mice of 10 different CC lines, an average of 39.2 mice per line. The mice were assigned in four different experimental groups, as was described earlier. Briefly, we maintained 391 F1 mice, males, and females, with mutant and wild-type Smad4 genotypes for 48 weeks (Table 1) average 9.8 mice per group. Approximately half of them were heterozygous for the knockout allele (51.9% versus 48.1% in the control group) Table 2, and 51.7% were females versus 48.3% males (Table 2).

Table 1. Summary table of the total number (N) of F1 mice from each CC line males and females separately assessed in each study group (Wildtype VS Mutant).

	Male		Female		
	wild type	Mutant	Wild type	Mutant	
IL72	8	11	14	8	41
IL188	5	8	3	8	24
IL2513	8	3	4	12	27
IL2750	6	7	6	11	30
IL3348	10	9	12	13	44
IL5000	14	10	9	12	45
IL5008	9	12	13	11	45
IL6009	6	10	10	5	31
IL6012	10	9	6	11	36
IL6018	16	18	19	15	69
Total					391

Table 2. Comparing the prevalence of wild-type mice to mutant mice in the SMAD4 gene and the predominance of female mice to that of male mice.

	Frequency	Percent (%)	Total
Wild Type	188	48.1	391 (100%)
Mutant	203	51.9	
Female	202	51.7	
Male	189	48.3	

3.2. Dynamics of body weight (BW) changes during the experimental period for F1 (Smad4x CC) mice

Figure 1 shows the BW measurements of 10 F1 CC-C57BL/6-Smad4 lines at 14 time points. At 6 weeks, control mice weighed 19.09 to 21.53 g (Figure 1b), while F1 CC-Smad4+/- mice weighed 17.43

to 21.72 g (Figure 1b) Smad4^{+/-} lines had significant ($p < 0.05$) body weight differences. Except for IL2513, the control and Smad4^{+/-} groups had no initial differences. Both groups and lines gained weight during the trial. Only males from line 6018 began losing weight around week 32 in both control and mutant groups, as shown in (Figure 2). Females reached nearly Plateau week 40 and 48, or with modest increase (Figure 2a). IL2513 weighed the least, while IL6012 weighed the most (Figure 2), even in Figure 1, which shows the dynamic BW for each line regardless of sex.

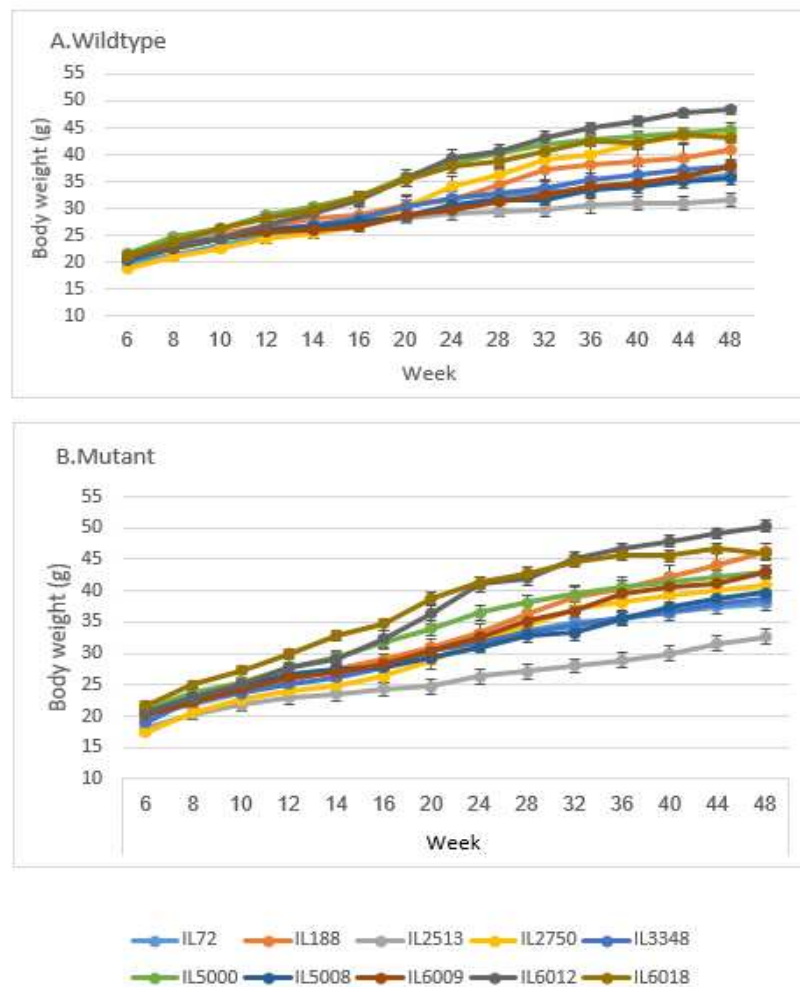


Figure 1. BW (\pm SE) A of the control and B Smad4^{+/-} groups at 14 time points (8, 10, 12, 14, 16, 20, 24, 28, 32, 36, 40, 44 and 48 weeks old) of 10 different sets of 10 different CC lines. The X-axis represents the time points (in weeks) while the Y-axis represents values of body weight (g). One-way ANOVA performed for statistical analysis, $p < 0.05$.

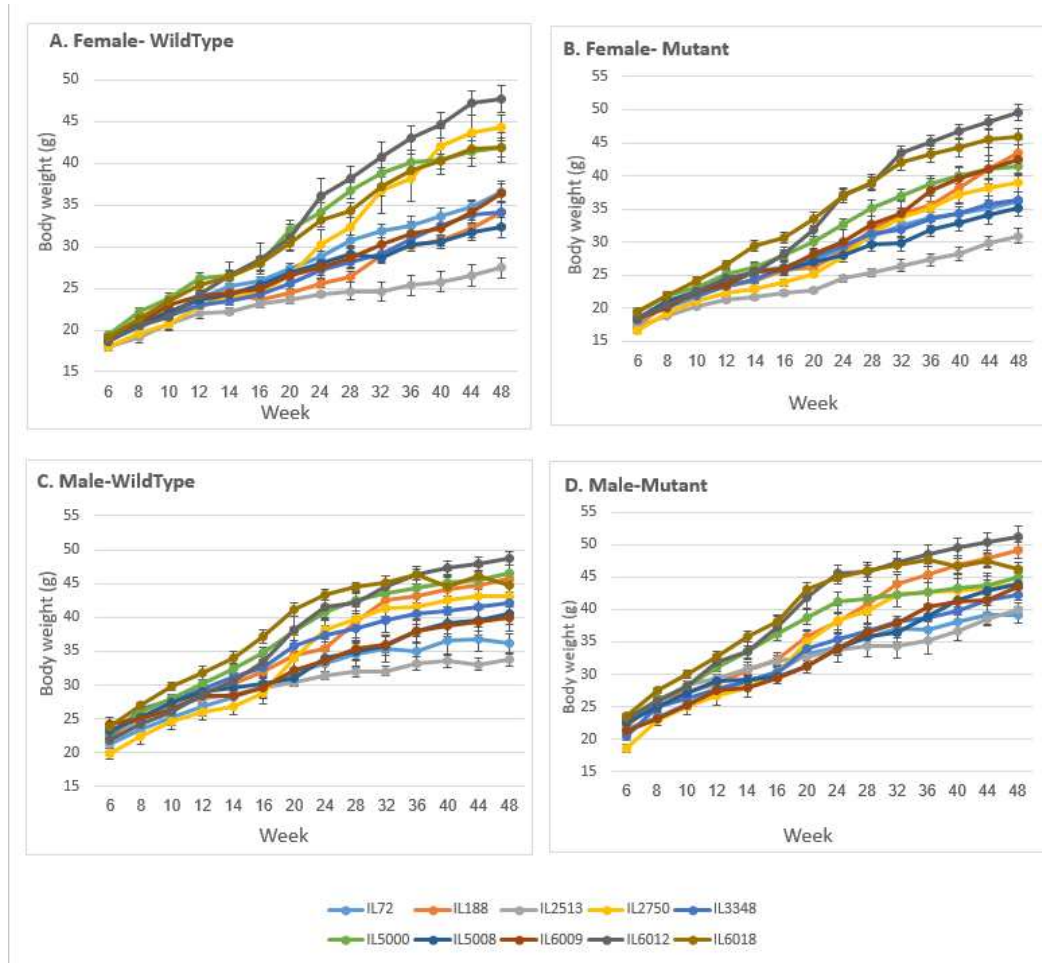


Figure 2. Body weight (\pm SE) of the A, C control and B, D *Smad4*^{+/-} groups at 14 time points (8, 10, 12, 14, 16, 20, 24, 28, 32, 36, 40, 44 and 48 weeks old) of 10 different set of CC lines. The X-axis represents the time points (in weeks) while the Y-axis represents values of body mass (g). One-way ANOVA performed for statistical analysis, $p < 0.05$.

The control lines' BWs ranged from 31.6750 (IL2513) to 48.3063 (IL6012) the end of the experimental weeks. While, F1 *Smad4*^{+/-} carriers weighed 32.7-50.31g. Both control and *Smad4*^{+/-} carriers had significantly different final body weights Figure 1.

3.3. Dynamics of BW Changes in grams over the 48-week experiment for each CC line

Each CC mouse line has a distinct pattern of BW gain in both control and mutant animals, and the weight difference between control and mutant mice is unique to each line. Except for lines IL2513, IL2750, IL3348, and IL5000, mutant mice in most lines and all analysis groups of each line weighed more than control mice. Line IL2513 has a unique weight gain pattern; in the overall group, control mice are heavier than mutant mice until the weight ratio reverses, which occurs at 44 weeks (Figure 3). From the start of the experiment to 20 weeks, the difference in BW between control and mutant mice is significant. In the female group, the point of reversal occurs at 20-week-old rounds, and the weight difference increases, but not significantly, during the experiment. Males have no bending point; mutant mice weighed more than control mice from the start of the experiment, but the weight difference grew over time and became significant only between weeks 44 and 48.

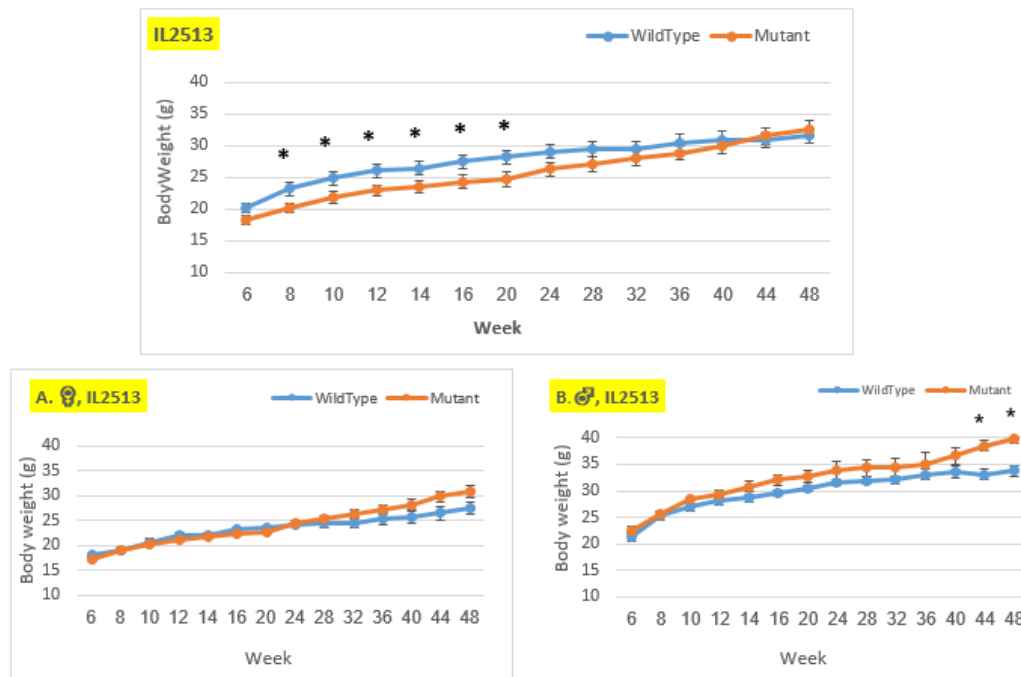


Figure 3. Body weight changes in wildtype and Smad4^{+/-} mice at 14 time points for line IL2513 (6-48 weeks old). The X-axis depicts time points (in weeks), while the Y-axis depicts body weight values. The weight increase trend in this line is distinctive and distinct from the other lines; in the overall group, the weight of control mice is greater than the weight of mutant mice up to a reversal point at which the weight ratio reverses, which occurs at 44 weeks. The difference is significant between control mice and mutant mice BW from the start of the experiment to age of 20 weeks in overall group. In the female group, the bending point begins at 20-week-old reversal point and grows somewhat until the end of the experiment, but not significantly. There is no bending point in males. Mutant mice weighed more than control mice from the start of the experiment, but the difference in the weights grew with time and became significant only around week 44 and 48.

In the overall and female groups, the control mice consistently weighed significantly more than the mutant mice in line IL2750 (Figure 4). Even at the end of the experiment, the control mice in the overall group beginning in week 44 and the female group beginning in week 40 weighed significantly more than the mutant mice. In the male group, there is no difference in BW. This could mean that females were more susceptible to the mutation and that it prevented them from gaining weight, just like control mice. Line IL5000 was found in all groups where the control mice were heavier than the mutant mice in a few weeks (Figure 5). This occurred in females from weeks 28 to 48, males from weeks 10-15 and 20-36, and the overall group throughout the experiment, but with no statistical significance. Line IL3348, in the male group, the control mice weighed more than the mutant mice from the start of the experiment until week 40, but there was no significant difference, and by weeks 44 and 48, their weight was nearly identical (Figure 6). The overall group and the female groups have a reversal point where the control mice weighed more than the mutation mice before it reversed. This occurs at week 10 in the female group and week 20 in the overall group.

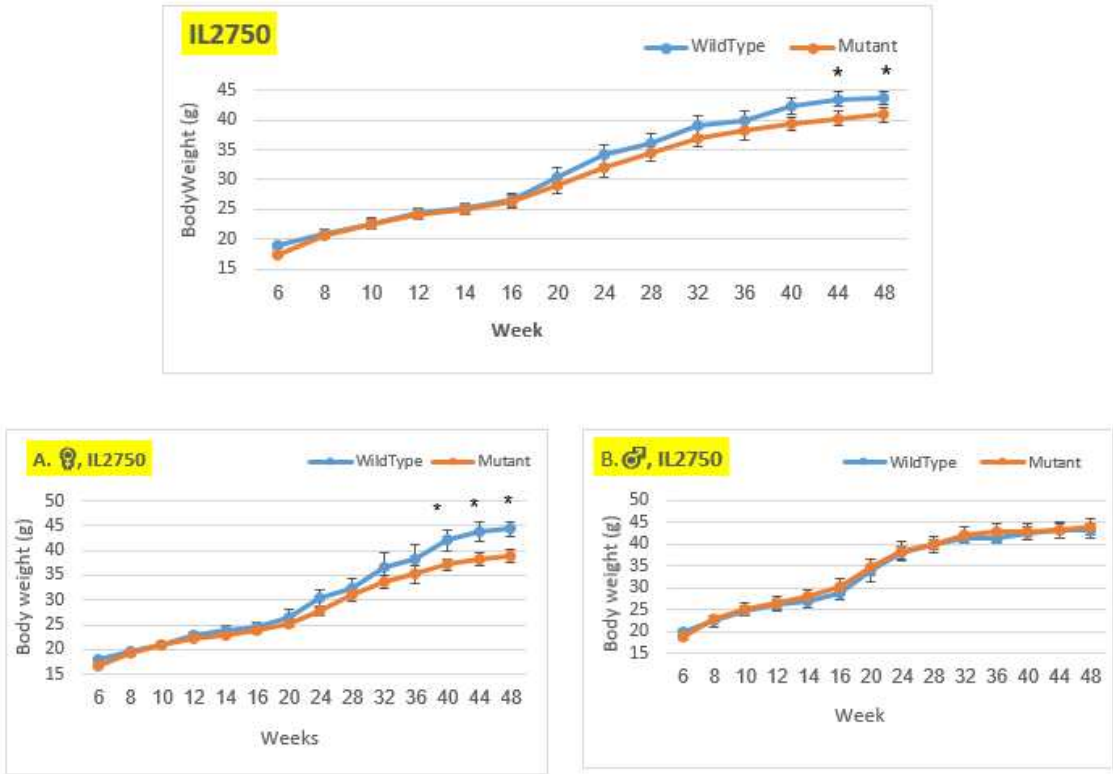


Figure 4. Body weight changes in wildtype and Smad4+/- mice at 14 time points for line IL2750 (6-48 weeks old). The X-axis depicts time points (in weeks), while the Y-axis depicts body weight values. From week 20 through the end of the period, we may observe a difference in the weight of mutant mice compared to control mice in both the female and overall groups. Only from week 40 demonstrates a significant difference. There is no difference in the male group.

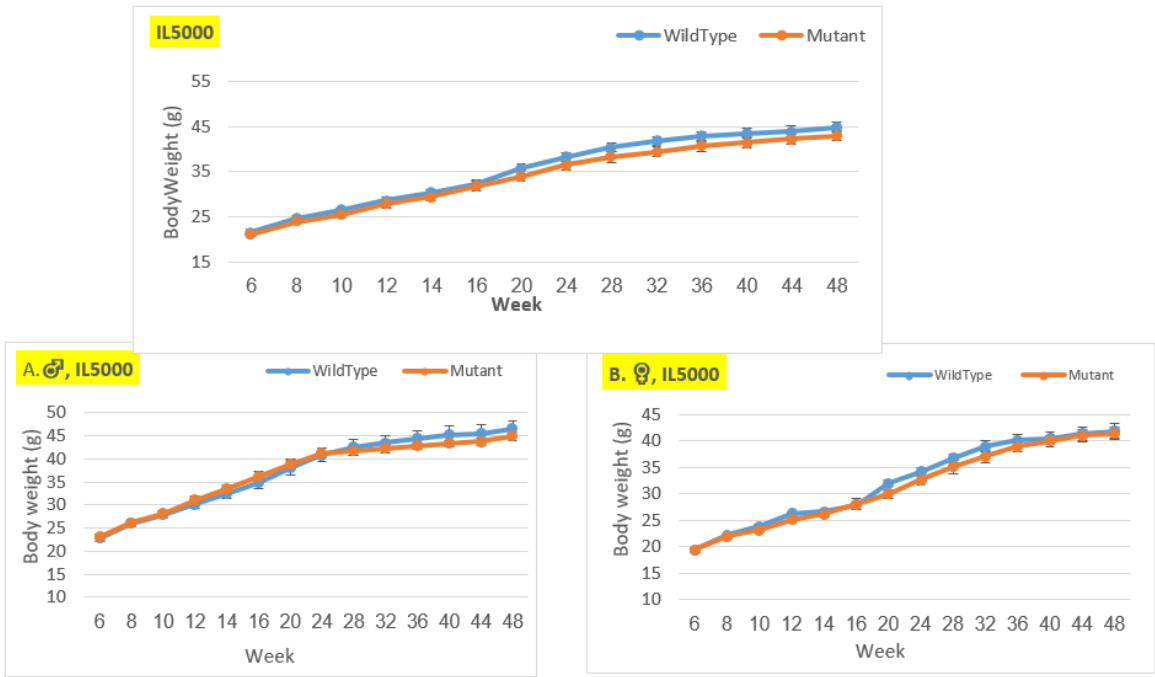


Figure 5. Body weigh changes in wildtype and Smad4^{+/-} mice at 14 time points for line IL5000 (6-48 weeks old). The X-axis depicts time points (in weeks), while the Y-axis depicts body weight values. Throughout the experiment, the control mice weighed more than the mutant mice, and the difference rose later in the experiment but did not achieve a statistically significant difference. A reversal points in the female group at week 24 where the mutant mice weighed more than the control mice and vice versa, but there was no significant difference. In the male group, control and mutant mice weighed the same from week 40 to week 48, whereas control mice weighed more before that. There was no discernible difference.

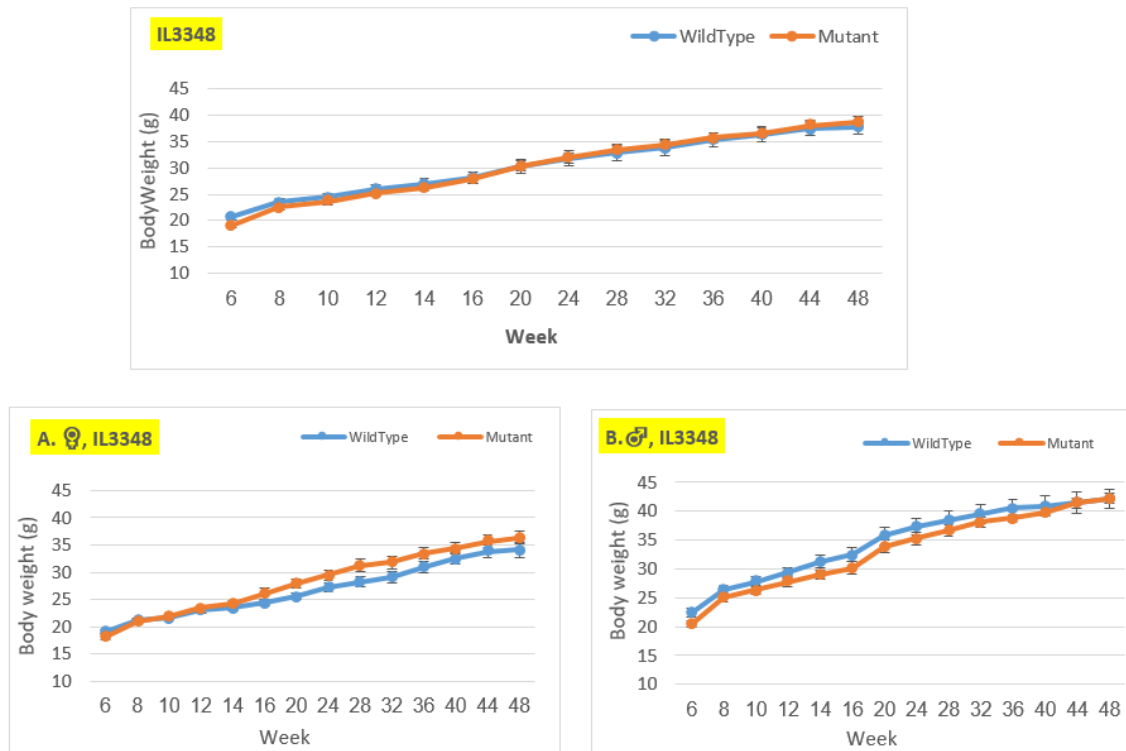


Figure 6. Body weight changes in wildtype and Smad4^{+/-} mice at 14 time points for line IL3348 (6-48 weeks old). The X-axis depicts time points (in weeks), while the Y-axis depicts body weight values. The mutant mice weigh about the same as the control mice throughout the experiment in the overall group. Nonetheless, when analyzed by gender, in the female group, the mutant mice weighed more than the control mice from around week 14 to the end of the experiment, but the difference is not statistically significant. In contrast, in the male group, the control mice weighed more than the mutant mice from the beginning of the experiment until week 40, but there was no significant difference, and by weeks 44 and 48, their weight was nearly identical.

Line IL72 is the most resistant to the body weight mutation, especially in the overall and female populations (Figure 7). Throughout the experiment, the mutant mice weigh nearly the same as the control mice. Throughout the experiment, there was an increase in weight difference favoring the mutant mice in the male group, but it was insignificant. Line IL188 in all groups, the mutant mice weigh more than the control mice. Among all experimental groups, the female groups are the most affected by weight from the mutation (Figure 8). At the start of the experiment, Line IL5008 was resistant to the effect of the mutation on weight; mice with the mutation weighed nearly as much as control mice (Figure 9). However, beginning at week 36, the weight difference between the mutant mice and the control mice in the female and male groups increased throughout the experiment but was not statistically significant. However, the difference in the overall group begins earlier, at week 28, and continues to rise throughout the experiment, reaching significant level at week 44.

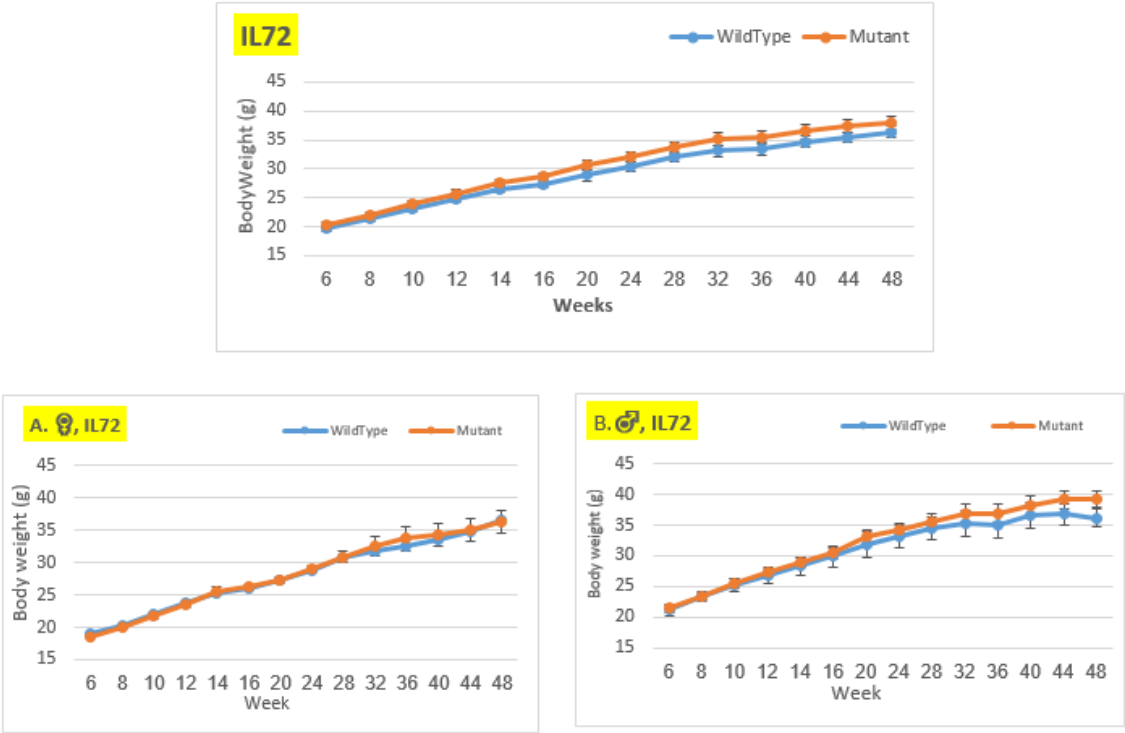


Figure 7. Body weigh changes in wildtype and Smad4+/- mice at 14 time points for line IL72 (6-48 weeks old). The X-axis depicts time points (in weeks), while the Y-axis depicts body mass values in grams (g). We can see a difference in the weight of mice with a mutation compared to control mice in the male group from week 28 until the conclusion of the period, although it is not significant. In females and the overall group, there is essentially no difference in BW between mutant and wild type mice.

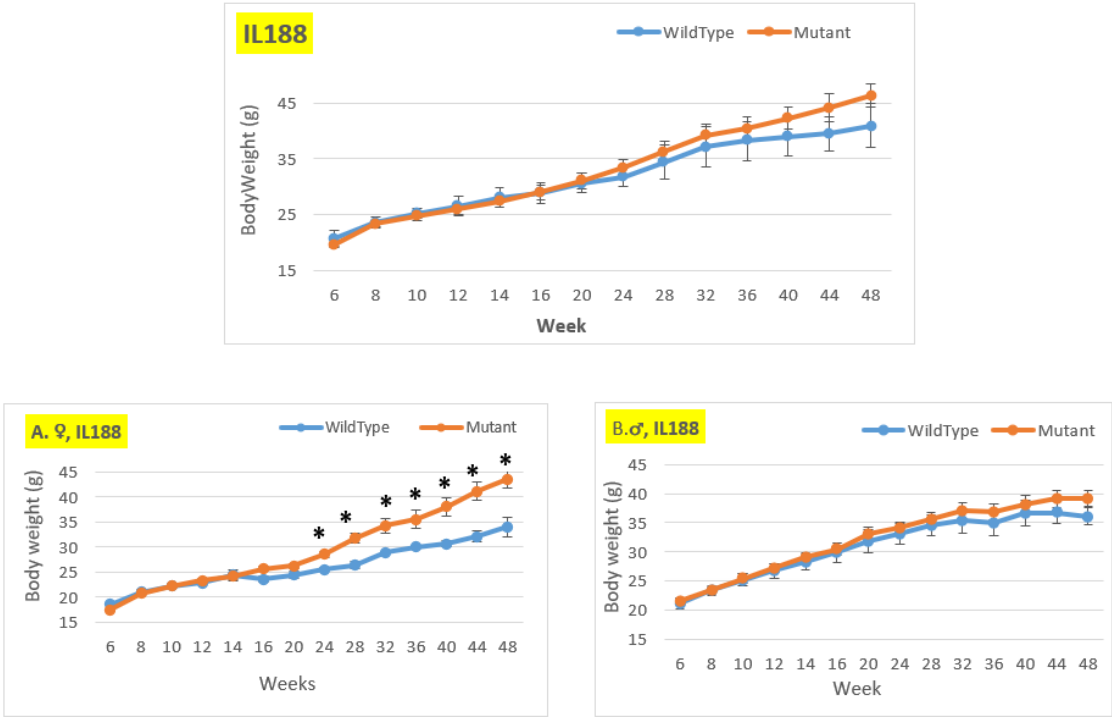


Figure 8. Body weight changing of the wildtype and Smad4^{+/-} mice for line IL188 at 14-time points (6-48 weeks old). The X-axis represents the time points (in weeks), while the Y-axis represents values of body mass. We can see a difference in the weight of mice with a mutation compared to control mice in all groups (males, females, and overall) from week 24 to the end of the period, although only the female group from week 24 shows a significant difference.

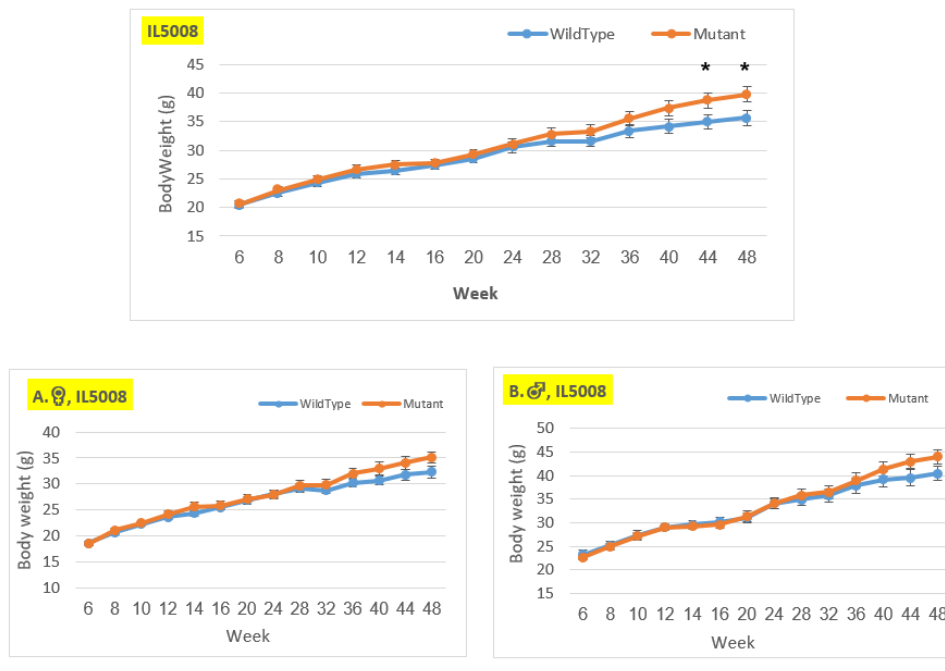


Figure 9. Body weight changes in wildtype and Smad4^{+/-} mice at 14 time points for line IL5008 (6-48 weeks old). The X-axis depicts time points (in weeks), while the Y-axis depicts body weight values. In this line, mutant mice weigh more than control mice in all three groups (males, females, and overall); however, the weight difference grew during the experiment and is only significant in the overall group at weeks 44 and 48.

Line IL6009, from week 14 through the end of the period, we may observe a difference in the weight of mutant mice compared to control mice in both the female and overall groups. Only from week 28 demonstrates a significant difference in the overall group, and from week 36 in the female group. In the male group, there was no significant difference until week 12 the control group weighed more than the mutant is, starting through the end of the experiment, the mutant mice weighed more than the control mice, also without significant (Figure 10). Line IL6018, a significant difference in the weight of mutant mice compared to control mice in both the female and overall groups from week 14 to the end of the period was observed, but no difference was observed between the male groups (Figure 11). In contrast, the male group demonstrated a difference at line IL6012 throughout the experiment, the mutant mice were heavier than the control mice, but this difference was insignificant. In the female and overall groups throughout the experiment, the weight differences were negligible (Figure 12).

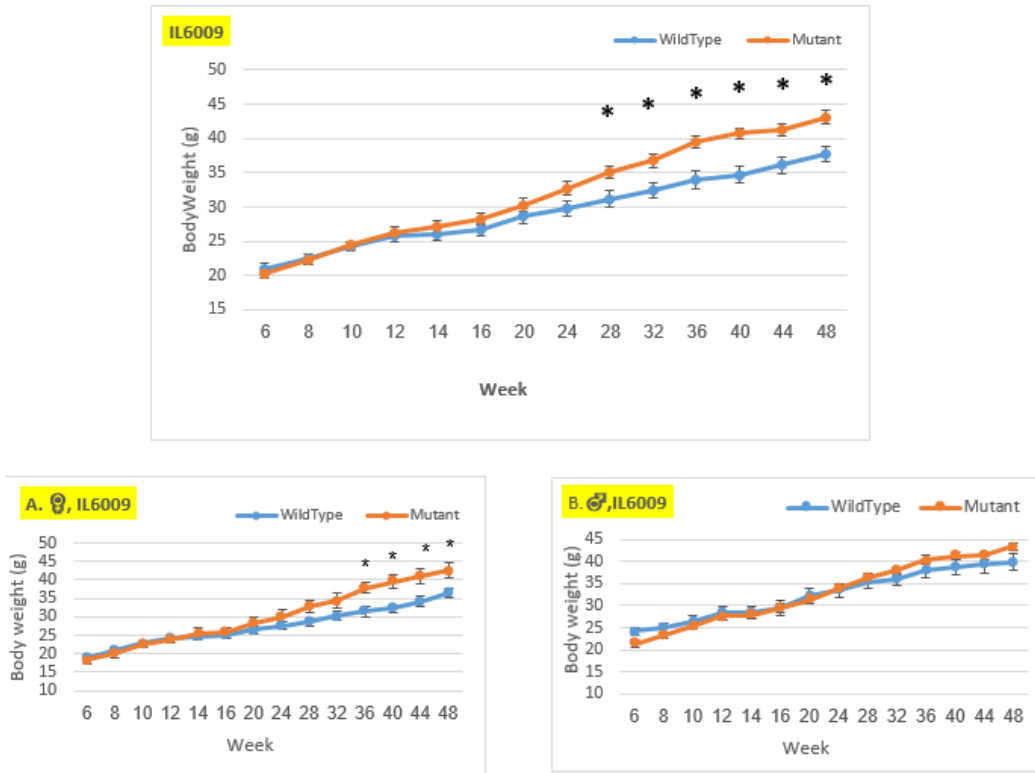


Figure 10. Body weight changes in wildtype and Smad4^{+/-} mice at 14 time points for line IL6009 (6-48 weeks old). The X-axis depicts time points (in weeks), while the Y-axis depicts body weight values. From week 14 through the end of the period, we may observe a difference in the weight of mutant mice compared to control mice in both the female and overall groups. Only from week 28 demonstrates a significant difference in the overall group and from week 36 a significant difference in the female group. In the male group, there is no significant change, until week 12, the control group weighed more than the mutant. Starting through the end of the experiment, the mutant mice weighed more than the control mice, but the difference was not significant.

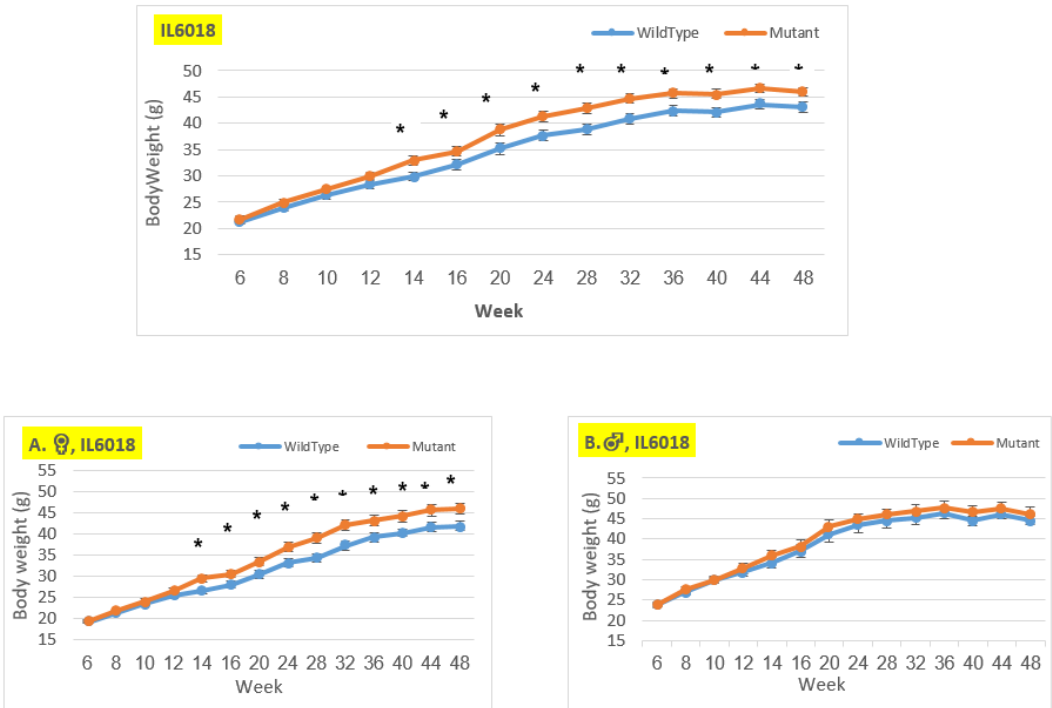


Figure 11. Body weight changes in wildtype and Smad4^{+/-} mice at 14 time points for line IL6018 (6-48 weeks old). The X-axis depicts time points (in weeks), while the Y-axis depicts body weight values. We may see a substantial difference in the weight of mutant mice compared to control mice in both the female and overall groups from week 14 to the end of the period. There is no difference between the male groups.

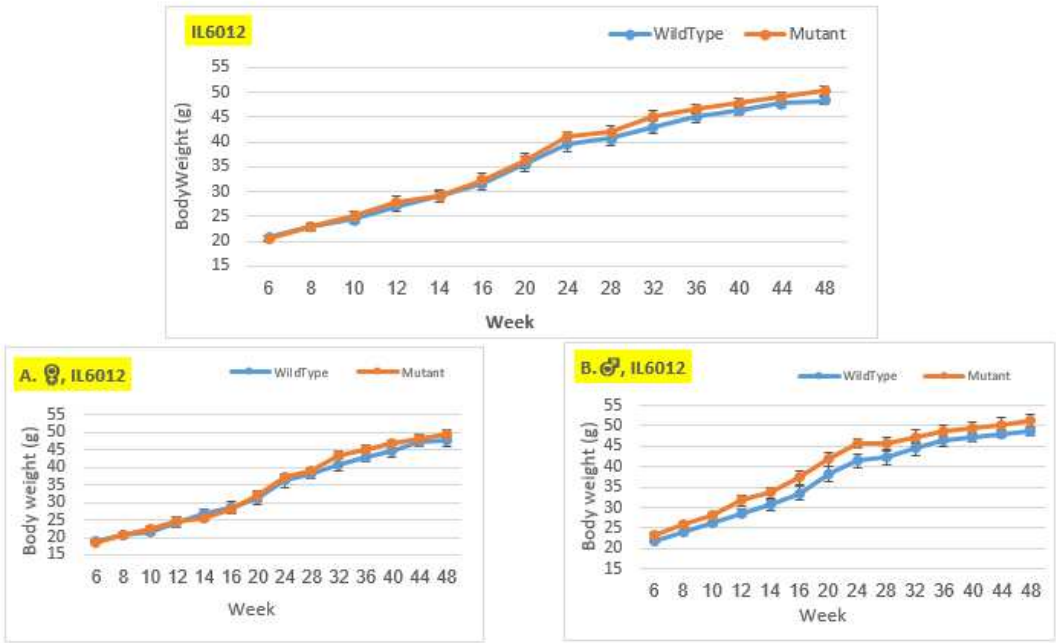


Figure 12. Body weigh changes in wildtype and Smad4^{+/−} mice at 14 time points for line IL6012 (6–48 weeks old). The X-axis depicts time points (in weeks), while the Y-axis depicts body weight values. In the overall group and the female group, the weight of the mutant mice was practically identical to that of the control mice from the beginning of the experiment until about week 20, after which the mutation mice weighed more than the control mice until the end of the experiment. In the male group, mutant mice consistently weighed more than control mice throughout the duration of the experiment, without any significant difference.

3.4. Percentage of BW gain (g) of CC lines after 48 weeks

Figure 13 shows different %BW responses between the CC lines, which may support our proposed study. F1 mice with the Smad4 mutation gained 47.211% more weight than wild-type F1 mice (41.753%) by week 16, 89.2% compared to 77.67% by week 32, 101.6 % versus 87.97% by week 40, and 111.8% versus 95.69% by week 48. IL6009 and IL3348 showed a significant pattern. For lines IL5008 and IL188, the difference started to be significant only towards week 40, and for IL2513 in week 48. Line IL5000 showed that the control mice gained more weight than the mutant mice from %BW 32–6 to %BW 48–6, and line IL2513 only in the first %BW 16–6, after which the trend reversed and the difference grew in favor of the mutant mice at %BW 48–6. 18 In line IL6018, the difference between mutant and control mice's weight gain is decreasing and no longer significant at %BW 40–6 and %BW 48–6. 18 In lines IL72 and IL2750 at the start of the experiment, mutant mice gain more weight than control mice, but by experiment's end, average weight gain is almost identical. Figures 14–17 show the %BW of mutation mice versus control mice, each line separately and the total mean of all the mice by gender, comparing males and females with and without a mutation. As the experiment progresses and the mice age, the weight differences between mutant and control mice, as well as males and females with and without a mutation, increase. Most of the lines show a distinct pattern throughout the experiment. Line IL188 shows a significant weight difference between female mice with and without a mutation, which increases over time. Only in %BW 6–48 does line IL2513 show a significant difference in mouse weight between mutant females and controls and mutant males and controls. Line IL2750 observes significant %BW differences between males with and without mutation at weeks 16–6 and 6–32, and these differences diminish as the trial progresses. A mutation eliminated a gender difference by week 48.

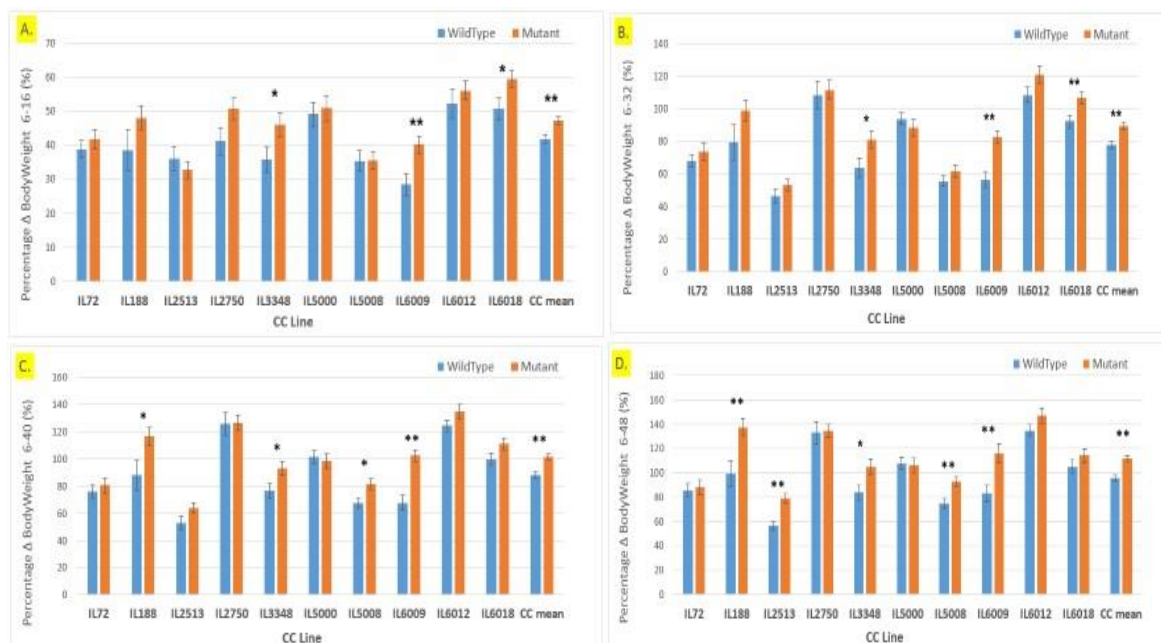


Figure 13. Percentage body weight gain (g) of ten different CC lines after 48 weeks of BW monitoring in the overall F1 population (males and females). (A) Shows BW percentage gain in 10 different lines of F1 population with mutation vs. wild type until week 16, BW percentage gain until week 32 is

shown in (B), (C) Shows the percentage increase in BW until week 40, and (D) the percentage increase in BW until week 48.

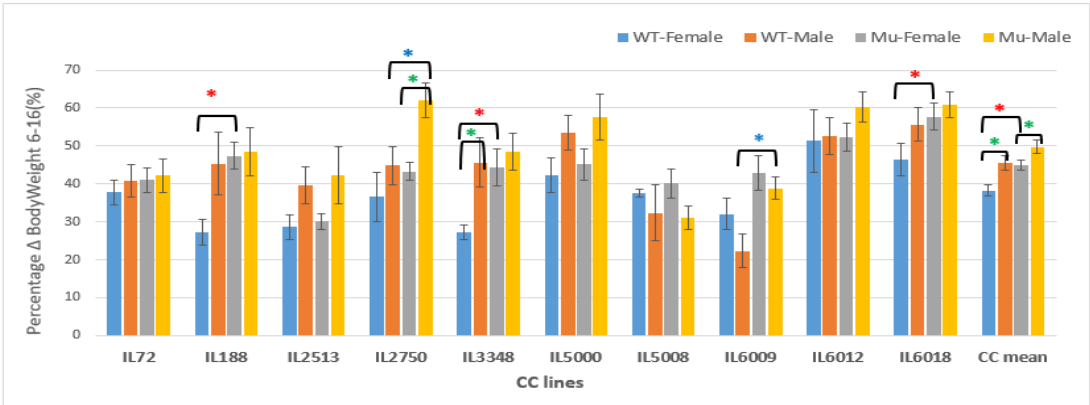


Figure 14. Percentage changes in Body weight in grams (g) between weeks 6-16 of ten CC lines, separately for females and males, having wildtype and mutant genotypes. The X-axis depicts the various CC lines, while the Y-axis depicts the percentage change in BW over weeks 6-16.

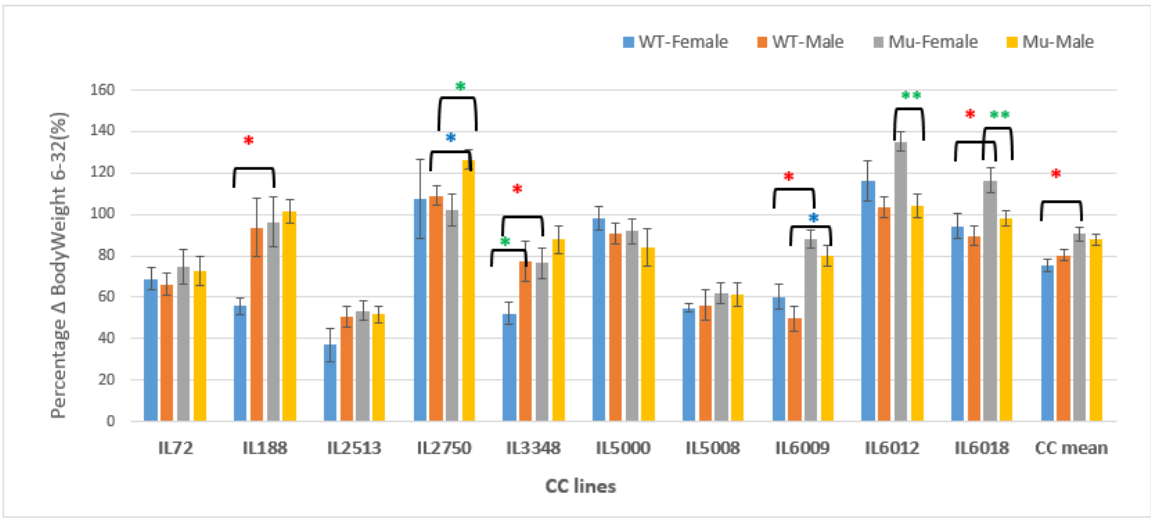


Figure 15. Percentage changes in body weight in grams (g) between weeks 6-32 of ten CC lines, separately for females and males, having wildtype and mutant genotypes. The X-axis depicts the various CC lines, while the Y-axis depicts the percentage change in BW over weeks 6-32.

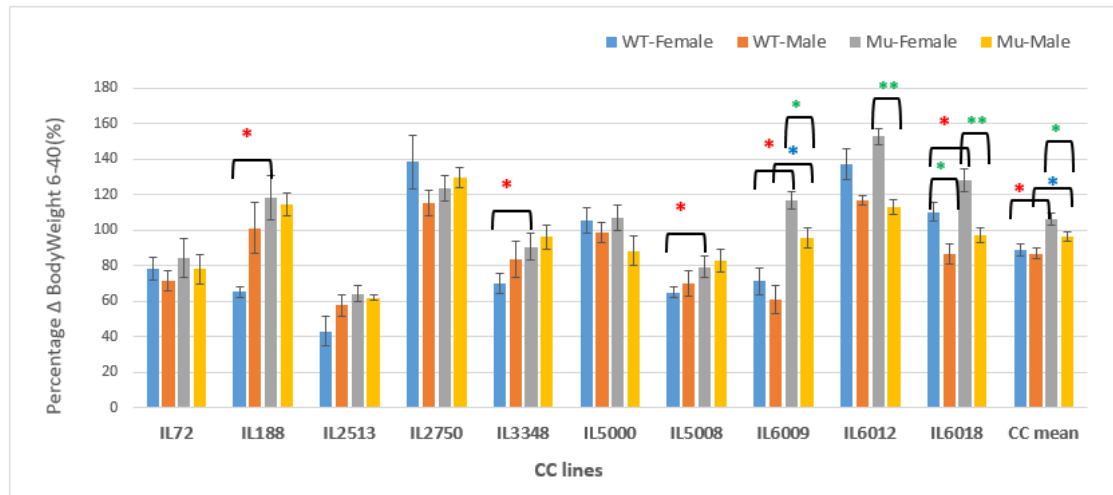


Figure 16. Changes in BW as a percentage in grams (g) between weeks 6-40 of ten CC lines, separately for females and males, having wildtype and mutant genotypes. The X-axis depicts the various CC lines, while the Y-axis depicts the percentage change in BW over weeks 6-40.

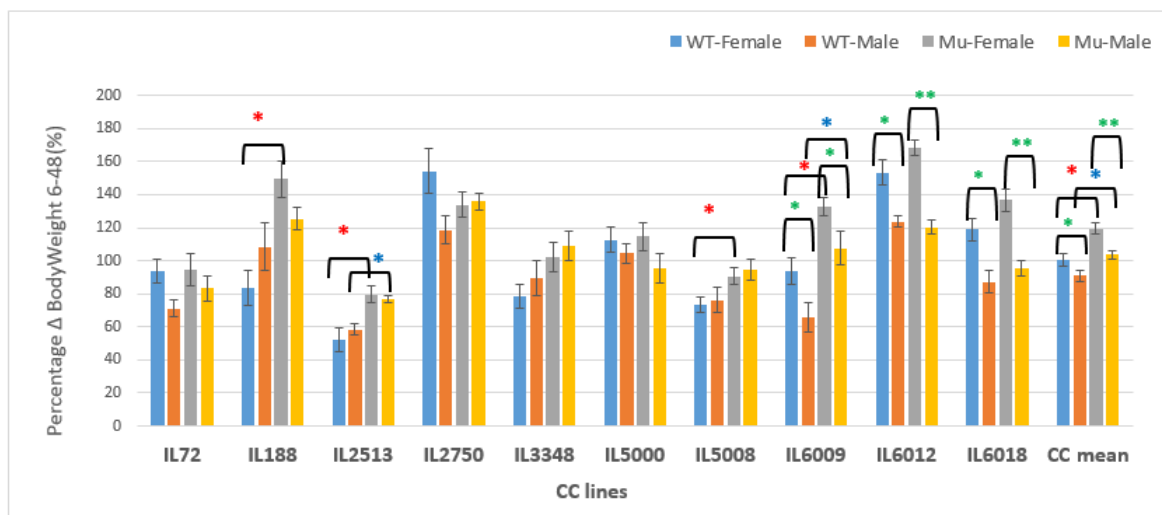


Figure 17. Changes in BW as a percentage in grams (g) between weeks 6-48 of ten CC lines, separately for females and males, having wildtype and mutant genotypes. The X-axis depicts the various CC lines, while the Y-axis depicts the percentage change in BW over weeks 6-48.

Line IL2750 shows significant differences in %BW only at %BW 16-6 and %BW 6-32, and these differences diminish as the experiment continues. In those weeks, males and females with a mutation gained weight differently; however, as the experiment progressed, both sexes gained weight similarly by week 48. Line IL3348 begins an experiment comparing mutation- and mutation-free females. The difference decreases but remains significant until %BW 40-6. In %BW 16-6 and %BW 6-32, male control mice are heavier than females, and the difference decreases over time. Line IL6009 has a large difference between mutant and nonmutant males. In %BW 6-32, mutation-free and mutation-carrying females differed significantly. Male and female mutant mice had significantly different %BW 6-40. Once a difference between male and female control mice is introduced, %BW 6-48 becomes significant in all groups.

Males and females of line IL6012 did not differ significantly between mutant and control mice. Neither male nor female IL6012 mice differed significantly between mutant and control mice. In the mutant group (6-32, 6-40, and 6-48) and the control group (%BW 48-6), females have a higher %BW than males. Line IL6018 showed a significant difference between mutant and control females early in

the experiment, but only until %BW 40-6. In the mutation group and control mice, %BW 48-6 favors females. Mutant mice started at %BW 6-32 and control mice at %BW 6-40.

Males gained %BW more than females at the beginning of the experiment in both groups, mutant mice and control mice in the cc mean, but this trend reversed over time, and by the end of the experiment, females gained significantly more weight in both groups. This is true for lines IL6000, IL6012, IL6018, and for lines IL72 and IL5000 but it is not significant. %BW line IL5008 began with females gaining more %BW than males, but reversed itself with no significant differences. IL3348 Males gained a greater %BW than females, and only the control group showed a significant difference in %BW 16-6 and %BW 32-6. Line IL2513 begins with a large difference in %BW between males and females and ends with a dramatic decrease.

3.5. Computational methods

3.5.1. Heritability

This study aimed to discover whether BW phenotypic variance has a genetic basis in Smad4-knockout F1 populations. Table 3 summarizes heritability (H^2) values calculated to answer this topic. One-way ANOVA was used to calculate the heritability of sex- and genotype-specific characteristics. Different traits calculated: %BW6-16, %BW6-32, %BW6-40, %BW6-48 for both sexes and genotypes. In both control and mutant groups, female %BW had a higher H^2 value than males. Wildtype sexes have higher heritability than mutants. In the fourth groups, %BW6-16 values were low and leaned toward 0. Both genotype groups had the same H^2 in %BW6-32 (0.52). In females of both genotypes, H^2 in %BW 6-32 is highest (0.62, 0.64 respectively).

Table 3. Exhibit the results of calculating heritability (H^2) values. The heritability was calculated using One-Way ANOVA for %BW6-16, %BW6-32, %BW6-40, %BW6-48 traits that were calculated separately by sex and genotype.

		Trait	df between	df within	n	MS between	MS within	VG	H^2
Female	Wild	% Δ BW16-6	10.00	87.00	8.00	582.11	170.34	51.47	0.23
		% Δ BW32-6	10.00	87.00	8.00	5380.42	464.75	614.46	0.57
		% Δ BW40-6	10.00	87.00	8.00	7035.67	457.41	822.28	0.64
		% Δ BW48-6	10.00	86.00	7.91	7986.66	512.58	792.01	0.61
	Mutant	% Δ BW16-6	10.00	97.00	8.91	682.42	154.55	59.25	0.28
		% Δ BW32-6	10.00	97.00	8.91	7073.21	497.56	738.08	0.60
		% Δ BW40-6	10.00	97.00	8.91	8974.52	510.82	831.03	0.62
		% Δ BW48-6	10.00	97.00	8.91	8705.48	614.00	908.23	0.60
Male	Wild	% Δ BW16-6	10.00	83.00	7.64	923.51	289.01	83.09	0.22
		% Δ BW32-6	10.00	83.00	7.64	3753.41	400.83	439.03	0.52
		% Δ BW40-6	10.00	83.00	7.64	3558.53	476.41	403.61	0.46
		% Δ BW48-6	10.00	83.00	7.64	4110.59	541.02	467.44	0.46
	Mutant	% Δ BW16-6	10.00	88.00	8.09	1230.39	201.05	127.22	0.39
		% Δ BW32-6	10.00	88.00	8.09	3524.05	361.56	390.87	0.52
		% Δ BW40-6	10.00	88.00	8.09	2554.56	403.16	265.90	0.40
		% Δ BW48-6	10.00	88.00	8.09	2597.15	536.89	254.64	0.32

3.5.2. Classification and Regression models

To predict the genotype (0/1) of a mouse based on its sex and 14 recorded weights, Table 4 displays the results of six distinct classification algorithms for each line. Most lines in this table have low AUC values. The Logistic Regression model predicted the genotype of four lines (IL188, IL2513, IL2750, and IL6009), making it the most effective. Line IL2513 of Logistic Regression had the greatest AUC (0.875). IL6009 has multiple high values, including SVC = 0.836, LR = 0.82, NaBa = 0.787, and KNN = 0.798. No model was able to predict the IL3348, IL5000, IL5008, and IL6012 lines with scores greater than 0.6.

Table 4. Displays the results of six different classification algorithms for each line to predict the genotype (0/1) of a mouse based on its sex and 14 recorded weights.

Line	IL72	IL188	IL2513	IL2750	IL3348	IL5000	IL5008	IL6009	IL6012	IL6018
N	41	24	27	29	44	45	45	31	36	69
6W	0.29	-0.30	0.64	0.08	0.30	0.35	0.38	0.23	0.33	0.44
8W	0.24	-0.21	0.63	0.13	0.37	0.37	0.56	0.15	0.27	0.53
10W	0.26	-0.03	0.60	0.28	0.57	0.39	0.53	0.08	0.37	0.51
12W	0.35	-0.20	0.64	0.44	0.52	0.32	0.52	0.13	0.27	0.48
14W	0.43	-0.32	0.59	0.41	0.54	0.25	0.54	0.03	0.32	0.61
16W	0.32	-0.40	0.56	0.33	0.53	0.21	0.58	-0.15	0.31	0.64
20W	0.47	-1.23	0.42	0.13	0.51	0.20	0.71	-0.12	0.25	0.61
24W	0.46	-1.58	0.55	0.15	0.66	0.06	0.74	-0.28	0.40	0.61

3.5.3. Model details

We predicted the BW of mice between weeks 8 and 48. The model was evaluated by means of 5-fold cross-validation. The mean score is presented per Line in Table 5. Lines IL6018, IL5008, IL2513 and IL3348 are the most predictable in terms of mouse weight, with line 5008 having the greatest predictive ability. Beginning in week 6, line IL2570 is highly predictable. Lines IL6009 and IL188 are the least predictable, while line IL72 results improved with more data but remained unpredictable.

Table 5. Displays the results of BW prediction algorithms for each line to predict the genotype (0/1) of a mouse based on its sex and 14 recorded weights.

Line	IL72	IL188	IL2513	IL2750	IL3348	IL5000	IL5008	IL6009	IL6012	IL6018
N	41	24	27	29	44	45	45	31	36	69
DT	0.504	0.5	0.652	0.513	0.41	0.497	0.543	0.58	0.494	0.433
NaBa	0.603	0.459	0.67	0.58	0.553	0.491	0.584	0.798	0.528	0.644
KNN	0.506	0.5	0.718	0.637	0.555	0.42	0.556	0.787	0.483	0.539
RF	0.543	0.619	0.747	0.546	0.571	0.512	0.457	0.718	0.518	0.561
SVC	0.418	0.438	0.682	0.386	0.407	0.472	0.325	0.836	0.459	0.576
LR	0.332	0.747	0.875	0.752	0.531	0.392	0.557	0.82	0.405	0.586

3.5.4. Correlation analysis between the studied traits

One of our proposed experiments was to evaluate if initial and BW gain correlate with adult weight in mice with or without a mutation. Pearson's correlations quantify BW correlations and direction (increase or decrease). +1 indicates a perfect positive relationship and correlation, -1 a perfect negative one, and 0 no relationship. Early and late mouse weights were correlated using heatmaps. Each line reflects mice's weekly weight. Fourth groups on each line are separate. Here are three examples of a heatmap for lines IL72, IL3348, and IL6009. The variations observed between different CC lines and within the same line have been presented in Figures 18–20.

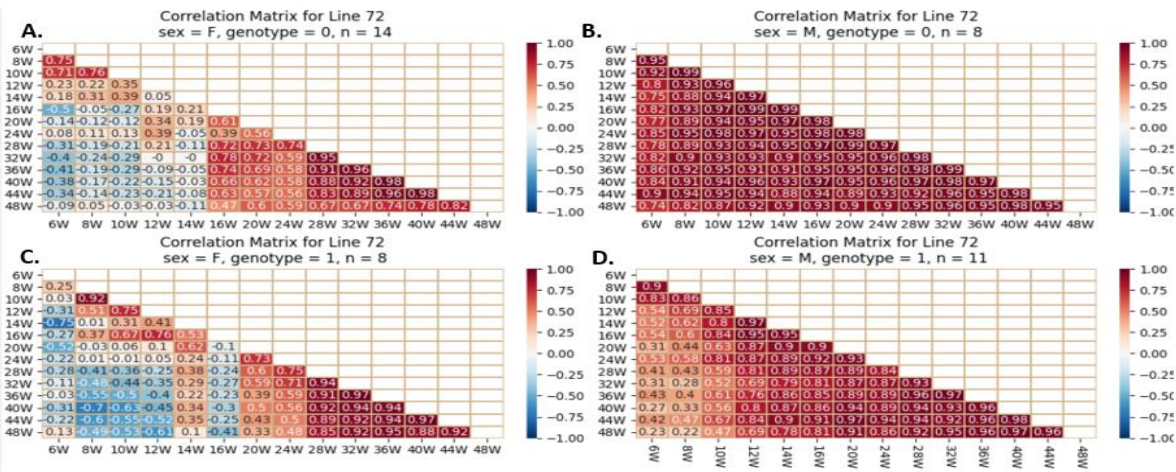


Figure 18. Heatmap and Pearson correlation between the body weight of female and male mice of line IL72, either wild-type or mutant, between various weeks. Figures (A) depict wildtype female, (B) wildtype male, (C) mutant female, and (D) mutant male. The r values ranged from minimum (-1) to maximum (1).

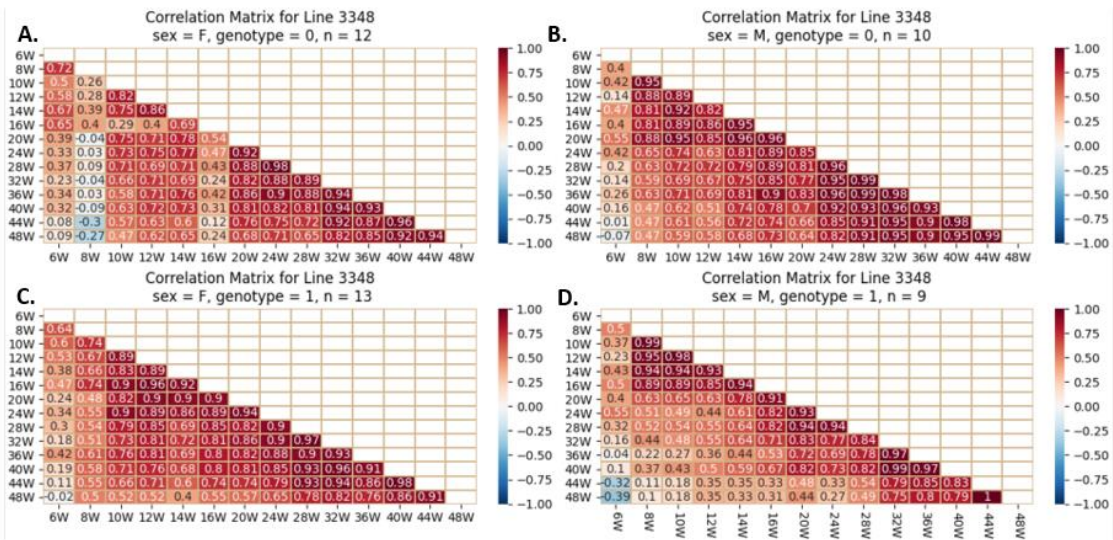


Figure 19. Heatmap and Pearson correlation between the body weight of female and male mice of line IL3348, either wild-type or mutant, between various weeks. Figures (A) depict wildtype female, (B) wildtype male, (C) mutant female, and (D) mutant male. The r values ranged from minimum (-1) to maximum (1).

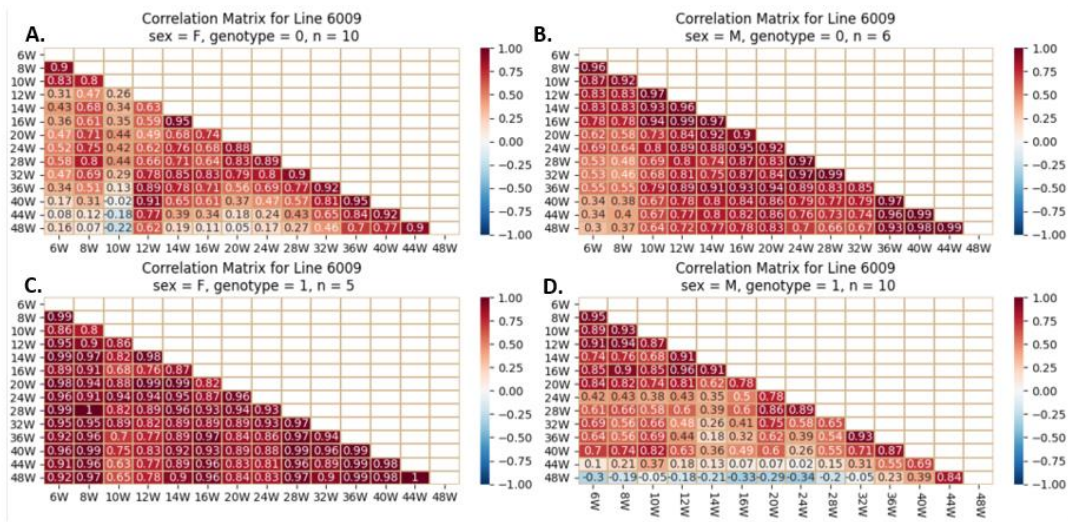


Figure 20. Heatmap and Pearson correlation between the body weight of female and male mice of line IL6009, either wild-type or mutant, between various weeks. Figures (A) depict wildtype female, (B) wildtype male, (C) mutant female, and (D) mutant male. The *r* values ranged from minimum (-1) to maximum (1).

In advanced weeks, the weight of mice with a positive and stronger correlation tends toward 1 with the weight in subsequent weeks; this correlation begins at week 28 and strengthens by week 44; however, this correlation is poorly investigated in mutant males of line IL6009 and is weaker in control females. Line IL72 shows a color difference between the four groups' initial and advanced weights. The relationship between initial weight and last week's weight is strong and positive in male controls, but weaker in mutants. Mutant females have a worse relationship than normal females. Line IL6009 mutant females show a strong and positive relationship throughout the experiment, beginning with week 6 weights, which correlate strongly with all other weeks. Week 10's correlation is lowest. The male control shows a strong positive correlation throughout the experiment. Week 10 weakens and strengthens it. Males with a mutation have a negative correlation between final weight at week 48 and weight from weeks 6 to 32.

4. Discussion

Although BW is a reliable, valid and easy-to-measure metric, its short-term dynamics are poorly understood because it is difficult and time-consuming to make frequent longitudinal measurements and previous studies estimating BW typically used infrequent measurements (e.g., every 6-12 months). Few studies have examined weekly or seasonal patterns in BW, although no study to our knowledge has estimated BW over the long term in setting of tumor suppressor knockout in mice model. Our results showed that BW varies significantly between CC strains and within the line depending on whether *Smad4* is knocked out, demonstrating that CC mice is an excellent platform for studying BW interactions and genotype. Additional lines with and without a mutation must be weighed for phenotype-genotype association analyses. Once acquired, such information can be used to forecast weight gain risk and build genetically based preventative and treatment techniques.

Animal welfare is routinely request of monitored using BW [40]. Typically, a loss of 20 percent of BW is used as the default endpoint [41]. However, these can be difficult to determine precisely in cancer models since an increase in tumor mass can mask the loss of overall BW [42]. Unexpected weight loss is frequently one of the earliest detectable signs of cancer. As the disease progresses, wasting and muscle loss, also known as cachexia [43], can worsen and contribute to death. On the other side, overweight-related cancers are becoming more common, making this disease spectrum a

public health priority. Obesity is a primary predictor of cancer's rising incidence and prevalence, and it may soon overtake smoking as the leading avoidable cause of cancer [6].

The heatmaps emphasize the differences between lines. There were different correlations between initial BW and final BW between lines and within lines regarding sex, and this is exacerbated in the presence of mutant tumor suppressors. Weight gain differs in populations with different genetic backgrounds as observed previously by Lone et al. [44]. Some lines gained more weight in the presence of Smad4 knockout, while others gained less. It also depended on the gender of the mice, but in general, the presence of the mutation resulted in overweight mice. The finding that null alleles of mouse genes frequently reduce but occasionally increase BW is not unique [45]. To our knowledge, no previous research has examined the overall effect of gene deletion on mice's BW because knockout studies are often done to answer specific gene research questions. However, Daniel et al. scanned the Mouse Genome Database for each knockout gene on six chromosomes to objectively analyze the effect of gene knockouts on BW [46]. They estimate that more than 6,000 genes contribute to mouse body size, and 30% of null alleles result in a mouse with lower BW, while 3% result in a mouse with higher BW [47]. The genome favors weight gain with ten times as many genes increasing size as decreasing. This supports the theory that mice are "hard-wired" to favor positive energy balance [48].

In this research, we carried out a comparative analysis of various ML algorithms for predicting the genotype of mice. It was shown that ML algorithm Logistic Regression provided better results than common regression algorithms for prediction. A successful classification of a genotype based solely on weight has undeniable clinical utility. Although the cost of sequencing has decreased significantly over time, it is still not a common practice. Line IL6009 exhibited the highest level of predictability, with an AUC > 0.75 as predicted by four different models. As demonstrated by the correlation matrices, this line's behavior differed between the two genotypes. In line IL6012, for instance, the heatmaps are extremely similar, and the classification results are also lower. The results clearly show differences between lines. In other words, an individual's entire genotype influences the predictability of this Smad4 knockout gene. The interactions of this protein with other proteins within the cell could be a possible explanation for this phenomenon.

To predict the BW of mice, the global outcomes are anticipated; the more information we provide the algorithm, the more accurately it predicts. However, even when we only give the algorithm week 8, the results are quite good. The line-separated results are quite unexpected. While some lines, such as IL5008 and IL2513, are highly predictable from week 8 or even week 6, others, such as IL188, are entirely unpredictable. In addition, we observe that adding more data to the algorithm hinders its performance for some lines, including lines IL5000 and IL6009. The outcomes clearly vary between lines. We believe that these variations are due to the different genotypes of the lines, as the behavior varies based on genotype. Some behaviors are more predictable than others, particularly when a linear regression model is used. From what we have seen, for most lines it seems like there are 2 linear behaviors with a specific changing point slope-wise. In the future, it could be interesting to try to predict the time of change in the slope based on the initial features alone. Studies in mouse strains with various genetic backgrounds and established genetic structures provide unparalleled opportunities to find and analyze trait-related genomic loci. Additional lines with and without a mutation must be weighed for phenotype-genotype association analyses. Once acquired, such information can be used to forecast weight gain risk and build genetically based preventative and treatment techniques.

Overweight and obesity not just increase the risk of CVD and T2D, but also of many different types of cancers [49]. Even though it is known that BW is a variable, efforts to stop the obesity epidemic have failed [50]. The global burden of cancer has increased in tandem with the prevalence of obesity, and 13 types of cancer have been linked to fat [51]. Breast (postmenopausal), colon-rectum, endometrium, ovary, pancreas, kidney, gallbladder, gastric cardia, liver, oesophagus (adenocarcinoma), meningioma, thyroid, and multiple myeloma are the cancers associated with overweight or obesity. However, there is a prominent research gap between weight gain and certain cancers linked to obesity. Consequently, there is a pressing need to monitor these mice and their

weight at older ages and to determine if there is a correlation between the percentage of weight gain in mutant mice compared to the control group and cancer susceptibility of various types.

In the future, it would be interesting to use the advantages of Smad4 knockout CC mice to investigate additional phenotypes of MetS, such as insulin resistance, diabetes, and lipid metabolism, in addition to cancer susceptibility. Understanding how genetic background affects metabolic profiles in a Smad4 knockout setting is essential for precision medicine and individualized pharmaceutical therapies.

Author Contributions: Conceptualization, F.A.I., A.N., and I.A.N.; Methodology, N.Q., I.M.L., A.K., A.H., O.Z. and N.B.N.; Validation, F.A.I.; Investigation, I.M.L., O.Z., K.M. and N.W.; Resources, F.A.I., P.P., and N.W.; Data Curation, N.Q., I.M.L., A.K., A.H., O.Z. and N.B.; Writing—Original Draft Preparation, N.Q., O.Z. and I.M.L.; Writing—Review and Editing, I.M.L., and F.A.I.; Supervision, F.A.I. and A.N.; Project Administration, F.A.I.; Funding Acquisition, F.A.I., and I.A.B.; All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by a core fund from Tel Aviv University, and the Department of Oral and Maxillofacial Surgery, Baruch Padeh Medical Center, Poriya, Israel

Institutional Review Board Statement: all animal experiments in this study were compliant with national standards for the care and use of laboratory animals, and the experiment was reviewed and approved by Tel Aviv University's Institutional Animal Care and Use Committee (IACUC), with an approved number (01-19-044).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: This study did not report any data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Safaei, M.; Sundararajan, E.A.; Driss, M.; Boulila, W.; Shapi'i, A. A systematic literature review on obesity: Understanding the causes & consequences of obesity and reviewing various machine learning approaches used to predict obesity. *Computers in biology and medicine*, **2021**, 136, 104754.
2. El-Sayed Moustafa, J.S.; Froguel, P. From obesity genetics to the future of personalized obesity therapy. *Nature Reviews Endocrinology*, **2013**, 9(7), 402-413.
3. Lone, I.M.; Midlej, K.; Ben Nun, N.; Iraqi, F.A. Intestinal Cancer Development in Response to Oral Infection with High-Fat Diet-Induced Type 2 Diabetes (T2D) in Collaborative Cross Mice under Different Host Genetic Background Effects. *Mamm. Genome*, **2023**, 34, 56–75. [CrossRef]
4. Ghnaim, A.; Lone, I.M.; Ben Nun, N.; Iraqi, F.A. Unraveling the Host Genetic Background Effect on Internal Organ Weight Influenced by Obesity and Diabetes Using Collaborative Cross Mice. *Int. J. Mol. Sci.* **2023**, 24, 8201. [CrossRef]
5. Da Luz, F.Q.; Hay, P.; Touyz, S.; Sainsbury, A. Obesity with comorbid eating disorders: associated health risks and treatment approaches. *Nutrients*, **2018**, 10(7), 829.
6. Ligibel, J.A.; Alfano, C.M.; Courneya, K.S.; Demark-Wahnefried, W.; Burger, R.A.; Chlebowski, R.T.; Fabian, C.J.; Gucalp, A.; Hershman, D.L.; Hudson, M.M.; Jones, L.W. American Society of Clinical Oncology position statement on obesity and cancer. *Journal of clinical oncology*, **2014**, 32(31), 3568.
7. Bhaskaran, K.; Douglas, I.; Forbes, H.; dos-Santos-Silva, I.; Leon, D.A.; Smeeth, L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5·24 million UK adults. *The Lancet*, **2014**, 384(9945), 755-765.
8. Yehia, R.; Lone, I.M.; Yehia, I.; Iraqi, F.A. Studying the Pharmagenomic Effect of Portulaca Oleracea Extract on Anti-Diabetic Therapy Using the Collaborative Cross Mice. *Phytomed. Plus*, **2023**, 3, 100394. [CrossRef]
9. Lone, I.M.; Iraqi, F.A. Genetics of Murine Type 2 Diabetes and Comorbidities. *Mamm. Genome*, **2022**, 33, 421–436. [CrossRef][PubMed]
10. Kamboh, M.I.; Demirci, F.Y.; Wang, X.; et al. Genome-wide association study of Alzheimer's disease. *Translational psychiatry*, **2012**, 2(5), e117-e117.
11. Thorleifsson, G.; Walters, G.B.; Gudbjartsson, D.F.; et al. Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nature genetics*, **2009**, 41(1), 18-24.

12. Kelishadi, R.; Poursafa, P. A review on the genetic, environmental, and lifestyle aspects of the early-life origins of cardiovascular disease. *Current problems in pediatric and adolescent health care*, **2014**, 44(3), 54-72.
13. Burza, M.A.; Spagnuolo, R.; Montalcini, T.; Doldo, P.; Pujia, A.; Romeo, S. Effect of excess body weight on the genetic susceptibility to cancer. *Journal of Clinical Gastroenterology*, **2014**, 48, S78-S79.
14. [14] Lone, I.M.; Zohud, O.; Nashef, A.; Kirschneck, C.; Proff, P.; Watted, N.; Iraqi, F.A. Dissecting the Complexity of Skeletal-Malocclusion-Associated Phenotypes: Mouse for the Rescue. *Int. J. Mol. Sci.* **2023**, 24, 2570. [CrossRef]
15. Hahn, S.A.; Schutte, M.; Shamsul Hoque, A.T.M.; et al. DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science*, **1996**, 271(5247), 350-353.
16. Takaku, K.; Oshima, M.; Miyoshi, H.; Matsui, M.; Seldin, M.F.; Taketo, M.M. Intestinal tumorigenesis in compound mutant mice of both Dpc4 (Smad4) and Apc genes. *Cell*, **1998** 92(5), 645-656.
17. Kitamura, T.; Kometani, K.; Hashida, H.; et al. SMAD4-deficient intestinal tumors recruit CCR1+ myeloid cells that promote invasion. *Nature genetics*, **2007**, 39(4), 467-475.
18. [18] Zohud, O.; Lone, I.M.; Midlej, K.; Obaida, A.; Masarwa S.; Schröder A.; Küchler EC.; Nashef A.; Kassem F.; Reiser V.; Chaushu G.; Mott, R.; Krohn, S.; Kirschneck, C.; Proff, P.; Watted, N.; Iraqi, F.A. Towards Genetic Dissection of Skeletal Class III Malocclusion: A Review of Genetic Variations Underlying the Phenotype in Humans and Future Directions. *J Clin Med.* 2023 Apr 29;12(9):3212. doi: 10.3390/jcm12093212. PMID: 37176653; PMCID: PMC10179046.
19. Izeradjene, K.; Combs, C.; Best, M.; et al. KrasG12D and Smad4/Dpc4 haploinsufficiency cooperate to induce mucinous cystic neoplasms and invasive adenocarcinoma of the pancreas. *Cancer cell*, **2007**, 11(3), 229-243.
20. Bardeesy, N.; et al. "Smad4 is dispensable for normal pancreas development yet critical in progression and tumor biology of pancreas cancer." *Genes & development* **2006**, 3130-3146.
21. Teng, Y.; Sun, A.N.; Pan, X.C.; Yang, G.; Yang, L.L.; Wang, M.R.; Yang, X. Synergistic function of Smad4 and PTEN in suppressing forestomach squamous cell carcinoma in the mouse. *Cancer research*, **2006**, 66(14), 6972-6981.
22. Xu, X.; Kobayashi, S.; Qiao, W.; Li, C.; Xiao, C.; Radaeva, S.; Stiles, B.; Wang, R.H.; Ohara, N.; Yoshino, T.; LeRoith, D. Induction of intrahepatic cholangiocellular carcinoma by liver-specific disruption of Smad4 and Pten in mice. *The Journal of clinical investigation* **2006**, 116(7), pp.1843-1852. doi: 10.1172/JCI27282.
23. Pühringer-Oppermann, F.; Sarbia, M.; Ott, N.; Brücher, B.L. The predictive value of genes of the TGF- β 1 pathway in multimodally treated squamous cell carcinoma of the esophagus. *Int J Colorectal Dis.* **2010**, 25, 515-521. doi: 10.1007/s00384-009-0867-z.
24. Zhu, Y.; Richardson, J.A.; Parada, L.F.; Graff, J.M. Smad3 mutant mice develop metastatic colorectal cancer. *Cell*, **1998**, 94(6), 703-714.
25. Tan, C.K.; Leuenberger, N.; Tan, M.J.; Yan, Y.W.; Chen, Y.; Kambadur, R.; Wahli, W.; Tan, N.S. Smad3 deficiency in mice protects against insulin resistance and obesity induced by a high-fat diet. *Diabetes*, **2011**, 60(2), 464-476.
26. Yadav, H.; Quijano, C.; Kamaraju, A.K.; Gavrilova, O. et al. Protection from obesity and diabetes by blockade of TGF- β /Smad3 signaling," *Cell Metab*, **2011**, 14(1), 67-79. doi: 10.1016/j.cmet.2011.04.013.
27. Li, J.; Sun, Y.B.Y.; Chen, W.; Fan, J.; Li, S.; Qu, X.; Chen, Q.; Chen, R.; Zhu, D.; Zhang, J.; Wu, Z. Smad4 promotes diabetic nephropathy by modulating glycolysis and OXPHOS. *EMBO Rep.* **2020**, 21(2), e48781. doi: 10.15252/embr.201948781.
28. Churchill, G.A.; Airey, D.C.; Allayee, H.; et al. The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nat Genet.* **2004**, 36(11):1133-1137. doi:10.1038/ng1104-1133
29. Aylor, D.L.; Valdar, W.; Foulds-Mathes, W.; et al. Genetic analysis of complex traits in the emerging Collaborative Cross. *Genome Res.* **2011**, 21(8):1213-1222. doi:10.1101/gr.111310.110
30. Chesler, E.J.; Miller, D.R.; Branstetter, L.R.; et al. The Collaborative Cross at Oak Ridge National Laboratory: developing a powerful resource for systems genetics. *Mamm Genome.* **2008**, 19(6):382-389. doi:10.1007/s00335-008-9135-8
31. Abu Toamih Atamni, H.; Nashef, A.; Iraqi F.A. The Collaborative Cross mouse model for dissecting genetic susceptibility to infectious diseases. *Mammalian Genome.* **2018**, 29:471-87.
32. Durrant, C.; Tayem, H.; Yalcin, B.; Cleak, J.; Goodstadt, L.; de Villena, F.P.; Mott, R.; Iraqi, F.A. Collaborative Cross mice and their power to map host susceptibility to *Aspergillus fumigatus* infection. *Genome research.* **2011**, 21(8):1239-48.

33. Atamni, H.J.; Mott, R.; Soller, M.; Iraqi, F.A. High-fat-diet induced development of increased fasting glucose levels and impaired response to intraperitoneal glucose challenge in the collaborative cross mouse genetic reference population. *BMC genetics*. **2016**, 17:1-9.
34. [34] Lone, I.M.; Zohud, O.; Midlej, K.; Proff, P.; Watted, N.; Iraqi, F.A. Skeletal Class II Malocclusion: From Clinical Treatment Strategies to the Roadmap in Identifying the Genetic Bases of Development in Humans with the Support of the Collaborative Cross Mouse Population. *J Clin Med*. 2023,12(15):5148. doi: 10.3390/jcm12155148. PMID: 37568550; PMCID: PMC10420085.
35. Keane, T.M.; Goodstadt, L.; Danecek, P.; White, M.A.; Wong, K.; Yalcin, B.; Heger, A.; Agam, A.; Slater, G.; Goodson, M.; Furlotte, N.A. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature*. **2011**, 477(7364):289-94.
36. Edgar, T.W.; Manz, D.O. Research methods for cyber security. *Syngress*. **2017**.
37. Nashef, A.; Qahaz, N.; El-Naaj, I.A.; Iraqi F.A. Systems genetics analysis of oral squamous cell carcinoma susceptibility using the mouse model: current position and new perspective. *Mammalian Genome*. **2021**;32(5):323-31..
38. Truett, G.E.; Heeger, P.; Mynatt, R.L.; Truett, A.A.; Walker, J.A.; Warman, M.L. Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). *Biotechniques*. **2000**, 29(1):52-4.
39. Soofi, A.A.; Awan, A. Classification techniques in machine learning: applications and issues. *Journal of Basic & Applied Sciences*. **2017**, 13(1):459-65.
40. Zou, Q.; Qu, K.; Luo, Y.; Yin, D.; Ju, Y.; Tang, H. Predicting diabetes mellitus with machine learning techniques. *Frontiers in genetics*. **2018**, 9:515.
41. D.B. Morton and P. H. Griffiths. "Guidelines on the recognition of pain, distress and discomfort in experimental animals and an hypothesis for assessment.," *Vet Rec*, vol. 116, no. 16, 1985, doi: 10.1136/vr.116.16.431.
42. D. B. Morton. "A systematic approach for establishing humane endpoints," *ILAR J*, vol. 41, no. 2, **2000**, doi: 10.1093/ilar.41.2.80.
43. Paster, E.V.; Villines, K.A.; Hickman, D.L. Endpoints for mouse abdominal tumor models: refinement of current criteria. *Comparative medicine*. **2009**,59(3):234-41.
44. Argilés, J.M.; Busquets, S.; Stemmler, B.; López-Soriano, F.J. Cancer cachexia: understanding the molecular basis. *Nature Reviews Cancer*. **2014**,14(11):754-62.
45. Snell, G.D. Inheritance in the house mouse, the linkage relations of short-ear, hairless, and naked. *Genetics*. **1931**,16(1):42.
46. Reed, D.R.; Lawler, M.P.; Tordoff, M.G. Reduced body weight is a common effect of gene knockout in mice. *BMC genetics*. **2008**, 9:1-6.
47. Schwartz, M.W.; Woods S.C.; Seeley, R.J.; Barsh G.S.; Baskin D.G.; Leibel R.L. Is the energy homeostasis system inherently biased toward weight gain? *Diabetes*. **2003**, 52(2):232-8.
48. Van Kruijsdijk, R.C.; Van Der Wall, E.; Visseren, F.L. Obesity and cancer: the role of dysfunctional adipose tissue. *Cancer epidemiology, biomarkers & prevention*. **2009**, 18(10):2569-78.
49. Lone, I.M.; Ben Nun, N.; Ghnaim, A.; Schaefer, A.S.; Hourri-Haddad, Y.; Iraqi, F.A. High-Fat Diet and Oral Infection Induced Type 2 Diabetes and Obesity Development under Different Genetic Backgrounds. *Anim. Model. Exp. Med*. **2023**, 6, 131–145. [CrossRef]
50. da Silva, M.; Weiderpass, E.; Licaj, I.; Lissner, L.; Rylander, C. Excess body weight, weight gain and obesity-related cancer risk in women in Norway: the Norwegian Women and Cancer study. *British journal of cancer*. **2018**, 119(5):646-56.
51. Global Burden of Disease Cancer Collaboration. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study. *JAMA oncology*. **2017**, 3(4):524.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.