



Uniwersytet Przyrodniczy w Poznaniu
Wydział Rolnictwa, Ogrodnictwa i Bioinżynierii

Adrian Cyplik

**Metody estymacji efektów interakcji wyższych rzędów genów
determinujących cechy ilościowe na podstawie obserwacji
fenotypowych i markerów molekularnych**

Methods of estimation the higher rank interactions of genes determining quantitative
traits based on phenotypic observations and molecular markers

Rozprawa doktorska w dziedzinie nauk rolniczych

w dyscyplinie rolnictwo i ogrodnictwo

Doctoral thesis in agricultural sciences

in the agriculture and horticulture discipline

Promotor:

Prof. UPP dr hab. Jan Bocianowski

Katedra Metod Matematycznych
i Statystycznych

Poznań, 2023

Składam serdeczne podziękowania

*Promotorowi, za przekazaną wiedzę, poświęcony czas i udzielanie cennych rad
oraz za wsparcie i życzliwość na każdym etapie powstawania niniejszej pracy*

Współautorom publikacji, za ich wkład w przeprowadzone badania

*Współpracownikom Katedry Metod Matematycznych i Statystycznych,
za liczne wskazówki i okazaną pomoc*

Rodzinie, za nieocenione wsparcie, cierpliwość i każde dobre słowo.

**Wykaz prac naukowych wchodzących w skład cyklu publikacji
rozprawy doktorskiej:**

Publikacja I. Cyplik, A., Bocianowski, J. (2022). **Analytical and numerical comparisons of two methods of estimation of additive × additive × additive interaction of QTL effects.** Journal of Applied Genetics 63(2), 213–221. DOI:10.1007/s13353-021-00676-7

Impact Factor 2021: 2,653

Punktacja wg MEiN: 140

Publikacja II. Cyplik, A., Sobiech, A., Tomkowiak, A., Bocianowski, J. (2022). **Genetic Parameters for Selected Traits of Inbred Lines of Maize (*Zea mays* L.).** Applied Sciences 12(14), 6961. DOI:10.3390/app12146961

Impact Factor 2021: 2,838

Punktacja wg MEiN: 100

Publikacja III. Cyplik, A., Czyczyło-Mysza, I.M., Jankowicz-Cieslak, J., Bocianowski, J. (2023). **QTL×QTL×QTL Interaction Effects for Total Phenolic Content of Wheat Mapping Population of CSDH Lines under Drought Stress by Weighted Multiple Linear Regression.** Agriculture 13(4), 850. DOI:10.3390/agriculture13040850

Impact Factor 2022: 3,6

Punktacja wg MEiN: 140

Publikacja IV. Cyplik, A., Bocianowski, J. (2023). **A Comparison of Methods to Estimate Additive-by-Additive-by-Additive of QTL×QTL×QTL Interaction Effects by Monte Carlo Simulation Studies.** International Journal of Molecular Sciences 24(12), 10043. DOI:10.3390/ijms241210043

Impact Factor 2022: 5,6

Punktacja wg MEiN: 140

Sumaryczny IF (wg roku opublikowania): 14,691

Sumaryczna liczba punktów ministerialnych (wg roku opublikowania): 520

Spis treści

ROZDZIAŁ I Autoreferat rozprawy	6
1. Wykaz stosowanych skrótów.....	7
2. Streszczenie	9
3. Summary.....	10
4. Wprowadzenie	11
4.1. Hodowla roślin.....	11
4.1.1. Cele hodowli roślin.....	11
4.1.2. Metody hodowli roślin.....	12
4.2. Cechy ilościowe.....	14
4.3. Interakcje genetyczne	16
5. Hipoteza badawcza i cel pracy doktorskiej	19
6. Materiały i metody.....	20
6.1. Metody estymacji.....	20
6.1.1. Oszacowanie na podstawie fenotypu.....	20
6.1.2. Oszacowanie na podstawie obserwacji genotypowych	21
6.1.2.1. Selekcja markerów.....	21
6.1.2.2. Regresja nieważona	21
6.1.2.3. Regresja ważona	23
6.2. Zbiory danych.....	23
6.2.1. Zbiór I (Publikacja I)	23
6.2.2. Zbiór II (Publikacja I).....	24
6.2.3. Zbiór III (Publikacja II)	25
6.2.4. Zbiór IV (Publikacja III).....	25
6.3. Badania symulacyjne	26
6.4. Oprogramowanie	28
7. Wyniki i dyskusja	29
7.1. Wyniki.....	29
7.1.1. Publikacja I	29
7.1.1.1. Porównanie analityczne	29
7.1.1.2. Porównanie numeryczne.....	29
7.1.2. Publikacja II.....	37
7.1.3. Publikacja III	39

7.1.3.1.	Oszacowanie na podstawie fenotypu	39
7.1.3.2.	Oszacowanie na podstawie genotypu	40
7.1.3.3.	Oszacowanie na podstawie genotypu - regresja nieważona	40
7.1.3.4.	Oszacowanie na podstawie genotypu - regresja ważona	43
7.1.4.	Publikacja IV	44
7.1.4.1.	Oszacowania na podstawie fenotypu	49
7.1.4.2.	Oszacowania na podstawie genotypu	49
7.1.4.3.	Oszacowania fenotypowe a genotypowe	50
7.1.4.4.	Błędy średniokwadratowe.....	50
7.1.4.5.	Współczynniki determinacji	51
7.2.	Dyskusja	52
8.	Podsumowanie	60
9.	Literatura.....	61
ROZDZIAŁ II Publikacje wchodzące w skład rozprawy		77
1.	Analytical and numerical comparisons of two methods of estimation of additive × additive × additive interaction of QTL effects.....	78
2.	Genetic Parameters for Selected Traits of Inbred Lines of Maize (<i>Zea mays</i> L.).	88
3.	QTL×QTL×QTL Interaction Effects for Total Phenolic Content of Wheat Mapping Population of CSDH Lines under Drought Stress by Weighted Multiple Linear Regression	96
4.	A Comparison of Methods to Estimate Additive–by–Additive–by–Additive of QTL×QTL×QTL Interaction Effects by Monte Carlo Simulation Studies	108
ROZDZIAŁ III Oświadczenia autorów.....		125

ROZDZIAŁ I

Autoreferat rozprawy

1. Wykaz stosowanych skrótów

AA – (ang. *alpha amylase*) α -amylaza

aa – interakcje addytywno×addytywne (epistatyczne)

aaa – interakcje addytywno×addytywno×addytywne

AIC – (ang. *Akaike Information Criterion*) kryterium informacyjne Akaike

cM – Centymorgan

CSDH – (ang. *Chinese SpringSQ1 doubled haploid*) linie podwojonych haploidów pszenicy *Chinese SpringSQ1*

DH – (ang. *doubled haploid*) podwojone haploidy

DP – (ang. *diastatic power*) siła diastatyczna

FWC – (ang. *field water capacity*) połowa pojemność wodna

GP – (ang. *grain protein*) zawartość białka w ziarnie

GS – (ang. *genomic selection*) selekcja genomowa

GWAS – (ang. *genome-wide association study*) badania asocjacyjne całego genomu

GY – (ang. *grain yield*) plon ziarna

H – (ang. *height*) wysokość rośliny

HD – (ang. *heading date*) termin kłoszenia

KW – (ang. *1000 kernel weight*) masa 1000 ziaren

L – (ang. *lodging*) wyleganie

LG – (ang. *Linkage Group*) grupa sprzężeń

MABC – (ang. *marker-assisted backcrossing*) krzyżowanie wsteczne wspomagane markerami

MAS – (ang. *marker assisted selection*) selekcja wspomagana markerami molekularnymi

ME – (ang. *malt extract*) ekstrakt słodu

NABGMP – (ang. *North American Barley Genome Mapping Project*)
Północnoamerykański Projekt Badania Genomu Jęczmienia

NH – (ang. *the number of days from planting until emergence of 50% of heads on main tillers*) termin kwitnienia 50% roślin

NM – (ang. *the number of days from planting until physiological maturity*) liczba dni od
zasadzenia do osiągnięcia dojrzałości fizjologicznej

QTL – (ang. *quantitative trait locus*) locus cechy ilościowej

RFLP – (ang. *restriction fragments length polymorphism*) polimorfizm długości
fragmentów restrykcyjnych

RI – (ang. *Recombinant Inbred*) linie wsobne

SNP – (ang. *Single-Nucleotide Polimorphism*) polimorfizm pojedynczego nukleotydu

TW – (ang. *test weight*) test wagowy

WG – (ang. *weight of grain harvested per unit area*) masa zebranego ziarna na
jednostkę powierzchni

2. Streszczenie

Celem procesu hodowlanego jest uzyskanie nowych genotypów o cechach ulepszonych w stosunku do form rodzicielskich. Parametry związane z efektem addytywnego działania genów, a także ich interakcjami (takimi jak efekt interakcji addytywno \times addytywnej i efekt interakcji addytywno \times addytywno \times addytywnej) mogą wpływać na decyzje dotyczące przydatności materiału hodowlanego. Zrozumienie architektury genetycznej złożonych cech jest głównym wyzwaniem w erze postgenomicznej, szczególnie w przypadku efektów locus cechy ilościowej (QTL), interakcji QTL \times QTL i interakcji QTL \times QTL \times QTL.

Prezentowana rozprawa miała na celu porównanie metod estymacji efektów interakcji wyższych rzędów genów determinujących cechy ilościowe z wykorzystaniem obserwacji fenotypowych i markerów molekularnych. Zaprezentowano dwie metody estymacji: metodę fenotypową oraz genotypową. W metodzie genotypowej wykorzystano regresję wielokrotną nieważoną oraz regresję wielokrotną ważoną w dwóch wariantach estymacji parametrów. Przedstawione metody estymacji całkowitego efektu interakcji *aaa* porównano analitycznie, numerycznie oraz symulacyjnie.

Materiał badawczy stanowiły cztery zbiory danych składające się ze: 150 linii podwojonych haploidów jęczmienia testowanych w szesnastu środowiskach; 145 linii podwojonych haploidów jęczmienia testowanych w dziewięciu środowiskach; 252 linii wsobnych kukurydzy badanych w dwóch środowiskach; 94 linii podwojonych haploidów pszenicy testowanych w jednym środowisku. W badaniach symulacyjnych poddano analizie 84 sytuacje eksperymentalne.

Wszystkie przedstawione metody okazały się użytecznymi narzędziami statystycznymi do charakterystyki QTL i oszacowania interakcji *aaa*. W porównaniu do metody fenotypowej, spośród metod genotypowych najdokładniejsze oceny uzyskano stosując ważoną regresję wielokrotną. W badaniach symulacyjnych, regresja ważona również pozwoliła uzyskać wyniki najbliższe wartości prawdziwej. Wyniki przedstawione w omówionych publikacjach stanowią pierwsze doniesienia dotyczące wykorzystania metod analitycznych, numerycznych i symulacyjnych estymacji interakcji *aaa* efektów QTL \times QTL \times QTL.

Słowa kluczowe: interakcje genetyczne wyższych rzędów, locus cechy ilościowej, symulacje Monte Carlo, analiza regresji, markery molekularne

3. Summary

The breeding process aims to obtain new genotypes with traits improved over the parental forms. Parameters related to the additive effect of genes, as well as their interactions (such as the additive \times additive interaction effect and the additive \times additive \times additive interaction effect), can influence decisions on the suitability of breeding material for this purpose. Understanding the genetic architecture of complex traits is a major challenge in the post-genomic era, especially for quantitative trait locus (QTL) effects, QTL \times QTL interactions and QTL \times QTL \times QTL interactions.

The presented dissertation aimed to compare methods of estimating the higher rank interactions of genes determining quantitative traits based on phenotypic observations and molecular markers. Two estimation methods were presented: the phenotypic method and the genotypic method. The genotypic method used unweighted multiple regression and weighted multiple regression in two variants of parameters selection. The presented methods for estimating the total interaction effect of *aaa* were compared analytically, numerically and in simulation.

The study material consisted of four data sets consisting of: 150 barley doubled haploid lines tested in sixteen environments; 145 barley doubled haploid lines tested in nine environments; 252 maize inbred lines tested in two environments; and 94 wheat doubled haploid lines tested in one environment. In the simulation study, 84 experimental situations were analyzed.

All of the methods presented proved to be useful statistical tools for QTL characterization and estimation of the *aaa* interaction. Compared to the phenotypic method, weighted multiple regression yielded the most accurate estimates among the genotypic methods. In simulation studies, weighted regression also produced results closest to the true value. The results presented in the papers discussed above are the first reports on the use of analytical, numerical and simulation methods for estimating *aaa* interactions of QTL \times QTL \times QTL effects.

Keywords: higher-order genetic interactions, quantitative trait locus, Monte Carlo simulation study, regression analysis, molecular markers

4. Wprowadzenie

4.1. Hodowla roślin

Hodowla roślin jest podstawową praktyką w rolnictwie i ogrodnictwie, mającą na celu opracowanie nowych i ulepszonych odmian roślin o pożądanym cechach. Poprzez staranną selekcję i kontrolowane krzyżowanie roślin rodzicielskich, hodowcy dążą do poprawy takich cech, jak: wielkość plonu, odporność na choroby, tolerancja na stresy środowiskowe i zawartość składników odżywczych. Proces ten zazwyczaj obejmuje wiele pokoleń roślin, aby zapewnić stabilność i jednolitość pożądanym cech. Tradycyjne metody hodowli roślin zostały uzupełnione o nowoczesne narzędzia biotechnologiczne (Moose, Mumm, 2008), takie jak inżynieria genetyczna (Nogué, Mara, Collonnier, Casacuberta, 2016) i selekcja wspomagana markerami (Fujino, Hirayama, Kaji, 2019), które przyspieszają proces hodowli. Hodowla roślin odgrywa kluczową rolę w podejmowaniu wyzwań związanych z bezpieczeństwem żywnościowym, dostosowywaniu upraw do zmieniającego się klimatu i promowaniu zrównoważonego rolnictwa w celu zaspokojenia potrzeb rosnącej światowej populacji.

4.1.1. Cele hodowli roślin

Cele hodowli roślin obejmują szeroki zakres działań prowadzących do poprawy wydajności, zdolności adaptacyjnych i zrównoważonego rozwoju upraw. Cele te różnią się w zależności od konkretnych potrzeb rolników, konsumentów i środowiska.

Jednym z głównych celów hodowli roślin jest zwiększenie wydajności upraw. Wybierając i hodując rośliny o wyższym potencjale plonowania, hodowcy przyczyniają się do zwiększenia produkcji żywności i zapewnienia globalnego bezpieczeństwa żywnościowego. Rozwój upraw o zwiększonej odporności na patogeny, wirusy, grzyby oraz inne szkodliwe mikroorganizmy pomaga zmniejszyć straty w uprawach spowodowane chorobami (Hussain, 2015). Przynosi to nie tylko korzyści ekonomiczne rolnikom, ale także minimalizuje zapotrzebowanie na chemiczne pestycydy, promując przyjazne dla środowiska praktyki rolnicze. Odmiany, które mogą wytrzymać różne stresy środowiskowe, takie jak susza, ekstremalne temperatury, zasolenie i złe warunki glebowe, zapewniają stabilne plony nawet w trudnych warunkach klimatycznych, przyczyniając się do adaptacji do zmian klimatu (Ceccarelli i in., 2010). Poprawa zawartości składników odżywczych w uprawach ma zasadnicze znaczenie dla walki

z niedożywieniem i poprawy zdrowia publicznego (Clarke, Wiseman, 2000). Hodowcy koncentrują się na zwiększaniu w uprawach poziomu niezbędnych składników odżywczych, takich jak witaminy, minerały i przeciwutleniacze. Hodowla roślin ma na celu poprawę ogólnych cech jakościowych upraw, między innymi smaku, konsystencji, wyglądu i trwałości (Civille, Oftedal, 2012). Cechy te mają kluczowe znaczenie dla akceptacji konsumentów i popytu. Oprócz upraw spożywczych, hodowla roślin koncentruje się również na rozwoju upraw do celów niespożywczych, takich jak produkcja biopaliw (Tammisola, 2010) i materiałów przemysłowych (van Soest, Mastebroek, de Meijer, 1992). Zrozumienie złożonych interakcji między roślinami, szkodnikami i czynnikami środowiskowymi pozwala hodowcom rozwijać uprawy oraz skutecznie zarządzać tymi interakcjami, zapewniając zdrowsze i bardziej produktywne rośliny.

Aby osiągnąć te cele w hodowli roślin, stosowane są zarówno konwencjonalne metody hodowlane, jak i nowoczesne podejścia biotechnologiczne (Bradshaw, 2017). Ostatecznym celem jest stworzenie zrównoważonych, odpornych i wydajnych odmian roślin uprawnych, które mogą zaspokoić różnorodne i zmieniające się potrzeby rolnictwa i społeczeństwa (Borlaug, 1983).

4.1.2. Metody hodowli roślin

Metody hodowli roślin można ogólnie podzielić na dwie kategorie: konwencjonalne i genetyczne. Metody genetyczne, znane również jako hodowla molekularna lub podejścia biotechnologiczne, obejmują wykorzystanie markerów molekularnych, inżynierii genetycznej i genomiki w celu przyspieszenia procesu hodowli i wprowadzenia określonych cech do roślin. Oto kilka kluczowych metod genetycznych stosowanych w hodowli roślin:

1. Selekcja wspomagana markerami molekularnymi (MAS). MAS obejmuje wykorzystanie markerów molekularnych powiązanych z określonymi genami lub cechami będącymi przedmiotem zainteresowania. Analizując DNA roślin, hodowcy mogą zidentyfikować osobniki niosące pożądane geny bez konieczności obserwowania rzeczywistej cechy. Metoda ta przyspiesza proces selekcji i zwiększa precyzję programów hodowlanych (Cobb, Biswas, Platten, 2019).

2. Selekcja genomowa (GS). GS wykorzystuje wysokoprzepustowe genotypowanie i fenotypowanie do przewidywania wydajności roślin w oparciu o ich

cały genom. Uwzględnia zbiorowy wkład wszystkich genów w złożone cechy, umożliwiając hodowcom wybór najlepszych kombinacji alleli w celu bardziej efektywnego osiągnięcia pożądanych cech (Robertsen, Hjortshøj, Janss, 2019).

3. Rośliny transgeniczne (modyfikowane genetycznie). Inżynieria genetyczna obejmuje bezpośrednią manipulację DNA organizmu poprzez wstawianie określonych genów pochodzących od innych gatunków. W hodowli roślin metody transgeniczne mogą wprowadzać nowe cechy do upraw, takie jak odporność na choroby, tolerancja na herbicydy lub ulepszona zawartość składników odżywczych (Collinge, Sarrocco, 2022).

4. Edycja genów (CRISPR-Cas9). Technologie edycji genów, takie jak CRISPR-Cas9, umożliwiają precyzyjne modyfikacje DNA rośliny bez wprowadzania obcych genów. Technika ta jest wykorzystywana do tworzenia ukierunkowanych zmian w określonych genach, naśladując naturalną zmienność genetyczną lub zakłócając niepożądane cechy (Zhang, Neik, Thomas, Batley, 2023).

5. Mutageneza. Mutageneza polega na wywoływaniu losowych mutacji w DNA roślin za pomocą związków chemicznych, promieniowania lub innych metod. Badając powstałe mutacje, hodowcy mogą odkryć osobniki o pożądanych cechach, które nie występowały w pierwotnej populacji (Shelake, Pramanik, Kim, 2019).

6. Produkcja haploidów i podwojonych haploidów. Techniki te umożliwiają szybkie tworzenie homozygotycznych roślin (roślin o identycznych allelach) poprzez indukowanie haploidalnych zarodków, które są następnie podwajane w celu przywrócenia pełnego zestawu chromosomów. Linie homozygotyczne są cenne dla hodowli, ponieważ wykazują spójną ekspresję cech (Tuveesson, Larsson, Ordon, 2021; Weyen, 2021).

7. Krzyżowanie wsteczne wspomagane markerami (MABC). MABC łączy precyzję markerów molekularnych z konwencjonalnym krzyżowaniem wstecznym. Pozwala hodowcom na wprowadzenie określonego genu z rośliny dawcy do rośliny biorcy, jednocześnie minimalizując przeciąganie niepożądanych genów od dawcy (Kim i in., 2021).

8. Sekwencjonowanie genomu i bioinformatyka. Postępy w sekwencjonowaniu genomu i bioinformatyce umożliwiają identyfikację genów kandydujących związanych z pożądanymi cechami. Informacje te pomagają hodowcom w wyborze osobników

o najkorzystniejszym składzie genetycznym do ukierunkowanego doskonalenia (Salgotra, Stewart, 2020).

Wykorzystanie tych metod genetycznych sprawiło, że hodowla roślin stała się bardziej wydajna, precyzyjna i ukierunkowana, umożliwiając hodowcom opracowanie ulepszonych odmian roślin uprawnych o pożądanym cechach, aby sprostać wyzwaniom współczesnego rolnictwa i zapewnić zrównoważoną produkcję żywności. Kluczowe znaczenie ma jednak uwzględnienie aspektów etycznych, środowiskowych i regulacyjnych związanych z wykorzystaniem metod genetycznych w hodowli roślin.

4.2. Cechy ilościowe

Cechy ilościowe, znane również jako złożone, są dziedzicznymi cechami roślin i zwierząt, które wykazują ciągłą zmienność i są determinowane przez wiele genów i czynników środowiskowych. Cechy te mają ogromne znaczenie w rolnictwie i genetyce, ponieważ regulują m.in. plon, wysokość, masę czy odporność na choroby. W przeciwieństwie do cech dyskretnych (np. kolor kwiatów) kontrolowanych przez pojedynczy gen, cechy ilościowe determinowane są przez połączone efekty wielu loci genetycznych i interakcji środowiskowych. Każdy gen może mieć niewielki wpływ na daną cechę, co sprawia, że ich indywidualny wpływ jest trudny do zidentyfikowania (Mackay, 2014; Members of the Complex Trait Consortium, 2003). Ponadto czynniki zewnętrzne, takie jak temperatura, wilgotność, składniki odżywcze gleby i stosowane zabiegi agrotechniczne mogą również wpływać na ekspresję tych cech.

Metody statystyczne, takie jak mapowanie loci cech ilościowych (QTL) i badania asocjacyjne całego genomu (GWAS), są wykorzystywane do identyfikacji regionów genomowych odpowiedzialnych za te cechy (Khan i in., 2021). Mapowanie QTL obejmuje tworzenie map genetycznych i identyfikację regionów związanych z określonymi cechami poprzez analizę wzorców dziedziczenia w segregujących populacjach. Z drugiej strony GWAS wykorzystuje duże zestawy markerów genetycznych do wykrywania powiązań między określonymi markerami a zmiennością cech w populacji niespokrewnionych osobników. Metody te pomagają badaczom wskazać geny kandydujące, które przyczyniają się do zmienności obserwowanej w cechach ilościowych, zapewniając cenny wgląd w podstawową architekturę genetyczną.

Zrozumienie genetycznych podstaw cech ilościowych ma kluczowe znaczenie dla nowoczesnych wysiłków w zakresie hodowli roślin, ponieważ umożliwia hodowcom podejmowanie świadomych decyzji i opracowywanie ulepszonych odmian o zwiększonej wydajności i zdolności adaptacji do różnych warunków środowiskowych. Identyfikując czynniki genetyczne, które wpływają na cechy takie jak plon, zawartość składników odżywczych i tolerancja na stres, hodowcy mogą selektywnie krzyżować rośliny z korzystnymi odmianami w celu uzyskania potomstwa o pożądanym kombinacjach cech. Takie ukierunkowane podejście przyspiesza proces hodowli, prowadząc do opracowania nowych odmian roślin uprawnych, które wykazują lepszą wydajność i lepiej spełniają zmieniające się wymagania rolnictwa i produkcji żywności.

Co więcej, analiza cech ilościowych jest przydatna również w hodowli odmian o złożonych cechach o niskiej odziedziczalności (Viana, e Silva, Mundim, Azevedo, Jan, 2016). Dzięki dokładnemu zrozumieniu podstawowych mechanizmów genetycznych, hodowcy mogą stosować bardziej zaawansowane techniki, takie jak selekcja genomowa i hodowla wspomagana markerami, w celu poprawy dokładności i skuteczności selekcji roślin o pożądanym cechach.

Badanie cech ilościowych i zastosowanie narzędzi genetycznych zrewolucjonizowało dziedzinę hodowli roślin, umożliwiając hodowcom poruszanie się po złożoności dziedziczenia cech i znacząco przyczyniając się do zrównoważonego rolnictwa, globalnego bezpieczeństwa żywnościowego i rozwoju odpornych odmian roślin uprawnych zdolnych do wytrzymania stresów środowiskowych i stale zmieniających się krajobrazów rolniczych.

Większość ważnych cech roślin to cechy ilościowe, na które wpływa wiele genów i loci cech ilościowych (QTL). Wpływ tych genów na ekspresję cech ilościowych był rozważany u różnych gatunków, takich jak jęczmień (Sayed i in., 2021), kukurydza (Ren, Zhang, Li, Wu, 2022), rzepak (Bocianowski, Kozak, Liersch, Bartkowiak-Broda, 2011; Botero-Ramírez i in., 2020; Gacek i in., 2021), ryż (Kabange i in., 2020; Kwon i in., 2021), burak cukrowy (Pegot-Espagnet i in., 2019) i pszenica (Lephuthing, Khumalo, Tolmay, Dube, Tsilo, 2022). Wiele loci wykazujących skumulowane niewielkie efekty jest często eliminowanych w rutynowych analizach QTL. Jednak w interakcji z innymi loci mogą one znacząco wpływać na obserwowaną cechę

ilościową. W ostatnich latach wpływ epistazy (podwójnej interakcji) na obserwowaną cechę był rozważany i opisywany między innymi dla jęczmienia (Beheshtizadeh i in., 2018), kukurydzy (Ku i in., 2012), rzepaku (Yusuf, Richter, Möllers, 2022), grochu (Krajewski i in., 2012), ryżu (Ali i in., 2022), soi (Han i in., 2017), pomidora (Smeda i in., 2018), pszenżyta (Dhariwal i in., 2018) i pszenicy (Pundir i in., 2022). Opracowano kilka metod lokalizacji genów i szacowania efektu epistazy (Bocianowski, 2012c; Carlborg, Andersson, Kinghorn, 2000; Chase, Adler, Lark, 1997; Holland, 1998; Kao, Zeng, Teasdale, 1999; Sen, Churchill, 2001; Zeng, Kao, Basten, 1999).

4.3. Interakcje genetyczne

Interakcja genetyczna odnosi się do złożonego współdziałania między wieloma genami w genomie organizmu, wpływając na ekspresję cech i określając ogólny fenotyp. W tej skomplikowanej relacji czynników genetycznych wpływ jednego genu może być modyfikowany lub maskowany przez obecność innych genów. Skutkuje to połączonym wynikiem, który odbiega od tego, który każdy gen wytworzyłby niezależnie. Interakcje te mogą być synergistyczne, gdy łączny efekt jest większy niż suma indywidualnych efektów, lub antagonistyczne, gdy wpływ jednego genu przeciwdziała lub zmniejsza wpływ innego. Interakcje genetyczne odgrywają kluczową rolę w kształtowaniu różnorodności organizmów żywych, wpływając na ich podatność na choroby, reakcję na bodźce środowiskowe i trajektorie ewolucyjne. Badanie interakcji genetycznych jest niezbędne do zrozumienia złożoności genetyki i zapewnia cenny wgląd w podstawowe mechanizmy funkcjonowania organizmów (Hartman, Garvik, Hartwell, 2001; Manolio i in., 2009).

Istnieją różne rodzaje interakcji genetycznych:

1. Epistaza. Epistaza występuje, gdy jeden gen maskuje lub modyfikuje działanie innego genu. Pierwszy z tych genów jest epistatyczny, a drugi hipostatyczny. Ta interakcja może prowadzić do nieoczekiwanych i różnorodnych zmian fenotypowych. Na przykład, obecność jednego genu może tłumić ekspresję drugiego genu, prowadząc do tego, że określona cecha nie jest obserwowana (Li, Song, Zhao, Wang, Fu, 2014; Luo i in., 2009).

2. Interakcje synergistyczne. W tych interakcjach dwa lub więcej genów współpracuje ze sobą w celu zwiększenia lub wzmocnienia ekspresji określonej cechy. Połączony efekt jest większy niż to, co każdy gen wytworzyłby indywidualnie.

Interakcje synergiczne mogą prowadzić do powstania nowych i silnych fenotypów, które zapewniają selektywną przewagę w niektórych środowiskach (Elena, Lenski, 1997).

3. Interakcje antagonistyczne. Interakcje antagonistyczne występują, gdy obecność jednego genu utrudnia lub zmniejsza działanie innego genu. Ten rodzaj interakcji może prowadzić do kompromisów między różnymi cechami, ponieważ ekspresja jednej cechy może odbywać się kosztem innej (Machida, Nakagawa, Kojima, Takahashi, Machida, 2015).

4. Mutacje supresorowe. Mutacja supresorowa to zmiana genetyczna, która kompensuje szkodliwe skutki mutacji pierwotnej. Może ona zniwelować negatywny wpływ początkowej mutacji, przywracając poprawny fenotyp (Bautista, Carr, Mitchell, 2021).

5. Warunkowa letalność. Niektóre interakcje genów są warunkowe, co oznacza, że ich efekty zależą od określonych warunków środowiskowych. Określona kombinacja genów może być śmiertelna w jednym środowisku, ale nie wykazywać takiej własności w innym (Larrimore, Rancati, 2019).

Zrozumienie interakcji genetycznych ma kluczowe znaczenie dla różnych dziedzin biologii i może mieć znaczące implikacje w badaniach genetycznych, rolnictwie, medycynie i biologii ewolucyjnej. Pomaga naukowcom zrozumieć podstawowe mechanizmy odpowiedzialne za złożone cechy i choroby. Dodatkowo, badanie interakcji genetycznych może prowadzić do głębszego zrozumienia złożoności ewolucji i sposobu, w jaki organizmy dostosowują się do środowisk w czasie.

Interakcje genetyczne wyższego rzędu odnoszą się do złożonych relacji między trzema lub więcej genami w genomie organizmu (Bateson, Mendel, 1902). Podczas gdy pojęcie interakcji genetycznej zazwyczaj obejmuje pary genów, interakcje wyższego rzędu uwzględniają połączone efekty wielu genów. Interakcje te dodają kolejną warstwę złożoności do zrozumienia regulacji genetycznej i ekspresji fenotypowej.

W interakcjach genetycznych wyższego rzędu efekty wielu genów są ze sobą powiązane i nie można ich w pełni wyjaśnić, badając każdy gen z osobna. Wraz ze wzrostem liczby zaangażowanych genów, potencjalne wyniki fenotypowe stają się bardziej zróżnicowane i trudne do przewidzenia. Interakcje wyższego rzędu mogą

skutkować nieoczekiwanymi i nieaddytywnymi efektami, w których łączny wpływ nie jest jedynie sumą indywidualnego wkładu każdego genu (Taylor, Ehrenreich, 2015).

Zrozumienie interakcji genetycznych wyższego rzędu jest kluczowe z kilku powodów. Wiele cech biologicznych i podatności na choroby jest regulowanych przez połączone działanie wielu genów. Interakcje wyższego rzędu odgrywają znaczącą rolę w kształtowaniu ekspresji tych złożonych cech, co sprawia, że konieczne jest poznanie tych relacji w celu ich kompleksowego zrozumienia. Interakcje wyższego rzędu mogą również ujawnić redundancję genetyczną, w której wiele genów ma podobne funkcje, a utrata jednego genu może być kompensowana przez inne. I odwrotnie, mogą również ujawnić buforowanie genetyczne, w którym wpływ mutacji na konkretny gen jest tłumiony lub maskowany przez inne geny (Hartman, Garvik, Hartwell, 2001). W kontekście medycyny, interakcje wyższego rzędu są niezbędne do identyfikacji potencjalnych celów dla produkcji nowych leków. Ukierunkowanie na wiele genów jednocześnie może okazać się niezbędne do skutecznego leczenia złożonych chorób, na które wpływają te interakcje (Lahey, Van Hulle, Singh, Waldman i Rathouz, 2011). Interakcje wyższego rzędu mogą mieć także znaczący wpływ na zdolność organizmu do ewolucji i adaptacji do zmieniającego się środowiska. Zrozumienie, w jaki sposób interakcje genetyczne kształtują ewolucję populacji, jest konieczne do badania bioróżnorodności i specjacji (Weinreich, Lan, Wylie, Heckendorn, 2013).

Interakcje wyższego rzędu są często pomijane w badaniach, chociaż koncepcja interakcji genetycznych jest znana od ponad wieku (Bateson, Mendel, 1902). Wynika to z wymagań zaawansowanych technik eksperymentalnych i obliczeniowych (Hartman i in., 2001; Manolio i in., 2009). Ekran genetyczny (ang. *genetic screen*), sekwencjonowanie o wysokiej przepustowości i modelowanie obliczeniowe to tylko niektóre z podejść stosowanych do analizy złożonych interakcji między wieloma genami. Wraz z postępem badań genetycznych naukowcy odkrywają coraz bardziej skomplikowane i dynamiczne interakcje między genami. Odkrycia te nie tylko pogłębiają nasze zrozumienie genetycznych podstaw życia, ale także obiecują postęp w medycynie spersonalizowanej, rolnictwie i innych dziedzinach, w których manipulacje i interwencje genetyczne mają kluczowe znaczenie. Badając interakcje genetyczne wyższego rzędu, zyskujemy dokładniejsze zrozumienie złożoności i różnorodności żywych organizmów oraz ich podstawowych architektur genetycznych.

5. Hipoteza badawcza i cel pracy doktorskiej

Jak wykazano we wprowadzeniu, badanie interakcji genetycznych wyższego rzędu przyczynia się do poznania złożonych relacji genetycznych, które są często pomijane w tradycyjnych analizach. W kontekście hodowli roślin, spostrzeżenia te oferują możliwość znacznej poprawy upraw. W związku z powyższym hipoteza badawcza zakłada, że wykorzystując obserwacje fenotypowe oraz markery molekularne, można estymować efekty interakcji wyższych rzędów genów determinujących cechy ilościowe.

Badania obejmowały:

1. Opracowanie estymatorów parametru związanego z efektem potrójnej interakcji genów determinujących cechy ilościowe, na podstawie obserwacji tylko fenotypowych oraz z uwzględnieniem obserwacji markerów molekularnych.
2. Porównanie analityczne dwóch metod (fenotypowej i genotypowej) estymacji całkowitego efektu interakcji *aaa*.
3. Porównanie numeryczne dwóch metod (fenotypowej i genotypowej) estymacji całkowitego efektu interakcji *aaa*.
4. Wykorzystanie regresji wielokrotnej nieważonej z uwzględnieniem interakcji wyższego rzędu (*aaa*) ze stopniową selekcją cech według kryteriów informacyjnych Akaike (AIC) do analizy danych.
5. Wykorzystanie regresji wielokrotnej nieważonej z uwzględnieniem interakcji wyższego rzędu (*aaa*) ze stopniową selekcją cech według procedury regresji krokowej do analizy danych.
6. Wykorzystanie ważonej regresji wielokrotnej do poprawy estymacji parametru związanego z interakcją wyższego rzędu (*aaa*).
7. Porównanie symulacyjne dwóch metod (fenotypowej i trzech wariantów metody genotypowej) estymacji całkowitego efektu interakcji *aaa*.

Przeprowadzone badania miały na celu opracowanie i porównanie metod wykorzystujących obserwacje fenotypowe oraz markery molekularne do estymowania efektów interakcji wyższych rzędów genów, a także ich wpływu na cechy ilościowe. Cel pracy został zrealizowany przy użyciu danych zawierających obserwacje fenotypowe oraz markery molekularne z wykorzystaniem metod statystycznych.

6. Materiały i metody

6.1. Metody estymacji

6.1.1. Oszacowanie na podstawie fenotypu

Zakładając, że w doświadczeniu zaobserwowano n homozygotycznych linii (podwojonych haploidów, DH; wsobnych, RI) roślin, otrzymujemy n -wymiarowy wektor fenotypowych średnich obserwacji $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n]'$ i q n -wymiarowych wektorów obserwacji genotypu markera \mathbf{m}_l , $l=1, 2, \dots, q$. Element i -ty ($i=1, 2, \dots, n$) wektora \mathbf{m}_l jest równy -1 lub 1, w zależności od genotypu rodzica wykazywanego przez i -tą linię.

Oszacowanie efektu addytywno \times addytywno \times addytywnej interakcji (interakcji potrójnej, *aaa*) na podstawie obserwacji tylko fenotypowych \mathbf{y} wymaga identyfikacji grup linii ekstremalnych, tj. linii o minimalnej i maksymalnej ekspresji obserwowanej cechy (Choo, Reinbergs, 1982). Grupa linii minimalnych składa się z linii, które teoretycznie zawierają tylko allele zmniejszające wartość cechy. Analogicznie, grupa linii maksymalnych zawiera linie, które zawierają tylko allele zwiększające wartość cechy. W zaproponowanej metodzie identyfikujemy grupy linii ekstremalnych jako, odpowiednio, linie o wartościach minimalnych i maksymalnych empirycznego rozkładu średnich. Całkowity efekt potrójnej interakcji *aaa* można oszacować za pomocą wzoru (Publikacja I: Cyplik, Bocianowski, 2022):

$$\widehat{aaa}_p = \frac{1}{2}(\bar{L}_{max} + \bar{L}_{min}) - \bar{L}, \quad (1)$$

gdzie \bar{L}_{min} i \bar{L}_{max} oznaczają, odpowiednio, średnie dla grup linii minimalnych i maksymalnych, \bar{L} oznacza średnią dla wszystkich linii. Liczbę genów (liczbę czynników efektywnych) uzyskaną na podstawie obserwacji jedynie fenotypowych obliczono według wzoru przedstawionego przez Kaczmarek, Surma, Adamski (1988):

$$\widehat{K} = \frac{(\bar{L}_{max} - \bar{L}_{min})^2}{4V_L}, \quad (2)$$

gdzie V_L jest wariancją addytywną.

Statystyki testowe do weryfikacji hipotez o wartościach parametrów genetycznych różnych od zera są podane przez (Publikacja II: Cyplik i in., 2022):

$$F_{aaa} = \frac{MS_{aaa}}{MS_e}, \quad (3)$$

gdzie MS_{aaa} i MS_e są średnimi kwadratami dla, odpowiednio, parametrów aaa i błędu.

6.1.2. Oszacowanie na podstawie obserwacji genotypowych

6.1.2.1. Selekcja markerów

Wariant I (Publikacje I i IV)

Do selekcji markerów dla modelu (4) zastosowano stopniową selekcję według kryteriów informacyjnych Akaike (AIC) (Akaike, 1998). Procedura składała się z dwóch etapów: najpierw podzielono markery na grupy w oparciu o chromosomy, na których się znajdowały i przeprowadzono stopniową selekcję według AIC; następnie połączono wybrane markery w jedną grupę i powtórzono selekcję jak powyżej. Wszystkie wybrane markery zostały połączone w jedną grupę, a ostatnia selekcja została przeprowadzona na modelu uwzględniającym efekt interakcji aaa . Aby przeciwdziałać problemowi wielokrotnych porównań, zastosowano korektę Bonferroniego.

Wariant II (Publikacje III i IV)

Selekcji markerów wybranych do modelu (4) przeprowadzono za pomocą procedury regresji krokowej (Bocianowski, Krajewski, 2009). W tym przypadku zastosowano algorytm trzystopniowy: najpierw dokonano selekcji krokowej wstecznej niezależnie we wszystkich grupach sprzężeń; następnie wybrane w ten sposób markery umieszczono w jednej grupie i poddano drugiej selekcji wstecznej (Jansen, Stam, 1994); w trzecim etapie uwzględniono sytuacje, w których wybrane markery znajdowały się na chromosomie bardzo blisko siebie (bliżej niż 5 cM). Ponieważ markery te są prawdopodobnie powiązane z jednym QTL, tylko marker, o największej wartości statystyki testowej, został wybrany do modelu. Krytyczny poziom istotności wynoszący 0,001, uzyskany z poprawki Bonferroniego (Province, 2001), został użyty dla pierwszego i drugiego etapu.

6.1.2.2. Regresja nieważona

Oszacowanie aaa bazuje na założeniu, że geny determinujące cechę są ściśle powiązane z obserwowanymi markerami molekularnymi. Wybierając spośród

wszystkich obserwowanych markerów p , możemy wyjaśnić zmienność cechy i modelować obserwacje dla linii jako za pomocą modelu (Publikacja I: Cyplik, Bocianowski, 2022):

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{W}\boldsymbol{\delta} + \mathbf{e}, \quad (4)$$

gdzie $\mathbf{1}$ oznacza n -wymiarowy wektor jedynek, μ oznacza średnią ogólną, \mathbf{X} oznacza $(n \times p)$ -wymiarową macierz postaci $\mathbf{X} = [\mathbf{m}_{l_1} \ \mathbf{m}_{l_2} \ \cdots \ \mathbf{m}_{l_p}]$, $l_1, l_2, \dots, l_p \in \{1, 2, \dots, q\}$, $\boldsymbol{\beta}$ oznacza p -wymiarowy wektor nieznanych parametrów postaci $\boldsymbol{\beta}' = [a_{l_1} \ a_{l_2} \ \cdots \ a_{l_p}]$, \mathbf{Z} oznacza macierz, której kolumny są iloczynami niektórych kolumn macierzy \mathbf{X} , $\boldsymbol{\gamma}$ oznacza wektor nieznanych parametrów postaci $\boldsymbol{\gamma}' = [aa_{l_1l_2} \ aa_{l_1l_3} \ \cdots \ aa_{l_{p-1}l_p}]$, \mathbf{W} oznacza macierz, której kolumny są iloczynami trzech wybranych kolumn macierzy \mathbf{X} , $\boldsymbol{\delta}$ oznacza wektor nieznanych parametrów postaci $\boldsymbol{\delta}' = [aaa_{l_1l_2l_3} \ aaa_{l_1l_2l_4} \ \cdots \ aaa_{l_{p-2}l_{p-1}l_p}]$, \mathbf{e} oznacza n -wymiarowy wektor zmiennych losowych takich, że $E(e_i)=0$, $Cov(e_i, e_j)=0$, dla $i \neq j$, $i, j = 1, 2, \dots, n$. Parametry $a_{l_1}, a_{l_2}, \dots, a_{l_p}$ są addytywnymi efektami genów kontrolujących cechę, parametry $aa_{l_1l_2}, aa_{l_1l_3}, \dots, aa_{l_{p-1}l_p}$ są efektami interakcji addytywno \times addytywnej, a parametry $aaa_{l_1l_2l_3}, aaa_{l_1l_2l_4}, \dots, aaa_{l_{p-2}l_{p-1}l_p}$ są efektami interakcji addytywno \times addytywno \times addytywnej. Zakładamy, że efekty interakcji epistatycznych i potrójnych wykazują tylko loci z istotnymi addytywnymi efektami działania genów. Założenie to znacznie zmniejsza liczbę potencjalnych znaczących efektów i sprawia, że model regresji jest bardziej użyteczny.

Oznaczając przez $\boldsymbol{\alpha}' = [\mu \ \boldsymbol{\beta}' \ \boldsymbol{\gamma}' \ \boldsymbol{\delta}']$ i $\mathbf{G} = [\mathbf{1} \ \mathbf{X} \ \mathbf{Z} \ \mathbf{W}]$ otrzymujemy model w postaci macierzowej:

$$\mathbf{y} = \mathbf{G}\boldsymbol{\alpha} + \mathbf{e}. \quad (5)$$

Jeśli \mathbf{G} jest macierzą pełnego rzędu, to oszacowanie $\boldsymbol{\alpha}$ jest dane przez (Searle, 1982):

$$\hat{\boldsymbol{\alpha}} = (\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\mathbf{y}. \quad (6)$$

Całkowity efekt interakcji potrójnej aaa genów wpływających na cechę można estymować wg wzoru (Publikacja I: Cyplik, Bocianowski, 2022):

$$\hat{aaa}_g = \sum_{k=1}^{p-2} \sum_{\substack{k'=2 \\ k' \neq k}}^{p-1} \sum_{\substack{k''=3 \\ k'' \neq k, k'}}^p \hat{aaa}_{l_k l_{k'} l_{k''}}. \quad (7)$$

6.1.2.3. Regresja ważona

Zmodyfikowana wersja regresji danych markerowych na cechę ilościową jest rozważana poprzez zastosowanie ważonej wielokrotnej regresji liniowej. Macierz diagonalna \mathbf{W} nieznanych wariancji może być estymowana przez empiryczne wariancje obserwacji linii. W tym modelu oszacowanie $\boldsymbol{\alpha}_w$ wynosi:

$$\hat{\boldsymbol{\alpha}}_w = (\mathbf{G}'\mathbf{W}^{-1}\mathbf{G})^{-1}\mathbf{G}'\mathbf{W}^{-1}\mathbf{y}, \quad (8)$$

gdzie $\mathbf{W} = (w_{ii})$, w_{ii} jest oszacowaną wariancją dla i -tej linii, $i = 1, 2, \dots, n$. Wybór markerów do modelu przy zastosowaniu regresji ważonej jest dokonywany tą samą metodą, jak opisano dla przypadku regresji nieważonej (wariant II).

Całkowity efekt interakcji potrójnej aaa_{gw} genów wpływających na cechę ilościową z uwzględnieniem ważonej wielokrotnej regresji liniowej można estymować jako:

$$\hat{aaa}_{gw} = \sum_{k=1}^{p-2} \sum_{\substack{k'=2 \\ k' \neq k}}^{p-1} \sum_{\substack{k''=3 \\ k'' \neq k, k'}}^p \hat{aaa}_{l_k l_{k'} l_{k''}}. \quad (9)$$

Współczynniki determinacji zostały zastosowane do zmierzenia jakości dopasowania obu modeli (uzyskanych przy użyciu metod regresji nieważonej i regresji ważonej) do danych i określały ilość wariancji fenotypowej wyjaśnionej łącznie przez wszystkie rozważane parametry.

6.2. Zbiory danych

6.2.1. Zbiór I (Publikacja I)

Dane pochodziły z Północnoamerykańskiego Projektu Badania Genomu Jęczmienia (ang. *North American Barley Genome Mapping Project*, NABGMP)

i składały się ze 150 linii podwojonych haploidów (DH) jęczmienia testowanych w szesnastu środowiskach [Crookston, Minnesota, 1992; Ithaca, Nowy Jork, 1992; Guelph, Ontario, 1992; Pullman, Waszyngton, 1992; Brandon, Manitoba, 1992; Outlook, Saskatchewan, 1992; Goodale, Saskatchewan, 1992; Saskatoon, Saskatchewan, 1992; Tetonia, Idaho, 1992; Bozeman, Montana (nawadniane), 1992; Bozeman, Montana (suche), 1992; Aberdeen, Idaho, 1991; Klamath Falls, Oregon, 1991; Pullman, Waszyngton, 1991; Bozeman, Montana (nawadniane), 1991; Bozeman, Montana (suche), 1991]. Linie DH zostały opracowane przez Oregon State University Barley Breeding Program poprzez skrzyżowanie odmian jęczmienia Steptoe i Morex (Kleinhofs i in., 1993; Romagosa, Ullrich, Han, Hayes, 1996; Yao i in., 2022) metodą *Hordeum bulbosum*. Zastosowana mapa sprzężeń składała się z 223 markerów molekularnych, głównie RFLP, ze średnią odległością między markerami równą 5,66 cM. Linie DH badano pod względem ośmiu cech fenotypowych [zawartość α -amylazy, AA; moc diastatyczna, DP; zawartość białka w ziarnie, GP; plon ziarna, GY; wysokość rośliny, H; termin kłoszenia, HD; wyleganie, L; ekstrakt słodu, ME; (Hayes i in., 1993)]. Brakujące wartości markerów zostały oszacowane na podstawie danych markerów flankujących (Martinez, Curnow, 1994), a dane dotyczące cech GP, L i ME zostały przekształcone według wzoru: $\arcsin\sqrt{x/100}$.

6.2.2. Zbiór II (Publikacja I)

Dane, również pochodzące z projektu NABGMP, składały się ze 145 linii podwojonych haploidów (DH) jęczmienia (kombinacja dwurzędowych odmian Harrington \times TR306) analizowanych pod względem siedmiu cech fenotypowych (masa zebranego ziarna na jednostkę powierzchni, WG; termin kwitnienia 50% roślin [liczba dni], NH; liczba dni od posadzenia do osiągnięcia dojrzałości fizjologicznej, NM; wysokość roślin, H; wyleganie przekształcone na stopnie Blissa $\arcsin\sqrt{x/100}$, L; masa 1000 ziaren, KW; test wagowy, TW) w pięciu środowiskach (w czterech środowiskach obserwacje prowadzono przez dwa lata: Brandon, Manitoba, 1992 i 1993; Ailsa Craig, Ontario, 1992 i 1993; Elora, Ontario, 1992 i 1993; Outlook, Saskatchewan, 1992 i 1993; Ste-Anne-de-BeUevue, Quebec, 1993) (Tinker i in., 1996; Yao i in., 2022). Wykorzystano mapę złożoną ze 127 markerów molekularnych (głównie RFLP) ze średnią odległością między markerami równą 10,62 cM.

6.2.3. Zbiór III (Publikacja II)

Trzeci zbiór danych stanowiła kolekcja 252 linii wsobnych (RI) kukurydzy. 151 linii RI pochodziło z Hodowli Roślin Smolice Sp. z o.o. Grupa Instytutu Hodowli i Aklimatyzacji Roślin-Państwowego Instytutu Badawczego, natomiast 101 linii RI pochodziło z Małopolskiej Hodowli Roślin Sp. z o.o. Doświadczenie polowe z 252 liniami RI zostało założone na poletkach o powierzchni 10 m², w układzie bloków losowanych kompletnych w trzech powtórzeniach, w dwóch miejscowościach oddalonych od siebie o 120 km: Smolice (51°42'12" N, 17°10'10" E) i Kobierzyce (50°58'1" N, 16°55'50" E). Po zbiorach, w pierwszej połowie listopada, wykonano pomiary następujących cech: długość kolby, średnicę kolby, długość rdzenia, średnicę rdzenia, liczbę rzędów ziarna, liczbę ziaren w rzędzie, masę tysiąca ziaren (TKW) oraz plon. Pomiary przeprowadzono na dziesięciu losowo wybranych kolbach w każdym z powtórzeń. Gleby w Smolicach powstały z wapienia i kredy. Dla tej gleby przyjęto klasę bonitacyjną III B i kompleks pszenney dobry. Kukurydzę wysiano ręcznie. Nawożenie mineralne dostosowano do potrzeb kukurydzy uprawianej na zbiór ziarna. Zastosowano 350 kg·ha⁻¹ polifoska 6 i 160 kg·ha⁻¹ mocznika. Zawartość składników odżywczych była następująca: przed zbiorem ozimin: 70 kg P₂O₅·ha⁻¹, 21 kg N·ha⁻¹, 105 kg K₂O·ha⁻¹ w formie nawozu polifoska, a przed siewem 73,6 kg N·ha⁻¹ w formie mocznika. Chemiczne zwalczanie chwastów przeprowadzono za pomocą opryskiwacza.

6.2.4. Zbiór IV (Publikacja III)

Materiałem roślinnym była populacja mapująca składająca się z 94 linii podwojonych haploidów pszenicy (ang. *Chinese SpringSQ1 doubled haploid*, CSDH). Mapa genetyczna populacji, o łącznej długości około 4040 cM, zawierała 920 markerów molekularnych. Szczegółowy opis populacji mapującej i mapy przedstawiono w Czyżyło-Mysza i in. (2013). 94 linie CSDH poddano wernalizacji przez siedem tygodni w temperaturze 4°C z zachowaniem czasu naświetlania 8 godzin światła/16 godzin ciemności (dzień krótki). Wykiełkowane sadzonki wysadzono następnie do doniczek o pojemności 3 l (średnica 15 cm i wysokość 20 cm) w trzech powtórzeniach dla każdej linii i dla każdego zabiegu (564 rośliny na każdy rok, 3 dla kontroli i 3 dla suszy). Rośliny były uprawiane w doniczkach w tunelu wegetacyjnym i chronione przed deszczem w warunkach zbliżonych do naturalnych, od wiosny do jesieni (maj-wrzesień). Ograniczone nawadnianie utrzymywano przez cztery tygodnie,

począwszy od fazy krzewienia. 20-25% polowej pojemności wodnej (ang. *field water capacity*, FWC) przyjęto jako silną suszę (grupa testowa), a 65-70% FWC jako dobrze nawodnione (grupa kontrolna). Zawartość wody w glebie mierzono za pomocą HydroSense Soil Water System CS630 (Campbell Scientific, Logan, UT, USA) przy użyciu 12 cm sond i metody wagowej. Co trzy dni sprawdzano ciężar losowo wybranych doniczek w celu określenia ilości wody do podlewania. W ostatnim dniu suszy pobrano próbki liści flagowych w celu oceny całkowitej zawartości fenoli. Próbki liści homogenizowano w etanolu; całkowitą zawartość fenoli oznaczano metodą Folin-Ciocalteu Singletona i Rossiego (Singleton, Rossi, 1965). Absorbancję przy 765 nm odczytano za pomocą czytnika mikroplętek (Synergy 2, BioTek, Winooski, VT, USA), a kwas chlorogenowy zastosowano jako standard. Doświadczenie zostało powtórzone podczas trzech sezonów uprawy w latach 2010-2012. Warunki wzrostu i wszystkie procedury eksperymentalne zostały zachowane. Analizy przeprowadzono oddzielnie dla każdego roku, dla grupy testowej i kontrolnej. Ostatecznie uzyskano sześć podzbiorów danych.

6.3. Badania symulacyjne

W badaniach symulacyjnych Monte Carlo porównujących fenotypowe i genotypowe (nieważone i dwie modyfikacje ważne) oszacowania efektów interakcji *aaa* QTL przyjęto następujące warianty parametrów. Założono, że prawdziwa wartość parametru związanego z efektem interakcji *aaa* jest równa 15 ($aaa = 15$), a średnia ogólna cechy ilościowej wynosi 100. Analizowano 500 linii homozygotycznych. Każda linia była reprezentowana przez dziesięć roślin. Założono 1000 markerów molekularnych rozmieszczonych w dziesięciu chromosomach o długości 100 cM każdy. Liczba markerów molekularnych na każdym chromosomie była równa 100. Odległość między markerami wynosiła 1 cM. Liczba QTL wpływających na cechę wynosiła 14 (każdy z efektem addytywnym równym dwa). Zakładano, że każdy QTL był dokładnie sprzężony z markerem, tzn. odległość QTL od markera wynosiła 0 cM. Poszczególne QTL rozmieszono następująco: jeden QTL na chromosomie 1 (Q1), po dwa QTL na chromosomach 3 (Q2 i Q3) i 5 (Q4 i Q5) oraz po trzy QTL na chromosomach 7 (Q6, Q7 i Q8), 9 (Q9, Q10 i Q11) i 10 (Q12, Q13 i Q14) (Tabela 1). Założono działanie pięciu par epistatycznych: Q1-Q2, Q4-Q5, Q6-Q7, Q8-Q11 i Q1-Q14 (Rysunek 1). W ten sposób zdefiniowano działanie dwóch par epistatycznych na tym samym chromosomie:

Q4 i Q5 (na chromosomie 5) oraz Q6 i Q7 (na chromosomie 7), a także trzech par epistatycznych na różnych chromosomach.

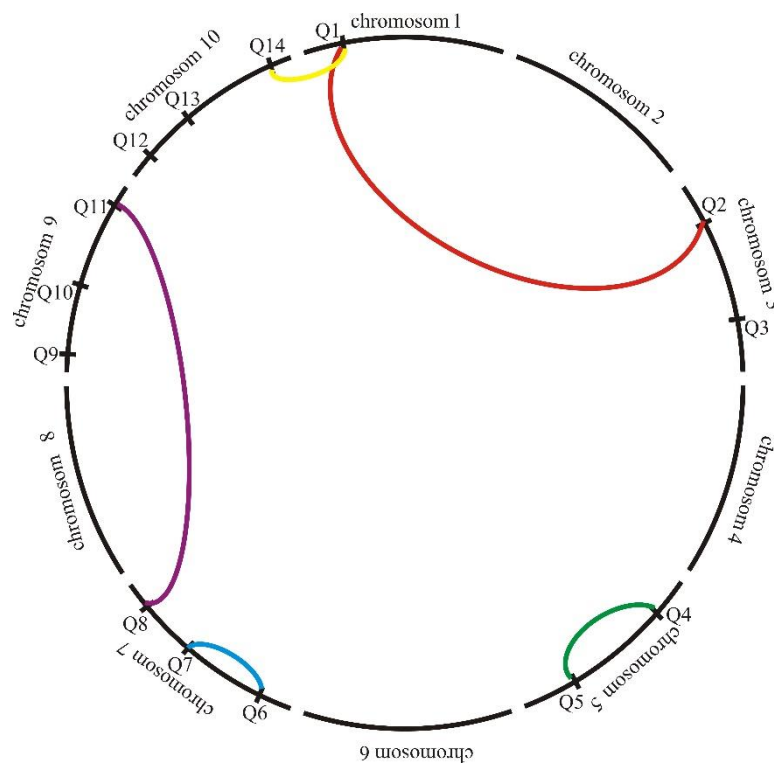
Tabela 1. Liczba założonych trójek QTL-QTL-QTL, ich oznaczenia i lokalizacje na chromosomach oraz kodowanie poszczególnych wariantów

Kod wariantu	Liczba trójek	Trójki QTL-QTL-QTL	Lokalizacja trójek na chromosomach
C01	1	Q6Q7Q8	wszystkie na jednym chromosomie
C02	1	Q1Q2Q3	na dwóch różnych chromosomach
C03	1	Q1Q2Q4	na trzech różnych chromosomach
C04	2	Q6Q7Q8, Q9Q10Q11	wszystkie na jednym chromosomie
C05	2	Q1Q2Q3, Q4Q5Q6	na dwóch różnych chromosomach
C06	2	Q1Q2Q4, Q3Q5Q7	na trzech różnych chromosomach
C07	3	Q6Q7Q8, Q9Q10Q11, Q12Q13Q14	wszystkie na jednym chromosomie
C08	3	Q1Q2Q3, Q4Q5Q6, Q7Q10Q11	na dwóch różnych chromosomach
C09	3	Q1Q2Q4, Q3Q5Q7, Q4Q6Q9	na trzech różnych chromosomach
C10	5	Q6Q7Q8, Q9Q10Q11, Q12Q13Q14, Q1Q2Q3, Q1Q4Q5	trzy trójki na jednym chromosomie i dwie trójki na dwóch chromosomach
C11	5	Q1Q2Q3, Q2Q4Q5, Q4Q6Q7, Q1Q9Q10, Q1Q12Q14	na dwóch różnych chromosomach
C12	5	Q1Q2Q4, Q3Q5Q6, Q1Q9Q12, Q2Q10Q14, Q5Q10Q12	na trzech różnych chromosomach

Źródło: Publikacja IV: Cyplik, Bocianowski (2023)

Założono równe efekty epistatyczne dla wszystkich pięciu par w dwóch wariantach: $aa = 2$ lub $aa = 0$ (brak epistazy). Przyjęto, że liczba trójek QTL-QTL-QTL z efektami interakcji aaa , które miały wpływ na cechę, wynosiła jeden, dwa, trzy lub pięć. QTL z efektami interakcji aaa były zlokalizowane na jednym chromosomie (trzy trójki na jednym chromosomie i dwie trójki na dwóch chromosomach), dwóch chromosomach lub trzech chromosomach (każdy QTL znajdował się na innym chromosomie) (Tabela 1). Założono, że efekty poszczególnych trójek genów były: (i) równe dla wszystkich trójek lub (ii) jeden efekt potrójny QTL-QTL-QTL był znacznie większy od drugiego (dla dwóch trójek: 10 i 5; dla trzech trójek: 9, 3 i 3; dla pięciu trójek: 7, 2, 2, 2 i 2). Zakładana wariancja błędu była równa 5 lub 10. Dla każdej kombinacji parametrów przeprowadzono 10000 generacji. Otrzymano zestawy danych zawierających wektor obserwacji fenotypowych i wektory obserwacji genotypów

markerowych. Dla każdego zestawu danych oszacowano wariancje dla linii i efekt interakcji aaa za pomocą metody fenotypowej (\widehat{aaa}_{jp}) przedstawionej w rozdziale 6.1.1., oraz nieważonej (\widehat{aaa}_{jgu}) i ważonej (\widehat{aaa}_{jgw1} , \widehat{aaa}_{jgw2}), gdzie $gw1$ i $gw2$ oznaczają, odpowiednio, selekcję markerów według wariantu I i II (rozdział 6.1.2.1.) ($j = 1, 2, \dots, 10000$) wersji metody genotypowej przedstawionej w rozdziale 6.1.2. Dodatkowo oszacowano współczynniki determinacji R_j^2 . Następnie obliczono średnie wartości oszacowań parametrów $\overline{\widehat{aaa}}_p$, $\overline{\widehat{aaa}}_{gu}$, $\overline{\widehat{aaa}}_{gw1}$ oraz $\overline{\widehat{aaa}}_{gw2}$ dla każdej serii, wraz z błędami średniokwadratowymi. Obliczono średnie wartości R^2 .



Rysunek 1. Lokalizacja na chromosomach założonych 14 QTL i pięciu par epistatycznych. Źródło: Publikacja IV: Cyplik, Bocianowski (2023)

6.4. Oprogramowanie

Do przeprowadzenia analiz wykorzystano oprogramowanie statystyczne GenStat v. 22 (VSN International, 2022), oprogramowanie statystyczne R v. 4.2.2 (R Core Team, 2022) oraz RStudio (RStudio Team, 2020).

7. Wyniki i dyskusja

7.1. Wyniki

7.1.1. Publikacja I

7.1.1.1. Porównanie analityczne

Estymatory (1) i (7) potrójnej interakcji *aaa* mogą być analizowane i porównywane przy upraszczających założeniach, że: (i) markery są niesprężone oraz (ii) segregacja każdego markera jest zgodna z modelem genetycznym odpowiednim dla analizowanej populacji, co w prezentowanym przypadku oznacza, że prawdopodobieństwo zaobserwowania „1” jest takie samo jak zaobserwowania „-1”. Założenia te zostaną spełnione, jeśli model (4) traktuje obserwacje markerów jako stałe. W rzeczywistości wektory \mathbf{m}_l ($l = 1, 2, \dots, q$) stanowią obserwacje pewnych zmiennych losowych. Gdy dane markera spełniają dokładnie założenia (i) i (ii), otrzymuje się (Publikacja I: Cyplik, Bocianowski, 2022):

$$\widehat{aaa}_g = \sum_{k=1}^{p-2} \sum_{\substack{k'=2 \\ k' \neq k}}^{p-1} \sum_{\substack{k''=3 \\ k'' \neq k, k'}}^p \left[\frac{1}{2} \left(\bar{y}^{(l_k l_{k'} l_{k''}, +)} + \bar{y}^{(l_k l_{k'} l_{k''}, -)} \right) - \bar{y} \right], \quad (10)$$

gdzie $\bar{y}^{(l_k l_{k'} l_{k''}, -)}$ oraz $\bar{y}^{(l_k l_{k'} l_{k''}, +)}$ oznaczają średnie dla linii z obserwacjami k -tego, k' -tego i k'' -tego markera równe odpowiednio -1 i 1.

W rzeczywistości dane markerowe nie spełniają dokładnie warunków dla modelu (10). Biorąc pod uwagę, że markery wybrane do modelu (4) znajdują się daleko od siebie na mapie sprzężeń, założenie (i) jest prawdziwe. Do sprawdzenia założenia (ii), test χ^2 jest stosowany przed wykonaniem każdej analizy.

7.1.1.2. Porównanie numeryczne

Uzyskane wyniki oszacowań całkowitego efektu interakcyjnego *aaa* przedstawiono w tabelach 2-5. Tabele 2 i 3 zawierają wyniki analizy metodami, odpowiednio, fenotypową i genotypową dla 150 linii podwojonych haploidów jęczmienia z kombinacji Steptoe \times Morex; tabele 4 i 5 – dla 145 linii DH jęczmienia z kombinacji Harrington \times TR306. Rysunki 2 i 3 przedstawiają względne porównanie fenotypowych i genotypowych oszacowań całkowitego efektu interakcji *aaa* w postaci

wykresu pudełkowego wartości $(\widehat{aaa}_g/\widehat{aaa}_p) \cdot 100$, sklasyfikowanych według obserwowanych cech fenotypowych.

Wyniki pokazują, że w 90 przypadkach (70%) zaobserwowano statystycznie istotne efekty interakcji *aaa* dla metody fenotypowej (Tabela 2). Taką samą liczbę interakcji stwierdzono dla metody genotypowej (z uwzględnieniem obserwacji markerowych), ale tylko w 72 przypadkach wyniki były istotne statystycznie (Tabela 3). Porównania genotypowych i fenotypowych oszacowań całkowitego efektu interakcji *aaa* pokazują, że w większości przypadków (79%) efekt był mniejszy dla metody genotypowej niż całkowity efekt interakcji *aaa* obliczony na podstawie obserwacji jedynie fenotypowych (Rysunek 2). Zakres obliczonych względnych porównań był stosunkowo szeroki i wynosił od -1590,91% dla HD do 1800,00% dla H w tym samym środowisku (WA92). W sumie w pięciu przypadkach zaobserwowano wartości względnych porównań wyższe niż |1000|%. Najmniejszy zakres względnych porównań zaobserwowano dla cechy DP. Liczba genów (czynników efektywnych) wynosiła od 3 do 10, ze średnią 3,4 (Tabela 2). Minimalna liczba markerów uwzględnionych w modelu wynosiła 12, a maksymalna 32, ze średnią 19,5 markerów na model. Liczba potrójnych interakcji wynosiła od 0 do 35, ze średnią 8,3 (Tabela 3).

Tabela 2. Oszacowania fenotypowe całkowitego efektu interakcji addytywno × addytywno × addytywnej dla 150 linii podwojonych haploidów jęczmienia uzyskanych z kombinacji Steptoe × Morex

Środowisko [#]	Cecha [§]							
	AA	DP	GY	GP	HD	H	L	ME
ID91	1,56 @ (5)	11,36 (8)	-0,31 (6)	0,08 (4)	-0,05 (4)	0,90 (7)	-	0,45 (5)
ID92	1,72 (5)	31,17 (5)	-0,78 (7)	0,21 (5)	-0,75 (4)	6,42 (6)	-	0,37 (5)
MA92	-	-	-0,22 (4)	-	-0,79 (3)	-0,48 (4)	3,14 (7)	-
MN92	3,13 (6)	5,16 (5)	-0,54 (5)	0,24 (4)	0,93 (3)	1,33 (4)	-	0,61 (6)
MTd91	-	-	-0,12 (4)	-	0,32 (3)	-0,67 (5)	-	-
MTd92	2,82 (6)	28,11 (6)	-0,74 (5)	0,50 (5)	0,68 (3)	6,25 (5)	14,50 (7)	-0,33 (3)
MTi91	4,15 (4)	26,29 (10)	-0,37 (3)	0,33 (5)	1,31 (4)	0,68 (5)	-	-0,25 (5)
MTi92	1,03 (3)	20,32 (9)	-0,13 (4)	1,30 (3)	-0,73 (3)	3,55 (5)	4,93 (4)	-0,73 (5)
NY92	-	-	0,26 (5)	-	0,31 (3)	-1,33 (5)	20,87 (9)	-
ON92	-	-	-0,14 (4)	-	0,98 (4)	1,03 (4)	11,80 (7)	-
OR91	3,82 (6)	17,59 (5)	0,54 (5)	0,46 (5)	-2,66 (4)	-4,64 (5)	-	-0,10 (3)
SKg92	-	-	-0,23 (4)	-	0,43 (4)	5,91 (4)	-	-
SKk92	-	-	-0,34 (5)	-	0,55 (3)	0,55 (4)	-	-
SKo92	-	-	0,56 (6)	-	1,05 (3)	-3,94 (5)	-4,00 (3)	-
WA91	2,67 (5)	13,05 (4)	-0,08 (4)	-0,20 (5)	2,14 (5)	-5,28 (5)	-	0,02 (6)
WA92	1,88 (4)	27,41 (9)	-0,34 (6)	0,69 (3)	0,11 (4)	0,33 (3)	-	0,35 (4)

[#]ID91 – Aberdeen, Idaho, 1991; ID92 – Tetonia, Idaho, 1992; MA92 – Brandon, Manitoba, 1992; MN92 – Crookston, Minnesota, 1992; MTd91 – Bozeman, Montana (suche), 1991; MTd92 – Bonzeman, Montana (suche), 1992; MTi91 – Bozeman, Montana (nawadniane), 1991; MTi92 – Bozeman, Montana (nawadniane), 1992; NY92 – Ithaca, Nowy Jork, 1992; ON92 – Guelph, Ontario, 1992; OR91 – Klamath Falls, Oregon, 1991; Kg92 – Goodlae, Saskatchewan, 1992; SKk92 – Kcfr, Saskatchewan, 1992; SKo92 – Outlook, Saskatchewan, 1992; WA91 – Pullman, Waszyngton, 1991; WA92 – Pullman, Waszyngton, 1992

[§]AA – α -amylaza; DP – siła diastatyczna; GP – zawartość białka w ziarnie; GY – plon ziarna; H – wysokość rośliny; HD – termin kłoszenia; L – wyleganie; ME – ekstrakt słodu

@liczba genów (liczba czynników efektywnych) uzyskana wyłącznie na podstawie obserwacji fenotypowych

Źródło: Publikacja I: Cyplik, Bocianowski (2022)

Tabela 3. Oszacowania genotypowe całkowitego addytywno × addytywno × addytywnego efektu interakcji dla 150 linii podwojonych haploidów linii jęczmienia uzyskanych z kombinacji Steptoe × Morex

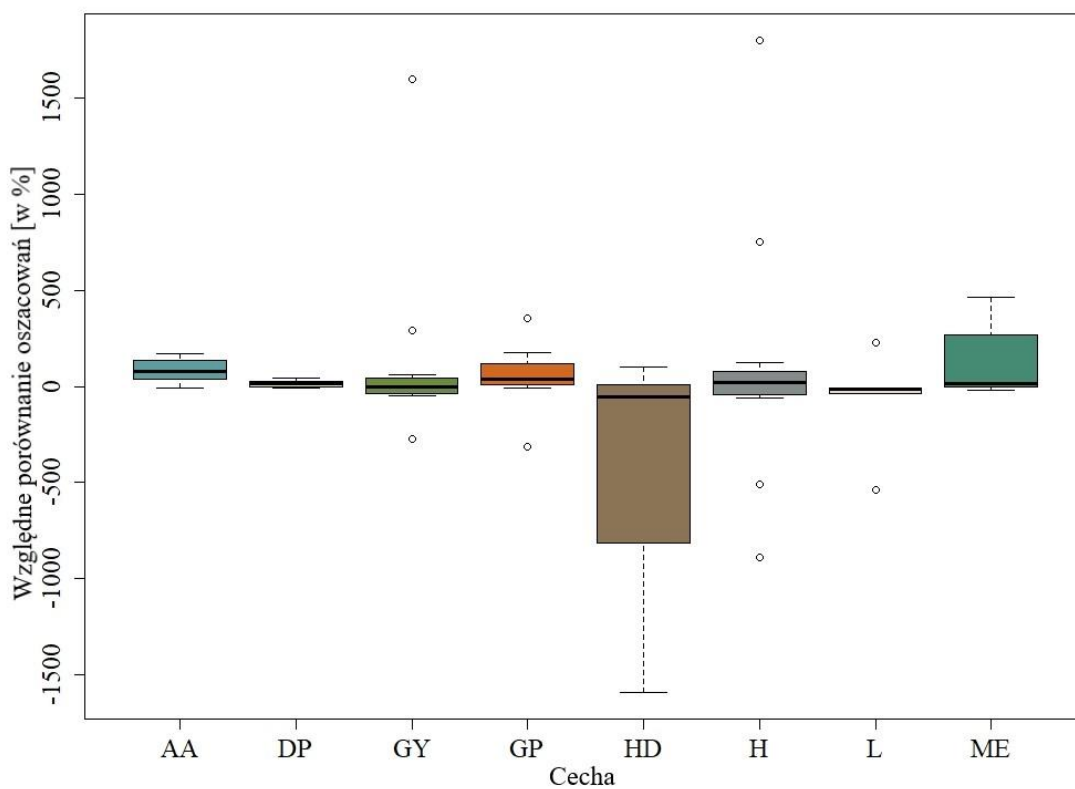
Środowisko [#]	Cecha [§]							
	AA	DP	GY	GP	HD	H	L	ME
ID91	*NS **(17 0)	-1,25 (32 1)	0,07 (22 8)	0,14 (17 25)	0,67 (20 11)	-0,53 (24 10)	-	0,02 (27 1)
ID92	2,89 (23 13)	NS (14 0)	0,31 (22 9)	0,08 (21 1)	0,19 (27 3)	1,61 (25 5)	-	-0,07 (27 3)
MA92	-	-	-0,09 (18 1)	-	1,27 (20 16)	-3,60 (20 10)	-16,82 (15 23)	-
MN92	2,24 (26 2)	2,13 (19 17)	-0,34 (18 7)	0,85 (16 23)	-2,70 (19 23)	0,17 (17 20)	-	1,18 (22 14)
MTd91	-	-	0,33 (19 20)	-	0,06 (15 2)	3,41 (22 14)	-	-
MTd92	-0,01 (29 2)	-0,18 (24 6)	0,02 (20 12)	-0,04 (27 4)	-0,28 (24 4)	1,24 (23 4)	-2,39 (26 6)	-0,04 (31 1)
MTi91	-0,25 (18 21)	NS (15 0)	-1,08 (15 23)	0,08 (19 1)	NS (13 0)	NS (14 0)	-	NS (18 0)
MTi92	1,59 (18 7)	2,14 (16 28)	-2,08 (20 16)	NS (13 0)	NS (12 0)	NS (18 0)	11,33 (21 11)	0,07 (15 2)
NY92	-	-	0,02 (22 10)	-	-4,37 (18 25)	-1,67 (16 21)	-2,19 (22 6)	-
ON92	-	-	0,00 (25 7)	-	-0,68 (25 4)	-9,17 (20 11)	-4,60 (24 13)	-
OR91	NS (15 0)	4,22 (15 5)	-0,19 (15 1)	-1,45 (17 30)	NS (15 0)	-1,13 (16 1)	-	-0,34 (22 9)
SKg92	-	-	NS (21 0)	-	NS (15 0)	-1,46 (16 1)	-	-
SKk92	-	-	-0,01 (16 4)	-	0,56 (16 30)	NS (17 0)	-	-
SKo92	-	-	-0,13 (21 8)	-	0,37 (13 1)	-1,32 (22 10)	0,44 (21 8)	-
WA91	3,20 (20 10)	NS (16 0)	NS (13 0)	-0,12 (18 1)	0,07 (14 3)	NS (17 0)	-	NS (13 0)
WA92	1,44 (22 8)	3,88 (20 8)	0,16 (20 14)	0,25 (19 8)	-1,75 (27 4)	5,94 (15 35)	-	1,63 (19 18)

[#]Opis środowisk w Tabeli 2

[§]Opis cech w Tabeli 2

*NS – nieistotny; **(x | y): x – liczba uwzględnionych markerów, y – liczba istotnych interakcji *aaa*; “-” – nie znaleziono interakcji *aaa*.

Źródło: Publikacja I: Cyplik, Bocianowski (2022)



Rysunek 2. Względne porównanie fenotypowych i genotypowych oszacowań całkowitego efektu interakcji potrójnej dla 150 linii DH jęczmienia uzyskanych z kombinacji Steptoe × Morex: wykres pudełkowy wartości $(\widehat{aaa}_g / \widehat{aaa}_p) \cdot 100$, sklasyfikowanych według obserwowanych cech fenotypowych (AA – α -amylaza; DP – moc diastatyczna; GP – zawartość białka w ziarnie; GY – plon ziarna; H – wysokość rośliny; HD – termin kłoszenia; L – wyleganie; ME – ekstrakt słoðu)

Źródło: Publikacja I: Cyplik, Bocianowski (2022)

Dla linii DH z kombinacji Harrington × TR306 wyniki pokazują, że w 63 przypadkach (100%) uzyskano statystycznie istotne efekty interakcji *aaa* przy estymacji metodą fenotypową (Tabela 4).

Tabela 4. Fenotypowe oszacowania całkowitego efektu interakcji addytywno × addytywno × addytywnej dla 145 linii podwojonych haploidów jęczmienia uzyskanych z kombinacji Harrington × TR306

Środowisko [#]	Cecha ^s						
	WG	NH	NM	H	L	KW	TW
ON92a	-6,02 @ (10)	0,11 (5)	-1,34 (8)	1,87 (4)	9,24 (3)	1,28 (5)	-0,77 (11)
ON93a	12,12 (7)	0,33 (6)	0,42 (8)	-0,76 (9)	14,43 (3)	0,03 (6)	-1,97 (2)
ON92b	6,21 (5)	0,27 (10)	0,08 (4)	0,26 (3)	-0,34 (3)	0,81 (3)	-0,60 (9)
ON93b	-5,67 (7)	0,25 (6)	0,22 (3)	0,64 (6)	15,65 (6)	0,55 (5)	-0,39 (8)
MB92	-9,00 (5)	0,29 (9)	1,23 (11)	4,48 (4)	-0,51 (4)	0,00 (8)	-2,09 (3)
MB93	-26,10 (6)	0,89 (11)	-0,10 (4)	0,94 (9)	-3,41 (8)	-0,89 (6)	-1,63 (13)
QC93	-9,14 (5)	0,77 (7)	-0,60 (3)	-1,03 (3)	18,30 (5)	-0,71 (3)	-0,96 (5)
SK92a	61,75 (2)	1,15 (7)	0,12 (5)	1,78 (3)	-9,58 (3)	-2,11 (3)	-2,93 (0)
SK93a	-3,39 (7)	-0,54 (4)	-0,87 (5)	0,61 (3)	4,96 (2)	0,71 (7)	-0,68 (8)

[#]ON92a – Ailsa Craig, Ontario, 1992; ON93a – Ailsa Craig, Ontario, 1993; ON92b – Elora, Ontario, 1992; ON93b – Elora, Ontario, 1993; MB92 – Brandon, Manitoba, 1992; MB93 – Brandon, Manitoba, 1993; QC93 – Ste-Anne-de-Bellevue, Quebec, 1993; SK92a – Outlook, Saskatchewan, 1992; SK93a – Outlook, Saskatchewan, 1992

^sWG – masa zebranego ziarna na jednostkę powierzchni; NH – termin kwitnienia 50% roślin [liczba dni]; NM – liczba dni od zasadzenia do osiągnięcia dojrzałości fizjologicznej; H – wysokość rośliny; L – wyleganie; KW – masa 1000 ziaren; TW – test wagowy.

@liczba genów (liczba czynników efektywnych) uzyskana wyłącznie na podstawie obserwacji fenotypowych

Źródło: Publikacja I: Cyplik, Bocianowski (2022)

Taką samą liczbę interakcji stwierdzono dla metody uwzględniającej obserwacje markerowe, ale tylko w 35 przypadkach wyniki były istotne statystycznie (Tabela 5).

Porównanie genotypowych i fenotypowych oszacowań całkowitego efektu interakcji addytywno \times addytywno \times addytywnej pokazuje, że w większości przypadków (79%) efekt uzyskany metodą genotypową był mniejszy niż całkowity efekt interakcji *aaa* otrzymany na podstawie obserwacji jedynie fenotypowych (Rysunek 3). Zakres obliczonych względnych porównań był bardzo duży i wynosi od -2194,31% (dla WG w środowisku QC93) do 2866,67% (dla KW w ON93a). W czterech przypadkach zaobserwowano wartości względnych porównań wyższe niż |1000|%. Najmniejszy zakres względnych porównań zaobserwowano dla cechy NM. Liczba genów (efektywnych czynników) wynosiła od 0 do 13, ze średnią 5,6 (Tabela 4). Minimalna liczba markerów uwzględnionych w modelu wynosiła 7, a maksymalna 21, ze średnią 13,9 markerów na model. Liczba potrójnych interakcji wynosiła od 0 do 36, ze średnią 4,8 (Tabela 3).

Tabela 5. Oszacowania genotypowe całkowitego efektu interakcji addytywno \times addytywno \times addytywnej dla 145 linii podwojonych haploidów jęczmienia uzyskanych z kombinacji Harrington \times TR306

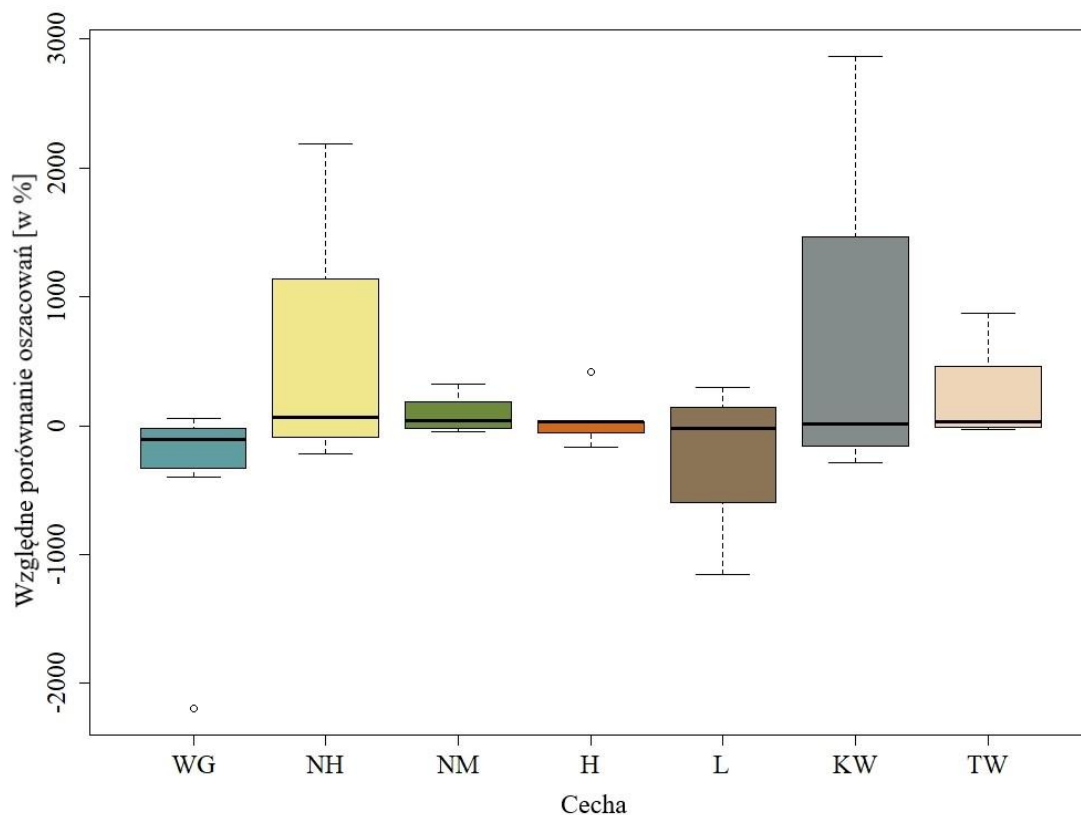
Środowisko [#]	Cecha [§]						
	WG	NH	NM	H	L	KW	TW
ON92a	23,94 **(21 7)	NS (16 0)	-0,34 (16 4)	0,56 (14 1)	NS (13 0)	0,72 (13 1)	-0,46 (12 7)
ON93a	5,26 (16 1)	0,10 (12 1)	NS (15 0)	1,26 (15 1)	NS (12 0)	0,86 (13 1)	-0,16 (7 1)
ON92b	-5,54 (20 1)	NS (16 0)	0,15 (14 1)	NS (12 0)	3,93 (12 1)	-0,21 (12 2)	-5,21 (17 18)
ON93b	-3,08 (13 1)	5,46 (16 15)	NS (9 0)	2,69 (16 26)	-1,62 (15 21)	NS (15 0)	NS (13 0)
MB92	9,77 (14 1)	0,27 (13 4)	-0,54 (14 3)	1,42 (14 2)	-1,51 (13 1)	-0,71 (11 2)	NS (16 0)
MB93	67,68 (14 36)	-1,93 (14 24)	NS (14 0)	NS (10 0)	NS (12 0)	2,57 (18 12)	0,50 (15 3)
QC93	200,56 (17 20)	NS (13 0)	-1,91 (15 30)	NS (13 0)	NS (16 0)	NS (13 0)	NS (11 0)
SK92a	*NS (7 0)	NS (17 0)	-0,03 (12 2)	NS (11 0)	NS (13 0)	NS (12 0)	NS (16 0)
SK93a	NS (14 0)	NS (12 0)	-0,40 (17 21)	-0,34 (15 28)	-1,46 (14 2)	NS (13 0)	NS (16 0)

[#]Opis środowisk w Tabeli 4

[§]Opis cech w Tabeli 4

*NS – nieistotne; **(x | y): x – liczba uwzględnionych markerów, y – liczba istotnych interakcji *aaa*.

Źródło: Publikacja I: Cyplik, Bocianowski (2022)



Rysunek 3. Względne porównanie fenotypowych i genotypowych oszacowań całkowitego efektu interakcji potrójnej dla 145 linii DH jęczmienia uzyskanych z kombinacji Harrington × TR306: wykres pudełkowy wartości $(\widehat{aaa}_g / \widehat{aaa}_p) \cdot 100$, sklasyfikowanych według obserwowanych cech fenotypowych (H – wysokość roślin; KW – masa 1000 ziaren; L – wyleganie; NH – termin kwitnienia 50% roślin [liczba dni]; NM – liczba dni od posadzenia do osiągnięcia dojrzałości fizjologicznej; TW – test wagowy; WG – masa ziarna zebranego z jednostki powierzchni)

Źródło: Publikacja I: Cyplik, Bocianowski (2022)

W sumie, porównaniom numerycznym poddano 153 zestawy obserwacji, niezależnie dla każdej cechy i każdego środowiska. Dane ze zbiorów danych I i II były rozpatrywane oddzielnie.

7.1.2. Publikacja II

Wyniki oszacowań całkowitego efektu addytywnego działania genów (*a*), efektu epistazy (*aa*) i całkowitego efektu interakcji addytywno \times addytywno \times addytywnej (*aaa*) przedstawiono w Tabeli 6. Całkowite efekty addytywne były istotne statystycznie we wszystkich badanych przypadkach. W 67% (24 z 36) przypadków efekty interakcji potrójnej były statystycznie istotne. Wszystkie interakcje *aaa* były istotne dla liczby rzędów ziarna we wszystkich lokalizacjach/pochodzeniach. Statystycznie istotne efekty epistazy stwierdzono w 58% badanych przypadków (21 z 36), z których liczba rzędów ziarna była również jedyną cechą, charakteryzującą się istotnościami we wszystkich lokalizacjach. Zakres oszacowań interakcji potrójnej (*aaa*) był stosunkowo niski w większości przypadków, wynosząc od 0,58 do 1,38. Jednak w czterech przypadkach zakresy wartości *aaa* były znacznie wyższe i wynosiły 2,75, 7,62, 15,66 i 20,67 dla, odpowiednio, liczby rzędów ziarna, liczby ziaren w rzędzie, masy ziarna z kolby i masy tysiąca ziaren. Najmniejszy zakres oszacowań stwierdzono dla średnicy kolby. Podobny zakres estymacji otrzymano dla interakcji podwójnych, z analogicznymi najniższymi i najwyższymi wartościami dla poszczególnych cech. Najniższe wartości zakresu oszacowań *aa* stwierdzono dla średnicy kolby (0,30), a najwyższe dla liczby rzędów ziarna, masy ziarna z kolby i masy tysiąca ziaren (odpowiednio 2,45, 3,12, 11,59 i 15,45). Jedynie w przypadku liczby rzędów ziarna wpływ był zawsze dodatni, zarówno dla interakcji *aa*, jak i *aaa* dla poszczególnych lokalizacji/pochodzenia. Dla innych cech wartości estymatorów *aa* i *aaa* miały różny znak pomiędzy lokalizacjami/pochodzeniem. Jednakże w danej lokalizacji/pochodzeniu często był on taki sam, dodatni lub ujemny.

Tabela 6. Wartości średnie i sumaryczne efekty addytywne (*a*) oraz sumaryczne efekty interakcyjne *aa* i *aaa* dla 252 (151 pochodzących ze Smolic, 101 z Kobierzyce) linii wsobnych kukurydzy oszacowane na podstawie wartości jedynie fenotypowych

Lokalizacja	Pochodzenie	Parametr	Długość kolby	Średnica kolby	Długość rdzenia	Średnica rdzenia	Liczba rzędów ziarna	Liczba ziaren w rzędzie	Masa ziarna z kolby	Masa tysiąca ziaren	Plon
Kobierzyce	Smolice	średnia	15,55	3,97	15,44	2,16	15,73	27,91	114,8	265,15	4,59
		<i>a</i>	4,03 ***	0,92 ***	4,06 ***	0,71 ***	4,80 ***	6,97 ***	47,41 ***	91,79 ***	1,90 ***
		<i>aa</i>	0,09	0,09	0,06	0,12 **	1,07 **	-0,01	-3,19 *	-1,94	-0,13 *
		<i>aaa</i>	0,06	0,18 *	0,24	0,19 **	1,27 **	0,42	-4,19 *	0,47	-0,17 *
Kobierzyce	Kobierzyce	średnia	15,37	3,85	15,17	2,03	15,41	27,44	106,61	254,47	4,26
		<i>a</i>	3,83 ***	0,66 ***	3,82 ***	0,59 ***	4,56 ***	8,61 ***	49,44 ***	91,02 ***	1,98 ***
		<i>aa</i>	-0,38 *	-0,08	-0,52 *	-0,01	1,59 ***	-0,05	-1,88	3,02	-0,08
		<i>aaa</i>	-0,76 *	-0,06	-0,80 **	0,03	1,59 ***	0,39	1,04	-1,34	0,04
Smolice	Smolice	średnia	13,36	4,09	13,6	2,24	15,01	24,97	94,38	288,9	3,78
		<i>a</i>	3,40 ***	0,75 ***	3,28 ***	0,96 ***	4,40 ***	11,15 ***	42,72 ***	103,9 ***	1,71 ***
		<i>aa</i>	0,44 *	0,03	0,42 *	-0,39 ***	0,86 **	-3,11 ***	3,97 *	-4,06	0,16 *
		<i>aaa</i>	0,30 *	0,24 **	0,57 *	-0,39 ***	0,99 **	-7,20 ***	3,68 *	-3,56	0,15 *
Smolice	Kobierzyce	średnia	12,91	3,98	13,3	2,28	14,91	25,32	87,46	276,22	3,5
		<i>a</i>	4,28 ***	0,75 ***	3,56 ***	0,33 ***	4,11 ***	9,06 ***	43,81 ***	97,5 ***	1,75 ***
		<i>aa</i>	-0,47 *	-0,21 **	-0,08	0,04	1,20 **	-2,27 **	-7,62 **	11,39 *	-0,30 **
		<i>aaa</i>	-1,08 **	-0,37 **	-0,30 *	0,04	1,76 ***	-3,32 ***	-11,98 ***	17,11 **	-0,48 ***

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Źródło: Publikacja II: Cyplik i in. (2022)

7.1.3. Publikacja III

7.1.3.1. Oszacowanie na podstawie fenotypu

Fenotypowe oszacowania całkowitego efektu addytywno \times addytywno \times addytywnego (aaa_p) zawartości fenoli przedstawiono w Tabeli 7. W czterech z sześciu przypadków całkowity efekt aaa_p był dodatni. Ujemny efekt zaobserwowano dla grup kontrolnych z 2011 i 2012 roku. Najwyższy całkowity efekt aaa zaobserwowano dla grupy testowej w 2012 roku, a najniższy dla grupy kontrolnej w 2011 roku. Efekty potrójnej interakcji były istotne statystycznie dla wszystkich grup z wyjątkiem grupy kontrolnej w 2012 roku.

Tabela 7. Wartości minimalne i maksymalne średnich dla linii, wartości średnie całkowitej zawartości fenoli dla wszystkich linii podwojonych haploidów, współczynnik zmienności, fenotypowe oszacowania całkowitego efektu addytywno \times addytywno \times addytywnego (aaa_p) oraz liczba genów (liczba efektywnych czynników)

Rok	2010		2011		2012	
	Grupa kontrolna	Grupa testowa	Grupa kontrolna	Grupa testowa	Grupa kontrolna	Grupa testowa
Minimum	7,104	6,499	7,164	6,231	7,645	4,626
Maksimum	14,414	12,682	16,321	15,070	13,781	13,023
Średnia	10,450	9,435	11,993	10,214	10,736	8,204
Współczynnik zmienności	18,06%	18,72%	15,39%	15,51%	14,34%	23,63%
aaa_p	0,309 *	0,156 *	-0,250 *	0,437 **	-0,023	0,620 ***
Liczba genów (liczba efektywnych czynników)	4,617	3,824	5,050	6,021	4,164	4,846

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.
i in. (2023)

Źródło: Publikacja III: Cyplik

Wartości minimalne i maksymalne średnich dla linii CSDH, jak również średnie dla wszystkich linii podwojonych haploidów całkowitej zawartości fenoli były większe dla grup kontrolnych niż dla grup testowych we wszystkich trzech latach prowadzenia badań. Współczynniki zmienności były większe dla grup testowych niż dla grup kontrolnych we wszystkich przypadkach. Wartości współczynnika zmienności wynosiły od 14,34% dla grupy kontrolnej w 2012 r. do 23,63% dla grupy testowej w 2012 r., co wskazuje na niski rozrzut danych w stosunku do wartości średnich (Tabela 7). Liczba

genów (czynników efektywnych) różniła się między grupami, wynosząc od 3,824 dla grupy testowej w 2010 do 6,021 dla grupy testowej w 2011. Różnica między największą i najmniejszą liczbą genów we wszystkich grupach była równa 2,197. Między grupami kontrolnymi i grupami testowymi rozpatrywanymi w tym samym roku, najwyższą różnicę wynoszącą 0,971 zaobserwowano w 2011 roku (Tabela 7).

7.1.3.2. Oszacowanie na podstawie genotypu

Oszacowanie całkowitego efektu aaa obliczone metodą genotypową przy zastosowaniu nieważonej (aaa_{gu}) i ważonej (aaa_{gw}) wielokrotnej regresji liniowej przedstawiono w Tabeli 8.

Tabela 8. Oszacowania genotypowe całkowitych efektów addytywno \times addytywno \times addytywnych otrzymane przy zastosowaniu nieważonej (aaa_{gu}) i ważonej (aaa_{gw}) wielokrotnej regresji liniowej

Rok		2010		2011		2012	
	Czynnik stresowy	Grupa kontrolna	Grupa testowa	Grupa kontrolna	Grupa testowa	Grupa kontrolna	Grupa testowa
	Liczba QTL	6	5	4	6	7	6
	Liczba aaa_{gu}	8	1	0	2	14	8
Regresja nieważona	Min,	-0,288	0,350		-0,489	-0,410	-0,138
	Efekt aaa_{gu}						
	Max,	0,533	0,350		0,569	0,538	0,409
	Całkowity	-0,043	0,350		0,110	0,170	0,688
	R^2 [w %]	32,4	42,0		36,3	39,1	44,5
	Liczba QTL	16	15	20	14	26	14
	Liczba aaa_{gw}	3	5	9	5	6	6
Regresja ważona	Min,	-0,659	-0,494	-0,607	-0,499	-0,526	-1,124
	Efekt aaa_{gw}						
	Max,	0,598	0,612	0,548	0,368	0,557	1,182
	Całkowity	0,443	-0,265	1,178	-1,302	-0,184	-0,228
	R^2 [w %]	46,2	61,4	95,0	58,8	78,9	58,5

Źródło: Publikacja III: Cyplik i in. (2023)

7.1.3.3. Oszacowanie na podstawie genotypu - regresja nieważona

Liczba QTL wykrytych przy zastosowaniu tradycyjnej (nieważonej) metody lokalizacji genów (Tabela 8) była podobna do liczby czynników efektywnych (Tabela 7) i wynosiła od czterech (grupa kontrolna, 2011) do siedmiu (grupa kontrolna, 2012).

Liczba istotnych trójek genów QTL×QTL×QTL była bardzo zróżnicowana i wynosiła od zera (grupa kontrolna, 2011) i jednego (grupa testowa, 2010) do czternastu (grupa kontrolna, 2012). Zakres wartości efektu potrójnej interakcji, gdy obserwowana cecha była determinowana przez co najmniej dwie trójki, wynosił od 0,547 (grupa testowa, 2012) do 1,058 (grupa testowa, 2011) (Tabela 8 i Tabela 9). Całkowity efekt aaa_{gu} wynosił od -0,043 (grupa kontrolna, 2010) do 0,688 (grupa testowa, 2012) (Tabela 8). Zmienność fenotypowa całkowitej zawartości fenoli wyjaśniona przez całkowitą interakcję potrójną była podobna we wszystkich pięciu przypadkach, w których liczba interakcji potrójnych była większa od zera i wynosiła od 32,4% (grupa kontrolna, 2010) do 44,5% (grupa testowa, 2012) (Tabela 8).

Tabela 9. Oszacowania genotypowe efektów interakcji addytywno × addytywno × addytywnych dla poszczególnych trójek QTL×QTL×QTL otrzymane na podstawie nieważonej (aaa_{gu}) i ważonej (aaa_{gw}) wielokrotnej regresji liniowej

Rok			2010		2011		2012	
Czynnik stresowy			Grupa kontrolna	Grupa testowa	Grupa kontrolna	Grupa testowa	Grupa kontrolna	Grupa testowa
QTL1 (LG*)	QTL2 (LG)	QTL3 (LG)	aaa_{gu}	aaa_{gw}	aaa_{gu}	aaa_{gw}	aaa_{gu}	aaa_{gw}
dupw004b (1A)	wPt-3094 (3A)	wmc1 (3B)		0,471				
dupw004b (1A)	wPt-668160 (1A)	gwm635b (7A)		0,612				
wPt-3094 (3A)	gwm165.3 (4A)	wPt-0391 (4B)		0,350				
wPt-3094 (3A)	gwm635b (7A)	wPt-671748 (7D)		-0,494				
cfa2262 (3A)	dupw004a (4A)	blt101.t7 (4D)				-0,499		
blt4.1 (3B)	wPt-4048 (3B)	gwm269.2 (4A)					-0,138	
blt4.1 (3B)	gwm269.2 (4A)	gwm165.3 (4A)					0,049	
wPt-6239 (3B)	blt101.t7 (4D)	wPt-2697 (5A)				-0,337		
wPt-6239 (3B)	blt101.t7 (4D)	wmc83 (7A)				-0,348		
rPt-8896 (3B)	wmc1 (3B)	gwm52 (3D)		-0,376				
wPt-0021 (3B)	gwm52 (3D)	barc60 (4B)	-0,659					
gwm269.2 (4A)	gwm165.3 (4A)	gwm205 (5D)	0,008					
gwm269.2 (4A)	gwm165.3 (4A)	barc44 (5D)	-0,152					
gwm269.2 (4A)	gwm165.3 (4A)	m69p78.1 (7A)	-0,100					
gwm269.2 (4A)	gwm205 (5D)	barc44 (5D)	0,063					
gwm269.2 (4A)	gwm205 (5D)	m69p78.1 (7A)	-0,263					
gwm269.2 (4A)	barc44 (5D)	wPt-9834 (5A)	0,533					
gwm165.3 (4A)	gwm165.3 (4A)	wPt-667091 (7D)					0,158	
gwm165.3 (4A)	wPt-0391 (4B)	wPt-3883 (7A)					0,026	
gwm165.3 (4A)	wPt-0391 (4B)	wPt-667091 (7D)					-0,179	
gwm165.3 (4A)	wPt-0391 (4B)	wmc243b (7D)					0,247	
gwm165.3 (4A)	gwm205 (5D)	m69p78.1 (7A)	0,156					
gwm165.3 (4A)	barc44 (5D)	m69p78.1 (7A)	-0,288					
gwm165.3 (4A)	gwm174 (5D)	wPt-3883 (7A)					0,026	
gwm165.3 (4A)	gwm174 (5D)	wPt-667091 (7D)					0,031	
gwm165.3 (4A)	gwm174 (5D)	wmc243b (7D)					-0,161	

Tabela 9. Ciąg dalszy

Rok			2010		2011		2012	
Czynnik stresowy			Grupa kontrolna		Grupa testowa		Grupa kontrolna	
QTL1 (LG*)	QTL2 (LG)	QTL3 (LG)	<i>aaa_{gu}</i>	<i>aaa_{gw}</i>	<i>aaa_{gu}</i>	<i>aaa_{gw}</i>	<i>aaa_{gu}</i>	<i>aaa_{gw}</i>
gwm165.3 (4A)	wPt-3883 (7A)	wmc243b (7D)						-0,008
wPt-0391 (4B)	wPt-0391 (4B)	wPt-667091 (7D)						-0,410
wPt-0391 (4B)	gwm174 (5D)	wPt-3883 (7A)						0,128
wPt-0391 (4B)	gwm174 (5D)	wmc243b (7D)						0,538
wPt-0391 (4B)	wPt-3883 (7A)	wPt-667091 (7D)						-0,151
barc152 (1B)	wPt-6239 (3B)	gwm191b (3D)						-0,400
barc152 (1B)	m65p64.8_4B (4B)	wPt-8149 (7A)						-0,526
barc152 (1B)	m92p78.10 (2A)	m60p64.13_3B (3B)						0,557
barc152 (1B)	wPt-8072 (2B)	gwm165.3 (4A)						-0,493
m65p64.8a (5B)	gwm271 (5B)	m69p78.1 (7A)	0,504					
gwm174 (5D)	wPt-3883 (7A)	wPt-667091 (7D)						-0,130
gwm174 (5D)	wPt-3883 (7A)	wmc243b (7D)						0,055
psr648_1B (1B)	wmc181 (2A)	gwm285 (3B)	0,598					
m17p65.1 (1B)	cfid73 (2D)	wmc468 (4A)			-0,478			
wmc432 (1D)	wPt-3738 (1D)	psp2151.3 (2A)						-0,640
wPt-3738 (1D)	tPt-0202 (3A)	gwm269.2 (4A)						-1,124
wPt-3738 (1D)	tPt-0202 (3A)	dupw004a (4A)						1,182
wPt-3738 (1D)	gwm513 (4B)	wmc157 (7D)						0,387
wPt-3738 (1D)	wmc429 (1D)	tPt-0202 (3A)						0,354
wmc429 (1D)	cfid11 (2D)	wPt-10291 (3D)				-0,505		
wmc429 (1D)	gwm30.1 (2D)	tPt-0202 (3A)				0,481		
wmc429 (1D)	wPt-7466 (2D)	wPt-9749 (2D)				-0,607		
wmc429 (1D)	gwm269.2 (4A)	dupw004a (4A)						-0,387
wmc429 (1D)	wPt-6316 (1D)	wPt-7466 (2D)				0,548		
wmc429 (1D)	wPt-6316 (1D)	wPt-10291 (3D)				0,436		
wmc429 (1D)	wPt-6316 (1D)	gwm161 (3D)				0,493		
wmc429 (1D)	wPt-6316 (1D)	gwm165.3 (4A)				-0,511		
wmc429 (1D)	wPt-6316 (1D)	gwm639.2 (5B)				0,419		
wmc429 (1D)	wPt-732556 (1D)	gwm30.1 (2D)				0,424		
m92p78.10 (2A)	wPt-6239 (3B)	gwm161 (3D)						0,360
m92p78.10 (2A)	wPt-6239 (3B)	wPt-8149 (7A)						0,317
wmc453a (2A)	wPt-3883 (7A)	wPt-8919 (7B)				0,569		
wmc453a (2A)	wmc283.1 (7A)	wPt-8919 (7B)				-0,459		
psp2151.3 (2A)	blt4.1 (3B)	wPt-4048 (3B)						0,201
psp2151.3 (2A)	blt4.1 (3B)	gwm269.2 (4A)						0,136
psp2151.3 (2A)	blt4.1 (3B)	gwm165.3 (4A)						-0,095
psp2151.3 (2A)	wPt-4048 (3B)	gwm269.2 (4A)						0,409
psp2151.3 (2A)	wPt-4048 (3B)	gwm165.3 (4A)						0,091
psp2151.3 (2A)	gwm269.2 (4A)	gwm165.3 (4A)						0,035
wPt-3949 (2B)	wmc257 (2B)	gwm165.3 (4A)					0,368	
wmc257 (2B)	wPt-2697 (5A)	gwm292_5D (5D)						-0,485

* LG – grupa sprzężeń

Źródło: Publikacja III: Cyplik i in. (2023)

7.1.3.4. Oszacowanie na podstawie genotypu - regresja ważona

Zastosowanie regresji ważonej spowodowało zdecydowany wzrost liczby wykrytych QTL w porównaniu do klasycznej regresji nieważonej. Wzrost ten wynosił od 233% (z 6 do 14 w grupach testowych w 2011 i 2012 roku) do 500% (z 4 do 20, grupa kontrolna, 2011) (Tabela 8). Liczba trójek przy zastosowaniu regresji ważonej była wyższa niż dla regresji nieważonej w trzech przypadkach (grupa testowa, 2010; grupa kontrolna, 2011; grupa testowa, 2011). W pozostałych przypadkach (grupa kontrolna, 2010; grupa kontrolna, 2012; grupa testowa, 2012) była mniejsza (Tabela 8). Zakres oszacowanych wartości efektu *aaa* był większy we wszystkich przypadkach z wyjątkiem grupy testowej w 2011, gdzie zmniejszył się z 1,058 do 0,867 (Tabela 8 i Tabela 9). Współczynniki determinacji dla modelu regresji ważonej były znacznie wyższe niż dla regresji nieważonej i wynosiły od 46,2% (grupa kontrolna, 2010) do 95,0% (grupa kontrolna, 2011) (Tabela 8).

Genotypowe oszacowania efektów interakcji *aaa* dla poszczególnych trójek QTL×QTL×QTL otrzymane na podstawie nieważonej (*aaa_{gu}*) i ważonej (*aaa_{gw}*) wielokrotnej regresji liniową przedstawiono w Tabeli 9. Zaobserwowano 67 istotnych statystycznie potrójnych interakcji. Żadna z trójek QTL×QTL×QTL nie była istotna równocześnie przy zastosowaniu regresji nieważonej i ważonej. Trójki istotne statystycznie przy zastosowaniu regresji ważonej nie były istotne statystycznie przy zastosowaniu regresji nieważonej i odwrotnie (Tabela 9). Najczęściej występującymi QTL w interakcji potrójnej były gwm165.3 (22 razy), gwm269.2 (11 razy) i wmc429 (11 razy). Po zastosowaniu regresji nieważonej, genami najczęściej występującymi w interakcjach potrójnych były gwm165.3 (19 razy), gwm269.2 (11 razy) i wPt-0391 (9 razy) (Tabela 9). Stosując regresję ważoną, genami najczęściej obserwowanymi w interakcjach potrójnych były wmc429 (jedenaście razy), wPt-3738 (pięć razy), wPt-6239 (pięć razy) i wPt-6316 (również pięć razy) (Tabela 9).

Łącznie dla regresji ważonej i nieważonej, w 43 potrójnych interakcjach (64%) każdy z QTL znajdował się w innej grupie sprzężeń (ang. *Linkage Group*, LG). W 24 przypadkach (36%) dwa QTL (z trzech) znajdowały się w jednej LG. Nie zaobserwowano sytuacji, w której wszystkie trzy QTL z trójki byłyby zlokalizowane w jednej LG (Tabela 9). Przy zastosowaniu regresji nieważonej zaobserwowano 23 przypadki (34%), w których każdy z trzech QTL znajdował się w innej LG. Natomiast

w 10 przypadkach (15%) dwa QTL były zlokalizowane na jednym chromosomie. Stosując regresję ważoną, zaobserwowano 20 przypadków (30%), w których każdy QTL z istotnych trójek był zlokalizowany w innej grupie sprzężeń. Natomiast w 14 przypadkach (21%) dwa QTL były zlokalizowane w jednej LG (Tabela 9). Rozkład QTL stanowiących trójkę QTL×QTL×QTL w poszczególnych grupach sprzężeń był bardzo zróżnicowany i wynosił od 2 (na 7B) do 39 (na 4A) (Tabela 9).

7.1.4. Publikacja IV

W przeprowadzonych badaniach symulacyjnych Monte Carlo rozpatrywano różne sytuacje eksperymentalne. Tabele 10-17 przedstawiają wyniki analiz symulacyjnych przeprowadzonych w celu porównania oszacowań efektów interakcji *aaa* uzyskanych czterema metodami: metodą fenotypową oraz trzema metodami genotypowymi (jedna nieważona i dwie ważne). Tabele 10-13 pokazują wyniki uzyskane przy założonej wariancji błędu wynoszącej 5, podczas gdy tabele 14-17 prezentują wyniki, gdy wariancja błędu była równa 10. Tabele 10, 11, 14 i 15 zawierają wyniki otrzymane przy założeniu, że całkowity efekt epistazy był równy 0, a tabele 12, 13, 16 i 17 – przy założeniu, że efekt epistazy był znacząco różny od zera. Wyniki badań symulacyjnych przy założeniu, że wszystkie efekty trójek *aaa* były równe, przedstawiono w tabelach 10, 12, 14 i 16; natomiast tabele 11, 13, 15 i 17 zawierają wyniki uzyskane przy założeniu różnych efektów trójek *aaa*.

Tabela 10. Fenotypowe i genotypowe oszacowania całkowitego efektu addytywno \times addytywno \times addytywnego, błędy średniokwadratowe i średnie współczynniki determinacji R^2 uzyskane w badaniach symulacyjnych przy założeniach: wariancja błędu była równa 5, brak epistazy, równe efekty dla wszystkich trójek

Kod wariantu	Liczba efektów interakcji QTL \times QTL \times QTL	Liczba chromosomów z trójkami QTL	Oszacowania				Błąd średniokwadratowy dla				R^2
			\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	
C01	1	1	17,99	17,01	17,63	17,22	0,23	0,31	0,25	0,42	0,89
C02	1	2	18,01	17,09	17,71	17,32	0,27	0,33	0,28	0,41	0,87
C03	1	3	18,32	17,12	18,02	17,59	0,28	0,40	0,33	0,44	0,92
C04	2	1	17,55	16,82	17,30	16,91	0,33	0,42	0,38	0,43	0,84
C05	2	2	17,80	17,09	17,62	17,39	0,37	0,44	0,40	0,50	0,88
C06	2	3	18,27	17,62	17,93	17,52	0,37	0,47	0,42	0,55	0,95
C07	3	1	15,97	15,48	15,82	15,55	0,45	0,53	0,49	0,66	0,94
C08	3	2	16,43	15,66	16,21	16,09	0,47	0,59	0,51	0,68	0,87
C09	3	3	16,91	16,01	16,72	16,49	0,61	0,72	0,68	0,73	0,92
C10	5	1,2	15,57	14,42	15,02	14,77	0,55	0,71	0,60	0,82	0,90
C11	5	2	16,25	14,93	15,99	15,21	0,56	0,78	0,69	0,88	0,87
C12	5	3	16,66	15,82	16,49	16,31	0,62	0,82	0,77	1,01	0,91

p — metoda fenotypowa, gu — regresja nieważona, gw1 — regresja ważona: tylko ważne efekty potrójnej interakcji, gw2 — regresja ważona: ważne wszystkie efekty: QTL, interakcja epistatyczna i potrójna.

Źródło: Publikacja IV: Cyplik, Bocianowski (2023)

Tabela 11. Fenotypowe i genotypowe oszacowania całkowitego efektu addytywno \times addytywno \times addytywnego, błędy średniokwadratowe i średnie współczynniki determinacji R^2 uzyskane w badaniu symulacyjnym przy założeniach: wariancja błędu równa 5, brak epistazy, różne efekty dla trójek

Kod wariantu	Liczba efektów interakcji QTL \times QTL \times QTL	Liczba chromosomów z trójkami QTL	Oszacowania				Błąd średniokwadratowy dla				R^2
			\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	
C04	2	1	17,48	16,97	17,42	16,89	0,30	0,44	0,37	0,44	0,84
C05	2	2	17,89	17,32	17,59	17,42	0,41	0,47	0,48	0,52	0,95
C06	2	3	18,33	17,58	17,99	17,99	0,42	0,49	0,45	0,59	0,83
C07	3	1	15,92	15,47	15,89	15,73	0,47	0,51	0,48	0,59	0,90
C08	3	2	16,58	15,60	16,43	15,98	0,51	0,60	0,53	0,66	0,87
C09	3	3	17,02	16,21	16,69	16,55	0,59	0,77	0,63	0,89	0,88
C10	5	1,2	15,52	14,62	15,17	14,80	0,56	0,73	0,69	0,79	0,88
C11	5	2	16,19	15,23	15,89	15,70	0,59	0,74	0,63	0,90	0,92
C12	5	3	16,70	16,01	16,52	16,25	0,73	1,01	0,88	1,17	0,94

p — metoda fenotypowa, gu — regresja nieważona, gw1 — regresja ważona: tylko ważne efekty potrójnej interakcji, gw2 — regresja ważona: ważne wszystkie efekty: QTL, interakcja epistatyczna i potrójna.

Źródło: Publikacja IV: Cyplik, Bocianowski (2023)

Tabela 12. Fenotypowe i genotypowe oszacowania całkowitego efektu addytywno \times addytywno \times addytywnego, błędy średniokwadratowe i średnie współczynniki determinacji R^2 uzyskane w badaniu symulacyjnym przy założeniach: wariancja błędu równa 5, obecność epistazy, równe efekty dla wszystkich trójek

Kod wariantu	Liczba efektów interakcji QTL \times QTL \times QTL	Liczba chromosomów z Trójkami QTL	Oszacowania				Błąd średniokwadratowy dla				R^2
			$\overline{a\overline{a}}_p$	$\overline{a\overline{a}}_{gu}$	$\overline{a\overline{a}}_{gw1}$	$\overline{a\overline{a}}_{gw2}$	$\overline{a\overline{a}}_p$	$\overline{a\overline{a}}_{gu}$	$\overline{a\overline{a}}_{gw1}$	$\overline{a\overline{a}}_{gw2}$	
C01	1	1	17,72	16,53	17,39	16,71	0,19	0,28	0,21	0,39	0,89
C02	1	2	17,84	16,72	17,53	16,82	0,25	0,30	0,27	0,40	0,94
C03	1	3	18,01	16,89	17,72	17,19	0,25	0,37	0,32	0,42	0,90
C04	2	1	17,40	16,66	17,07	16,72	0,30	0,39	0,33	0,40	0,84
C05	2	2	17,71	16,79	17,55	17,01	0,36	0,41	0,40	0,49	0,90
C06	2	3	17,99	16,50	17,62	17,17	0,34	0,40	0,37	0,54	0,86
C07	3	1	15,79	15,03	15,39	15,19	0,40	0,49	0,42	0,58	0,89
C08	3	2	16,29	15,38	16,03	15,57	0,44	0,53	0,49	0,66	0,92
C09	3	3	16,77	15,55	16,59	16,02	0,58	0,69	0,66	0,70	0,89
C10	5	1,2	15,42	14,29	14,77	14,99	0,50	0,63	0,52	0,77	0,94
C11	5	2	16,09	14,70	15,81	15,33	0,54	0,75	0,62	0,81	0,95
C12	5	3	16,40	14,92	16,22	15,76	0,59	0,81	0,73	0,99	0,88

p — metoda fenotypowa, gu — regresja nieważona, gw1 — regresja ważona: tylko ważne efekty potrójnej interakcji, gw2 — regresja ważona: ważne wszystkie efekty: QTL, interakcja epistatyczna i potrójna.

Źródło: Publikacja IV: Cyplik, Bocianowski (2023)

Tabela 13. Fenotypowe i genotypowe oszacowania całkowitego efektu addytywno \times addytywno \times addytywnego, błędy średniokwadratowe i średnie współczynniki determinacji R^2 uzyskane w badaniu symulacyjnym przy założeniach: wariancja błędu równa 5, obecność epistazy, różne efekty dla trójek

Kod wariantu	Liczba efektów interakcji QTL \times QTL \times QTL	Liczba chromosomów z trójkami QTL	Oszacowania				Błąd średniokwadratowy dla				R^2
			$\overline{a\overline{a}}_p$	$\overline{a\overline{a}}_{gu}$	$\overline{a\overline{a}}_{gw1}$	$\overline{a\overline{a}}_{gw2}$	$\overline{a\overline{a}}_p$	$\overline{a\overline{a}}_{gu}$	$\overline{a\overline{a}}_{gw1}$	$\overline{a\overline{a}}_{gw2}$	
C04	2	1	17,19	16,53	16,98	16,77	0,33	0,40	0,37	0,48	0,95
C05	2	2	17,37	16,80	17,12	16,97	0,36	0,44	0,41	0,53	0,83
C06	2	3	17,42	16,99	17,28	17,17	0,36	0,45	0,39	0,61	0,85
C07	3	1	15,77	15,22	15,59	15,36	0,39	0,52	0,42	0,70	0,91
C08	3	2	15,90	15,39	16,22	15,78	0,42	0,59	0,51	0,68	0,95
C09	3	3	16,82	15,66	16,70	16,32	0,55	0,68	0,60	0,73	0,87
C10	5	1,2	14,99	14,38	14,73	14,55	0,51	0,63	0,53	0,79	0,89
C11	5	2	15,07	14,42	14,80	14,55	0,54	0,77	0,66	0,93	0,90
C12	5	3	15,18	14,66	14,98	14,81	0,58	0,83	0,69	0,97	0,88

p — metoda fenotypowa, gu — regresja nieważona, gw1 — regresja ważona: tylko ważne efekty potrójnej interakcji, gw2 — regresja ważona: ważne wszystkie efekty: QTL, interakcja epistatyczna i potrójna.

Źródło: Publikacja IV: Cyplik, Bocianowski (2023)

Tabela 14. Fenotypowe i genotypowe oszacowania całkowitego efektu addytywno \times addytywno \times addytywnego, błędy średniokwadratowe i średnie współczynniki determinacji R^2 uzyskane w badaniu symulacyjnym przy założeniach: wariancja błędu równa 10, brak epistazy, równe efekty dla wszystkich trójek

Kod wariantu	Liczba efektów interakcji QTL \times QTL \times QTL	Liczba chromosomów z trójkami QTL	Oszacowania				Błąd średniokwadratowy dla				R^2
			\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	
C01	1	1	18,03	17,09	17,77	17,39	0,44	0,57	0,49	0,73	0,91
C02	1	2	18,11	17,20	17,79	17,45	0,49	0,60	0,54	0,80	0,87
C03	1	3	18,39	17,22	18,19	17,66	0,54	0,77	0,63	0,91	0,95
C04	2	1	17,61	16,89	17,41	17,01	0,64	0,81	0,75	0,94	0,86
C05	2	2	17,88	17,31	17,77	17,52	0,66	0,79	0,69	0,94	0,87
C06	2	3	18,44	17,63	18,02	17,77	0,68	0,91	0,80	0,99	0,88
C07	3	1	16,02	15,54	15,91	15,62	0,77	0,91	0,84	1,02	0,90
C08	3	2	16,51	15,72	16,43	16,17	0,82	0,93	0,89	1,05	0,92
C09	3	3	16,98	16,29	16,55	16,40	0,83	0,93	0,88	1,11	0,93
C10	5	1,2	14,99	14,29	14,70	14,42	0,77	0,89	0,81	0,99	0,87
C11	5	2	15,17	14,40	15,01	14,62	0,83	0,99	0,91	1,12	0,84
C12	5	3	15,52	14,80	15,27	14,99	0,90	1,07	0,97	1,23	0,91

p — metoda fenotypowa, gu — regresja nieważona, gw1 — regresja ważona: tylko ważne efekty potrójnej interakcji, gw2 — regresja ważona: ważne wszystkie efekty: QTL, interakcja epistatyczna i potrójna.

Źródło: Publikacja IV: Cyplik, Bocianowski (2023)

Tabela 15. Fenotypowe i genotypowe oszacowania całkowitego efektu addytywno \times addytywno \times addytywnego, błędy średniokwadratowe i średnie współczynniki determinacji R^2 uzyskane w badaniu symulacyjnym przy założeniach: wariancja błędu równa 10, brak epistazy, różne efekty dla trójek

Kod wariantu	Liczba efektów interakcji QTL \times QTL \times QTL	Liczba chromosomów z trójkami QTL	Oszacowania				Błąd średniokwadratowy dla				R^2
			\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	
C04	2	1	17,51	17,00	17,49	17,23	0,55	0,77	0,61	0,80	0,90
C05	2	2	17,93	17,49	17,60	17,55	0,72	0,91	0,80	0,97	0,87
C06	2	3	18,50	17,69	18,23	18,01	0,79	0,99	0,88	1,07	0,94
C07	3	1	16,09	15,52	15,91	15,79	0,82	0,93	0,87	1,11	0,89
C08	3	2	16,22	15,63	16,09	15,99	0,91	1,21	0,99	1,33	0,86
C09	3	3	16,90	15,99	16,73	16,28	0,92	1,27	1,02	1,44	0,85
C10	5	1,2	15,29	14,47	15,03	14,71	0,88	0,99	0,92	1,03	0,90
C11	5	2	15,31	14,59	15,28	14,92	0,91	1,08	0,97	1,20	0,87
C12	5	3	15,57	14,66	15,02	15,23	0,97	1,22	1,12	1,40	0,88

p — metoda fenotypowa, gu — regresja nieważona, gw1 — regresja ważona: tylko ważne efekty potrójnej interakcji, gw2 — regresja ważona: ważne wszystkie efekty: QTL, interakcja epistatyczna i potrójna.

Źródło: Publikacja IV: Cyplik, Bocianowski (2023)

Tabela 16. Fenotypowe i genotypowe oszacowania całkowitego efektu addytywno \times addytywno \times addytywnego, błędy średniokwadratowe i średnie współczynniki determinacji R^2 uzyskane w badaniu symulacyjnym przy założeniach: wariancja błędu równa 10, obecność epistazy, równe efekty dla wszystkich trójek

Kod wariantu	Liczba efektów interakcji QTL \times QTL \times QTL	Liczba chromosomów z trójkami QTL	Oszacowania				Błąd średniokwadratowy dla				R^2
			\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	
C01	1	1	17,77	16,69	17,42	16,88	0,30	0,41	0,37	0,73	0,89
C02	1	2	17,90	16,79	17,66	16,89	0,42	0,55	0,49	0,79	0,94
C03	1	3	18,19	17,17	17,89	17,23	0,44	0,71	0,59	0,79	0,95
C04	2	1	17,42	16,70	17,11	16,99	0,51	0,77	0,57	0,82	0,88
C05	2	2	17,69	16,88	17,58	17,20	0,52	0,71	0,60	0,80	0,97
C06	2	3	18,08	16,91	17,71	17,44	0,58	0,77	0,66	0,84	0,86
C07	3	1	15,80	15,12	15,53	15,27	0,69	0,80	0,77	0,91	0,90
C08	3	2	16,32	15,44	16,42	16,33	0,73	0,92	0,81	0,99	0,92
C09	3	3	16,83	15,61	16,62	16,51	0,74	0,98	0,89	1,02	0,87
C10	5	1, 2	15,21	14,08	14,63	14,49	0,88	1,00	0,93	1,17	0,86
C11	5	2	15,44	14,32	14,77	14,57	0,92	1,19	1,04	1,31	0,91
C12	5	3	15,70	14,79	15,22	15,01	0,98	1,33	1,07	1,40	0,95

p — metoda fenotypowa, gu — regresja nieważona, gw1 — regresja ważona: tylko ważne efekty potrójnej interakcji, gw2 — regresja ważona: ważne wszystkie efekty: QTL, interakcja epistatyczna i potrójna.

Źródło: Publikacja IV: Cyplik, Bocianowski (2023)

Tabela 17. Fenotypowe i genotypowe oszacowania całkowitego efektu addytywno \times addytywno \times addytywnego, błędy średniokwadratowe i średnie współczynniki determinacji R^2 uzyskane w badaniu symulacyjnym przy założeniach: wariancja błędu równa 10, obecność epistazy, różne efekty dla trójek

Kod wariantu	Liczba efektów interakcji QTL \times QTL \times QTL	Liczba chromosomów z trójkami QTL	Oszacowania				Błąd średniokwadratowy dla				R^2
			\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	
C04	2	1	17,33	16,55	17,02	16,81	0,49	0,73	0,52	0,80	0,94
C05	2	2	17,49	17,01	17,19	16,99	0,55	0,80	0,58	0,92	0,91
C06	2	3	17,66	17,09	17,31	17,24	0,61	0,92	0,77	0,99	0,87
C07	3	1	15,80	15,49	15,76	15,50	0,66	0,88	0,79	0,92	0,85
C08	3	2	15,93	15,52	15,91	15,80	0,70	0,93	0,82	1,02	0,88
C09	3	3	16,99	15,72	16,82	16,55	0,82	1,09	0,97	1,25	0,90
C10	5	1, 2	14,73	14,21	14,58	14,39	0,79	0,98	0,85	1,07	0,95
C11	5	2	14,82	14,39	14,68	14,40	0,85	1,12	0,93	1,30	0,91
C12	5	3	15,01	14,52	14,95	14,77	0,98	1,21	1,07	1,44	0,92

p — metoda fenotypowa, gu — regresja nieważona, gw1 — regresja ważona: tylko ważne efekty potrójnej interakcji, gw2 — regresja ważona: ważne wszystkie efekty: QTL, interakcja epistatyczna i potrójna.

Źródło: Publikacja IV: Cyplik, Bocianowski (2023)

7.1.4.1. Oszacowania na podstawie fenotypu

Wyniki badań symulacyjnych wykazały, że oszacowanie fenotypowe aaa_p było zawsze większe niż 15, prawdziwa zakładana wartość, z wyjątkiem czterech przypadków dla pięciu trójek QTL×QTL×QTL: (1) liczba chromosomów zawierających QTL wynosiła 1, 2, wariancja błędu była równa 5, efekt epistazy był różny od zera, a efekty trójek aaa były różne (Tabela 13), (2) liczba chromosomów zawierających QTL wynosiła 1, 2, wariancja błędu była równa 10, efekt epistazy był równy zero, a efekty wszystkich trójek aaa były równe (Tabela 14), (3) i (4) liczba chromosomów zawierających QTL wynosiła odpowiednio 1, 2 (kod wariantu - C10) i 2 (kod wariantu - C11), wariancja błędu była równa 10, efekt epistazy był różny od zera, a efekty trójek aaa były różne (Tabela 17). Największe wartości aaa_p uzyskano dla jednej trójki QTL×QTL×QTL, natomiast najmniejsze dla pięciu trójek QTL×QTL×QTL (biorąc pod uwagę analogiczny rozkład trójek w grupie sprzężeń). Największe wartości aaa_p uzyskano w sytuacji, gdy QTL tworzące trójkę QTL×QTL×QTL były zlokalizowane na różnych chromosomach, natomiast najmniejsze dla sytuacji, gdy wszystkie trzy QTL tworzące trójkę QTL×QTL×QTL były zlokalizowane w jednej grupie sprzężeń. Oszacowania fenotypowe były większe przy założeniu braku epistazy niż przy założeniu wpływu epistazy na wartości obserwowanej cechy ilościowej. Rozważając wpływ efektów trójek, równych lub różnych, uzyskano podobne wartości oszacowań fenotypowych. Wartości aaa_p były większe, gdy wariancja błędu była równa 10 niż dla wariancji błędu wynoszącej 5, z wyjątkiem sytuacji, gdy założono pięć trójek.

7.1.4.2. Oszacowania na podstawie genotypu

Oszacowania genotypowe efektów aaa były większe niż prawdziwa zakładana wartość 15 przy założeniu wpływu na obserwowaną cechę 1, 2 i 3 trójek QTL×QTL×QTL. Przy założeniu determinacji cechy przez pięć trójek QTL×QTL×QTL, w większości przypadków (58,33%), oceny metodą genotypową były mniejsze niż 15. Przy założonej wariancji błędu równej 5, w 19 przypadkach (51,78%) oszacowania genotypowe były mniejsze niż prawdziwa zakładana wartość aaa równa 15 (Tabele 14-17). Największe wartości estymacji genotypowych (nieważone i ważne) uzyskano dla jednej zakładanej trójki QTL×QTL×QTL, a najmniejsze dla pięciu trójek QTL×QTL×QTL. Największe wartości uzyskano w sytuacji, gdy QTL stanowiące trójki były zlokalizowane na różnych chromosomach, z wyjątkiem przypadku, gdy założono

pięć trójek QTL×QTL×QTL o różnych efektach przy braku epistazy i wariancji błędu 10 (Tabela 16). W przeciwieństwie do tego, najmniejsze wartości *aaa* otrzymano dla przypadków, gdy wszystkie trzy QTL tworzące trójki zostały umieszczone w jednej grupie sprzężeń. W większości przypadków oszacowania genotypowe były większe dla różnych efektów niż dla równych efektów trójek. Estymacje genotypowe były większe przy założonej większej (10) wariancji błędu niż przy mniejszej (5) wariancji, w sytuacjach, gdy cecha ilościowa była determinowana przez 1, 2 lub 3 trójki QTL×QTL×QTL. W przeciwieństwie do tego, przy założeniu pięciu trójek QTL×QTL×QTL, wartości estymatorów otrzymane metodą genotypową były większe dla wariancji błędu 5 niż dla wariancji błędu 10.

7.1.4.3. Oszacowania fenotypowe a genotypowe

Różnice między oszacowaniami fenotypowymi a genotypowymi dla regresji nieważonej były zawsze dodatnie i były ogólnie mniejsze, gdy zakładano różne efekty trójek QTL×QTL×QTL. Zależności $\overline{aaa}_p > \overline{aaa}_{gw1} > \overline{aaa}_{gw2} > \overline{aaa}_{gu}$ nie wystąpiły jedynie w trzech przypadkach: (1) gdy dwie trójki QTL×QTL×QTL były zlokalizowane na trzech chromosomach, dla błędu wariancji równego 5, braku epistazy i różnych efektów trójek ($\overline{aaa}_{gw1} = \overline{aaa}_{gw2} = 17.99$, Tabela 11), (2) gdy trzy trójki QTL×QTL×QTL były zlokalizowane w dwóch grupach sprzężeń, błąd wariancji równy 5, brak epistazy i różne efekty trójek ($\overline{aaa}_{gw1} > \overline{aaa}_p$, Tabela 13) oraz (3) gdy trzy trójki QTL×QTL×QTL o równych efektach były zlokalizowane w dwóch grupach sprzężeń, a efekt epistazy był różny od zera: w którym to przypadku $\overline{aaa}_{gw1} > \overline{aaa}_{gw2} > \overline{aaa}_p$ (Tabela 16). Wartości estymatorów otrzymane przy zastosowaniu regresji ważonej były zawsze bliższe oszacowaniom fenotypowym niż wartości *aaa* otrzymane przy zastosowaniu regresji nieważonej. Metoda ważona (wariant I selekcji markerów) wykazała najlepsze (najbliższe prawdziwej zakładanej wartości 15) oceny trójek *aaa*.

7.1.4.4. Błędy średniokwadratowe

Ogólnie, degresji wartości estymatorów towarzyszył wzrost ich błędu średniokwadratowego (Tabele 10-17). Ogólnie, zaobserwowano zależności błędu średniokwadratowego dla ocen $\widehat{aaa}_p < \widehat{aaa}_{gw1} < \widehat{aaa}_{gu} < \widehat{aaa}_{gw2}$, z wyjątkiem sytuacji, gdy zakładano, że dwie trójki QTL×QTL×QTL o różnych efektach były zlokalizowane w dwóch grupach sprzężeń, w przypadku braku efektu epistatycznego

i błąd wariancji równego 5. W tym przypadku błąd średniokwadratowy dla $\widehat{aaa}_{gw1} > \widehat{aaa}_{gw2}$ (Tabela 11). Błąd średniokwadratowy był większy przy założonej wariancji błędu wynoszącej 10 niż przy mniejszej założonej wariancji błędu. Wartości efektu zakładanych trójek (równe lub różne) nie miały wpływu na wartości błędu średniokwadratowego. Z drugiej strony, założenie obecności lub braku efektu epistazy wpłynęło na wartości błędu średniokwadratowego dla \widehat{aaa} : większe wartości błędu średniokwadratowego uzyskano przy założeniu braku epistazy. Najmniejsze błędy średniokwadratowe uzyskano, gdy cecha była determinowana przez jedną trójkę QTL×QTL×QTL, natomiast największe uzyskano przy założeniu działania pięciu trójek QTL×QTL×QTL. Ogólnie, zaobserwowano zależności wartości błędu średniokwadratowego dla liczby założonych trójek: $1QQQ < 2QQQ < 3QQQ < 5QQQ$. Najmniejszy błąd średniokwadratowy uzyskano dla trójek QTL×QTL×QTL zlokalizowanych w jednej grupie sprzężeń, a największy uzyskano, gdy trójki QTL×QTL×QTL były zlokalizowane w trzech chromosomach.

7.1.4.5. Współczynniki determinacji

Wariancja wyjaśniona przez potrójne interakcje QTL×QTL×QTL wynosiła od 83% (Tabela 11 i 13) do 97% (Tabela 16). Współczynniki determinacji dla jednej potrójnej interakcji QTL×QTL×QTL wynosiły od 0,87 do 0,95, ze średnią 0,91. Współczynniki determinacji dla dwóch założonych trójek QTL×QTL×QTL wynosiły od 0,83 do 0,97, ze średnią 0,89. Współczynniki determinacji dla trzech i pięciu zakładanych trójek QTL×QTL×QTL były podobne i wynosiły od 0,85 i 0,84 do 0,95 i 0,95, ze średnią 0,90 i 0,90, odpowiednio, dla trzech i pięciu zakładanych trójek QTL×QTL×QTL.

7.2. Dyskusja

Celem procesu hodowlanego jest uzyskanie nowych genotypów o cechach ulepszonych w stosunku do form rodzicielskich (Cullis, Smith, Cocks, Butler, 2020). Parametry związane z efektem addytywnego działania genów, jak również ich interakcjami (takimi jak efekt epistazy addytywno \times addytywnej interakcji gen \times gen i efekt addytywno \times addytywno \times addytywny fekt interakcji gen \times gen \times gen) mogą wpływać na decyzje dotyczące przydatności materiału hodowlanego (Ali i in., 2020; Bocianowski, Warzecha, Nowosad, Bathelt, 2019; Bouvet, Makouanzi, Cros, Vigneron, 2016; Labroo, Studer, Rutkoski, 2021; Raffo i in., 2022; Voss-Fels, Cooper, Hayes, 2019; Wang i in., 2012). Zrozumienie architektury genetycznej złożonych cech jest głównym wyzwaniem w erze postgenomicznej, szczególnie w przypadku efektów QTL, interakcji QTL \times QTL (Yang, Zhu, Williams, 2007) i interakcji QTL \times QTL \times QTL.

Analityczne i numeryczne porównanie metod oszacowania całkowitych efektów interakcji *aaa* przedstawiono w Publikacji I. Porównanie numeryczne przeprowadzono na 153 zestawach obserwacji uzyskanych z dwóch populacji linii podwojonych haploidów jęczmienia.

Porównanie analityczne wykazało, że przy założeniu prawidłowej segregacji i braku sprzężeń między markerami, równania na estymatory fenotypowe i genotypowe są podobne. Ponadto porównanie wykazało, że efekt interakcji *aaa* każdej trójki QTL jest mniejszy niż efekt fenotypowy.

Numeryczne porównanie estymacji efektu interakcji *aaa* pokazało, że w większości przypadków (79%) genotypowe oszacowanie interakcji *aaa* było mniejsze niż fenotypowe. Wynika to z faktu, że ocena fenotypowa dotyczy sumy efektów interakcji *aaa* wszystkich trójek genów, w przeciwieństwie do oceny genotypowej, które uwzględnia sumę efektów tylko wybranych trójek genów. W pozostałych przypadkach (21%), gdy wartości oszacowań fenotypowych były mniejsze niż genotypowych, może to być wynikiem wysokiej różnorodności genetycznej przy mniejszej różnorodności fenotypowej linii DH. Wysokie zakresy różnic dla wartości estymatorów były najprawdopodobniej wynikiem wielu różnych wariantów eksperymentalnych, składających się z różnych cech, środowisk i sytuacji eksperymentalnych (Bocianowski, Krajewski, 2009). Liczba czynników efektywnych obliczona na podstawie obserwacji fenotypowych nie wpływała bezpośrednio na liczbę

istotnych markerów, jak również na liczbę interakcji *aaa* uwzględnionych w modelach genotypowych. Zarówno liczba czynników efektywnych, jak i liczba markerów uwzględnionych w modelu, były zgodne (w większości przypadków), co znajduje swoje uzasadnienie w przedstawionej metodzie, która miała na celu uwzględnienie maksymalnej liczby najlepiej dopasowanych czynników. Z kolei liczba interakcji *aaa* znacznie różniła się w poszczególnych modelach, co może być wynikiem pominięcia markerów, które same w sobie nie poprawiają modelu, ale mogą tworzyć trójki o istotnych efektach interakcyjnych. Markery te charakteryzują się małym efektem głównym, który sam w sobie nie determinuje istotnie obserwowanej cechy, lub jego brakiem.

Zastosowana krokowa selekcja parametrów według kryteriów informacyjnych Akaike pozwoliła uzyskać podobne wyniki do wyników otrzymanych w pracy wykorzystującej te same zbiory danych do estymacji epistazy, w której zastosowano regresję krokową wsteczną (Bocianowski, 2012b), a także do wyników estymacji interakcji podwójnych otrzymanych metodą mapowania interwałowego (ICIM) przedstawionych przez Li, Ribaut, Li, Wang (2008). Przedstawione wyniki pokazały, że uwzględnienie interakcji wyższego rzędu (*aaa*) w modelu regresji wielokrotnej może mieć znaczący wpływ na oszacowanie efektu QTL.

Ważnym założeniem przyjętym w badaniach jest fakt, że efekty interakcji *aaa* pokazują tylko loci związane z markerami o istotnych efektach głównych. Uwzględnienie dodatkowych markerów może ujawnić dodatkowe interakcje, jednak przy znacznym wzroście wymagań dotyczących ilości danych (Manolio i in., 2009).

Wyniki zawarte w drugiej publikacji (Cyplik, Sobiech, Tomkowiak, Bocianowski, 2022) pokazały, że w 67% przypadków całkowity efekt potrójnej interakcji, a w 58% przypadków efekt interakcji epistatycznych były istotne statystycznie. Wyniki te podkreślają znaczenie uwzględnienia interakcji wyższych rzędów w podobnych badaniach. Wysokie zakresy oszacowań miały swoje uzasadnienie w różnych sytuacjach eksperymentalnych, na które wpływ mogły mieć lokalizacje, pochodzenie czy warunki uprawy roślin (Bocianowski, Krajewski, 2009). Rodzaj wpływu, pozytywny lub negatywny, zarówno podwójnych, jak i potrójnych efektów interakcji był często analogiczny. Prowadzi to do wniosku, że kompleksowa analiza efektów QTL (addytywnych, interakcji epistatycznych *aa* i interakcji potrójnych *aaa*) jest niezbędna

dla celów programu hodowlanego. Wnioski płynące z Publikacji II dotyczące analizy fenotypowej są analogiczne do tych uzyskanych w Publikacji I i jednocześnie stanowią ich potwierdzenie przy wykorzystaniu innego zestawu danych eksperymentalnych.

W Publikacji I wykazano, że oszacowania efektów interakcji QTL×QTL×QTL uzyskane na podstawie obserwacji markerów molekularnych (przy użyciu tradycyjnej regresji nieważonej) miały mniejsze wartości bezwzględne niż te, które były oparte na obserwacjach wyłącznie fenotypowych. Uzyskane różnice były na tyle duże, że w Publikacji III zaproponowano nową metodę mającą na celu poprawę estymacji parametru związanego z interakcją QTL×QTL×QTL, którą było zastosowanie regresji ważonej.

Numeryczne porównanie trzech metod oszacowania efektów interakcji *aaa*, tj. metody fenotypowej i dwóch metod genotypowych (regresji nieważonej i regresji ważonej), przeprowadzono na 94 liniach podwojonych haploidów (DH) uzyskanych poprzez skrzyżowanie genotypów pszenicy heksaploidalnej (*Triticum aestivum* L.) Chinese Spring (CS) i SQ1 (linia hodowlana o wysokiej zawartości kwasu abscysynowego) (Quarrie i in. 2005). Analizie poddano oddzielnie grupy testowe i grupy kontrolne obserwowane przez trzy lata (2010-2012). Do wyznaczenia trójek QTL i estymacji efektów ich interakcji QTL×QTL×QTL zastosowano regresję ważoną, podczas gdy większość opublikowanych badań koncentruje się jedynie na analizie pojedynczych genów jako determinantów cech ilościowych (Botero-Ramírez i in., 2020; Gacek i in., 2021; Kabange i in., 2020; Kwon i in., 2021; Lephuthing i in., 2022; Pegot-Espagnet i in., 2019; Ren i in., 2022; Sayed i in., 2021). Regresja ważona była wcześniej zastosowana do oszacowania interakcji epistazy QTL×QTL (Bocianowski, Nowosad, 2015; Ful, Ritland, 1996; Nap, Conner, Mlynárová, Stiekema, Jansen, 1997; Routman, Cheverud, 1997; Slim, Chatelain, Azencott, Vert, 2020), jednak do tej pory interakcje wyższego rzędu były pomijane. Dlatego innowacyjność zaproponowanej metody polega na określeniu sposobu wyboru trójek QTL i oszacowaniu ich efektów za pomocą regresji ważonej.

Wyniki wykazały, że wykryte QTL×QTL×QTL pozytywnie wpłynęły na dokładność oszacowania estymatorów *aaa* dla badanych cech. Ważny aspekt stanowi założenie, że efekty interakcji *aaa* wykazują tylko loci całkowicie sprzężone

z markerami o istotnych efektach. Należy jednak podkreślić, że selekcja markerów o istotnych efektach została przeprowadzona przy zastosowaniu regresji ważonej.

Zastosowanie regresji nieważonej przyniosło poprawę (co do wartości bezwzględnej) efektu aaa_{gu} w porównaniu z oszacowaniami bazującymi na obserwacjach tylko fenotypach w trzech przypadkach (grupa testowa w 2010 r., grupa kontrolna w 2012 r. i grupa testowa w 2012 r.). Natomiast regresja ważona spowodowała poprawę (co od wartości bezwzględnej) oceny parametru aaa_{gw} w porównaniu do aaa_p w pięciu przypadkach, z wyjątkiem grupy testowej w 2012 r. (Tabela 7 i Tabela 8).

Uwzględnienie wpływu potrójnej interakcji na ekspresję cech ilościowych może być również istotne w kontekście hipostazy, czyli sytuacji, w której ekspresja genu jest maskowana przez działanie innego genu lub przez efekt interakcji epistatycznej (Rieger, Michaelis, Green, 2012). Nieuwzględnienie wpływu interakcji wyższego rzędu w genetyce ilościowej może skutkować niedokładnym oszacowaniem cech ilościowych. W Publikacji II badając linie wsobne kukurydzy, zaobserwowano istotność potrójnej interakcji przy braku efektu epistazy. Zjawisko to wystąpiło w trzech sytuacjach: dla średnicy kolby (dwukrotnie) i dla długości rdzenia. Odwrotna sytuacja (istotna epistaza i nieistotna potrójna interakcja) nie wystąpiła w żadnym przypadku.

Symulacje Monte Carlo są narzędziem badawczym bardzo szeroko wykorzystywanym do rozwiązywania różnych problemów poprzez uzyskiwanie przybliżonych wyników. Szacowanie efektów epistazy z wykorzystaniem badań symulacyjnych zostało przedstawione na przykład przez Bocianowskiego (2012a), Ahsan i in. (2019), de los Campos, Sorensen, Toro (2019), Wang i in. (2020) oraz Sharma, Sadeghian Tehrani, Kumar, Song (2022). Porównanie metod oszacowania addytywnych efektów interakcji QTL×QTL×QTL za pomocą badań symulacyjnych Monte Carlo, nie zostało dotychczas przeprowadzone. W Publikacji IV opisano, po raz pierwszy, badania symulacyjne porównujące metody estymacji całkowitego efektu aaa interakcji $gen \times gen \times gen$. Do oszacowania efektów interakcji QTL×QTL×QTL zastosowano ważoną wielokrotną regresję liniową, w dwóch zaproponowanych wariantach (rozdziały 6.1.2.1. oraz 6.3.).

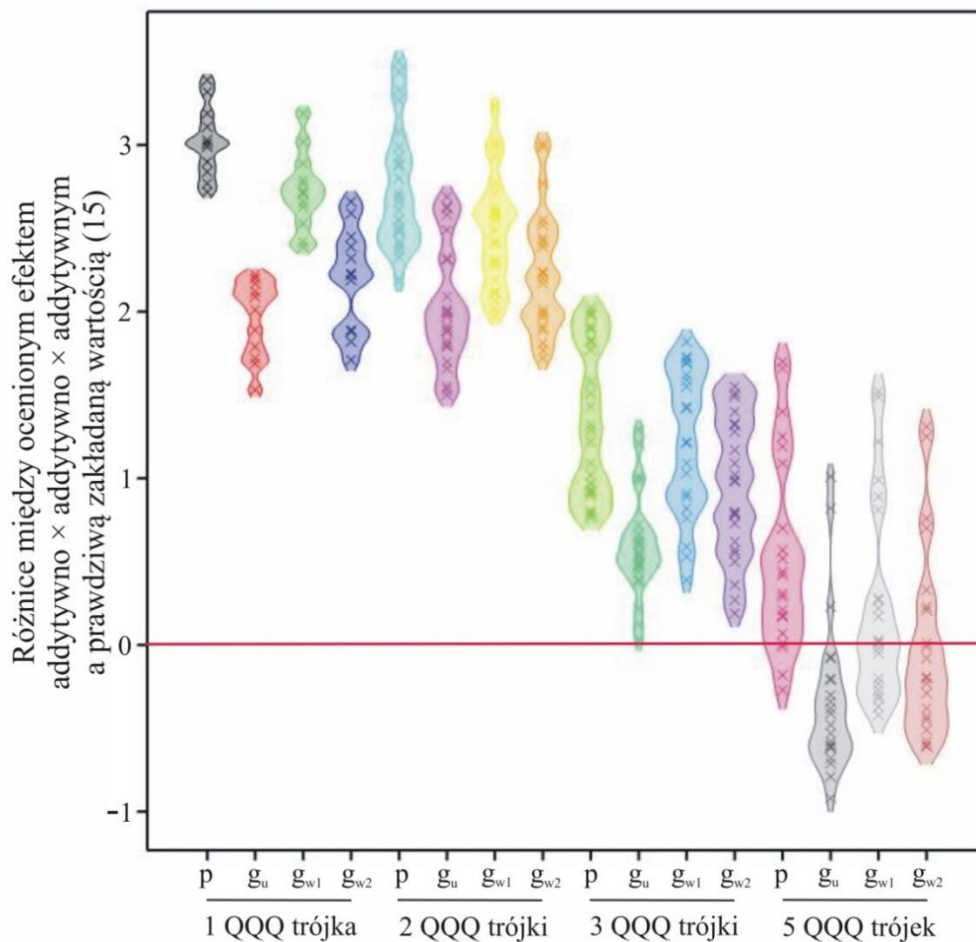
Badania symulacyjne Monte Carlo przeprowadzone w celu porównania metod estymacji nie uwzględniały wszystkich możliwych sytuacji eksperymentalnych.

Symulacje nie mogły w pełni odzwierciedlić złożoności interakcji genetycznych w świecie rzeczywistym. Wyniki należy więc interpretować z ostrożnością aż do czasu przeprowadzenia dalszej walidacji empirycznej. Jednakże kombinacje parametrów użyte w przedstawionych analizach odpowiadają przypadkom najczęściej spotykanym w rzeczywistych badaniach QTL (Avery i in., 2019; Balestre, de Souza, 2016; Berke, Rocheford, 1999; Burke, Tang, Knapp, Rieseberg, 2002; Chattopadhyay i in., 2021; Crawford, Zeng, Mukherjee, Zhou, 2017; De Leon, Pruthi, Jampala, Borjas, Subudhi, 2020; Flint-Garcia, Jampatong, Darrah, McMullen, 2003; Goutelle i in., 2009; Johnson i in., 2021; Rajurkar, Muthukumar, Ayyenar, Thomas, Chandra Babu, 2021; Ronin, Korol, Nevo, 1999; Shi i in., 2020; Sorojsrisom, Haller, Ambrose, Eaton, 2022; Viana, Souza, 2023; Wang, Xie, Li, Zhu, Xing, 2022; Wu, 1999; Zimnoch-Guzowska i in., 2000).

W badaniach symulacyjnych liczbę analizowanych linii najczęściej przyjmuje się na poziomie od 500 do 2000 (Ronin i in., 1999). Viana, Souza (2023) przeprowadzili analizę 400 roślin, genotypowanych dla 975 polimorfizmów pojedynczego nukleotydu (ang. *Single-Nucleotide Polimorphism*, SNP) rozmieszczonych na 10 chromosomach o długości 100 cM. Tysiąc markerów SNP uwzględnili w swoich badaniach Crawford i in. (2017). Występowanie dziesięciu chromosomów jest charakterystyczne dla takich gatunków jak: *Anthoxanthum odoratum*, *Brassica campestris*, *Sorghum sudanense*, *Zea mize*. W praktycznych doświadczeniach liczba genów zlokalizowanych na poszczególnych chromosomach jest bardzo zróżnicowana (Berke, Rocheford, 1999; Burke i in., 2002; Flint-Garcia i in., 2003; Zimnoch-Guzowska i in., 2000). Liczba QTL założona w niniejszym badaniu wynosiła 14. Taką samą liczbę QTL przedstawili w swoich badaniach Wu (1999) oraz Balestre, de Souza (2016). Założona liczba pięciu par epistatycznych była wynikiem zaobserwowanym w wielu doświadczeniach praktycznych (Chattopadhyay i in., 2021; De Leon i in., 2020; Rajurkar i in., 2021; Shi i in., 2020). Badania symulacyjne bazujące na 10000 generacji zostały również przedstawione m.in. przez Goutelle i in. (2009), Avery i in. (2019), Johnson i in. (2021), Sorojsrisom i in. (2022) oraz Wang i in. (2022). Kombinacje parametrów przyjęte w prezentowanych badaniach symulacyjnych reprezentowały 84 różne sytuacje eksperymentalne. Wyniki uzyskane z badań symulacyjnych Monte Carlo wskazały na stabilność właściwości prezentowanych metod oszacowania efektu interakcji *aaa* na różnych typach materiału genetycznego. Niewielki wpływ wariancji błędu na oszacowanie efektów interakcji *aaa* wszystkimi czterema metodami, a także na wnioski

dotyczące porównania proponowanych metod szacowania, wskazuje na dobre perspektywy zastosowania przedstawionych wniosków dla różnych gatunków roślin.

Stwierdzono, że liczba zakładanych trójek QTL×QTL×QTL, ich położenie na jednym, dwóch lub trzech chromosomach oraz ich efekty (nierówne lub równe) wpłynęły na wartości estymatorów i ich porównanie. Oszacowanie metodą fenotypową zmniejszało się w kierunku prawdziwej wartości wraz ze wzrostem liczby zakładanych trójek QTL×QTL×QTL (Rysunek 4).



Rysunek 4. Różnice między ocenionym efektem addytywno × addytywno × addytywnym a prawdziwą zakładaną wartością (15) dla poszczególnych metod estymacji i różnej liczby zakładanych trójek QTL×QTL×QTL, gdzie p — metoda fenotypowa, g_u — regresja nieważona, g_{w1} — regresja ważona: ważne tylko efekty potrójnej interakcji, g_{w2} — regresja ważona: ważne wszystkie efekty: QTL, interakcja epistatyczna i potrójna.

Źródło: Publikacja IV: Cyplik, Bocianowski (2023)

W przypadku największej liczby trójek QTL×QTL×QTL wartości aaa_p mogły być mniejsze niż zakładana wartość prawdziwa, ponieważ kombinacje genotypów generujące skrajne linie nie były odpowiednio reprezentowane w symulowanej próbie. Estymacja fenotypowa była bliższa wartości prawdziwej również wtedy, gdy liczba chromosomów z trójkami QTL×QTL×QTL zmniejszała się, tj. gdy nastąpiło nagromadzenie sprzężonych efektów indywidualnych. Podobnie, ocena fenotypowa dążyła do prawdziwej wartości, gdy zakładany efekt epistazy był różny od zera. Oszacowanie fenotypowe było zawyżone dla większej zakładanej wariancji błędu. Estymacja genotypowa zmniejszała swoją wartość w bardzo podobny sposób do fenotypowej. Różnica między oszacowaniem fenotypowym i genotypowym była najmniejsza, gdy założono pięć trójek QTL×QTL×QTL, czyli gdy zastosowano model zbliżony do prawdziwie poligenowego.

Różnica między fenotypowymi i genotypowymi oszacowaniami całkowitego efektu aaa , obserwowana zarówno w porównaniach numerycznych (Publikacja I), jak i symulacyjnych, jest częściowo spowodowana faktem, że metody genotypowe nie znajdują wszystkich trójek gen×gen×gen, które determinują cechę. Jednak zastosowanie regresji ważonej pokazuje, że różnica między szacunkami fenotypowymi i genotypowymi może zostać zmniejszona, a całkowite efekty genetyczne będą bliższe prawdziwej wartości. Podobne wyniki uzyskano oszacowując efekty potrójnej interakcji QTL×QTL×QTL w populacji mapującej linii podwojonych haploidów pszenicy (Publikacja III).

Jako bezpośrednia konsekwencja interakcji, w szczególności zaangażowania QTL w interakcje epistatyczne (Bocianowski, 2013) i interakcje wyższego rzędu, efekty pojedynczych loci QTL zależą od genotypów innych loci. Jak pokazują przedstawione tu wyniki, uwzględnienie obecności efektów epistazy QTL×QTL miało wpływ na ocenę efektu potrójnej interakcji QTL×QTL×QTL. Brak epistazy zakładany w badaniach symulacyjnych można rozpatrywać na dwa sposoby: (1) jako rzeczywisty brak efektu epistazy lub (2) jako sumaryczny zerowy efekt par QTL×QTL z efektami epistatycznymi poszczególnych par różnymi od zera. Próby wykorzystania QTL w programach hodowlanych muszą uwzględniać nie tylko efekty epistatyczne, ale także interakcje wyższego rzędu. Określenie udziału potrójnych interakcji QTL×QTL×QTL jest ważne dla zrozumienia genetycznych podstaw złożonych cech, dlatego modele genetyczne do mapowania QTL zakładające brak potrójnej interakcji QTL×QTL×QTL

mogą prowadzić do niepełnego oszacowania parametrów QTL i epistazy QTL×QTL (Bocianowski, 2013). Wskazują na to wyniki przedstawione w Publikacji II, gdzie analizowano parametry genetyczne dla wybranych cech linii wsobnych kukurydzy. Chociaż podejście przedstawione w Publikacji IV opiera się między innymi na mapie genetycznej, proponowane metody estymacji addytywnych efektów interakcji QTL×QTL×QTL mogą być stosowane w mapowaniu asocjacyjnym, jako uzupełnienie analizy efektów głównych genów i epistazy (Rakoczy-Trojanowska i in., 2017a, 2017b; Tomkowiak i in., 2019, 2021).

W sumie QTL z efektami addytywnymi, *aa* efektami epistatycznymi QTL×QTL i *aaa* efektami potrójnej interakcji QTL×QTL×QTL wyjaśniły ponad 83% zmienności fenotypowej. Informacje uzyskane w tym badaniu będą przydatne do zastosowania QTL w hodowli roślin poprzez selekcję wspomaganą markerami.

8. Podsumowanie

Na podstawie wyników badań zaprezentowanych w rozprawie doktorskiej zaproponowano następujące wnioski:

1. Wszystkie przedstawione metody były użytecznymi narzędziami statystycznymi do charakterystyki QTL i oszacowania interakcji addytywno \times addytywno \times addytywnej *aaa*.
2. Wyznaczanie efektów epistatycznych umożliwia lepsze zrozumienie interakcji poszczególnych genów i pozwala na bardziej precyzyjne estymowanie efektów poszczególnych genów.
3. Efekty interakcji epistatycznych *aa* i efekty interakcji *aaa* są często podobne zarówno pod względem rodzaju wpływu (pozytywny lub negatywny), jak i samej wartości oszacowania.
4. Zastosowanie regresji ważonej, niezależnie od zaproponowanego wariantu, powoduje, że uzyskane estymacje są bliższe tym uzyskanym metodą fenotypową, co stanowi poprawę w porównaniu do oszacowań uzyskanych metodą regresji nieważonej.
5. W badaniach praktycznych należy oczekiwać, że oszacowania całkowitego efektu *aaa* interakcji gen \times gen \times gen będą mniejsze niż całkowite oszacowanie na podstawie jedynie fenotypu. Stwierdzenie innej sytuacji, należy dokładnie przeanalizować, sprawdzając, czy spełnione są założenia genetyczne dotyczące prawidłowej segregacji i braku sprzężeń między markerami.
6. Wyniki przedstawione w omówionych publikacjach stanowią pierwsze doniesienia dotyczące wykorzystania metod analitycznych, numerycznych i symulacyjnych estymacji interakcji *aaa* efektów QTL \times QTL \times QTL.
7. Interakcje wyższego rzędu, choć powszechnie pomijane w badaniach, mają istotny wpływ na cechy fenotypowe. Nieuwzględnienie efektu interakcji wyższego rzędu w genetyce ilościowej może skutkować znacznym niedoszacowaniem addytywnych efektów QTL.
8. Konieczne są dalsze badania nad interakcjami wyższego rzędu oraz metodami ich oszacowania.

9. Literatura

1. Ahsan, A., Monir, M., Meng, X., Rahaman, M., Chen, H., Chen, M. (2019). Identification of epistasis loci underlying rice flowering time by controlling population stratification and polygenic effect. *DNA Research*, 26(2), 119–130. <https://doi.org/10.1093/dnares/dsy043>
2. Akaike, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle. In: E. Parzen, K. Tanabe, G. Kitagawa (Red.), *Selected Papers of Hirotugu Akaike* (s. 199–213). Springer. https://doi.org/10.1007/978-1-4612-1694-0_15
3. Ali, F., Chen, W., Faiz, S., Wang, Y., Wei, X., Xie, L., Jiao, G., Shao, G., Hu, S., Tang, S., Sheng, Z., Hu, P. (2022). QTL mapping for grain appearance quality traits using doubled haploid population of rice under different environments. *Pakistan Journal of Botany*, 54(4), 1365–1275. [https://doi.org/10.30848/PJB2022-4\(4\)](https://doi.org/10.30848/PJB2022-4(4))
4. Ali, M., Zhang, L., DeLacy, I., Arief, V., Dieters, M., Pfeiffer, W.H., Wang, J., Li, H. (2020). Modeling and simulation of recurrent phenotypic and genomic selections in plant breeding under the presence of epistasis. *The Crop Journal*, 8(5), 866–877. <https://doi.org/10.1016/j.cj.2020.04.002>
5. Avery, L., Rotondi, N., McKnight, C., Firestone, M., Smylie, J., Rotondi, M. (2019). Unweighted regression models perform better than weighted regression techniques for respondent-driven sampling data: Results from a simulation study. *BMC Medical Research Methodology*, 19(1), 202. <https://doi.org/10.1186/s12874-019-0842-5>
6. Balestre, M., de Souza, C.L. (2016). Bayesian reversible-jump for epistasis analysis in genomic studies. *BMC Genomics*, 17(1), 1012. <https://doi.org/10.1186/s12864-016-3342-6>
7. Bateson, W., Mendel, G. (1902). *Mendel's Principles of Heredity: A Defence, with a Translation of Mendel's Original Papers on Hybridisation*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511694462>

8. Bautista, D.E., Carr, J.F., Mitchell, A.M. (2021). Suppressor Mutants: History and Today's Applications. *EcoSal Plus*, 9(2), eESP-0037-2020. <https://doi.org/10.1128/ecosalplus.ESP-0037-2020>
9. Beheshtizadeh, H., Fakheri, B.A., Aghnoum, R., Mahdinezhad, N., Pourdad, S.S., Masoudi, B. (2018). QTL mapping of grain yield and its components under normal and drought stress conditions in barley (*Hordeum vulgare* L.). *Indian Journal of Genetics and Plant Breeding*, 78(01), 01.
10. Berke, T.G., Rocheford, T.R. (1999). Quantitative Trait Loci for Tassel Traits in Maize. *Crop Science*, 39(5), 1439–1443. <https://doi.org/10.2135/cropsci1999.3951439x>
11. Bocianowski, J. (2012a). A comparison of two methods to estimate additive-by-additive interaction of QTL effects by a simulation study. *Journal of Theoretical Biology*, 308, 20–24. <https://doi.org/10.1016/j.jtbi.2012.05.018>
12. Bocianowski, J. (2012b). Analytical and numerical comparisons of two methods of estimation of additive \times additive interaction of QTL effects. *Scientia Agricola*, 69, 240–246. <https://doi.org/10.1590/S0103-90162012000400002>
13. Bocianowski, J. (2012c). The use of weighted multiple linear regression to estimate QTL-by-QTL epistatic effects. *Genetics and Molecular Biology*, 35, 802–809. <https://doi.org/10.1590/S1415-47572012005000071>
14. Bocianowski, J. (2013). Epistasis interaction of QTL effects as a genetic parameter influencing estimation of the genetic additive effect. *Genetics and Molecular Biology*, 36(1), 93–100. <https://doi.org/10.1590/S1415-47572013000100013>
15. Bocianowski, J., Kozak, M., Liersch, A., Bartkowiak-Broda, I. (2011). A heuristic method of searching for interesting markers in terms of quantitative traits. *Euphytica*, 181(1), 89–100. <https://doi.org/10.1007/s10681-011-0424-z>
16. Bocianowski, J., Krajewski, P. (2009). Comparison of the genetic additive effect estimators based on phenotypic observations and on molecular marker data. *Euphytica*, 165, 113–122. <https://doi.org/10.1007/s10681-008-9770-x>

17. Bocianowski, J., Nowosad, K. (2015). Mixed linear model approaches in mapping QTLs with epistatic effects by a simulation study. *Euphytica*, 202(3), 459–467. <https://doi.org/10.1007/s10681-014-1329-4>
18. Bocianowski, J., Warzecha, T., Nowosad, K., Bathelt, R. (2019). Genotype by environment interaction using AMMI model and estimation of additive and epistasis gene effects for 1000-kernel weight in spring barley (*Hordeum vulgare* L.). *Journal of Applied Genetics*, 60(2), 127–135. <https://doi.org/10.1007/s13353-019-00490-2>
19. Borlaug, N.E. (1983). Contributions of Conventional Plant Breeding to Food Production. *Science*, 219(4585), 689–693. <https://doi.org/10.1126/science.219.4585.689>
20. Botero-Ramírez, A., Laperche, A., Guichard, S., Jubault, M., Gravot, A., Strelkov, S.E., Manzanares-Dauleux, M.J. (2020). Clubroot Symptoms and Resting Spore Production in a Doubled Haploid Population of Oilseed Rape (*Brassica napus*) Are Controlled by Four Main QTLs. *Frontiers in Plant Science*, 11, 604527. <https://www.frontiersin.org/articles/10.3389/fpls.2020.604527>
21. Bouvet, J.-M., Makouanzi, G., Cros, D., Vigneron, P. (2016). Modeling additive and non-additive effects in a hybrid population using genome-wide genotyping: Prediction accuracy implications. *Heredity*, 116(2), 2. <https://doi.org/10.1038/hdy.2015.78>
22. Bradshaw, J.E. (2017). Plant breeding: Past, present and future. *Euphytica*, 213(3), 60. <https://doi.org/10.1007/s10681-016-1815-y>
23. Burke, J.M., Tang, S., Knapp, S.J., Rieseberg, L.H. (2002). Genetic Analysis of Sunflower Domestication. *Genetics*, 161(3), 1257–1267. <https://doi.org/10.1093/genetics/161.3.1257>
24. Carlborg, Ö., Andersson, L., Kinghorn, B. (2000). The Use of a Genetic Algorithm for Simultaneous Mapping of Multiple Interacting Quantitative Trait Loci. *Genetics*, 155(4), 2003–2010. <https://doi.org/10.1093/genetics/155.4.2003>
25. Ceccarelli, S., Grando, S., Maatougui, M., Michael, M., Slash, M., Haghparast, R., Rahmanian, M., Taheri, A., Al-Yassin, A., Benbelkacem, A., Labdi, M., Mimoun, H., Nachit, M. (2010). Plant breeding and climate changes. *The Journal*

- of *Agricultural Science*, 148(6), 627–637.
<https://doi.org/10.1017/S0021859610000651>
26. Chase, K., Adler, F.R., Lark, K.G. (1997). Epistat: A computer program for identifying and testing interactions between pairs of quantitative trait loci. *Theoretical and Applied Genetics*, 94(6), 724–730.
<https://doi.org/10.1007/s001220050471>
 27. Chattopadhyay, K., Mohanty, S.K., Vijayan, J., Marndi, B.C., Sarkar, A., Molla, K.A., Chakraborty, K., Ray, S., Sarkar, R.K. (2021). Genetic Dissection of Component Traits for Salinity Tolerance at Reproductive Stage in Rice. *Plant Molecular Biology Reporter*, 39(2), 386–402. <https://doi.org/10.1007/s11105-020-01257-4>
 28. Choo, T.M., Reinbergs, E. (1982). Analyses of Skewness and Kurtosis for Detecting Gene Interaction in a Doubled Haploid Population. *Crop Science*, 22(2), 231–235. <https://doi.org/10.2135/cropsci1982.0011183X002200020008x>
 29. Civile, G.V., Oftedal, K.N. (2012). Sensory evaluation techniques—Make “good for you” taste “good”. *Physiology & Behavior*, 107(4), 598–605. <https://doi.org/10.1016/j.physbeh.2012.04.015>
 30. Clarke, E.J., Wiseman, J. (2000). Developments in plant breeding for improved nutritional quality of soya beans I. Protein and amino acid content. *The Journal of Agricultural Science*, 134(2), 111–124.
<https://doi.org/10.1017/S0021859699007431>
 31. Cobb, J.N., Biswas, P.S., Platten, J.D. (2019). Back to the future: Revisiting MAS as a tool for modern plant breeding. *Theoretical and Applied Genetics*, 132(3), 647–667. <https://doi.org/10.1007/s00122-018-3266-4>
 32. Collinge, D.B., Sarrocco, S. (2022). Transgenic approaches for plant disease control: Status and prospects 2021. *Plant Pathology*, 71(1), 207–225. <https://doi.org/10.1111/ppa.13443>
 33. Crawford, L., Zeng, P., Mukherjee, S., Zhou, X. (2017). Detecting epistasis with the marginal epistasis test in genetic mapping studies of quantitative traits. *PLOS Genetics*, 13(7), e1006869. <https://doi.org/10.1371/journal.pgen.1006869>

34. Cullis, B.R., Smith, A.B., Cocks, N.A., Butler, D.G. (2020). The Design of Early-Stage Plant Breeding Trials Using Genetic Relatedness. *Journal of Agricultural, Biological and Environmental Statistics*, 25(4), 553–578. <https://doi.org/10.1007/s13253-020-00403-5>
35. Czyczyło-Mysza, I., Tyrka, M., Marcińska, I., Skrzypek, E., Karbarz, M., Dziurka, M., Hura, T., Dziurka, K., Quarrie, S.A. (2013). Quantitative trait loci for leaf chlorophyll fluorescence parameters, chlorophyll and carotenoid contents in relation to biomass and yield in bread wheat and their chromosome deletion bin assignments. *Molecular Breeding*, 32(1), 189–210. <https://doi.org/10.1007/s11032-013-9862-8>
36. De Leon, T.B., Pruthi, R., Jampala, B., Borjas, A.H., Subudhi, P.K. (2020). Genetic determinants for agronomic and yield-related traits localized on a GBS-SNP linkage map from a japonica × indica cross in rice. *Plant Gene*, 24, 100249. <https://doi.org/10.1016/j.plgene.2020.100249>
37. de los Campos, G., Sorensen, D.A., Toro, M.A. (2019). Imperfect Linkage Disequilibrium Generates Phantom Epistasis (& Perils of Big Data). *G3 Genes|Genomes|Genetics*, 9(5), 1429–1436. <https://doi.org/10.1534/g3.119.400101>
38. Dhariwal, R., Fedak, G., Dion, Y., Pozniak, C., Laroche, A., Eudes, F., Randhawa, H.S. (2018). High Density Single Nucleotide Polymorphism (SNP) Mapping and Quantitative Trait Loci (QTL) Analysis in a Biparental Spring Triticale Population Localized Major and Minor Effect Fusarium Head Blight Resistance and Associated Traits QTL. *Genes*, 9(1), 1. <https://doi.org/10.3390/genes9010019>
39. Elena, S.F., Lenski, R.E. (1997). Test of synergistic interactions among deleterious mutations in bacteria. *Nature*, 390(6658), 395–398. <https://doi.org/10.1038/37108>
40. Flint-Garcia, S.A., Jampatong, C., Darrah, L.L., McMullen, M.D. (2003). Quantitative Trait Locus Analysis of Stalk Strength in Four Maize Populations. *Crop Science*, 43(1), 13–22. <https://doi.org/10.2135/cropsci2003.1300a>
41. Fujino, K., Hirayama, Y., Kaji, R. (2019). Marker-assisted selection in rice breeding programs in Hokkaido. *Breeding Science*, 69(3), 383–392. <https://doi.org/10.1270/jsbbs.19062>

42. Ful, Y.-B., Ritland, K. (1996). Marker-Based Inferences About Epistasis for Genes Influencing Inbreeding Depression. *Genetics*, *144*(1), 339–348. <https://doi.org/10.1093/genetics/144.1.339>
43. Gacek, K., Bayer, P.E., Anderson, R., Severn-Ellis, A.A., Wolko, J., Łopatyńska, A., Matuszczak, M., Bocianowski, J., Edwards, D., Batley, J. (2021). QTL Genetic Mapping Study for Traits Affecting Meal Quality in Winter Oilseed Rape (*Brassica Napus* L.). *Genes*, *12*(8), 8. <https://doi.org/10.3390/genes12081235>
44. Goutelle, S., Bourguignon, L., Maire, P.H., Van Guilder, M., Conte, J.E., Jelliffe, R.W. (2009). Population Modeling and Monte Carlo Simulation Study of the Pharmacokinetics and Antituberculosis Pharmacodynamics of Rifampin in Lungs. *Antimicrobial Agents and Chemotherapy*, *53*(7), 2974–2981. <https://doi.org/10.1128/aac.01520-08>
45. Han, Y., Tan, Y., Hu, H., Chang, W., Dong, L., Wang, Z., Zhao, X., Li, W., Teng, W. (2017). Quantitative trait loci with additive and epistatic effects underlying resistance to two HG types of soybean cyst nematode. *Plant Breeding*, *136*(5), 720–727. <https://doi.org/10.1111/pbr.12484>
46. Hartman, J.L., Garvik, B., Hartwell, L. (2001). Principles for the Buffering of Genetic Variation. *Science*, *291*(5506), 1001–1004. <https://doi.org/10.1126/science.1056072>
47. Hayes, P.M., Liu, B.H., Knapp, S.J., Chen, F., Jones, B., Blake, T., Franckowiak, J., Rasmusson, D., Sorrells, M., Ullrich, S.E., Wesenberg, D., Kleinjohs, A. (1993). Quantitative trait locus effects and environmental interaction in a sample of North American barley germ plasm. *Theoretical and Applied Genetics*, *87*(3), 392–401. <https://doi.org/10.1007/BF01184929>
48. Holland, J. (1998). Computer note. EPISTACY: A SAS program for detecting two-locus epistatic interactions using genetic marker information. *Journal of Heredity*, *89*(4), 374–375. <https://doi.org/10.1093/jhered/89.4.374>
49. Hussain, B. (2015). Modernization in plant breeding approaches for improving biotic stress resistance in crop plants. *Turkish Journal of Agriculture and Forestry*, *39*(4), 515–530. <https://doi.org/10.3906/tar-1406-176>

50. Jansen, R.C., Stam, P. (1994). High resolution of quantitative traits into multiple loci via interval mapping. *Genetics*, *136*(4), 1447–1455. <https://doi.org/10.1093/genetics/136.4.1447>
51. Johnson, M.S., Gopalakrishnan, S., Goyal, J., Dillingham, M.E., Bakerlee, C.W., Humphrey, P.T., Jagdish, T., Jerison, E.R., Kosheleva, K., Lawrence, K.R., Min, J., Moulana, A., Phillips, A.M., Piper, J.C., Purkanti, R., Rego-Costa, A., McDonald, M.J., Nguyen Ba, A.N., Desai, M.M. (2021). Phenotypic and molecular evolution across 10,000 generations in laboratory budding yeast populations. *eLife*, *10*, e63910. <https://doi.org/10.7554/eLife.63910>
52. Kabange, N.R., Park, S.-Y., Shin, D., Lee, S.-M., Jo, S.-M., Kwon, Y., Cha, J.-K., Song, Y.-C., Ko, J.-M., Lee, J.-H. (2020). Identification of a Novel QTL for Chlorate Resistance in Rice (*Oryza sativa* L.). *Agriculture*, *10*(8), 8. <https://doi.org/10.3390/agriculture10080360>
53. Kaczmarek, Z., Surma, M., Adamski, T. (1988). Epistatic effects in estimation of the number of genes on the basis of doubled haploid lines. *Genetica Polonica*, *29*(3–4), 353–359. <http://agro.icm.edu.pl/agro/element/bwmeta1.element.agro-8914751e-b8aa-4cb6-a38b-1ccb566bca53>
54. Kao, C.-H., Zeng, Z.-B., Teasdale, R.D. (1999). Multiple Interval Mapping for Quantitative Trait Loci. *Genetics*, *152*(3), 1203–1216. <https://doi.org/10.1093/genetics/152.3.1203>
55. Khan, S.U., Saeed, S., Khan, M.H.U., Fan, C., Ahmar, S., Arriagada, O., Shahzad, R., Branca, F., Mora-Poblete, F. (2021). Advances and Challenges for QTL Analysis and GWAS in the Plant-Breeding of High-Yielding: A Focus on Rapeseed. *Biomolecules*, *11*(10), 10. <https://doi.org/10.3390/biom11101516>
56. Kim, M.-S., Yang, J.-Y., Yu, J.-K., Lee, Y., Park, Y.-J., Kang, K.-K., Cho, Y.-G. (2021). Breeding of High Cooking and Eating Quality in Rice by Marker-Assisted Backcrossing (MABc) Using KASP Markers. *Plants*, *10*(4), 4. <https://doi.org/10.3390/plants10040804>
57. Kleinhofs, A., Kilian, A., Saghai Maroof, M.A., Biyashev, R.M., Hayes, P., Chen, F.Q., Lapitan, N., Fenwick, A., Blake, T.K., Kanazin, V., Ananiev, E., Dahleen, L., Kudrna, D., Bollinger, J., Knapp, S.J., Liu, B., Sorrells, M., Heun, M., Franckowiak, J.D., Hoffman, D., Skadsen, R., Steffenson, B.J. (1993). A

- molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. *Theoretical and Applied Genetics*, 86(6), 705–712. <https://doi.org/10.1007/BF00222660>
58. Krajewski, P., Bocianowski, J., Gawłowska, M., Kaczmarek, Z., Pniewski, T., Świącicki, W., Wolko, B. (2012). QTL for yield components and protein content: A multienvironment study of two pea (*Pisum sativum* L.) populations. *Euphytica*, 183(3), 323–336. <https://doi.org/10.1007/s10681-011-0472-4>
59. Ku, L.X., Sun, Z.H., Wang, C.L., Zhang, J., Zhao, R.F., Liu, H.Y., Tai, G.Q., Chen, Y.H. (2012). QTL mapping and epistasis analysis of brace root traits in maize. *Molecular Breeding*, 30(2), 697–708. <https://doi.org/10.1007/s11032-011-9655-x>
60. Kwon, Y.-H., Kabange, N.-R., Lee, J.-Y., Lee, S.-M., Cha, J.-K., Shin, D.-J., Cho, J.-H., Kang, J.-W., Ko, J.-M., Lee, J.-H. (2021). Novel QTL Associated with Shoot Branching Identified in Doubled Haploid Rice (*Oryza sativa* L.) under Low Nitrogen Cultivation. *Genes*, 12(5), Article 5. <https://doi.org/10.3390/genes12050745>
61. Labroo, M.R., Studer, A.J., Rutkoski, J.E. (2021). Heterosis and Hybrid Crop Breeding: A Multidisciplinary Review. *Frontiers in Genetics*, 12, 643761. <https://www.frontiersin.org/articles/10.3389/fgene.2021.643761>
62. Lahey, B.B., Van Hulle, C.A., Singh, A.L., Waldman, I.D., Rathouz, P.J. (2011). Higher-Order Genetic and Environmental Structure of Prevalent Forms of Child and Adolescent Psychopathology. *Archives of General Psychiatry*, 68(2), 181–189. <https://doi.org/10.1001/archgenpsychiatry.2010.192>
63. Larrimore, K.E., Rancati, G. (2019). The conditional nature of gene essentiality. *Current Opinion in Genetics & Development*, 58–59, 55–61. <https://doi.org/10.1016/j.gde.2019.07.015>
64. Lephuthing, M.C., Khumalo, T.P., Tolmay, V.L., Dube, E., Tsilo, T.J. (2022). Genetic Mapping of Quantitative Trait Loci Associated with Plant Height and Yield Component Traits in a Wheat (*Triticum aestivum* L.) Doubled Haploid Population Derived from Tugela-DN × Elands. *Agronomy*, 12(10), 10. <https://doi.org/10.3390/agronomy12102283>

65. Li, C.-Q., Song, L., Zhao, H.-H., Wang, Q.-L., Fu, Y.-Z. (2014). Identification of quantitative trait loci with main and epistatic effects for plant architecture traits in Upland cotton (*Gossypium hirsutum* L.). *Plant Breeding*, 133(3), 390–400. <https://doi.org/10.1111/pbr.12161>
66. Li, H., Ribaut, J.-M., Li, Z., Wang, J. (2008). Inclusive composite interval mapping (ICIM) for digenic epistasis of quantitative traits in biparental populations. *Theoretical and Applied Genetics*, 116(2), 243–260. <https://doi.org/10.1007/s00122-007-0663-5>
67. Luo, X., Fu, Y., Zhang, P., Wu, S., Tian, F., Liu, J., Zhu, Z., Yang, J., Sun, C. (2009). Additive and Over-dominant Effects Resulting from Epistatic Loci Are the Primary Genetic Basis of Heterosis in Rice. *Journal of Integrative Plant Biology*, 51(4), 393–408. <https://doi.org/10.1111/j.1744-7909.2008.00807.x>
68. Machida, C., Nakagawa, A., Kojima, S., Takahashi, H., Machida, Y. (2015). The complex of ASYMMETRIC LEAVES (AS) proteins plays a central role in antagonistic interactions of genes for leaf polarity specification in *Arabidopsis*. *WIREs Developmental Biology*, 4(6), 655–671. <https://doi.org/10.1002/wdev.196>
69. Mackay, T.F.C. (2014). Epistasis and quantitative traits: Using model organisms to study gene–gene interactions. *Nature Reviews Genetics*, 15(1), 22–33. <https://doi.org/10.1038/nrg3627>
70. Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., Cho, J.H., Guttmacher, A.E., Kong, A., Kruglyak, L., Mardis, E., Rotimi, C.N., Slatkin, M., Valle, D., Whittemore, A.S., Boehnke, M., Clark, A.G., Eichler, E.E., Gibson, G., Haines, J.L., Mackay, T.F.C., McCarroll, S.A., Visscher, P.M. (2009). Finding the missing heritability of complex diseases. *Nature*, 461(7265), 747–753. <https://doi.org/10.1038/nature08494>
71. Martinez, O., Curnow, R.N. (1994). Missing markers when estimating quantitative trait loci using regression mapping. *Heredity*, 73, 198–206.
72. Members of the Complex Trait Consortium. (2003). The nature and identification of quantitative trait loci: A community’s view. *Nature Reviews Genetics*, 4(11), 911–916. <https://doi.org/10.1038/nrg1206>

73. Moose, S.P., Mumm, R.H. (2008). Molecular Plant Breeding as the Foundation for 21st Century Crop Improvement. *Plant Physiology*, 147(3), 969–977. <https://doi.org/10.1104/pp.108.118232>
74. Nap, J.-P., Conner, A.J., Mlynárová, L., Stiekema, W.J., Jansen, R.C. (1997). Dissection of a Synthesized Quantitative Trait to Characterize Transgene Interactions. *Genetics*, 147(1), 315–320. <https://doi.org/10.1093/genetics/147.1.315>
75. Nogué, F., Mara, K., Collonnier, C., Casacuberta, J.M. (2016). Genome engineering and plant breeding: Impact on trait discovery and development. *Plant Cell Reports*, 35(7), 1475–1486. <https://doi.org/10.1007/s00299-016-1993-z>
76. Pegot-Espagnet, P., Guillaume, O., Desprez, B., Devaux, B., Devaux, P., Henry, K., Henry, N., Willems, G., Goudemand, E., Mangin, B. (2019). Discovery of interesting new polymorphisms in a sugar beet (elite × exotic) progeny by comparison with an elite panel. *Theoretical and Applied Genetics*, 132(11), 3063–3078. <https://doi.org/10.1007/s00122-019-03406-0>
77. Province, M.A. (2001). 30 Sequential methods of analysis for genome scans. In: *Advances in Genetics* (T. 42, s. 499–514). Academic Press. [https://doi.org/10.1016/S0065-2660\(01\)42039-6](https://doi.org/10.1016/S0065-2660(01)42039-6)
78. Pundir, S., Sharma, R., Kumar, D., Singh, V.K., Chaturvedi, D., Kanwar, R.S., Röder, M.S., Börner, A., Ganai, M.W., Gupta, P.K., Sharma, S., Sharma, S. (2022). QTL mapping for resistance against cereal cyst nematode (*Heterodera avenae* Woll.) in wheat (*Triticum aestivum* L.). *Scientific Reports*, 12(1), 1. <https://doi.org/10.1038/s41598-022-12988-7>
79. Quarrie, S.A., Steed, A., Calestani, C., Semikhodskii, A., Lebreton, C., Chinoy, C., Steele, N., Pljevljakusić, D., Waterman, E., Weyen, J., Schondelmaier, J., Habash, D.Z., Farmer, P., Saker, L., Clarkson, D.T., Abugalieva, A., Yessimbekova, M., Turuspekov, Y., Abugalieva, S., Tuberosa, R., Sanguineti, M.-C., Hollington, P.A., Aragués, R., Royo, A., Dodig, D. (2005). A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring × SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theoretical and Applied Genetics*, 110(5), 865–880. <https://doi.org/10.1007/s00122-004-1902-7>

80. R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
81. Raffo, M.A., Sarup, P., Guo, X., Liu, H., Andersen, J.R., Orabi, J., Jahoor, A., Jensen, J. (2022). Improvement of genomic prediction in advanced wheat breeding lines by including additive-by-additive epistasis. *Theoretical and Applied Genetics*, 135(3), 965–978. <https://doi.org/10.1007/s00122-021-04009-4>
82. Rajurkar, A.B., Muthukumar, C., Ayyenar, B., Thomas, H.B., Chandra Babu, R. (2021). Mapping consistent additive and epistatic QTLs for plant production traits under drought in target populations of environment using locally adapted landrace in rice (*Oryza sativa* L.). *Plant Production Science*, 24(3), 388–403. <https://doi.org/10.1080/1343943X.2020.1862680>
83. Rakoczy-Trojanowska, M., Krajewski, P., Bocianowski, J., Schollenberger, M., Wakuliński, W., Milczarski, P., Masojć, P., Targońska-Karasek, M., Banaszak, Z., Banaszak, K., Brukwiński, W., Orczyk, W., Kilian, A. (2017a). Identification of Single Nucleotide Polymorphisms Associated with Brown Rust Resistance, α -Amylase Activity and Pre-harvest Sprouting in Rye (*Secale cereale* L.). *Plant Molecular Biology Reporter*, 35(3), 366–378. <https://doi.org/10.1007/s11105-017-1030-6>
84. Rakoczy-Trojanowska, M., Orczyk, W., Krajewski, P., Bocianowski, J., Stochmal, A., Kowalczyk, M. (2017b). *ScBx* gene based association analysis of hydroxamate content in rye (*Secale cereale* L.). *Journal of Applied Genetics*, 58(1), 1–9. <https://doi.org/10.1007/s13353-016-0356-3>
85. Ren, J., Zhang, X., Li, Z., Wu, P. (2022). Genetic Analysis of Maternal Haploid Inducibility for In Vivo Haploid Induction in Maize. *Agriculture*, 12(6), 6. <https://doi.org/10.3390/agriculture12060845>
86. Rieger, R., Michaelis, A., Green, M.M. (2012). *Glossary of Genetics and Cytogenetics: Classical and Molecular*. Springer Science & Business Media.
87. Robertsen, C.D., Hjortshøj, R.L., Janss, L.L. (2019). Genomic Selection in Cereal Breeding. *Agronomy*, 9(2), 2. <https://doi.org/10.3390/agronomy9020095>
88. Romagosa, I., Ullrich, S.E., Han, F., Hayes, P.M. (1996). Use of the additive main effects and multiplicative interaction model in QTL mapping for adaptation in

- barley. *Theoretical and Applied Genetics*, 93(1), 30–37. <https://doi.org/10.1007/BF00225723>
89. Ronin, Y.I., Korol, A.B., Nevo, E. (1999). Single- and Multiple-Trait Mapping Analysis of Linked Quantitative Trait Loci: Some Asymptotic Analytical Approximations. *Genetics*, 151(1), 387–396. <https://doi.org/10.1093/genetics/151.1.387>
 90. Routman, E.J., Cheverud, J.M. (1997). Gene effects on a quantitative trait: two-locus epistatic effects measured at microsatellite markers and at estimated QTL. *Evolution*, 51(5), 1654–1662. <https://doi.org/10.1111/j.1558-5646.1997.tb01488.x>
 91. RStudio Team. (2020). *RStudio: Integrated Development for R*. RStudio, PBC. <http://www.rstudio.com/>
 92. Salgotra, R.K., Stewart, C.N. (2020). Functional Markers for Precision Plant Breeding. *International Journal of Molecular Sciences*, 21(13), 13. <https://doi.org/10.3390/ijms21134792>
 93. Sayed, M.A., Nassar, S.M., Moustafa, E.S., Said, M.T., Börner, A., Hamada, A. (2021). Genetic Mapping Reveals Novel Exotic and Elite QTL Alleles for Salinity Tolerance in Barley. *Agronomy*, 11(9), Article 9. <https://doi.org/10.3390/agronomy11091774>
 94. Searle, S.R. (1982). *Matrix algebra useful for statistics*. Wiley.
 95. Sen, Ś., Churchill, G.A. (2001). A Statistical Framework for Quantitative Trait Mapping. *Genetics*, 159(1), 371–387. <https://doi.org/10.1093/genetics/159.1.371>
 96. Sharma, R., Sadeghian Tehrani, Z., Kumar, S., Song, M. (2022). Detecting genetic epistasis by differential departure from independence. *Molecular Genetics and Genomics*, 297(4), 911–924. <https://doi.org/10.1007/s00438-022-01893-3>
 97. Shelake, R.M., Pramanik, D., Kim, J.-Y. (2019). Evolution of plant mutagenesis tools: A shifting paradigm from random to targeted genome editing. *Plant Biotechnology Reports*, 13(5), 423–445. <https://doi.org/10.1007/s11816-019-00562-z>
 98. Shi, H., Guan, W., Shi, Y., Wang, S., Fan, H., Yang, J., Chen, W., Zhang, W., Sun, D., Jing, R. (2020). QTL mapping and candidate gene analysis of seed vigor-

- related traits during artificial aging in wheat (*Triticum aestivum*). *Scientific Reports*, 10(1), 1. <https://doi.org/10.1038/s41598-020-75778-z>
99. Singleton, V.L., Rossi, J.A. (1965). Colorimetry of Total Phenolics with Phosphomolybdic-Phosphotungstic Acid Reagents. *American Journal of Enology and Viticulture*, 16(3), 144–158. <https://doi.org/10.5344/ajev.1965.16.3.144>
 100. Slim, L., Chatelain, C., Azencott, C.-A., Vert, J.-P. (2020). Novel methods for epistasis detection in genome-wide association studies. *PLOS ONE*, 15(11), e0242927. <https://doi.org/10.1371/journal.pone.0242927>
 101. Smeda, J.R., Schillmiller, A.L., Anderson, T., Ben-Mahmoud, S., Ullman, D.E., Chappell, T.M., Kessler, A., Mutschler, M.A. (2018). Combination of Acylglucose QTL reveals additive and epistatic genetic interactions and impacts insect oviposition and virus infection. *Molecular Breeding*, 38(1), 3. <https://doi.org/10.1007/s11032-017-0756-z>
 102. Sorojsrisom, E.S., Haller, B.C., Ambrose, B.A., Eaton, D.A.R. (2022). Selection on the gametophyte: Modeling alternation of generations in plants. *Applications in Plant Sciences*, 10(2), e11472. <https://doi.org/10.1002/aps3.11472>
 103. Tammisola, J. (2010). Towards much more efficient biofuel crops—Can sugarcane pave the way? *GM Crops*, 1(4), 181–198. <https://doi.org/10.4161/gmcr.1.4.13173>
 104. Taylor, M.B., Ehrenreich, I.M. (2015). Higher-order genetic interactions and their contribution to complex traits. *Trends in Genetics*, 31(1), 34–40. <https://doi.org/10.1016/j.tig.2014.09.001>
 105. Tinker, N.A., Mather, D.E., Rossnagel, B.G., Kasha, K.J., Kleinhofs, A., Hayes, P. M., Falk, D.E., Ferguson, T., Shugar, L.P., Legge, W.G., Irvine, R.B., Choo, T.M., Briggs, K.G., Ullrich, S.E., Franckowiak, J.D., Blake, T.K., Graf, R.J., Dofing, S.M., Maroof, M.A.S., Scoles, G.J., Hoffman, D., Dahleen, L.S., Kilian, A., Chen, F., Biyashev, R.M., Kudrna, D.A., Steffenson, B.J. (1996). Regions of the genome that affect agronomic performance in two-row barley. *Crop Science*, 36(4), 1053–1062. <https://doi.org/10.2135/cropsci1996.0011183X003600040040x>
 106. Tomkowiak, A., Bocianowski, J., Spychała, J., Grynia, J., Sobiech, A., Kowalczewski, P.Ł. (2021). DArTseq-Based High-Throughput SilicoDArT and

- SNP Markers Applied for Association Mapping of Genes Related to Maize Morphology. *International Journal of Molecular Sciences*, 22(11), 11. <https://doi.org/10.3390/ijms22115840>
107. Tomkowiak, A., Bocianowski, J., Wolko, Ł., Adamczyk, J., Mikołajczyk, S., Kowalczewski, P.Ł. (2019). Identification of Markers Associated with Yield Traits and Morphological Features in Maize (*Zea mays* L.). *Plants*, 8(9), 9. <https://doi.org/10.3390/plants8090330>
 108. Tuvešson, S.D., Larsson, C.-T., Ordon, F. (2021). Use of Molecular Markers for Doubled Haploid Technology: From Academia to Plant Breeding Companies. W J. M. Segui-Simarro (Red.), *Doubled Haploid Technology: Volume 2: Hot Topics, Apiaceae, Brassicaceae, Solanaceae* (s. 49–72). Springer US. https://doi.org/10.1007/978-1-0716-1335-1_3
 109. van Soest, L.J.M., Mastebroek, H.D., de Meijer, E.P.M. (1992). Genetic resources and breeding: A necessity for the success of industrial crops. *Industrial Crops and Products*, 1(2), 283–288. [https://doi.org/10.1016/0926-6690\(92\)90029-U](https://doi.org/10.1016/0926-6690(92)90029-U)
 110. Viana, J.M.S., e Silva, F.F., Mundim, G.B., Azevedo, C.F., Jan, H.U. (2016). Efficiency of low heritability QTL mapping under high SNP density. *Euphytica*, 213(1), 13. <https://doi.org/10.1007/s10681-016-1800-5>
 111. Viana, J.M.S., Souza, C.A.S. (2023). Efficiency of mapping epistatic quantitative trait loci. *Heredity*, 131(1), 1. <https://doi.org/10.1038/s41437-023-00618-5>
 112. Voss-Fels, K.P., Cooper, M., Hayes, B.J. (2019). Accelerating crop genetic gains with genomic selection. *Theoretical and Applied Genetics*, 132(3), 669–686. <https://doi.org/10.1007/s00122-018-3270-8>
 113. VSN International. (2022). *Genstat Reference Manual (Release 22), Part 2 Directives*. VSN International. <https://repository.rothamsted.ac.uk/item/85y50/genstat-5-reference-manual>
 114. Wang, D., Salah El-Basyoni, I., Stephen Baenziger, P., Crossa, J., Eskridge, K.M., Dweikat, I. (2012). Prediction of genetic values of quantitative traits with epistatic effects in plant breeding populations. *Heredity*, 109(5), 5. <https://doi.org/10.1038/hdy.2012.44>

115. Wang, D., Tang, H., Liu, J.-F., Xu, S., Zhang, Q., Ning, C. (2020). Rapid epistatic mixed-model association studies by controlling multiple polygenic effects. *Bioinformatics*, 36(19), 4833–4837. <https://doi.org/10.1093/bioinformatics/btaa610>
116. Wang, X., Xie, W., Li, L., Zhu, J., Xing, F. (2022). Molecular Simulation Study on Mechanical Properties of Microcapsule-Based Self-Healing Cementitious Materials. *Polymers*, 14(3), 3. <https://doi.org/10.3390/polym14030611>
117. Weinreich, D.M., Lan, Y., Wylie, C.S., Heckendorn, R.B. (2013). Should evolutionary geneticists worry about higher-order epistasis? *Current Opinion in Genetics & Development*, 23(6), 700–707. <https://doi.org/10.1016/j.gde.2013.10.007>
118. Weyen, J. (2021). Applications of Doubled Haploids in Plant Breeding and Applied Research. In: J.M. Segui-Simarro (Red.), *Doubled Haploid Technology: Volume 1: General Topics, Alliaceae, Cereals* (s. 23–39). Springer US. https://doi.org/10.1007/978-1-0716-1315-3_2
119. Wu, R.L. (1999). Mapping Quantitative Trait Loci by Genotyping Haploid Tissues. *Genetics*, 152(4), 1741–1752. <https://doi.org/10.1093/genetics/152.4.1741>
120. Yang, J., Zhu, J., Williams, R.W. (2007). Mapping the genetic architecture of complex traits in experimental populations. *Bioinformatics*, 23(12), 1527–1536. <https://doi.org/10.1093/bioinformatics/btm143>
121. Yao, E., Blake, V.C., Cooper, L., Wight, C.P., Michel, S., Cagirici, H.B., Lazo, G.R., Birkett, C.L., Waring, D.J., Jannink, J.-L., Holmes, I., Waters, A.J., Eickholt, D.P., Sen, T.Z. (2022). GrainGenes: A data-rich repository for small grains genetics and genomics. *Database*, 2022, baac034. <https://doi.org/10.1093/database/baac034>
122. Yusuf, A.O., Richter, J.-C., Möllers, C. (2022). Genetic variation and QTL analysis of saturated fatty acids in two doubled haploid populations of oilseed rape (*Brassica napus* L.). *Euphytica*, 218(7), 88. <https://doi.org/10.1007/s10681-022-03043-7>

123. Zeng, Z.-B., Kao, C.-H., Basten, C.J. (1999). Estimating the genetic architecture of quantitative traits. *Genetics Research*, 74(3), 279–289. <https://doi.org/10.1017/S0016672399004255>
124. Zhang, F., Neik, T.X., Thomas, W.J.W., Batley, J. (2023). CRISPR-Based Genome Editing Tools: An Accelerator in Crop Breeding for a Changing Future. *International Journal of Molecular Sciences*, 24(10), 8623. <https://doi.org/10.3390/ijms24108623>
125. Zimnoch-Guzowska, E., Marczewski, W., Lebecka, R., Flis, B., Schäfer-Pregl, R., Salamini, F., Gebhardt, C. (2000). QTL Analysis of New Sources of Resistance to *Erwinia carotovora* ssp. *atroseptica* in Potato Done by AFLP, RFLP, and Resistance-Gene-Like Markers. *Crop Science*, 40(4), 1156–1167. <https://doi.org/10.2135/cropsci2000.4041156x>

ROZDZIAŁ II

Publikacje wchodzące w skład rozprawy

1. Analytical and numerical comparisons of two methods of estimation of additive \times additive \times additive interaction of QTL effects

Cyplik, A., Bocianowski, J.

Journal of Applied Genetics, 2022, 63(2), 213–221

DOI:10.1007/s13353-021-00676-7

Impact Factor 2021: 2,653; Punktacja wg MEiN: 140



Analytical and numerical comparisons of two methods of estimation of additive \times additive \times additive interaction of QTL effects

Adrian Cyplik¹ · Jan Bocianowski¹

Received: 6 August 2021 / Revised: 7 November 2021 / Accepted: 13 December 2021 / Published online: 23 December 2021
© The Author(s) 2021

Abstract

This paper presents the analytical and numerical comparison of two methods of estimation of additive \times additive \times additive (*aaa*) interaction of QTL effects. The first method takes into account only the plant phenotype, while in the second we also included genotypic information from molecular marker observation. Analysis was made on 150 doubled haploid (DH) lines of barley derived from cross Steptoe \times Morex and 145 DH lines from Harrington \times TR306 cross. In total, 153 sets of observation was analyzed. In most cases, *aaa* interactions were found with an exert effect on QTL. Results also show that with molecular marker observations, obtained estimators had smaller absolute values than phenotypic estimators.

Keywords Doubled haploid (DH) lines · Barley · QTL interaction · Genetic interactions · Statistical methods

Introduction

The analysis of inheritance of quantitative traits, due to their polygenic nature, requires the use of appropriate statistical and genetic methods. Among these methods, the most interesting are those that enable the determination of the mode of action of genes in the studied population.

The concept of genetic interactions is known for more than a hundred years (Bateson and Mendel 1902). Considering that a complex phenotype may be the effect of a combination of multiple loci, various statistical methods have been developed for identifying genetic epistasis effects (Chen et al. 2011). Most studies are focused on single locus analysis, which directly tests the association between individual genes and phenotypic variants. Pairwise interactions are often used in modern genetics (Brem et al. 2005; Jarvis and Cheverud 2011; Gaertner et al. 2012), but higher-order interactions are often neglected. This kind of more complex interaction requires complete, precise data to be successfully included, but this type of data was rarely available

since more recent times (Carlborg et al. 2006; Cordell 2009). There is no denying that we do not fully understand all of the mechanics of heritability and the higher-order interactions may be the missing element of explaining the relationship between genotype and phenotype (Hartman et al. 2001; Manolio et al. 2009).

Quantitative traits are not only one of the most important in the viewpoint of breeding programs but also can be influenced by a multiplicity of polymorphic genes, environmental conditions, and genetic interactions, making them extremely difficult to fully understand (Members of the Complex Trait Consortium 2003; Mackay 2014).

The purpose of the research reported in this article is to compare two methods of estimation of the parameter connected with the additive \times additive \times additive (*aaa*) interaction gene effect: the phenotypic method and the genotypic method. The comparison was made by analytical methods and with analyses of data sets of barley doubled haploid lines. To our knowledge, this is the first report about *aaa* interaction.

Material and methods

If in the experiment we observed n homozygous (doubled haploid, DH; recombinant inbred, RI) plant lines, we get an n -vector of phenotypic mean observations $\mathbf{y} = [y_1 y_2 \dots y_n]'$ and q n -vectors of marker genotype observations $\mathbf{m}_l, l = 1, 2, \dots, q$.

Communicated by Izabela Pawłowicz

✉ Jan Bocianowski
jan.bocianowski@up.poznan.pl

¹ Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Wojska Polskiego 28, 60-637 Poznań, Poland

The i -th element ($i = 1, 2, \dots, n$) of vector \mathbf{m}_i is equal -1 or 1 , depending on the parent's genotype exhibited by the i -th line.

Estimation based on the phenotype

Estimation of the additive \times additive \times additive interaction of homozygous loci (three-way epistasis) effect aaa on the basis of phenotypic observations \mathbf{y} requires identification of groups of extreme lines, i.e., lines with the minimal and maximal expression of the observed trait (Choo and Reinbergs 1982). The group of minimal lines consists of the lines which contain, theoretically, only alleles decreasing the value of the trait. Analogously, the group of maximal lines contains the lines which have only alleles increasing the trait value. In this paper, we identify the groups of extreme lines as minimal and maximal, respectively, lines of the empirical distribution of means. The total three-way epistasis interaction effect aaa can be estimated by the following formula:

$$\widehat{aaa}_p = \frac{1}{2} (\bar{L}_{\max} + \bar{L}_{\min}) - \bar{L}, \quad (1)$$

where \bar{L}_{\min} and \bar{L}_{\max} denote the means for the groups of minimal and maximal lines, respectively, \bar{L} denotes the mean for all lines. The number of genes (number of effective factors) obtained on the basis of phenotypic observations only was calculated using the formula presented by Kaczmarek et al. (1988).

Estimation based on the genotypic observations

Estimation of aaa is based on the assumption that the genes responsible for the trait are closely linked to the observed molecular marker. By choosing from all observed markers p , we can explain the variability of the trait, and model observations for the lines as follows:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{W}\boldsymbol{\delta} + \mathbf{e}, \quad (2)$$

where $\mathbf{1}$ denotes the n -dimensional vector of ones, μ denotes the general mean, \mathbf{X} denotes $(n \times p)$ -dimensional matrix of the form $\mathbf{X} = \begin{bmatrix} \mathbf{m}_{l_1} & \mathbf{m}_{l_2} & \dots & \mathbf{m}_{l_p} \end{bmatrix}$, $l_1, l_2, \dots, l_p \in \{1, 2, \dots, q\}$, $\boldsymbol{\beta}$ denotes the p -dimensional vector of unknown parameters of the form $\boldsymbol{\beta}' = \begin{bmatrix} a_{l_1} & a_{l_2} & \dots & a_{l_p} \end{bmatrix}$, \mathbf{Z} denotes matrix which columns are products of some columns of matrix \mathbf{X} , $\boldsymbol{\gamma}$ denotes the vector of unknown parameters of the form $\boldsymbol{\gamma}' = \begin{bmatrix} aa_{l_1 l_2} & aa_{l_1 l_3} & \dots & aa_{l_{p-1} l_p} \end{bmatrix}$, \mathbf{W} denotes matrix which columns are three-way products of some columns of matrix \mathbf{X} , $\boldsymbol{\delta}$ denotes the vector of unknown parameters of the form $\boldsymbol{\delta}' = \begin{bmatrix} aaa_{l_1 l_2 l_3} & aaa_{l_1 l_2 l_4} & \dots & aaa_{l_{p-2} l_{p-1} l_p} \end{bmatrix}$, and \mathbf{e} denotes the n -dimensional vector of random variables such that $E(e_i) = 0$, $Cov(e_i, e_j) = 0$ for $i \neq j$, $i, j = 1, 2, \dots, n$. The parameters $a_{l_1}, a_{l_2}, \dots, a_{l_p}$ are the additive effects of the genes controlling the

trait, parameters $aa_{l_1 l_2}, aa_{l_1 l_3}, \dots, aa_{l_{p-1} l_p}$ are the additive \times additive interaction effects and parameters $aaa_{l_1 l_2 l_3}, aaa_{l_1 l_2 l_4}, \dots, aaa_{l_{p-2} l_{p-1} l_p}$ are the additive \times additive \times additive interaction effects. We assume that the epistatic and three-way epistatic interaction effects show only loci with significant additive gene action effects. This assumption significantly decreases the number of potential significant effects and causes the regression model to be more useful.

Denoting by $\boldsymbol{\alpha}' = [\mu \ \boldsymbol{\beta}' \ \boldsymbol{\gamma}' \ \boldsymbol{\delta}']$ and $\mathbf{G} = [\mathbf{1} \ \mathbf{X} \ \mathbf{Z} \ \mathbf{W}]$ we obtain the model

$$\mathbf{y} = \mathbf{G}\boldsymbol{\alpha} + \mathbf{e}. \quad (3)$$

If \mathbf{G} is of full rank, the estimate of $\boldsymbol{\alpha}$ is given by (Searle 1982)

$$\hat{\boldsymbol{\alpha}} = (\mathbf{G}'\mathbf{G})^{-1} \mathbf{G}'\mathbf{y}. \quad (4)$$

The total three-way epistasis aaa effect of genes influencing the trait can be found as follows:

$$\widehat{aaa}_g = \sum_{k=1}^{p-2} \sum_{\substack{k'=2 \\ k' \neq k}}^{p-1} \sum_{\substack{k''=3 \\ k'' \neq k, k'}}^p \widehat{aaa}_{l_k l_{k'} l_{k''}}. \quad (5)$$

For the marker selection of model (2), we used a stepwise feature selection by Akaike information criteria (Akaike 1998). The procedure consisted of two steps: first, we divided markers into groups based on chromosomes they were located on and performed stepwise feature selection by AIC; after that, we combined the remaining markers into one group and we repeated selection as above. All of the remaining markers were combined into the final group and the last feature selection was performed on a model with additive \times additive \times additive interaction effect included. To counteract the multiple comparisons problem, we used the Bonferroni correction.

Examples

To compare the estimates of aaa obtained by different methods, the following data sets were used.

Example 1

The first set of data we used in our experiment comes from North American Barley Genome Mapping Project (NAB-GMP) and consists of 150 doubled haploid (DH) lines of barley tested in sixteen environments [Crookston, MN, 1992; Ithaca, NY, 1992; Guelph, Ontario, 1992; Pullman, WA, 1992; Brandon, Manitoba, 1992; Outlook, Saskatchewan, 1992; Goodale, Saskatchewan, 1992; Saskatoon, Saskatchewan, 1992; Tetonia, ID, 1992; Bozeman, MT (irrigated), 1992; Bozeman, MT (dryland), 1992; Aberdeen, ID, 1991; Klamath Falls, OR, 1991; Pullman, WA, 1991; Bozeman,

MT (irrigated), 1991; and Bozeman, MT (dryland), 1991]. Steptoe × Morex cross was developed by the Oregon State University Barley Breeding Program by crossing “Steptoe” and “Morex” barley varieties (Kleinhofs et al. 1993; Romagosa et al. 1996; <http://wheat.pw.usda.gov/ggpages/SxM>). The linkage map used consisted of 223 molecular markers, mostly RFLP, with mean distance between markers equal to 5.66 cM. Lines were analyzed for eight phenotypic traits (alpha amylase, AA; diastatic power, DP; grain protein, GP; grain yield, GY; height, H; heading date, HD; lodging, L; malt extract, ME; Hayes et al. 1993). Missing marker values were estimated with non-missing data of flanking markers (Martinez and Curnow 1994) and GP, L, and ME traits data were transformed by $\arcsin\sqrt{x/100}$.

Example 2

The second data set also comes from the NABGM project and consist of 145 doubled haploid (DH) lines of barley (cross of two-rowed varieties Harrington × TR306) analyzed for seven phenotypic traits (weight of grain harvested per unit area, WG; number of days from planting until emergence of 50% of heads on main tillers, NH; number of days from planting until physiological maturity, NM; plant height, H; lodging transformed by $\arcsin\sqrt{x/100}$, L; 1000 kernel weight, KW; test weight, TW) and tested in five environments (in four environments, observations were made over two years: Brandon, Manitoba, 1992 and 1993; Ailsa Craig, Ontario, 1992 and 1993; Elora, Ontario, 1992 and 1993; Outlook, Saskatchewan, 1992 and 1993; Ste-Anne-de-BeUevue, Quebec, 1993) (Tinker et al. 1996, <http://wheat.pw.usda.gov/ggpages/HxT>). We used the map composed of 127 molecular markers (mostly RFLP) with the mean distance between markers equal to 10.62 cm.

Considering that each trait and environment was classified as an independent variable in both cases, in total of 153 sets of observations were deemed. Trait data was transformed to achieve normal distribution of the observed features. In all cases, transformation was successful and normal distribution was obtained.

Results

Analytical comparison

The estimators, (1) and (5), of the three-way epistasis effect aaa can be analyzed and compared under simplified assumptions: (i) that the markers are unlinked and (ii) that the segregation of each marker is compatible with the genetic model appropriate for the analyzed population, which in our case means that the probability of observing “1” is the same as observing “–1”. This is true if we consider that model (2)

treats the marker observations as fixed. In fact, the vectors \mathbf{m}_l , $l = 1, 2, \dots, q$, constitute observations of some random variables. If the marker data satisfied exactly assumptions (i) and (ii) we would have

$$\widehat{aaa}_g = \sum_{k=1}^{p-2} \sum_{\substack{k'=2 \\ k' \neq k}}^{p-1} \sum_{\substack{k''=3 \\ k'' \neq k, k'}}^p \left[\frac{1}{2} \left(\bar{y}^{(l_k l_{k'} l_{k''}, +)} + \bar{y}^{(l_k l_{k'} l_{k''}, -)} \right) - \bar{y} \right], \quad (6)$$

where $\bar{y}^{(l_k l_{k'} l_{k''}, -)}$ and $\bar{y}^{(l_k l_{k'} l_{k''}, +)}$ denote the means for lines with observations of k -th, k' -th, and k'' -th markers equal – 1 and 1, respectively.

In practice, the marker data do not accurately meet the following conditions for model (6). Taking into consideration that markers chosen for model (2) are far apart from each other on the linkage map, assumption (i) is true. To test the assumption (ii) χ^2 , the test is used before any analysis is performed.

Numerical comparison

Obtained results for estimates of total additive × additive × additive interaction effect was presented in Tables 1, 2, 3, and 4. Tables 1 and 2 contain phenotypic and genotypic analysis, respectively, for the 150 doubled haploid lines of barley from the Steptoe × Morex cross; Tables 3 and 4 for the 145 doubled haploid lines of barley from the Harrington × TR306 cross. Figures 1 and 2 show the relative comparison of phenotypic and genotypic estimates of the total additive × additive × additive interaction effect in the form of a box-and-whisker diagram of the values $(\widehat{aaa}_g / \widehat{aaa}_p) \cdot 100$, classified by the observed phenotypic traits.

Results show that in 90 cases (70%) we found statistically significant additive × additive × additive interaction effects (Table 1). The same amount of interactions was found for marker observation, but only in 72 cases, where we confirmed results statistically (Table 2). Comparisons of genotypic and phenotypic estimates of the total additive × additive × additive interaction effect show that in the majority of cases (79%), the effect was smaller than the total aaa interaction effect from phenotypic observations alone (Fig. 1). However, the scope of calculated estimates is quite large ranging from – 1590.91% for HD to 1800.00% for H in the same environment (WA92). In a total of five cases, we observed estimate values higher than 1000%. The smallest range of estimates was observed for the trait DP. Number of genes (effective factors) ranged from 3–10 with average of 3.4 (Table 1). Minimal number of included markers equals 12, where maximum number was 32, with an average of 19.5 markers per model. The number of three-way interactions ranged from 0–35 with an average of 8.3 (Table 2).

For the Harrington × TR306, cross results show that in 63 cases (100%), we found statistically significant additive × additive × additive interaction effects (Table 3). The same amount

Table 1 Phenotypic estimates of the total additive × additive × additive interaction effect for the 150 doubled haploid lines of barley obtained from the Steptoe × Morex cross

Environment	Trait							
	AA [§]	DP	GY	GP	HD	H	L	ME
ID91 [#]	1.56 (5 [@])	11.36 (8)	−0.31 (6)	0.08 (4)	−0.05 (4)	0.90 (7)	-	0.45 (5)
ID92	1.72 (5)	31.17 (5)	−0.78 (7)	0.21 (5)	−0.75 (4)	6.42 (6)	-	0.37 (5)
MA92	-	-	−0.22 (4)	-	−0.79 (3)	−0.48 (4)	3.14 (7)	-
MN92	3.13 (6)	5.16 (5)	−0.54 (5)	0.24 (4)	0.93 (3)	1.33 (4)	-	0.61 (6)
MTd91	-	-	−0.12 (4)	-	0.32 (3)	−0.67 (5)	-	-
MTd92	2.82 (6)	28.11 (6)	−0.74 (5)	0.50 (5)	0.68 (3)	6.25 (5)	14.50 (7)	−0.33 (3)
MTi91	4.15 (4)	26.29 (10)	−0.37 (3)	0.33 (5)	1.31 (4)	0.68 (5)	-	−0.25 (5)
MTi92	1.03 (3)	20.32 (9)	−0.13 (4)	1.30 (3)	−0.73 (3)	3.55 (5)	4.93 (4)	−0.73 (5)
NY92	-	-	0.26 (5)	-	0.31 (3)	−1.33 (5)	20.87 (9)	-
ON92	-	-	−0.14 (4)	-	0.98 (4)	1.03 (4)	11.80 (7)	-
OR91	3.82 (6)	17.59 (5)	0.54 (5)	0.46 (5)	−2.66 (4)	−4.64 (5)	-	−0.10 (3)
SKg92	-	-	−0.23 (4)	-	0.43 (4)	5.91 (4)	-	-
SKk92	-	-	−0.34 (5)	-	0.55 (3)	0.55 (4)	-	-
SKo92	-	-	0.56 (6)	-	1.05 (3)	−3.94 (5)	−4.00 (3)	-
WA91	2.67 (5)	13.05 (4)	−0.08 (4)	−0.20 (5)	2.14 (5)	−5.28 (5)	-	0.02 (6)
WA92	1.88 (4)	27.41 (9)	−0.34 (6)	0.69 (3)	0.11 (4)	0.33 (3)	-	0.35 (4)

ID91, Aberdeen, ID, 1991; *ID92*, Tetonia, ID, 1992; *MA92*, Brandon, Manitoba, 1992; *MN92*, Crookston, MN, 1992; *MTd91*, Bozeman, MT, dry, 1991; *MTd92*, Bozeman, MT, dry, 1992; *MTi91*, Bozeman, MT, irrigated, 1991; *MTi92*, Bozeman, MT, irrigated, 1992; *NY92*, Ithaca, NY, 1992; *ON92*, Guelph, Ontario, 1992; *OR91*, Klamath Falls, OR, 1991; *Kg92*, Goodlae, Saskatchewan, 1992; *SKk92*, Kcfr, Saskatchewan, 1992; *SKo92*, Outlook, Saskatchewan, 1992; *WA91*, Pullman, WA, 1991; *WA92*, Pullman, WA, 1992. [§]AA, alpha amylase; *DP*, diastatic power; *GP*, grain protein; *GY*, grain yield; *H*, height; *HD*, heading date; *L*, lodging; *ME*, malt extract. [@]The number of genes (number of effective factors) obtained on the basis of phenotypic observations only

of interactions was found for marker observation, but only in 35 cases, where we confirmed results statistically (Table 4). Comparisons of genotypic and phenotypic estimates of the total additive × additive × additive interaction effect show that in majority of cases (79%), the effect was smaller than the total *aaa* interaction effect from phenotypic observations alone (Fig. 2). Same as above, the scope of calculated estimates is quite large ranging from −2194.31% for WG in environment QC93 to 2866.67% for KW in ON93a. In a total of four cases, we observed estimate values higher than 11000%. The smallest range of estimates was observed for the trait NM. The number of genes (effective factors) ranged from 0–13 with an average of 5.6 (Table 3). A minimal number of included markers equals 7, where the maximum number was 21, with an average of 13.9 markers per model. The number of three-way interactions ranged from 0–36 with an average of 4.8 (Table 2).

In total, we analyzed 153 sets of observations, independently for each trait and each environment. Both examples were considered separately.

Discussion

Breeding programs aim to enhance the most desirable traits. Actions based solely on phenotypic observations and gene effects are likely to miss the potentially huge impact of interaction and higher-order interaction effects (Taylor and Ehrenreich 2015). Analytical and numerical comparisons of methods of estimation of the total additive × additive × additive interaction effects are presented in this paper. The numerical comparison was conducted on 153 sets of observations from two examples of barley doubled haploid lines.

The analytic comparison shows that, under the assumption of correct segregation and no linkage between markers, the formulae for the phenotypic and genotypic estimators are comparable and that the additive × additive × additive interaction effect of each QTLs triad is smaller than the phenotypic effect.

Table 2 Genotypic estimates of the total additive × additive × additive interaction effect for the 150 doubled haploid lines of barley obtained from the Steptoe × Morex cross

Environment	Trait							
	AA ^S	DP	GY	GP	HD	H	L	ME
ID91 [#]	*NS **(17 0)	−1.25 (32 1)	0.07 (22 8)	0.14 (17 25)	0.67 (20 11)	−0.53 (24 10)	-	0.02 (27 1)
ID92	2.89 (23 13)	NS (14 0)	0.31 (22 9)	0.08 (21 1)	0.19 (27 3)	1.61 (25 5)	-	−0.07 (27 3)
MA92	-	-	−0.09 (18 1)	-	1.27 (20 16)	−3.60 (20 10)	−16.82 (15 23)	-
MN92	2.24 (26 2)	2.13 (19 17)	−0.34 (18 7)	0.85 (16 23)	−2.70 (19 23)	0.17 (17 20)	-	1.18 (22 14)
MTd91	-	-	0.33 (19 20)	-	0.06 (15 2)	3.41 (22 14)	-	-
MTd92	−0.01 (29 2)	−0.18 (24 6)	0.02 (20 12)	−0.04 (27 4)	−0.28 (24 4)	1.24 (23 4)	−2.39 (26 6)	−0.04 (31 1)
MTi91	−0.25 (18 21)	NS (15 0)	−1.08 (15 23)	0.08 (19 1)	NS (13 0)	NS (14 0)	-	NS (18 0)
MTi92	1.59 (18 7)	2.14 (16 28)	−2.08 (20 16)	NS (13 0)	NS (12 0)	NS (18 0)	11.33 (21 11)	0.07 (15 2)
NY92	-	-	0.02 (22 10)	-	−4.37 (18 25)	−1.67 (16 21)	−2.19 (22 6)	-
ON92	-	-	0.00 (25 7)	-	−0.68 (25 4)	−9.17 (20 11)	−4.60 (24 13)	-
OR91	NS (15 0)	4.22 (15 5)	−0.19 (15 1)	−1.45 (17 30)	NS (15 0)	−1.13 (16 1)	-	−0.34 (22 9)
SKg92	-	-	NS (21 0)	-	NS (15 0)	−1.46 (16 1)	-	-
SKk92	-	-	−0.01 (16 4)	-	0.56 (16 30)	NS (17 0)	-	-
SKo92	-	-	−0.13 (21 8)	-	0.37 (13 1)	−1.32 (22 10)	0.44 (21 8)	-
WA91	3.20 (20 10)	NS (16 0)	NS (13 0)	−0.12 (18 1)	0.07 (14 3)	NS (17 0)	-	NS (13 0)
WA92	1.44 (22 8)	3.88 (20 8)	0.16 (20 14)	0.25 (19 8)	−1.75 (27 4)	5.94 (15 35)	-	1.63 (19 18)

[#]ID91, Aberdeen, ID, 1991; ID92, Tetonia, ID, 1992; MA92, Brandon, Manitoba, 1992; MN92, Crookston, MN, 1992; MTd91, Bozeman, MT, dry, 1991; MTd92, Bozeman, MT, dry, 1992; MTi91, Bozeman, MT, irrigated, 1991; MTi92, Bozeman, MT, irrigated, 1992; NY92, Ithaca, NY, 1992; ON92, Guelph, Ontario, 1992; OR91, Klamath Falls, OR, 1991; Kg92, Goodlae, Saskatchewan, 1992; SKk92, Kcfr, Saskatchewan, 1992; SKo92, Outlook, Saskatchewan, 1992; WA91, Pullman, WA, 1991; WA92, Pullman, WA, 1992. ^SAA, alpha amylase; DP, diastatic power; GP, grain protein; GY, grain yield; H, height; HD, heading date; L, lodging; ME, malt extract. *NS, non significant; ***(x | y): x, number of included markers, y, number of significant *aaa* interactions; “−”, *aaa* interaction not found

The numerical comparison of estimates of additive × additive × additive interaction effect shows that in most cases (79% for both examples), genotypic estimate of *aaa* interaction is smaller than the phenotypic. This sentence is true due to the reason that phenotypic estimate consists of total additive × additive × additive interaction effects of all genes, unlike the genotypic estimate which includes only selected genes. For the rest of the cases that show lower values of phenotypic than genotypic estimates, it may be the result of a high genetic diversity with a lesser phenotypic diversity of the DH lines. High ranges of differences for the calculated estimates are most likely the

result of a lot of different experimental variants such as different traits, environments, and experimental situations (Bocianowski and Krajewski 2009). The number of genes (effective factors) in phenotypic estimation does not directly influence the number of markers, as well as the number of *aaa* interaction included in genotypic models. Both the number of effective factors and number of markers are pretty consistent with few outliers, which makes sense considering that our method tries to include the maximum amount of best-fitted factors. On the contrary, the number of *aaa* interactions ranged quite widely which

Table 3 Phenotypic estimates of the total additive \times additive \times additive interaction effect for the 145 doubled haploid lines of barley obtained from the cross Harrington \times TR306

Environment	Trait						
	WG ^S	NH	NM	H	L	KW	TW
ON92a [#]	-6.02 (10 [@])	0.11 (5)	-1.34 (8)	1.87 (4)	9.24 (3)	1.28 (5)	-0.77 (11)
ON93a	12.12 (7)	0.33 (6)	0.42 (8)	-0.76 (9)	14.43 (3)	0.03 (6)	-1.97 (2)
ON92b	6.21 (5)	0.27 (10)	0.08 (4)	0.26 (3)	-0.34 (3)	0.81 (3)	-0.60 (9)
ON93b	-5.67 (7)	0.25 (6)	0.22 (3)	0.64 (6)	15.65 (6)	0.55 (5)	-0.39 (8)
MB92	-9.00 (5)	0.29 (9)	1.23 (11)	4.48 (4)	-0.51 (4)	0.00 (8)	-2.09 (3)
MB93	-26.10 (6)	0.89 (11)	-0.10 (4)	0.94 (9)	-3.41 (8)	-0.89 (6)	-1.63 (13)
QC93	-9.14 (5)	0.77 (7)	-0.60 (3)	-1.03 (3)	18.30 (5)	-0.71 (3)	-0.96 (5)
SK92a	61.75 (2)	1.15 (7)	0.12 (5)	1.78 (3)	-9.58 (3)	-2.11 (3)	-2.93 (0)
SK93a	-3.39 (7)	-0.54 (4)	-0.87 (5)	0.61 (3)	4.96 (2)	0.71 (7)	-0.68 (8)

[#]ON92a, Ailsa Craig, Ontario, 1992; ON93a, Ailsa Craig, Ontario, 1993; ON92b, Elora, Ontario, 1992; ON93b, Elora, Ontario, 1993; MB92, Brandon, Manitoba, 1992; MB93, Brandon, Manitoba, 1993; QC93, Ste-Anne-de-Bellevue, Quebec, 1993; SK92a, Outlook, Saskatchewan, 1992; SK93a, Outlook, Saskatchewan, 1992. ^SWG, weight of grain harvested per unit area; NH, number of days from planting until emergence of 50% of heads on main tillers; NM, number of days from planting until physiological maturity; H, plant height; L, lodging; KW, 1000 kernel weight; TW, test weight. [@]The number of genes (number of effective factors) obtained on the basis of phenotypic observations only

Table 4 Genotypic estimates of the total additive \times additive \times additive interaction effect for the 145 doubled haploid lines of barley obtained from the cross Harrington \times TR306

Environment	Trait						
	WG ^S	NH	NM	H	L	KW	TW
ON92a [#]	23.94 **(21 7)	NS (16 10)	-0.34 (16 4)	0.56 (14 1)	NS (13 0)	0.72 (13 1)	-0.46 (12 7)
ON93a	5.26 (16 1)	0.10 (12 1)	NS (15 0)	1.26 (15 1)	NS (12 0)	0.86 (13 1)	-0.16 (7 1)
ON92b	-5.54 (20 1)	NS (16 0)	0.15 (14 1)	NS (12 0)	3.93 (12 1)	-0.21 (12 2)	-5.21 (17 18)
ON93b	-3.08 (13 1)	5.46 (16 15)	NS (9 0)	2.69 (16 26)	-1.62 (15 21)	NS (15 0)	NS (13 0)
MB92	9.77 (14 1)	0.27 (13 4)	-0.54 (14 3)	1.42 (14 2)	-1.51 (13 1)	-0.71 (11 2)	NS (16 0)
MB93	67.68 (14 36)	-1.93 (14 24)	NS (14 0)	NS (10 0)	NS (12 0)	2.57 (18 12)	0.50 (15 3)
QC93	200.56 (17 20)	NS (13 0)	-1.91 (15 30)	NS (13 0)	NS (16 0)	NS (13 0)	NS (11 0)
SK92a	*NS (7 0)	NS (17 0)	-0.03 (12 2)	NS (11 0)	NS (13 0)	NS (12 0)	NS (16 0)
SK93a	NS (14 0)	NS (12 0)	-0.40 (17 21)	-0.34 (15 28)	-1.46 (14 2)	NS (13 0)	NS (16 0)

[#]ON92a, Ailsa Craig, Ontario, 1992; ON93a, Ailsa Craig, Ontario, 1993; ON92b, Elora, Ontario, 1992; ON93b, Elora, Ontario, 1993; MB92, Brandon, Manitoba, 1992; MB93, Brandon, Manitoba, 1993; QC93, Ste-Anne-de-Bellevue, Quebec, 1993; SK92a, Outlook, Saskatchewan, 1992; SK93a, Outlook, Saskatchewan, 1992. ^SWG, weight of grain harvested per unit area; NH, number of days from planting until emergence of 50% of heads on main tillers; NM, number of days from planting until physiological maturity; H, plant height; L, lodging; KW, 1000 kernel weight; TW, test weight. *NS, non significant; **(*x* | *y*): *x*, number of included markers, *y*, number of significant *aaa* interactions

may be the result of omitting markers that by themselves do not improve the model but can create the best threes.

In this paper, stepwise feature selection by Akaike information criteria was used. We received comparable results to the previous paper using the same datasets (Bocianowski 2012) with backward stepwise regression as well as to the

method of inclusive interval mapping (ICIM) (described by Li et al. 2008). The presented results show that the inclusion of higher-order (*aaa*) interactions in multiple regression models can have an exert influence on QTL effect.

An important assumption to make is that *aaa* interaction effects show only loci connected to markers with

Fig. 1 Relative comparison of phenotypic and genotypic estimates of the total additive × additive × additive interaction effect for the 150 doubled haploid lines of barley obtained from the Steptoe × Morex cross: box-and-whisker diagram of the values $(\widehat{aaa}_g / \widehat{aaa}_p) \cdot 100$, classified by the observed phenotypic traits (AA, alpha amylase; DP, diastatic power; GP, grain protein; GY, grain yield; H, height; HD, heading date; L, lodging; ME, malt extract)

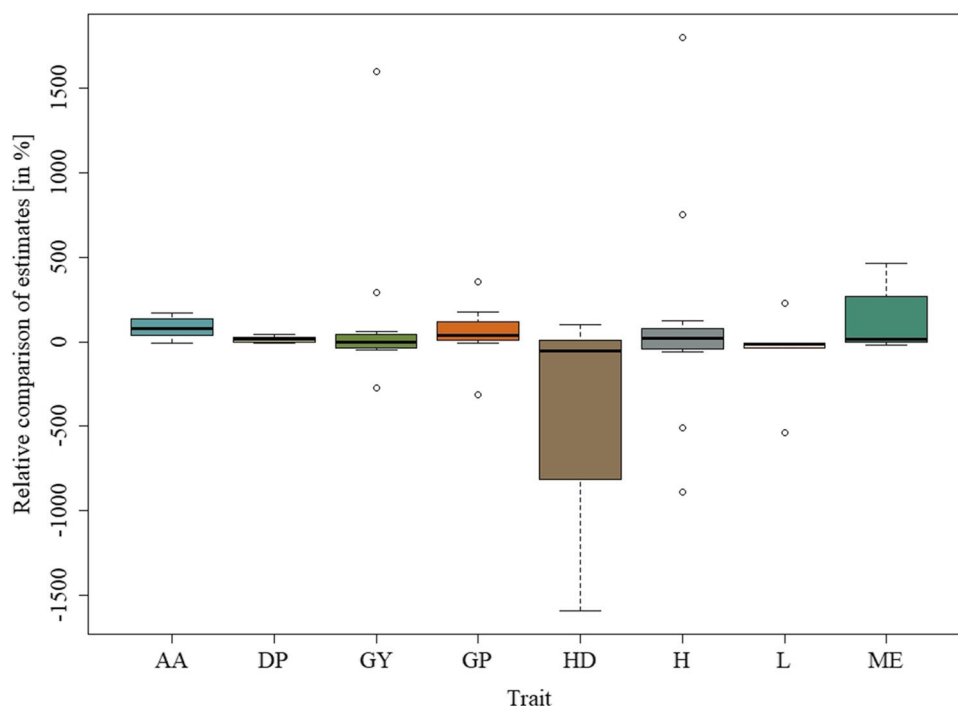
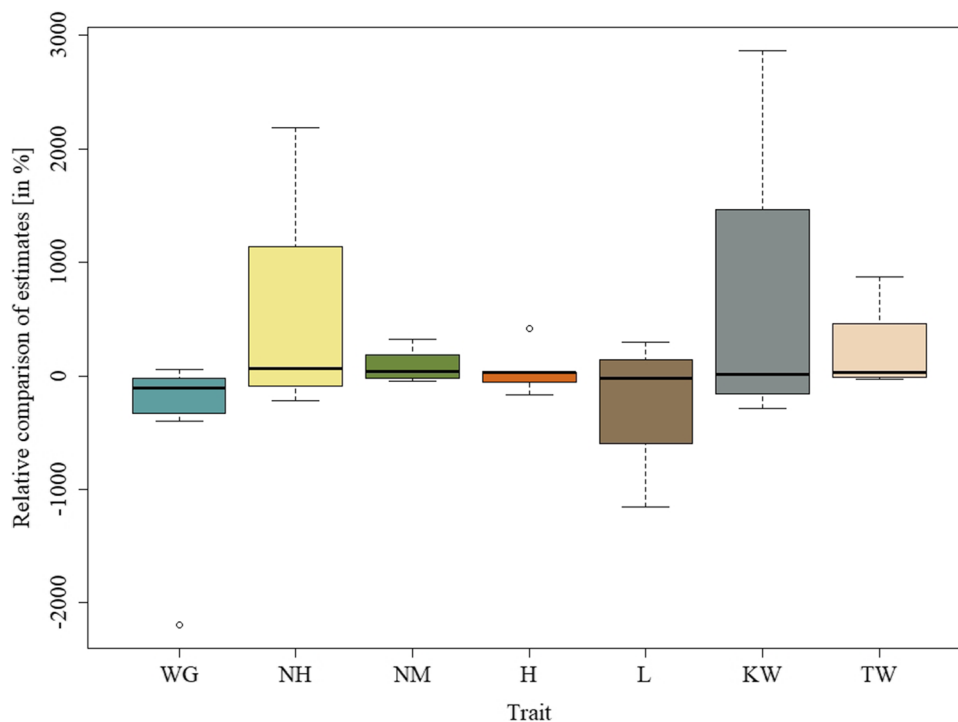


Fig. 2 Relative comparison of phenotypic and genotypic estimates of the total additive × additive × additive interaction effect for the 145 doubled haploid lines of barley obtained from the cross Harrington × TR306: box-and-whisker diagram of the values $(\widehat{aaa}_g / \widehat{aaa}_p) \cdot 100$, classified by the observed phenotypic traits (H, plant height; KW, 1000 kernel weight; L, lodging; NH, number of days from planting until emergence of 50% of heads on main tillers; NM, number of days from planting until physiological maturity; TW, test weight; WG, weight of grain harvested per unit area)



significant effects. Including additional markers may reveal additional interaction but with significant increase of data quantity requirement (Manolio et al. 2009). Further studies are necessary with respect to additive × additive × additive interaction effects conducted by machine learning methods and by simulation analysis that would

make possible consideration of different experimental situations. Current data was not sufficient enough to use machine learning for feature selection. For data containing more markers, we suggest the use of LASSO and SHAP values methods.

Conclusions

Higher-order interactions are usually neglected due to extensive data requirements, although this does not mean they are irrelevant, on the contrary — higher-order interactions occur often and can have a huge impact on phenotype.

The presented methods were useful statistical tools for QTL characteristics and allow estimating *aaa* interactions.

On the basis of available literature, this is the first report concerning the presence of analytical and numerical comparisons of two methods of estimation of additive \times additive \times additive interaction of QTL effects.

Further studies of higher-order interactions and methods of their estimation are necessary.

Author contribution Conceptualization, JB; methodology, AC and JB; software, AC; validation, AC and JB; formal analysis, AC; investigation, AC and JB; resources, AC and JB; data curation, AC and JB; writing—original draft preparation, AC; writing—review and editing, AC and JB; visualization, AC; supervision, JB; all authors have read and agreed to the published version of the manuscript.

Availability of data and material The data presented in this study are available on request from the corresponding authors.

Declarations

Ethics approval This article does not contain any studies with human participants or animals performed by any of the author.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflicts of interest Authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Akaike H (1998) Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K, Kitagawa G (Eds.) Selected Papers of Hirotugu Akaike. Springer Series in Statistics

- (Perspectives in Statistics). Springer: New York, NY, pp. 199–213. https://doi.org/10.1007/978-1-4612-1694-0_15
- Bateson W, Mendel G (1902) Mendel's Principles of Heredity: a defence, with a translation of Mendel's original papers on hybridisation. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511694462>
- Bocianowski J (2012) Analytical and numerical comparisons of two methods of estimation of additive \times additive interaction of QTL effects. *Sci Agric* 69:240–246. <https://doi.org/10.1590/S0103-90162012000400002>
- Bocianowski J, Krajewski P (2009) Comparison of the genetic additive effect estimators based on phenotypic observations and on molecular marker data. *Euphytica* 165:113–122. <https://doi.org/10.1007/s10681-008-9770-x>
- Brem RB, Storey JD, Whittle J, Kruglyak L (2005) Genetic interactions between polymorphisms that affect gene expression in yeast. *Nature* 436:701–703. <https://doi.org/10.1038/nature03865>
- Carlborg Ö, Jacobsson L, Åhlgren P, Siegel P, Andersson L (2006) Epistasis and the release of genetic variation during long-term selection. *Nat Genet* 38(4):418–420. <https://doi.org/10.1038/ng1761>
- Chen CCM, Schwender H, Keith J, Nunkesser R, Mengersen K, Macrossan P (2011) Methods for identifying SNP interactions: a review on variations of Logic Regression, Random Forest and Bayesian logistic regression. *IEEE/ACM Trans Comput Biol Bioinform* 8(6):1580–1591. <https://doi.org/10.1109/TCBB.2011.46>
- Choo TM, Reinbergs E (1982) Analyses of skewness and kurtosis for detecting gene interaction in a doubled haploid population. *Crop Sci* 22(2):231–235. <https://doi.org/10.2135/cropsci1982.0011183X002200020008x>
- Cordell HJ (2009) Detecting gene–gene interactions that underlie human diseases. *Nat Rev Genet* 10(6):392–404. <https://doi.org/10.1038/nrg2579>
- Gaertner BE, Parmenter MD, Rockman MV, Kruglyak L, Phillips PC (2012) More than the sum of its parts: a complex epistatic network underlies natural variation in thermal preference behavior in *Caenorhabditis elegans*. *Genetics* 192(4):1533–1542. <https://doi.org/10.1534/genetics.112.142877>
- Hartman JL, Garvik B, Hartwell L (2001) Principles for the buffering of genetic variation. *Science* 291:1001–1004. <https://doi.org/10.1126/science.1056072>
- Hayes PM, Liu BH, Knapp SJ, Chen F, Jones B, Blake T, Franckowiak J, Rasmusson D, Sorrells M, Ullrich SE, Wesenberg D, Kleinhofs A (1993) Quantitative trait locus effects and environmental interaction in a sample of North American barley germ plasm. *Theor Appl Genet* 87(3):392–401. <https://doi.org/10.1007/BF01184929>
- Jarvis JP, Cheverud JM (2011) Mapping the epistatic network underlying murine reproductive fatpad variation. *Genetics* 187(2):597–610. <https://doi.org/10.1534/genetics.110.123505>
- Kaczmarek Z, Surma M, Adamski T (1988) Epistatic effects in estimation of the number of genes on the basis of doubled haploid lines. *Genetica Polonica* 29:353–359
- Kleinhofs A, Kilian A, Saghai Maroof MA, Biyashev RM, Hayes P, Chen FQ, Lapitan N, Fenwick A, Blake TK, Kanazin V, Ananiev E, Dahleen L, Kudrna D, Bollinger J, Knapp SJ, Liu B, Sorrells M, Heun M, Franckowiak JD, Hoffman D, Skadsen R, Steffenson BJ (1993) A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. *Theor Appl Genet* 86(6):705–712. <https://doi.org/10.1007/BF00222660>
- Li H, Ribaut JM, Li Z, Wang J (2008) Inclusive composite interval mapping (ICIM) for digenic epistasis of quantitative traits in biparental populations. *Theor Appl Genet* 116(2):243–260. <https://doi.org/10.1007/s00122-007-0663-5>
- Mackay TFC (2014) Epistasis and quantitative traits: using model organisms to study gene–gene interactions. *Nat Rev Genet* 15(1):22–33. <https://doi.org/10.1038/nrg3627>

- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TFC, McCarroll SA, Visscher PM (2009) Finding the missing heritability of complex diseases. *Nature* 461(7265):747–753. <https://doi.org/10.1038/nature08494>
- Martinez O, Curnow RN (1994) Missing markers when estimating quantitative trait loci using regression mapping. *Heredity* 73:198–206. <https://doi.org/10.1038/hdy.1994.120>
- Members of the Complex Trait Consortium (2003) The nature and identification of quantitative trait loci: a community's view. *Nat Rev Genet* 4(11):911–916. <https://doi.org/10.1038/nrg1206>
- Romagosa I, Ullrich SE, Han F, Hayes PM (1996) Use of the additive main effects and multiplicative interaction model in QTL mapping for adaptation in barley. *Theor Appl Genet* 93(1):30–37. <https://doi.org/10.1007/BF00225723>
- Searle SR (1982) *Matrix algebra useful for statistics*. Wiley, New York
- Taylor MB, Ehrenreich IM (2015) Higher-order genetic interactions and their contribution to complex traits. *Trends Genet* 31(1):34–40. <https://doi.org/10.1016/j.tig.2014.09.001>
- Tinker NA, Mather DE, Rossnagel BG, Kasha KJ, Kleinhofs A, Hayes PM, Falk DE, Ferguson T, Shugar LP, Legge WG, Irvine RB, Choo TM, Briggs KG, Ullrich SE, Franckowiak JD, Blake TK, Graf RJ, Dofing SM, Saghai Maroof MA, Scoles GJ, Hoffman D, Dahleen LS, Kilian A, Chen F, Biyashev RM, Kudrna DA, Steffenson BJ (1996) Regions of the genome that affect agronomic performance in two-row barley. *Crop Sci* 36(4):1053–1062. <https://doi.org/10.2135/cropsci1996.0011183X003600040040x>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2. Genetic Parameters for Selected Traits of Inbred Lines of Maize (*Zea mays* L.)

Cyplik, A., Sobiech, A., Tomkowiak, A., Bocianowski, J.

Applied Sciences, 2022, 12(14), 6961

DOI:10.3390/app12146961

Impact Factor 2021: 2,838; Punktacja wg MEiN: 100

Genetic Parameters for Selected Traits of Inbred Lines of Maize (*Zea mays* L.)

Adrian Cyplik ¹, Aleksandra Sobiech ², Agnieszka Tomkowiak ² and Jan Bocianowski ^{1,*}

¹ Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Wojska Polskiego 28, 60-637 Poznań, Poland; adrian.cyplik@up.poznan.pl

² Department of Genetics and Plant Breeding, Poznań University of Life Sciences, Dojazd 11, 60-632 Poznań, Poland; aleksandra.sobiech@up.poznan.pl (A.S.); agnieszka.tomkowiak@up.poznan.pl (A.T.)

* Correspondence: jan.bocianowski@up.poznan.pl; Tel.: +48-61-8487143

Abstract: This paper presents an estimation of the parameters connected with the additive (*a*) effect, additive by additive (*aa*) epistatic effect, and additive by additive by additive (*aaa*) interaction gene effect for nine quantitative traits of maize (*Zea mays* L.) inbred lines. To our knowledge, this is the first report about *aaa* interaction of maize inbred lines. An analysis was performed on 252 lines derived from Plant Breeding Smolice Ltd. (Smolice, Poland)—Plant Breeding and Acclimatization Institute-National Research Institute Group (151 lines) and Małopolska Plant Breeding Ltd. (Kobierzyce, Poland) (101 lines). The total additive effects were significant for all studied cases. Two-way and three-way significant interactions were found in most analyzed cases with a considerable impact on phenotype. Omitting the inclusion of higher-order interactions effect in quantitative genetics may result in a substantial underestimation of additive QTL effects. Expanding models with that information may also be helpful in future homozygous line crossing projects.

Keywords: inbred lines; maize; additive effect; epistasis; three-way epistasis



Citation: Cyplik, A.; Sobiech, A.; Tomkowiak, A.; Bocianowski, J. Genetic Parameters for Selected Traits of Inbred Lines of Maize (*Zea mays* L.). *Appl. Sci.* **2022**, *12*, 6961. <https://doi.org/10.3390/app12146961>

Academic Editor: Alexios Polidoros

Received: 22 June 2022

Accepted: 8 July 2022

Published: 9 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Currently, maize, along with wheat and rice, are among the economically most essential grain species [1]. Maize is one of the most important crops that allow for the continuous growth of the world's population. The importance of maize is greatly influenced by its high production potential and the possibility of extensive use of crops (food source, industrial raw material, various types of animal feed). In recent years, the area of maize cultivation in the world ranges from 140 to 145 million hectares, while its production exceeds 600 million tons of grain. The continuous increase in the range of maize cultivation is related to the breeding progress, which includes the use of heterosis and the creation of hybrids with fewer climate requirements [2–4]. It is also crucial to have access to more and more modern breeding methods and cultivation technologies. The demand for new varieties of maize is constantly growing, which makes it the subject of intensive genetic and breeding research.

From the breeding point of view, in the case of quantitative traits, the interaction of genes in the creation of a given trait is critical. It is known that each of the genes that determine a polygenic trait is inherited according to Mendel's laws. Three independent genes and two alleles at the locus are enough for the frequency distribution of the genotypes to approach the normal distribution. However, such a situation occurs in a very simplified case, i.e., when two-allelic loci with an identical, linear phenotypic effect interact additively [5].

In fact, the phenotypic effect of different loci may be different. Loci can be multiallelic. There can be partial or full dominance. We can also deal with overdomination, couplings, and the effect of interallelic interaction. Modifier genes may also appear, modifying the functions of other genes [6,7].

The traits with minor polygenic effects of each genotype overlap with environmental effects, i.e., the variability of the polygenic feature is influenced by the environment. Consequently, it is impossible to identify the genotype based on the phenotype [8].

Therefore, one of the primary tasks of applied genetics is to explain the relationship between genotype and phenotype. Thus, the phenotypic observation of a quantitative trait includes the environmental, genetic, as well as genotype by environment interaction [9,10].

Genetic analyses of various maize genotypes conducted for years have shown that gene interactions—dominance, epistasis, and pleiotropy—play an essential role in the evolution of the maize phenotype. It has long been noticed that many of the loci that differentiate maize and teosinte are pleiotropic. The latest examination of the regulatory architecture of *tb1* provides a detailed understanding of pleiotropy for a single domestication gene. The *tb1* is pleiotropic across many traits such as apical dominance, length of lateral branches, growth of leaves on the lateral branches, pedicellate spikelet development, and root architecture [11,12].

Genetic analyses are of the most importance for understanding the variability of phenotypic traits. The influence on those traits by genes and their pairwise interactions are well known [1,13,14], but considering higher-order interactions is still a novelty in this kind of studies [15].

The aim of this paper was an estimation of the three parameters connected with gene effects: additive (*a*), additive by additive (*aa*) epistatic, and additive by additive by additive (*aaa*) interaction for nine quantitative traits of maize inbred lines. Acquiring this new insight, especially on the effect of the higher-order interaction, was considered for the potential homozygous line crossing. On the basis of available literature, this is the first report concerning the presence of estimation of the additive \times additive \times additive interaction gene effect of maize inbred lines.

2. Materials and Methods

2.1. Plant Material

A collection of 252 inbred maize lines was evaluated. A total of 151 lines were derived from Plant Breeding Smolice Ltd. Plant Breeding and Acclimatization Institute-National Research Institute Group; however, 101 lines were derived from Małopolska Plant Breeding Ltd. These lines (252) were deployed in two localities, 120 km apart, Smolice (51°42'12" N, 17°10'10" E) and Kobierzyce (50°58'1" N, 16°55'50" E).

2.2. Experimental Field Trails

A field experiment with 252 inbred lines was set up on plots of 10 m², in a randomized complete block design (RCBD) in three replicates, in two locations. During the experiments, observations of quantitative traits were carried out. After the harvest, in the first half of November, biometric measurements were made. We observed the following features: cob length, cob diameter, core length, core diameter, the number of rows of grain, the number of grains in a row, thousand kernel weight (TKW), and yield. Measurements were carried out on ten randomly selected cob per replicate.

The soils in Smolice were made of limestone and chalk formations. For this solid valuation III B, and wheat-good complex by quality class, and maize can yield high both on wheat-beet and rye soils, class IV b.

Maize was sown by hand. Mineral fertilization was adapted to the needs of maize grown for grain harvesting. We applied 350 kg·ha⁻¹ polifoska 6 and 160 kg·ha⁻¹ urea. The nutrient content was as follows: before winter grind: 70 kg P₂O₅·ha⁻¹, 21 kg N·ha⁻¹, 105 kg K₂O·ha⁻¹ in polifoska form, and before sowing 73.6 kg N·ha⁻¹ in urea form. Chemical weed control was performed with a plot sprayer.

2.3. Genetic Parameters

The following formula estimated the additive (*a*) gene effect:

$$\hat{a} = \frac{1}{2}(\bar{L}_{max} - \bar{L}_{min}), \quad (1)$$

where \bar{L}_{min} and \bar{L}_{max} are the means for the extreme groups (minimal and maximal lines, respectively). The group of minimal (maximal) lines consists of the lines which contain, theoretically, only alleles reducing (increasing) the value of the trait [16]. Groups of extreme lines were identified by the quantile method [17] in which lines with mean values smaller (bigger) than 0.03 (0.97) quantile of the empirical distribution of means are assumed as minimal (maximal) lines.

The following formula estimated the additive \times additive (aa) epistasis interaction effect [18,19]:

$$\hat{aa} = \frac{1}{2}(\bar{L}_{max} + \bar{L}_{min}) - \bar{L}, \quad (2)$$

where \bar{L} is the mean of all inbred lines.

The additive \times additive \times additive (aaa) three-way epistasis interaction effect was estimated by the following formula [15]:

$$\hat{aaa} = \frac{1}{2}(L_{max} + L_{min}) - \bar{L}, \quad (3)$$

where L_{min} and L_{max} are the lines with minimal and maximal mean value, respectively.

The test statistics to verify hypotheses about genetic parameters different than zero are given by:

$$F_a = \frac{MS_a}{MS_e}, F_{aa} = \frac{MS_{aa}}{MS_e} \text{ and } F_{aaa} = \frac{MS_{aaa}}{MS_e}, \quad (4)$$

where MS_a , MS_{aa} , MS_{aaa} , and MS_e are mean squares for parameter a , epistasis aa , epistasis aaa , and residual, respectively.

3. Results

The obtained results for estimates of the total additive (a) gene effect, the total additive \times additive \times additive (aaa), alongside the two-way epistasis interaction effect, were presented in Table 1. The total additive effects were significant for all studied cases. Results show that 67% (24 out of 36 cases) of the total three-way epistasis effect was statistically significant. All of the aaa interactions were significant for the number of rows of grain in all locations/origins. Significant two-way interactions were found in 58% of studied cases (21 out of 36), and as above, the number of rows of grain was the only one with all locations significant. Between traits, the three-way interactions estimates range was relatively low for most cases, ranging between 0.58 and 1.38. In four cases, however, those values were much higher: the number of rows of grain, the number of grains in a row, mass of grain from the cob and thousand kernel weight (estimates scopes respectively: 2.75, 7.62, 15.66, 20.67). The smallest scope of estimates was found for the cob diameter. The same situation occurred for pairwise interactions, with the same lowest and highest traits. The lowest values of the estimates range were found for cob diameter (0.30) and the highest for the number of rows of grain, a mass of grain from the cob and thousand kernel weight (2.45, 3.12, 11.59, 15.45, respectively). Only for the number of rows of grain, the influence was always positive both in aa and aaa interactions for location/origin. In other cases, influences were mixed but often paired, having the same positive or negative impact on aa and aaa interaction effect.

Table 1. Mean values and the total additive (*a*) effects as well as the total interaction *aa* and *aaa* effects for the 252 (151 from Smolice, 101 from Kobierzyce) inbred lines of maize estimated on the basis of only phenotypic values.

Location	Origin	Parameter	Cob Length	Cob Diameter	Core Length	Core Diameter	The Number of Rows of Grain	The Number of Grains in A Row	Mass of Grain from the Cob	Thousand Kernel Weight	Yield
Kobierzyce	Smolice	mean	15.55	3.97	15.44	2.16	15.73	27.91	114.8	265.15	4.59
		<i>a</i>	4.03 ***	0.92 ***	4.06 ***	0.71 ***	4.80 ***	6.97 ***	47.41 ***	91.79 ***	1.90 ***
		<i>aa</i>	0.09	0.09	0.06	0.12 **	1.07 **	−0.01	−3.19 *	−1.94	−0.13 *
		<i>aaa</i>	0.06	0.18 *	0.24	0.19 **	1.27 **	0.42	−4.19 *	0.47	−0.17 *
Kobierzyce	Kobierzyce	mean	15.37	3.85	15.17	2.03	15.41	27.44	106.61	254.47	4.26
		<i>a</i>	3.83 ***	0.66 ***	3.82 ***	0.59 ***	4.56 ***	8.61 ***	49.44 ***	91.02 ***	1.98 ***
		<i>aa</i>	−0.38 *	−0.08	−0.52 *	−0.01	1.59 ***	−0.05	−1.88	3.02	−0.08
		<i>aaa</i>	−0.76 *	−0.06	−0.80 **	0.03	1.59 ***	0.39	1.04	−1.34	0.04
Smolice	Smolice	mean	13.36	4.09	13.6	2.24	15.01	24.97	94.38	288.9	3.78
		<i>a</i>	3.40 ***	0.75 ***	3.28 ***	0.96 ***	4.40 ***	11.15 ***	42.72 ***	103.9 ***	1.71 ***
		<i>aa</i>	0.44 *	0.03	0.42 *	−0.39 ***	0.86 **	−3.11 ***	3.97 *	−4.06	0.16 *
		<i>aaa</i>	0.30 *	0.24 **	0.57 *	−0.39 ***	0.99 **	−7.20 ***	3.68 *	−3.56	0.15 *
Smolice	Kobierzyce	mean	12.91	3.98	13.3	2.28	14.91	25.32	87.46	276.22	3.5
		<i>a</i>	4.28 ***	0.75 ***	3.56 ***	0.33 ***	4.11 ***	9.06 ***	43.81 ***	97.5 ***	1.75 ***
		<i>aa</i>	−0.47 *	−0.21 **	−0.08	0.04	1.20 **	−2.27 **	−7.62 **	11.39 *	−0.30 **
		<i>aaa</i>	−1.08 **	−0.37 **	−0.30 *	0.04	1.76 ***	−3.32 ***	−11.98 ***	17.11 **	−0.48 ***

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

4. Discussion

Maize is a model plant, so it is an excellent research object [20]. This species, along with rice, is the most commonly grown crop for human and animal consumption. It is a species grown for herb and forage [21]. According to USDA, the world's maize production in 2019/20 was 1116 million tons, and in 2020/21 increased by 42 million tons. Climate change and the massive demand for maize, conditioned by the natural increase, lead the research on genomic regions, which are significant for agronomy [22]. The modern approaches to maize cultivation predominately crosses plants between inbred lines. These hybrids test white vigor (heterosis) and more broad agronomic properties than parents. However, the heterosis effect cannot be estimated based on the phenotype parents line because its basis is unknown. The hypothesis is that heterosis based on completed damaged homozygous alleles is hidden in one line in inbred parents [23].

Heritability, in all its complexity, always gives us an answer in the end. However, fully predicting the outcome before it happens is still out of our reach. One of the missing pieces of information may be found in higher-order genetic interaction studies [24]. The main effects of genes in maize were analyzed by Chaikam et al. [25]. However, epistasis was estimated by Mihaljevic et al. [26], Blanc et al. [27], and Stange et al. [28]. A different, heuristic approach to assessing the effects of genes has been presented by Bocianowski et al. [29].

One of the best-known applications of heritability in quantitative genetic studies of traits is its predictive role, which helps determine the reliability of phenotypic value as an indicator of breeding value. High genetic progress and high heritability scores provide the best conditions for selection [30].

Genotype can be assessed by phenotype, representing the genotypic value in the dynamic environment. Gene expression results from additivity, dominance, and epistasis, affecting the expression of quantitative traits in a population. Inheritance results from the interaction of genotype and phenotype, the variation of which is useful in the selection process [31]. In the analysis shown by Amegbor et al. [32], phenotypic and genetic correlations showed similar trends for the traits analyzed.

According to Gazala et al. [33], a combinations of genes with the diverse large of effects and modes of action (additive, dominance, and epistasis) are involved in the inheritance of complex quantitative traits such as grain yield and the features of the yield structure, including a significant non-crossover/crossover interaction with the environment [33].

Complex maize traits are defined by polygenes exhibiting additive, dominance, epistasis, and gene–environment interaction effects. An example of such traits can be the ones of the leaf. Many statistical methods such as the additive and full genetic models have been developed to discover complex plant traits. However, most genetic association studies focus on additive effects, ignoring non-additive epistasis, environment–gene interactions, and interaction effects [34]. It leads to the absence of the heritability problem, which can significantly impact phenotypic variation [35].

Ma et al. [36] presented that the values of the additive effect of the genes were greater than the non-additive effects. Plant height and ear height were easily affected by environmental factors. The broad-sense heritability and narrow-sense heritability of kernel depth are higher. They are significantly affected by non-additive effects so that they can be selected in the late generation. Similar results for 1000-kernel weight in spring barley hulled and hull-less lines obtained Bocianowski et al. [10].

Results show that 67% of the total three-way epistasis effect and 58% of two-way interactions were statistically significant in the studied cases. These outcomes highlight the importance of including such information in similar studies. High ranges of estimates can be explained by divergent experimental situations with varying locations, origins, and growing conditions of plants [17]. The type of influence, positive or negative, of both two and three-way interaction effects was often paired. This leads to a conclusion that comprehensive analysis of QTLs effects (additive, additive by additive epistatic, and additive by additive by additive interaction) is necessary in the breeding program purposes. Obtained results can be the basis for selection in breeding program. It is the most suitable

method for plant breeding applications, and its main benefit is faster progress in yield and other traits.

5. Conclusions

A significant additive gene action effect in a specific population means that selection beginning in the early generations gives hope for obtaining transgressive homozygous lines.

The detection of epistatic effects enables a better understanding of the interaction of individual genes and allows for a more precise estimation of the effects of individual genes.

Higher-order interactions, although commonly neglected, often occur with significant influence on phenotypic traits. Presented results show that two-way interaction and three-way interaction effects are often similar in influence type (positive or negative) and estimate value itself. Omitting the inclusion of higher-order interactions effect in quantitative genetics may result in a considerable underestimation of additive QTL effects.

Further studies of and with the inclusion of higher-order genetic interactions are necessary.

Author Contributions: Conceptualization, A.C. and J.B.; methodology, A.C. and J.B.; software, A.C. and J.B.; validation, A.C., A.S., A.T. and J.B.; formal analysis, A.C. and J.B.; investigation, A.C., A.S., A.T. and J.B.; resources, A.S. and A.T.; data curation, A.S. and A.T.; writing—original draft preparation, A.C., A.S., A.T. and J.B.; writing—review and editing, A.C., A.S., A.T. and J.B.; supervision, J.B.; project administration, J.B.; funding acquisition, J.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

- Zhong, Y.; Zhao, X.; Li, W.; Zhou, W. Identification and meta-analysis of QTLs for four leaf shape traits in two maize populations under contracting watering environments. *Genet. Resour. Crop Evol.* **2021**, *68*, 1557–1575. [[CrossRef](#)]
- Shehata, A.I.; Al-Ghethar, H.A.; Al-Homaidan, A.A. Application of simple sequence repeat (SSR) markers for molecular diversity and heterozygosity analysis in maize inbred lines. *Saudi J. Biol. Sci.* **2009**, *16*, 57–62. [[CrossRef](#)] [[PubMed](#)]
- Berilli, A.P.C.G.; Pereira, M.G.; Gonçalves, L.S.A.; da Cunha, K.S.; Ramos, H.C.C.; Souza Filho, G.A.; do Amaral Júnior, A.T. Use of molecular markers in reciprocal recurrent selection of maize increases heterosis effects. *Genet. Mol. Res.* **2011**, *10*, 2589–2596. [[CrossRef](#)] [[PubMed](#)]
- Herzog, E.; Falke, K.C.; Presterl, T.; Scheuermann, D.; Ouzunova, M.; Frisch, M. Selection Strategies for the Development of Maize Introgression Populations. *PLoS ONE* **2014**, *9*, e92429. [[CrossRef](#)]
- Jiang, Y.; Schmidt, R.H.; Zhao, Y.; Reif, J.C. A quantitative genetic framework highlights the role of epistatic effects for grain-yield heterosis in bread wheat. *Nat. Genet.* **2017**, *49*, 1741–1746. [[CrossRef](#)] [[PubMed](#)]
- Luo, X.; Fu, Y.; Zhang, P.; Wu, S.; Tian, F.; Liu, J.; Zhu, Z.; Yang, J.; Sun, C. Additive and Over-dominant Effects Resulting from Epistatic Loci Are the Primary Genetic Basis of Heterosis in Rice. *J. Integr. Plant Biol.* **2009**, *51*, 393–408. [[CrossRef](#)]
- Li, C.-Q.; Song, L.; Zhao, H.-H.; Wang, Q.-L.; Fu, Y.-Z. Identification of quantitative trait loci with main and epistatic effects for plant architecture traits in Upland cotton (*Gossypium hirsutum* L.). *Plant Breed.* **2014**, *133*, 390–400. [[CrossRef](#)]
- Bisawas, A.; Sarker, U.; Banik, B.R.; Rohman, M.M.; Talukder, M.Z.A. Genotype × environment interaction for grain yield of maize (*Zea mays* L.) inbreds under salinity stress. *Bangladesh J. Agr. Res.* **2014**, *39*, 293–301. [[CrossRef](#)]
- Acosta-Pech, R.; Crossa, J.; de los Campos, G.; Pérez-Rodríguez, P. Genomic models with genotype × environment interaction for predicting hybrid performance: An application in maize hybrids. *Theor. Appl. Genet.* **2017**, *130*, 1431–1440. [[CrossRef](#)]
- Bocianowski, J.; Warzecha, T.; Nowosad, K.; Bathelt, R. Genotype by environment interaction using AMMI model and estimation of additive and epistasis gene effects for 1000-kernel weight in spring barley (*Hordeum vulgare* L.). *J. Appl. Genet.* **2019**, *60*, 127–135. [[CrossRef](#)]
- Hubbard, L.; McSteen, P.; Doebley, J.; Hake, S. Expression patterns and mutant phenotype of teosinte branched1 correlate with growth suppression in maize and teosinte. *Genetics* **2002**, *162*, 1927–1935. [[CrossRef](#)] [[PubMed](#)]
- Gaudin, A.C.; McClymont, S.A.; Soliman, S.S.; Raizada, M.N. The effect of altered dosage of a mutant allele of teosinte branched 1 (tb1-ref) on the root system of modern maize. *BMC Genet.* **2014**, *15*, 23. [[CrossRef](#)] [[PubMed](#)]

13. Xiao, Y.; Jiang, S.; Cheng, Q.; Wang, X.; Yan, J.; Zhang, R.; Qiao, F.; Ma, C.; Luo, J.; Li, W.; et al. The genetic mechanism of heterosis utilization in maize improvement. *Genome Biol.* **2021**, *22*, 148. [[CrossRef](#)]
14. Chen, Z.; Hu, K.; Yin, Y.; Tang, D.; Ni, J.; Li, P.; Wang, L.; Rong, T.; Liu, J. Identification of a major QTL and genome-wide epistatic interactions for single vs. paired spikelets in a maize-teosinte F₂ population. *Mol. Breed.* **2022**, *42*, 9. [[CrossRef](#)]
15. Cyplik, A.; Bocianowski, J. Analytical and numerical comparisons of two methods of estimation of additive \times additive \times additive interaction of QTL effects. *J. Appl. Genet.* **2022**, *63*, 213–221. [[CrossRef](#)]
16. Choo, T.M.; Reinbergs, E. Analyses of Skewness and Kurtosis for Detecting Gene Interaction in a Doubled Haploid Population. *Crop Sci.* **1982**, *22*, 231–235. [[CrossRef](#)]
17. Bocianowski, J.; Krajewski, P. Comparison of the genetic additive effect estimators based on phenotypic observations and on molecular marker data. *Euphytica* **2009**, *165*, 113–122. [[CrossRef](#)]
18. Bocianowski, J. A comparison of two methods to estimate additive-by-additive interaction of QTL effects by a simulation study. *J. Theor. Biol.* **2012**, *308*, 20–24. [[CrossRef](#)] [[PubMed](#)]
19. Bocianowski, J. Analytical and numerical comparisons of two methods of estimation of additive \times additive interaction of QTL effects. *Sci. Agric.* **2012**, *69*, 240–246. [[CrossRef](#)]
20. Andorf, C.; Beavis, W.D.; Hufford, M.; Smith, S.; Suza, W.P.; Wang, K.; Woodhouse, M.; Yu, J.; Lübberstedt, T. Technological advances in maize breeding: Past, present and future. *Theor. Appl. Genet.* **2019**, *132*, 817–849. [[CrossRef](#)]
21. Yadav, M.K.; Lipsa, D.; Sandeep, R. Evaluation of Host Plant Resistance Traits in Certain Genotypes of Maize (*Zea Mays* L.) against Maize Spotted Stem Borer (*Chilo Partellus* Swinhoe, Crambidae). *Int. J. Entomol. Res.* **2021**, *6*, 64–69.
22. Unterseer, S.; Bauer, E.; Haberer, G.; Seidel, M.; Knaak, C.; Ouzunova, M.; Meitingner, T.; Strom, T.M.; Fries, R.; Pausch, H.; et al. A powerful tool for genome analysis in maize: Development and evaluation of the high density 600 k SNP genotyping array. *BMC Genom.* **2014**, *15*, 823. [[CrossRef](#)] [[PubMed](#)]
23. Yang, J.; Mezouk, S.; Baumgarten, A.; Buckler, E.S.; Guill, K.E.; McMullen, M.D.; Mumm, R.H.; Ross-Ibarra, J. Incomplete dominance of deleterious alleles contributes substantially to trait variation and heterosis in maize. *PLoS Genet.* **2017**, *13*, e1007019. [[CrossRef](#)] [[PubMed](#)]
24. Taylor, M.B.; Ehrenreich, I.M. Higher-order genetic interactions and their contribution to complex traits. *Trends Genet.* **2015**, *31*, 34–40. [[CrossRef](#)]
25. Chaikam, V.; Molenaar, W.; Melchinger, A.E.; Boddupalli, P.M. Doubled haploid technology for line development in maize: Technical advances and prospects. *Theor. Appl. Genet.* **2019**, *132*, 3227–3243. [[CrossRef](#)] [[PubMed](#)]
26. Mihaljevic, R.; Utz, H.F.; Melchinger, A.E. No evidence for epistasis in hybrid and per se performance of elite European flint maize inbreds from generation means and QTL analyses. *Crop Sci.* **2005**, *45*, 2605–2613. [[CrossRef](#)]
27. Blanc, G.; Charcosset, A.; Mangin, B.; Gallais, A.; Moreau, L. Connected populations for detecting quantitative trait loci and testing for epistasis: An application in maize. *Theor. Appl. Genet.* **2006**, *113*, 206–224. [[CrossRef](#)]
28. Stange, M.; Schrag, T.A.; Utz, H.F.; Riedelshheimer, C.; Bauer, E.; Melchinger, A.E. High-density linkage mapping of yield components and epistatic interactions in maize with doubled haploid lines from four crosses. *Mol. Breed.* **2013**, *32*, 533–546. [[CrossRef](#)]
29. Bocianowski, J.; Kozak, M.; Liersch, A.; Bartkowiak-Broda, I. A heuristic method of searching for interesting markers in terms of quantitative traits. *Euphytica* **2011**, *181*, 89–100. [[CrossRef](#)]
30. Jiang, L.; Rong, M.; Wang, M.; Chen, D.; Yu, H. Combining Ability Analysis of Relevant Characters of Maize Inbred Lines Suitable for Machine Harvest. *Comput. Intell. Neurosci.* **2022**, *2022*, 2480801. [[CrossRef](#)]
31. Łacka, A.; Nowosad, K.; Bocianowski, J.; Szulc, P. Heredity of chloroplast pigments parameters of maize (*Zea mays* L.). *Maydica* **2021**, *66*, 8.
32. Amegbor, I.K.; van Biljon, A.; Shargie, N.; Tarekegne, A.; Labuschagne, M.T. Heritability and Associations among Grain Yield and Quality Traits in Quality Protein Maize (QPM) and Non-QPM Hybrids. *Plants* **2022**, *11*, 713. [[CrossRef](#)] [[PubMed](#)]
33. Gazala, P.; Gangappa, E.; Ramesh, S.; Vedashree, M.S.; Raghu, R. Genetics of grain yield and its component traits in maize. *Electron. J. Plant Breed.* **2022**, *13*, 132–136. [[CrossRef](#)]
34. Bocianowski, J. Epistasis interaction of QTL effects as a genetic parameter influencing estimation of the genetic additive effect. *Genet. Mol. Biol.* **2013**, *36*, 93–100. [[CrossRef](#)] [[PubMed](#)]
35. Khatun, M.; Monir, M.M.; Lou, X.; Zhu, J.; Xu, H. Genome-wide association studies revealed complex genetic architecture and breeding perspective of maize ear traits. *BMC Plant Biol.* **2022**. [[CrossRef](#)]
36. Ma, X.Q.; Tang, J.H.; Teng, W.T.; Yan, J.B.; Meng, Y.J.; Li, J.S. Epistatic interaction is an important genetic basis of grain yield and its components in maize. *Mol. Breed.* **2007**, *20*, 41–51. [[CrossRef](#)]

3. QTL×QTL×QTL Interaction Effects for Total Phenolic Content of Wheat Mapping Population of CSDH Lines under Drought Stress by Weighted Multiple Linear Regression

Cyplik, A., Czyczyło-Mysza, I.M., Jankowicz-Cieslak, J., Bocianowski, J.




Agriculture, 2023, 13(4), 850

DOI:10.3390/agriculture13040850

Impact Factor 2022: 3,6; Punktacja wg MEiN: 140

Article

QTL×QTL×QTL Interaction Effects for Total Phenolic Content of Wheat Mapping Population of CSDH Lines under Drought Stress by Weighted Multiple Linear Regression

Adrian Cyplik ^{1,*}, Ilona Mieczysława Czyczyło-Mysza ², Joanna Jankowicz-Cieslak ³ and Jan Bocianowski ^{2,*}

¹ Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, 60-637 Poznań, Poland

² Polish Academy of Sciences, The Franciszek Górski Institute of Plant Physiology, Niezapominajek 21, 30-239 Kraków, Poland

³ Plant Breeding and Genetics Laboratory, Joint FAO/IAEA Centre of Nuclear Techniques in Food and Agriculture, IAEA Laboratories Seibersdorf, International Atomic Energy Agency, Vienna International Centre, P.O. Box 100, A-1400 Vienna, Austria

* Correspondence: adrian.cyplik@up.poznan.pl (A.C.); jan.bocianowski@up.poznan.pl (J.B.)

Abstract: This paper proposes the use of weighted multiple linear regression to estimate the triple interaction (additive×additive×additive) of quantitative trait loci (QTLs) effects. The use of unweighted regression yielded an improvement (in absolute value) in the QTL×QTL×QTL interaction effects compared to assessment based on phenotypes alone in three cases (severe drought in 2010, control in 2012 and severe drought in 2012). In contrast, weighted regression yielded an improvement (in absolute value) in the evaluation of the aaa_{gw} parameter compared to aaa_p in five cases, with the exception of severe drought in 2012. The results show that by using weighted regression on marker observations, the obtained estimates are closer to the ones obtained by the phenotypic method. The coefficients of determination for the weighted regression model were significantly higher than for the unweighted regression and ranged from 46.2% (control in 2010) to 95.0% (control in 2011). Considering this, it is clear that a three-way interaction had a significant effect on the expression of quantitative traits.

Keywords: three-way epistasis; weighted regression; doubled haploid lines; water deprivation stress; *Triticum aestivum*



Citation: Cyplik, A.; Czyczyło-Mysza, I.M.;

Jankowicz-Cieslak, J.; Bocianowski, J. QTL×QTL×QTL Interaction Effects for Total Phenolic Content of Wheat Mapping Population of CSDH Lines under Drought Stress by Weighted Multiple Linear Regression.

Agriculture **2023**, *13*, 850. <https://doi.org/10.3390/agriculture13040850>

Academic Editor: Dongcheng Liu

Received: 26 January 2023

Revised: 4 April 2023

Accepted: 10 April 2023

Published: 11 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Approximately 80–95% of the fresh biomass of a plant is water, which plays an essential role in the plant growth, development and metabolism [1]. Drought is the most severe of all the environmental stresses affecting crop productivity and it belongs to the main factors restricting food production worldwide [2]. It is the single most significant threat to the future of global food security and has been a major famine trigger in the past. The unpredictable nature of drought depends on many factors: uneven and unreliable precipitation distribution, evapotranspiration and water-holding capacity around the rhizosphere [3,4]. Plants have evolved innate adaptations to stress conditions through a series of biochemical and physiological interventions involving the function of many stress-related genes.

Average growth is hampered by drought, which also alters the water cycle and makes plants less efficient water-users. This phenomenon is more complicated in plants since they exhibit a wide range of physiological and biochemical reactions at the cellular and organismal levels. Stomatal closure, membrane damage and disrupted activity of numerous enzymes, particularly those involved in adenosine triphosphate (ATP) generation, are the key factors reducing the rate of photosynthesis. Plants have a variety of defense mechanisms to tolerate drought. In response to drought stress, plants activate drought response mechanisms such as morphological and structural changes, the expression of

drought-tolerance genes, synthesis of hormones and synthesis of osmoregulators to alleviate drought stress [5]. Those actions primarily aim to minimize further oxidative stress, water loss and other unfavorable effects of drought on plants. Secondary metabolites bind reactive oxygen species (ROS) to protect plants from oxidative damage under drought stress [6].

Phenols play multiple roles in plants, as they are structural components of cell walls, contribute significantly to growth and development, and improve plant tolerance to abiotic stresses [7]. These phenolic compounds are electron and proton sources of reactive oxygen species. Drought stress is associated with increased phenol accumulation and antioxidant activity. The accumulation of phenols under abiotic stress belongs to one of the defense mechanisms [8]. Salicylic acid (SA) is an essential phenolic compound currently recognized as a plant growth regulator that enhances plant responses to drought stress. SA is vital for transpiration, ion uptake, solute transport, photosynthesis, flower induction and protein synthesis [9].

Recent advances in genomics have enabled whole-genome sequencing of any genotype. These genomic references allow us to pinpoint each single nucleotide polymorphism's (SNP's) precise chromosomal location and the location is generated by sequence-based genotyping (GBS) [10]. This large number of SNPs is used in genome-wide association studies (GWAS) and quantitative trait locus (QTL) mapping to identify genomic regions or genes that may regulate target traits such drought tolerance, among others. By analyzing the genetics of complex traits, it is possible to better understand abiotic stress defense mechanisms and consequently create varieties with enhanced drought stress tolerance. The further inclusion of genetic interactions in such models can significantly increase the accuracy of such predictions [11].

Most of the important plant traits are quantitative and are influenced by many genes of quantitative trait loci (QTLs). The effect of these genes on the expression of the quantitative trait has been considered in various species, such as barley [12], maize [13], oilseed rape [14–16], rice [17,18], sugar beet [19] and wheat [20]. Many loci showing cumulative minor effects are frequently being weeded out in routine QTL analyses. However, when interacting with other loci, they can significantly affect the observed quantitative trait. In recent years, the effect of epistasis (double interaction) on the observed trait has been considered and reported for barley [21], maize [22], oilseed rape [23], pea [24], rice [25], soybean [26], tomato [27], *Triticale* [28] and wheat [29], among others. Several methods have been developed for gene localization and estimation of the epistasis effect [30–36]. Higher-order interactions (additive×additive×additive, additive×additive×dominance, additive×dominance×additive, dominance×additive×additive, additive×dominance×dominance, dominance×additive×dominance, dominance×dominance×additive and dominance×dominance×dominance) are overlooked, although the concept of genetic interaction has been known for more than a century [37]. It is difficult to imagine that a quantitative trait is not determined by the interaction of more than two genes.

In the case of homozygous material, only additive×additive×additive interaction can be estimated. The research reported in this article aims to estimate parameters connected with the additive×additive×additive (QTL×QTL×QTL) interaction effects for total phenolic content of wheat mapping population from the cross Chinese Spring×SQ1 doubled haploid (CSDH) lines under drought stress by weighted multiple linear regression. The phenotypic and genotypic methods are presented and the selected methods offer a novel approach to determining abiotic stress responses in plants.

2. Materials and Methods

2.1. Plant Material

The plant material used was the mapping population consisting of 94 doubled haploid lines of wheat (CSDH). The genetic map of the population, with a total length of about 4040 cM, contains 920 molecular markers. A detailed description of the mapping population and map is presented in Czyczyło-Mysza et al. [38]. The 94 CSDH lines were vernalized for

7 weeks at 4 °C with an 8/16 h light/dark photoperiod (short day). Germinated seedlings were then planted in 3 L pots (diameter 15 cm and height 20 cm) in 3 replicates per each line and per treatment (564 plants per each year, 3 for control and 3 for drought). Plants were grown in pots in the vegetation tunnel and protected against rain in conditions similar to natural, from spring to autumn (May–September). Limited irrigation was maintained for four weeks starting at the stage of tillering. A 20–25% field water capacity (FWC) was adopted as a severe drought, and 65–70% FWC as well-watered (control). Soil water content was measured with HydroSense Soil Water System CS630 (Campbell Scientific, Logan, UT, USA) using 12 cm probes and weight methods. Every three days, the weight of randomly selected pots was checked to determine the water volume for watering. In the last day of drought treatment, flag leaves were sampled for evaluation of total phenolics contents. Sampled leaves were homogenized in ethanol; the total phenolic contents were determined using the Folin–Ciocalteu method of Singleton and Rossi [39]. The absorbance at 765 nm was read using a microplate reader (Synergy 2, BioTek, Winooski, VT, USA) and chlorogenic acid was used as a standard. The analyses were carried out separately for all years for severe drought and control groups. The experiment was repeated during three planting seasons between 2010 and 2012. The growth conditions and all experimental procedures were maintained.

2.2. Estimation of the QTL×QTL×QTL Interaction Effects

2.2.1. Estimation Based on the Phenotype

The total additive×additive×additive interaction of the homozygous loci (three-way epistasis) effect based on phenotypic (aaa_p) observations of total phenolic content was estimated using the formula [11]:

$$\widehat{aaa}_p = \frac{1}{2}(L_{max} + L_{min}) - \bar{L}, \quad (1)$$

where L_{min} and L_{max} are the doubled haploid lines with minimal and maximal mean value, respectively; \bar{L} is the mean of all DH lines.

The test statistic to verify the hypothesis about aaa_p different than zero is given by [40]

$$F_{aaa_p} = \frac{MS_{aaa_p}}{MS_e}, \quad (2)$$

where MS_{aaa_p} and MS_e are mean squares for epistasis aaa_p and residual, respectively. The number of genes (number of effective factors, \widehat{K}) obtained, based on phenotypic observations only, was calculated using the formula presented by Kaczmarek et al. [41]:

$$\widehat{K} = \frac{(L_{max} - L_{min})^2}{4V_L}, \quad (3)$$

where V_L is additive variance.

2.2.2. Estimation Based on the Genotype

The estimation of aaa_g was based on the assumption that genes responsible for the total phenolic content were completely linked to observed molecular markers. After the selection of p markers (among all observed), determining the total phenolic content, the phenotypic observation model for DH lines is given as

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{A}\boldsymbol{\beta} + \mathbf{E}\boldsymbol{\gamma} + \mathbf{T}\boldsymbol{\delta} + \mathbf{e}, \quad (4)$$

where \mathbf{y} are the mean values of total phenolic content, $\mathbf{1}$ is the n -vector of ones, μ is the general mean, \mathbf{A} is the $(n \times p)$ -matrix of the form $\mathbf{A} = [m_{l_1} \ m_{l_2} \ \dots \ m_{l_p}]$, $l_1, l_2, \dots, l_p \in \{1, 2, \dots, q\}$, $\boldsymbol{\beta}$ is the p -vector of unknown parameters of the form $\boldsymbol{\beta}' = [a_{l_1} \ a_{l_2} \ \dots \ a_{l_p}]$, \mathbf{E} is the matrix for which columns are products of some columns of matrix \mathbf{A} , $\boldsymbol{\gamma}$ is the vector of

unknown parameters of the form $\gamma' = [aa_{l_1l_2} \ aa_{l_1l_3} \ \dots \ aa_{l_{p-1}l_p}]$, T is the matrix for which columns are three-way products of some columns of matrix A , δ is the vector of unknown parameters of the form $\delta' = [aaa_{l_1l_2l_3} \ aaa_{l_1l_2l_4} \ \dots \ aaa_{l_{p-2}l_{p-1}l_p}]$, and e is the n -vector of random variables such that $E(e_i) = 0$, $Cov(e_i, e_j) = 0$ for $i \neq j$, $i, j = 1, 2, \dots, n$. The parameters $a_{l_1}, a_{l_2}, \dots, a_{l_p}$ are the additive effects of the genes controlling the total phenolic content, parameters $aa_{l_1l_2}, aa_{l_1l_3}, \dots, aa_{l_{p-1}l_p}$ are the additive \times additive interaction effects and parameters $aaa_{l_1l_2l_3}, aaa_{l_1l_2l_4}, \dots, aaa_{l_{p-2}l_{p-1}l_p}$ are the additive \times additive \times additive interaction effects. It was assumed that epistatic and three-way epistatic interaction effects show only loci with significant additive gene action effects. Consequently, this assumption results in a decrease in the number of possible significant effects and makes the regression model more useful. Loci with significant additive effects of genes were located and estimated previously by Czyczyło-Mysza et al. [42].

Unweighted Regression

The selection of markers chosen for model (4) can be made, e.g., by a stepwise regression procedure [43]. Here, a three-stage algorithm was used: first, the selection was made by a backward stepwise search independently inside all linkage groups; then, markers chosen in this way were put in one group and subjected to the second backward selection [44]; finally, the third stage considered situations where chosen markers were located on the chromosome very close to each other (closer than 5 cM). Because these markers are probably linked to one QTL, only the marker with the largest test statistic value was retained in the set. A critical significance level of 0.001, derived from the Bonferroni correction [45], was used for the first and second stages.

Denoting by $\alpha' = [\mu \ \beta' \ \gamma' \ \delta']$ and $G = [1 \ A \ E \ T]$, the model was obtained:

$$y = G\alpha + e. \tag{5}$$

If G is of full rank, the estimate of α_u from traditional (unweighted) multiple linear regression model is given by [46]

$$\widehat{\alpha}_u = (G'G)^{-1}G'y. \tag{6}$$

The total three-way epistasis aaa_{gu} effect of genes influencing the total phenolic content from a traditional (unweighted) multiple linear regression model can be found as

$$\widehat{aaa}_{gu} = \sum_{k=1}^{p-2} \sum_{k' \neq k}^{p-1} \sum_{k'' \neq k, k'}^p \widehat{aaa}_{l_k l_{k'} l_{k''}}. \tag{7}$$

Weighted Regression

The modified version of total phenolic content regression on marker data is considered by taking a weighted multiple linear regression, that is, a regression with a diagonal matrix W of unknown variances of observations, which, however, may be empirically found by estimation. The selection of markers for the weighted regression is made by a similar method described for the unweighted case but concerns the weights for lines. In this model, the estimate of α_w is

$$\widehat{\alpha}_w = (G'W^{-1}G)^{-1}G'W^{-1}y, \tag{8}$$

where $W = (w_{ii})$ with w_{ii} is the estimated variance for i -th DH line, $i = 1, 2, \dots, n$. The selection of markers for the weighted regression is made by the same method as described for the unweighted case.

The total three-way epistasis aaa_{gw} effect of genes influencing the total phenolic content from the weighted multiple linear regression model can be found as

$$\widehat{aaa}_{gw} = \sum_{k=1}^{p-2} \sum_{\substack{k'=1 \\ k' \neq k}}^{p-1} \sum_{\substack{k''=1 \\ k'' \neq k, k'}}^p \widehat{aaa}_{l_k l_{k'} l_{k''}} \tag{9}$$

The coefficients of determination were used to measure how both models (unweighted and weighted) fitted the data and were the amounts of the phenotypic variance explained by the total of three interactive models.

All analyses were conducted using the GenStat v. 22 statistics software.

3. Results

3.1. Estimation Based on the Phenotype

Phenotypic estimates of the total additive × additive × additive effect (aaa_p) are presented in Table 1. In four of the six cases, the total aaa_p effect was positive. A negative effect was observed for the 2011 and 2012 control groups. The highest total additive × additive × additive effect was observed for the 2012 severe drought group and the lowest for the 2011 control group. Triples interaction effects were statistically significant for all groups except the control group in 2012.

Table 1. Minimal and maximal values of average for lines, means for all doubled haploid lines for total phenolic content, coefficient of variation, phenotypic estimates of the total additive × additive × additive effect (aaa_p) and the number of genes (the number of effective factors).

Year	2010		2011		2012	
	Control	Severe Drought	Control	Severe Drought	Control	Severe Drought
Minimal	7.104	6.499	7.164	6.231	7.645	4.626
Maximal	14.414	12.682	16.321	15.070	13.781	13.023
Mean	10.450	9.435	11.993	10.214	10.736	8.204
Coefficient of variation	18.06	18.72	15.39	15.51	14.34	23.63
aaa_p	0.309 *	0.156 *	−0.250 *	0.437 **	−0.023	0.620 ***
The number of genes (the number of effective factors)	4.617	3.824	5.050	6.021	4.164	4.846

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Minimal and maximal values of averages for CSDH lines as well as means for all doubled haploid lines for total phenolic content were larger for the control groups than for the severe drought groups in all three years of study. The coefficients of variation were larger for severe drought groups than for the control groups in all cases. The coefficient of variation values ranged from 14.34 for the 2012 control groups to 23.63 for the 2012 severe drought group, indicating low data spread relative to the mean values (Table 1). The number of genes (effective factors) varied between groups, ranging from 3.824 for the 2010 severe drought group to 6.021 for the 2011 severe drought group. The difference between the highest and the lowest number of genes in all groups was equal to 2.197; between the control and severe drought groups measured in the same year, the highest difference was observed in the 2011 group valued at 0.971 (Table 1).

3.2. Estimation Based on the Genotype

Genotypic estimates of the total additive × additive × additive effect calculated based on unweighted (aaa_{gu}) and weighted (aaa_{gw}) multiple linear regression are presented in Table 2.

Table 2. Genotypic estimates of the total additive×additive×additive effects estimated based on unweighted (*aaa_{gu}*) and weighted (*aaa_{gw}*) multiple linear regression.

Year			2010		2011		2012	
Stress			Control	Severe Drought	Control	Severe Drought	Control	Severe Drought
Unweighted	QTLs number		6	5	4	6	7	6
	Number of <i>aaa_{gu}</i>		8	1	0	2	14	8
	<i>aaa_{gu}</i> effects	Min.	−0.288	0.350		−0.489	−0.410	−0.138
		Max.	0.533	0.350		0.569	0.538	0.409
		Total	−0.043	0.350		0.110	0.170	0.688
	R ² [in %]		32.4	42.0		36.3	39.1	44.5
Weighted	QTLs number		16	15	20	14	26	14
	Number of <i>aaa_{gw}</i>		3	5	9	5	6	6
	<i>aaa_{gw}</i> effects	Min.	−0.659	−0.494	−0.607	−0.499	−0.526	−1.124
		Max.	0.598	0.612	0.548	0.368	0.557	1.182
		Total	0.443	−0.265	1.178	−1.302	−0.184	−0.228
	R ² [in %]		46.2	61.4	95.0	58.8	78.9	58.5

3.2.1. Unweighted Regression

The number of QTLs detected using the traditional (unweighted) gene localization method (Table 2) was similar to the number of effective factors (Table 1) and ranged from four (control, 2011) to seven (control, 2012). The number of significant threes of QTL×QTL×QTL genes varied widely and ranged from zero (control, 2011) and one (severe drought, 2010) to fourteen (control, 2012). The range of effect values when the observed trait was determined by at least two threes was 0.547 (severe drought, 2012) to 1.058 (severe drought, 2011 (Tables 2 and 3)). The total *aaa_{gu}* effect ranged from −0.043 (control, 2010) to 0.688 (severe drought, 2012) (Table 2). The phenotypic variation of total phenolic content explained by the total triples interaction was similar in all five cases where the number of triples interactions was greater than zero and ranged from 32.4% (control, 2010) to 44.5% (severe drought, 2012) (Table 2).

Table 3. Genotypic estimates of the additive×additive×additive interaction effects for the particular QTL×QTL×QTL threes based on unweighted (*aaa_{gu}*) and weighted (*aaa_{gw}*) multiple linear regression.

Year			2010		2011		2012	
Stress			Control	Severe Drought	Control	Severe Drought	Control	Severe Drought
QTL1 (LG)	QTL2 (LG)	QTL3 (LG)	<i>aaa_{gu}</i>	<i>aaa_{gw}</i>	<i>aaa_{gu}</i>	<i>aaa_{gw}</i>	<i>aaa_{gu}</i>	<i>aaa_{gw}</i>
dupw004b (1A)	wPt-3094 (3A)	wmc1 (3B)				0.471		
dupw004b (1A)	wPt-668160 (1A)	gwm635b (7A)				0.612		
wPt-3094 (3A)	gwm165.3 (4A)	wPt-0391 (4B)		0.350				
wPt-3094 (3A)	wPt-671748 (7D)	wPt-0391 (4B)						
cfa2262 (3A)	dupw004a (4A)	blt101.t7 (4D)					−0.499	
blt4.1 (3B)	wPt-4048 (3B)	gwm269.2 (4A)						−0.138
blt4.1 (3B)	gwm269.2 (4A)	gwm165.3 (4A)						0.049
wPt-6239 (3B)	blt101.t7 (4D)	wPt-2697 (5A)					−0.337	
wPt-6239 (3B)	blt101.t7 (4D)	wmc83 (7A)					−0.348	
rPt-8896 (3B)	wmc1 (3B)	gwm52 (3D)						
wPt-0021 (3B)	gwm52 (3D)	barc60 (4B)		−0.659				
gwm269.2 (4A)	gwm165.3 (4A)	gwm205 (5D)	0.008					
gwm269.2 (4A)	gwm165.3 (4A)	barc44 (5D)	−0.152					
gwm269.2 (4A)	gwm165.3 (4A)	m69p78.1 (7A)	−0.100					
gwm269.2 (4A)	gwm205 (5D)	barc44 (5D)	0.063					
gwm269.2 (4A)	gwm205 (5D)	m69p78.1 (7A)	−0.263					
gwm269.2 (4A)	barc44 (5D)	wPt-9834 (5A)	0.533					
gwm165.3 (4A)	gwm165.3 (4A)	wPt-667091 (7D)					0.158	
gwm165.3 (4A)	wPt-0391 (4B)	wPt-3883 (7A)					0.026	
gwm165.3 (4A)	wPt-0391 (4B)	wPt-667091 (7D)					−0.179	
gwm165.3 (4A)	wPt-0391 (4B)	wmc243b (7D)					0.247	
gwm165.3 (4A)	gwm205 (5D)	m69p78.1 (7A)	0.156					
gwm165.3 (4A)	barc44 (5D)	m69p78.1 (7A)	−0.288					
gwm165.3 (4A)	gwm174 (5D)	wPt-3883 (7A)					0.026	
gwm165.3 (4A)	gwm174 (5D)	wPt-667091 (7D)					0.031	
gwm165.3 (4A)	gwm174 (5D)	wmc243b (7D)					−0.161	
gwm165.3 (4A)	wPt-3883 (7A)	wmc243b (7D)					−0.008	
gwm165.3 (4A)	wPt-0391 (4B)	wPt-667091 (7D)					−0.410	
wPt-0391 (4B)	gwm174 (5D)	wPt-3883 (7A)					0.128	
wPt-0391 (4B)	gwm174 (5D)	wmc243b (7D)					0.538	
wPt-0391 (4B)	wPt-3883 (7A)	wPt-667091 (7D)					−0.151	
barc152 (1B)	wPt-6239 (3B)	gwm191b (3D)						−0.400

Table 3. Cont.

Year			2010				2011				2012			
Stress			Control		Severe Drought		Control		Severe Drought		Control		Severe Drought	
QTL1 (LG)	QTL2 (LG)	QTL3 (LG)	<i>aaa_{gu}</i>	<i>aaa_{gw}</i>	<i>aaa_{gu}</i>	<i>aaa_{gw}</i>	<i>aaa_{gu}</i>	<i>aaa_{gw}</i>	<i>aaa_{gu}</i>	<i>aaa_{gw}</i>	<i>aaa_{gu}</i>	<i>aaa_{gw}</i>	<i>aaa_{gu}</i>	<i>aaa_{gw}</i>
barc152 (1B)	m65p64.8_4B (4B)	wPt-8149 (7A)											−0.526	
barc152 (1B)	m92p78.10 (2A)	m60p64.13_3B (3B)											0.557	
barc152 (1B)	wPt-8072 (2B)	gwm165.3 (4A)											−0.493	
m65p64.8a (5B)	gwm271 (5B)	m69p78.1 (7A)		0.504										
gwm174 (5D)	wPt-3883 (7A)	wPt-667091 (7D)									−0.130			
gwm174 (5D)	wPt-3883 (7A)	wmc243b (7D)									0.055			
psr648_1B (1B)	wmc181 (2A)	gwm285 (3B)		0.598										
m17p65.1 (1B)	cfid73 (2D)	wmc468 (4A)			−0.478									
wmc432 (1D)	wPt-3738 (1D)	psp2151.3 (2A)												−0.640
wPt-3738 (1D)	tPt-0202 (3A)	gwm269.2 (4A)												−1.124
wPt-3738 (1D)	tPt-0202 (3A)	dupw004a (4A)												1.182
wPt-3738 (1D)	gwm513 (4B)	wmc157 (7D)												0.387
wPt-3738 (1D)	wmc429 (1D)	tPt-0202 (3A)												0.354
wmc429 (1D)	cfid11 (2D)	wPt-10291 (3D)						−0.505						
wmc429 (1D)	gwm30.1 (2D)	tPt-0202 (3A)						0.481						
wmc429 (1D)	wPt-7466 (2D)	wPt-9749 (2D)						−0.607						
wmc429 (1D)	gwm269.2 (4A)	dupw004a (4A)												−0.387
wmc429 (1D)	wPt-6316 (1D)	wPt-7466 (2D)						0.548						
wmc429 (1D)	wPt-6316 (1D)	wPt-10291 (3D)						0.436						
wmc429 (1D)	wPt-6316 (1D)	gwm161 (3D)						0.493						
wmc429 (1D)	wPt-6316 (1D)	gwm165.3 (4A)						−0.511						
wmc429 (1D)	wPt-6316 (1D)	gwm639.2 (5B)						0.419						
wmc429 (1D)	wPt-732556 (1D)	gwm30.1 (2D)						0.424						
m92p78.10 (2A)	wPt-6239 (3B)	gwm161 (3D)										0.360		
m92p78.10 (2A)	wPt-6239 (3B)	wPt-8149 (7A)										0.317		
wmc453a (2A)	wPt-3883 (7A)	wPt-8919 (7B)							0.569					
wmc453a (2A)	wmc283.1 (7A)	wPt-8919 (7B)							−0.459					
psp2151.3 (2A)	bit4.1 (3B)	wPt-4048 (3B)												0.201
psp2151.3 (2A)	bit4.1 (3B)	gwm269.2 (4A)												0.136
psp2151.3 (2A)	bit4.1 (3B)	gwm165.3 (4A)												−0.095
psp2151.3 (2A)	wPt-4048 (3B)	gwm269.2 (4A)												0.409
psp2151.3 (2A)	wPt-4048 (3B)	gwm165.3 (4A)												0.091
psp2151.3 (2A)	gwm269.2 (4A)	gwm165.3 (4A)												0.035
wPt-3949 (2B)	wmc257 (2B)	gwm165.3 (4A)								0.368				
wmc257 (2B)	wPt-2697 (5A)	gwm292_5D (5D)								−0.485				

3.2.2. Weighted Regression

The use of weighted regression resulted in a definite increase in the number of QTLs detected compared to when classical unweighted regression was used. The increase ranged from 233% (from 6 to 14 in severe drought, 2011 and 2012) to 500% (from 4 to 20 in the control group in 2011) (Table 2). The number of threes using weighted regression was higher than using unweighted regression in three cases (severe drought, 2010; control 2011 and severe drought 2011). In the remaining cases (control, 2010; control, 2012 and severe drought 2012), it was smaller (Table 2). The spread of estimated effect values was larger in all cases except severe drought in 2011, where it decreased from 1.058 to 0.867 (Tables 2 and 3). The coefficients of determination for the weighted regression model were significantly higher than for the unweighted regression and ranged from 46.2% (control in 2010) to 95.0% (control in 2011) (Table 2).

Genotypic estimates of the additive×additive×additive interaction effects for the particular QTL×QTL×QTL threes based on unweighted (*aaa_{gu}*), and weighted (*aaa_{gw}*) multiple linear regression are presented in Table 3. Sixty-seven statistically significant triples interactions were observed. None of the QTL×QTL×QTL threes were significant in at least two cases: threes significant using weighted regression were not significant using unweighted regression and vice versa (Table 3). The QTLs most frequently found in triples interaction were gwm165.3 (22 times), gwm269.2 (11 times) and wmc429 (11 times). Using unweighted regression, the genes most frequently found in triples interaction were gwm165.3 (19 times), gwm269.2 (11 times) and wPt-0391 (9 times) (Table 3). Using weighted regression, the genes most frequently observed in triples interaction were wmc429 (eleven times), wPt-3738 (five times), wPt-6239 (five times) and wPt-6316 (also five times) (Table 3).

In 43 triples interactions, each QTL was located on a different linkage group (LG). In twenty-four cases, two QTLs (out of three) were located on one LG. There was no situation in which all three QTLs from a threesome were located on a single LG (Table 3). Using unweighted regression, 23 cases were observed with each of the QTLs of the 3 being at a different LG. In contrast, in ten cases, two QTLs were located on a single chromosome. Using weighted regression, 20 cases were observed where each QTL of the significant 3 was localized in a different linkage group. In contrast, in 14 cases, 2 QTLs were localized to

a single LG (Table 3). The distribution of QTLs included in the triples interaction on each linkage group varied widely, ranging from 2 (on 7B) to 39 (on 4A) (Table 3).

4. Discussion

The aim of a breeding process is to obtain new genotypes with improved traits compared to the parental forms. The parameter connected with the additive \times additive \times additive interaction gene action can influence decisions about the usefulness of the breeding material for that purpose. A numerical comparison of the three methods of estimation of the additive \times additive \times additive interaction effects are presented in this paper. The comparison was conducted on 94 doubled haploid lines (DHLs) generated from the cross between hexaploid wheat (*Triticum aestivum* L.) genotypes Chinese Spring (CS) and SQ1 (a high abscisic acid breeding line) according to Quarrie et al. [47]. The total phenolic contents were measured separately for three years (2010–2012) for severe drought and control groups. The present results demonstrated the use of weighted regression to determine the triplets of QTLs and estimate the effects of their QTL \times QTL \times QTL interaction. To the best of our knowledge, there is no published evidence on the use of weighted regression to assess the QTL \times QTL \times QTL triple interaction. Most studies focus on the analysis of single genes as determinants of a quantitative trait. Weighted regression was used to estimate QTL \times QTL epistasis interaction [48–52]. However, higher-order interactions are ignored. Therefore, the novelty of this proposal is in determining how to select triples of QTLs and estimate their effects using weighted regression.

Previously, in [11], we showed that QTL \times QTL \times QTL interaction effects assessments obtained with molecular marker observations (using traditional unweighted regression) had smaller absolute values than assessments based on phenotypic observations alone. The differences obtained were so large that it was decided to propose a new method with a chance to improve QTL \times QTL \times QTL assessments.

The present results obviously demonstrated that the detected QTL \times QTL \times QTL affected the total phenolic content. An important assumption to make is that *aaa* interaction effects show only loci connected to markers with significant effects. However, it should be emphasized that the selection of markers with significant effects was carried out using weighted regression. The use of unweighted regression yielded an improvement (in absolute value) in the effect of *aaa_{gu}* compared to assessment based on phenotypes alone in three cases (severe drought in 2010, control in 2012 and severe drought in 2012). In contrast, weighted regression yielded an improvement (in absolute value) in the evaluation of the *aaa_{gw}* parameter compared to *aaa_p* in five cases, with the exception of severe drought in 2012 (Tables 1 and 2).

Considering the effect of three-way interaction on the expression of quantitative traits may also be important in the context of hypostasis, a situation in which the phenotype is altered by the expression of an allele at a separate locus, in an epistasis event [53]. Genetic interaction analysis may be a powerful genetic strategy for analyzing the fitness and phenotypes of genotypes to dissect the interactions between genes, reason genes into biological pathways, and characterize genes of unknown operation. Omitting to consider the effect of higher-order interactions in quantitative genetics can result in a lack of completeness in determining a quantitative trait. Cyplik et al. [39], in their study of maize inbred lines, observed the significance of a triple interaction in the absence of the epistasis effect. This phenomenon occurred in three situations: for cob diameter (twice) and for core length. The reverse situation (significant epistasis and non-significant triple interaction) did not occur in any case.

5. Conclusions

The results show that by using weighted regression on marker observations, the obtained estimates are closer to the ones obtained by the phenotypic method. Considering this, it is clear that a three-way interaction had a significant effect on the expression of quantitative traits. The proposed weighted regression method for estimation of the

QTL×QTL×QTL interaction effect can be a bridge between the phenotypic method and the genotypic method.

Author Contributions: Conceptualization, J.B.; methodology, A.C., I.M.C.-M. and J.B.; software, A.C. and J.B.; validation, A.C., I.M.C.-M., J.J.-C. and J.B.; formal analysis, A.C. and J.B.; investigation, A.C., I.M.C.-M. and J.B.; resources, I.M.C.-M. and J.B.; data curation, I.M.C.-M.; writing—original draft preparation, A.C., I.M.C.-M., J.J.-C. and J.B.; writing—review and editing, A.C., I.M.C.-M., J.J.-C. and J.B.; supervision, J.B.; project administration, J.B.; funding acquisition, I.M.C.-M. and J.B. All authors have read and agreed to the published version of the manuscript.

Funding: The study was partly financed by the statutory investigations in The Franciszek Górski Institute of Plant Physiology, Polish Academy of Sciences, Kraków, Poland.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Brodersen, C.R.; Roddy, A.B.; Wason, J.W.; McElrone, A.J. Functional Status of Xylem Through Time. *Annu. Rev. Plant Biol.* **2019**, *70*, 407–433. [[CrossRef](#)] [[PubMed](#)]
2. Khan, M.A.; Iqbal, M.; Akram, M.; Ahmad, M.; Hassan, M.W.; Jamil, M. Recent Advances in Molecular Tool Development for Drought Tolerance Breeding in Cereal Crops: A Review. *Zemdirb.-Agric.* **2013**, *100*, 325–334. [[CrossRef](#)]
3. Seleiman, M.F.; Al-Suhaibani, N.; Ali, N.; Akmal, M.; Alotaibi, M.; Refay, Y.; Dindaroglu, T.; Abdul-Wajid, H.H.; Battaglia, M.L. Drought Stress Impacts on Plants and Different Approaches to Alleviate Its Adverse Effects. *Plants* **2021**, *10*, 259. [[CrossRef](#)]
4. Daryanto, S.; Wang, L.; Jacinthe, P.-A. Global Synthesis of Drought Effects on Cereal, Legume, Tuber and Root Crops Production: A Review. *Agric. Water Manag.* **2017**, *179*, 18–33. [[CrossRef](#)]
5. Lisar, S.Y.S.; Motafakkerazad, R.; Hossain, M.M.; Rahman, I.M.M.; Lisar, S.Y.S.; Motafakkerazad, R.; Hossain, M.M.; Rahman, I.M.M. *Water Stress in Plants: Causes, Effects and Responses*; IntechOpen: Rijeka, Croatia, 2012; ISBN 978-953-307-963-9.
6. Yadav, B.; Jogawat, A.; Rahman, M.S.; Narayan, O.P. Secondary Metabolites in the Drought Stress Tolerance of Crop Plants: A Review. *Gene Rep.* **2021**, *23*, 101040. [[CrossRef](#)]
7. Naikoo, M.I.; Dar, M.I.; Raghieb, F.; Jaleel, H.; Ahmad, B.; Raina, A.; Khan, F.A.; Naushin, F. Chapter 9—Role and Regulation of Plants Phenolics in Abiotic Stress Tolerance: An Overview. In *Plant Signaling Molecules*; Khan, M.I.R., Reddy, P.S., Ferrante, A., Khan, N.A., Eds.; Woodhead Publishing: Duxford, UK, 2019; pp. 157–168. ISBN 978-0-12-816451-8.
8. Hura, T.; Hura, K.; Ostrowska, A.; Grzesiak, M.; Dziurka, K. The Cell Wall-Bound Phenolics as a Biochemical Indicator of Soil Drought Resistance in Winter Triticale. *Plant Soil Environ.* **2013**, *59*, 189–195. [[CrossRef](#)]
9. Latif, F.; Ullah, F.; Mehmood, S.; Khattak, A.; Khan, A.U.; Khan, S.; Husain, I. Effects of Salicylic Acid on Growth and Accumulation of Phenolics in *Zea Mays* L. under Drought Stress. *Acta Agric. Scand. Sect. B* **2016**, *66*, 325–332. [[CrossRef](#)]
10. Hussain, W.; Baenziger, P.S.; Belamkar, V.; Guttieri, M.J.; Venegas, J.P.; Easterly, A.; Sallam, A.; Poland, J. Genotyping-by-Sequencing Derived High-Density Linkage Map and Its Application to QTL Mapping of Flag Leaf Traits in Bread Wheat. *Sci. Rep.* **2017**, *7*, 16394. [[CrossRef](#)]
11. Cyplik, A.; Bocianowski, J. Analytical and Numerical Comparisons of Two Methods of Estimation of Additive × Additive × Additive Interaction of QTL Effects. *J. Appl. Genet.* **2022**, *63*, 213–221. [[CrossRef](#)]
12. Sayed, M.A.; Nassar, S.M.; Moustafa, E.S.; Said, M.T.; Börner, A.; Hamada, A. Genetic Mapping Reveals Novel Exotic and Elite QTL Alleles for Salinity Tolerance in Barley. *Agronomy* **2021**, *11*, 1774. [[CrossRef](#)]
13. Ren, J.; Zhang, X.; Li, Z.; Wu, P. Genetic Analysis of Maternal Haploid Inducibility for *In Vivo* Haploid Induction in Maize. *Agriculture* **2022**, *12*, 845. [[CrossRef](#)]
14. Bocianowski, J.; Kozak, M.; Liersch, A.; Bartkowiak-Broda, I. A heuristic method of searching for interesting markers in terms of quantitative traits. *Euphytica* **2011**, *181*, 89–100. [[CrossRef](#)]
15. Botero-Ramírez, A.; Laperche, A.; Guichard, S.; Jubault, M.; Gravot, A.; Strelkov, S.E.; Manzanares-Dauleux, M.J. Clubroot Symptoms and Resting Spore Production in a Doubled Haploid Population of Oilseed Rape (*Brassica napus*) Are Controlled by Four Main QTLs. *Front. Plant Sci.* **2020**, *11*, 604527. [[CrossRef](#)]
16. Gacek, K.; Bayer, P.E.; Anderson, R.; Severn-Ellis, A.A.; Wolko, J.; Łopatyńska, A.; Matuszczak, M.; Bocianowski, J.; Edwards, D.; Batley, J. QTL Genetic Mapping Study for Traits Affecting Meal Quality in Winter Oilseed Rape (*Brassica Napus* L.). *Genes* **2021**, *12*, 1235. [[CrossRef](#)]
17. Kabange, N.R.; Park, S.-Y.; Shin, D.; Lee, S.-M.; Jo, S.-M.; Kwon, Y.; Cha, J.-K.; Song, Y.-C.; Ko, J.-M.; Lee, J.-H. Identification of a Novel QTL for Chlorate Resistance in Rice (*Oryza sativa* L.). *Agriculture* **2020**, *10*, 360. [[CrossRef](#)]

18. Kwon, Y.-H.; Kabange, N.-R.; Lee, J.-Y.; Lee, S.-M.; Cha, J.-K.; Shin, D.-J.; Cho, J.-H.; Kang, J.-W.; Ko, J.-M.; Lee, J.-H. Novel QTL Associated with Shoot Branching Identified in Doubled Haploid Rice (*Oryza sativa* L.) under Low Nitrogen Cultivation. *Genes* **2021**, *12*, 745. [[CrossRef](#)]
19. Pegot-Espagnet, P.; Guillaume, O.; Desprez, B.; Devaux, B.; Devaux, P.; Henry, K.; Henry, N.; Willems, G.; Goudemand, E.; Mangin, B. Discovery of interesting new polymorphisms in a sugar beet (elite × exotic) progeny by comparison with an elite panel. *Theor. Appl. Genet.* **2019**, *132*, 3063–3078. [[CrossRef](#)]
20. Lephuthing, M.C.; Khumalo, T.P.; Tolmay, V.L.; Dube, E.; Tsiolo, T.J. Genetic Mapping of Quantitative Trait Loci Associated with Plant Height and Yield Component Traits in a Wheat (*Triticum aestivum* L.) Doubled Haploid Population Derived from Tugela-DN × Elands. *Agronomy* **2022**, *12*, 2283. [[CrossRef](#)]
21. Beheshtizadeh, H.; Fakheri, B.A.; Aghnoum, R.; Mahdinezhad, N.; Pourdad, S.S.; Masoudi, B. QTL mapping of grain yield and its components under normal and drought stress conditions in barley (*Hordeum vulgare* L.). *Indian J. Genet. Plant Breed.* **2018**, *78*, 69–80. [[CrossRef](#)]
22. Ku, L.X.; Sun, Z.H.; Wang, C.L.; Zhang, J.; Zhao, R.F.; Liu, H.Y.; Tai, G.Q.; Chen, Y.H. QTL mapping and epistasis analysis of brace root traits in maize. *Mol. Breed.* **2012**, *30*, 697–708. [[CrossRef](#)]
23. Yusuf, A.O.; Richter, J.-C.; Möllers, C. Genetic variation and QTL analysis of saturated fatty acids in two doubled haploid populations of oilseed rape (*Brassica napus* L.). *Euphytica* **2022**, *218*, 88. [[CrossRef](#)]
24. Krajewski, P.; Bocianowski, J.; Gawłowska, M.; Kaczmarek, Z.; Pniewski, T.; Świącicki, W.; Wolko, B. QTL for yield components and protein content: A multi-environment study of two pea (*Pisum sativum* L.) populations. *Euphytica* **2012**, *183*, 323–336. [[CrossRef](#)]
25. Ali, F.; Chen, W.; Fiaz, S.; Wang, Y.; Wei, X.; Xie, L.; Jiao, G.; Shao, G.; Hu, S.; Tang, S.; et al. QTL Mapping for Grain Appearance Quality Traits Using Doubled Haploid Population of Rice Under Different Environments. *Pak. J. Bot.* **2022**, *54*, 1265–1275. [[CrossRef](#)] [[PubMed](#)]
26. Han, Y.; Tan, Y.; Hu, H.; Chang, W.; Dong, L.; Wang, Z.; Zhao, X.; Li, W.; Teng, W. Quantitative trait loci with additive and epistatic effects underlying resistance to two hg types of soybean cyst nematode. *Plant Breed.* **2017**, *136*, 720–727. [[CrossRef](#)]
27. Smeda, J.R.; Schillmiller, A.L.; Anderson, T.; Ben-Mahmoud, S.; Ullman, D.E.; Chappell, T.M.; Kessler, A.; Mutschler, M.A. Combination of Acylglucose QTL reveals additive and epistatic genetic interactions and impacts insect oviposition and virus infection. *Mol. Breed.* **2018**, *38*, 3. [[CrossRef](#)]
28. Dhariwal, R.; Fedak, G.; Dion, Y.; Pozniak, C.; Laroche, A.; Eudes, F.; Randhawa, H.S. High Density Single Nucleotide Polymorphism (SNP) Mapping and Quantitative Trait Loci (QTL) Analysis in a Biparental Spring Triticale Population Localized Major and Minor Effect *Fusarium* Head Blight Resistance and Associated Traits QTL. *Genes* **2018**, *9*, 19. [[CrossRef](#)]
29. Pundir, S.; Sharma, R.; Kumar, D.; Singh, V.K.; Chaturvedi, D.; Kanwar, R.S.; Röder, M.S.; Börner, A.; Ganai, W.M.; Gupta, P.K.; et al. QTL mapping for resistance against cereal cyst nematode (*Heterodera avenae* Woll.) in wheat (*Triticum aestivum* L.). *Sci. Rep.* **2022**, *12*, 9586. [[CrossRef](#)]
30. Chase, K.; Adler, F.R.; Lark, K.G. EPISTAT: A computer program for identifying and testing interaction between pairs of quantitative trait loci. *Theor. Appl. Genet.* **1997**, *94*, 724–730. [[CrossRef](#)]
31. Holland, J.B. Computer note. EPISTACY: A SAS program for detecting two-locus epistatic interaction using genetic marker information. *J. Hered.* **1998**, *89*, 374–375. [[CrossRef](#)]
32. Kao, C.-H.; Zeng, Z.-B.; Teasdale, R.D. Multiple interval mapping for quantitative trait loci. *Genetics* **1999**, *152*, 1203–1216. [[CrossRef](#)]
33. Zeng, Z.-B.; Kao, C.-H.; Batsen, C.J. Estimating the genetic architecture of quantitative traits. *Genet. Res.* **1999**, *74*, 279–289. [[CrossRef](#)]
34. Carlborg, Ö.; Andersson, L.; Kinghorn, B. The use of a genetic algorithm for simultaneous mapping interacting quantitative trait loci. *Genetics* **2000**, *155*, 2003–2010. [[CrossRef](#)]
35. Sen, S.; Churchill, G.A. A statistical framework for quantitative trait mapping. *Genetics* **2001**, *159*, 371–387. [[CrossRef](#)]
36. Bocianowski, J. The use of weighted multiple linear regression to estimate QTL-by-QTL epistatic effects. *Genet. Mol. Biol.* **2012**, *35*, 802–809. [[CrossRef](#)]
37. Bateson, W.; Mendel, G. *Mendel's Principles of Heredity: A Defence, with a Translation of Mendel's Original Papers on Hybridization*; Cambridge University Press: Cambridge, UK, 1902. [[CrossRef](#)]
38. Czyczyło-Mysza, I.; Tyrka, M.; Marcinska, I.; Skrzypek, E.; Karbarz, M.; Dziurka, M.; Hura, T.; Dziurka, K.; Quarrie, S.A. Quantitative trait loci for leaf chlorophyll fluorescence parameters, chlorophyll and carotenoid contents in relation to biomass and yield in bread wheat and their chromosome deletion bin assignments. *Mol. Breed.* **2013**, *32*, 189–210. [[CrossRef](#)]
39. Singleton, V.L.; Rossi, J.A. Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. *Am. J. Enol. Viticult.* **1965**, *16*, 144–158.
40. Cyplik, A.; Sobiech, A.; Tomkowiak, A.; Bocianowski, J. Genetic Parameters for Selected Traits of Inbred Lines of Maize (*Zea mays* L.). *Appl. Sci.* **2022**, *12*, 6961. [[CrossRef](#)]
41. Kaczmarek, Z.; Surma, M.; Adamski, T. Epistatic effects in estimation of the number of genes on the basis of doubled haploid lines. *Genet. Pol.* **1988**, *29*, 353–359.
42. Czyczyło-Mysza, I.M.; Cyganek, K.; Dziurka, K.; Quarrie, S.; Skrzypek, E.; Marcińska, I.; Myśków, B.; Dziurka, M.; Warchoń, M.; Kapłoniak, K.; et al. Genetic Parameters and QTLs for Total Phenolic Content and Yield of Wheat Mapping Population of CSDH Lines under Drought Stress. *Int. J. Mol. Sci.* **2019**, *20*, 6064. [[CrossRef](#)]

43. Bocianowski, J.; Krajewski, P. Comparison of the genetic additive effect estimators based on phenotypic observations and on molecular marker data. *Euphytica* **2009**, *165*, 113–122. [[CrossRef](#)]
44. Jansen, R.C.; Stam, P. High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* **1994**, *136*, 1447–1455. [[CrossRef](#)] [[PubMed](#)]
45. Province, M.A. 30 Sequential methods of analysis for genome scan. *Adv. Genet.* **2001**, *42*, 499–514. [[CrossRef](#)] [[PubMed](#)]
46. Searle, S.R. *Matrix Models for Unbalanced Data*; John Wiley & Sons, Inc.: New York, NY, USA, 1982; pp. 1–154.
47. Quarrie, S.A.; Steed, A.; Calestani, C.; Semikhodskii, A.; Lebreton, C.; Chinoy, C.; Steele, N.; Pljevljakusic, D.; Waterman, E.; Weyen, J.; et al. A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring × SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theor. Appl. Genet.* **2005**, *110*, 865–880. [[CrossRef](#)] [[PubMed](#)]
48. Fu, Y.B.; Ritland, K. Marker-Based Inferences About Epistasis for Genes Influencing Inbreeding Depression. *Genetics* **1996**, *144*, 339–348. [[CrossRef](#)]
49. Nap, J.P.; Canner, A.J.; Mlynarova, L.; Stiekema, W.J.; Jansen, R.C. Dissection of a Synthesized Quantitative Trait to Characterize Transgene Interactions. *Genetics* **1997**, *147*, 315–320. [[CrossRef](#)]
50. Routman, E.J.; Cheverud, J.M. Gene effects on a quantitative trait: Two-locus epistatic effects measured at microsatellite markers and at estimated QTL. *Evolution* **1997**, *51*, 1654–1662. [[CrossRef](#)]
51. Bocianowski, J.; Nowosad, K. Mixed linear model approaches in mapping QTLs with epistatic effects by a simulation study. *Euphytica* **2015**, *202*, 459–467. [[CrossRef](#)]
52. Slim, L.; Chatelain, C.; Azencott, C.A.; Vert, J.P. Novel methods for epistasis detection in genome-wide association studies. *PLoS ONE* **2020**, *15*, e0242927. [[CrossRef](#)]
53. Rieger, R.; Michaelis, A.; Green, M.M. *A Glossary of Genetics and cytogenetics: Classical and Molecular*; Springer: New York, NY, USA, 1968; ISBN 9780387076683.

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.

**4. A Comparison of Methods to Estimate Additive–by–
Additive–by–Additive of QTL×QTL×QTL Interaction
Effects by Monte Carlo Simulation Studies**

Cyplik, A., Bocianowski, J.

International Journal of Molecular Sciences, 2023, 24(12), 10043

DOI:10.3390/ijms241210043

Impact Factor 2022: 5,6; Punktacja wg MEiN: 140



Article

A Comparison of Methods to Estimate Additive-by-Additive-by-Additive of QTL \times QTL \times QTL Interaction Effects by Monte Carlo Simulation Studies

Adrian Cyplik and Jan Bocianowski *

Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Wojska Polskiego 28, 60-637 Poznań, Poland; adrian.cyplik@up.poznan.pl

* Correspondence: jan.bocianowski@up.poznan.pl; Tel.: +48-61-8487143

Abstract: The goal of the breeding process is to obtain new genotypes with traits improved over the parental forms. Parameters related to the additive effect of genes as well as their interactions (such as epistasis of gene-by-gene interaction effect and additive-by-additive-by-additive of gene-by-gene-by-gene interaction effect) can influence decisions on the suitability of breeding material for this purpose. Understanding the genetic architecture of complex traits is a major challenge in the post-genomic era, especially for quantitative trait locus (QTL) effects, QTL-by-QTL interactions and QTL-by-QTL-by-QTL interactions. With regards to the comparing methods for estimating additive-by-additive-by-additive of QTL \times QTL \times QTL interaction effects by Monte Carlo simulation studies, there are no publications in the open literature. The parameter combinations assumed in the presented simulation studies represented 84 different experimental situations. The use of weighted regression may be the preferred method for estimating additive-by-additive-by-additive of QTL-QTL-QTL triples interaction effects, as it provides results closer to the true values of total additive-by-additive-by-additive interaction effects than using unweighted regression. This is also indicated by the obtained values of the determination coefficients of the proposed models.

Keywords: biometrical genetics; homozygous lines; Monte Carlo simulation study; phenotypic observations; quantitative trait loci; regression analysis; three-way interaction; weighted regression



Citation: Cyplik, A.; Bocianowski, J. A Comparison of Methods to Estimate Additive-by-Additive-by-Additive of QTL \times QTL \times QTL Interaction Effects by Monte Carlo Simulation Studies. *Int. J. Mol. Sci.* **2023**, *24*, 10043. <https://doi.org/10.3390/ijms241210043>

Academic Editor: Maoteng Li

Received: 25 April 2023

Revised: 5 June 2023

Accepted: 11 June 2023

Published: 12 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The goal of breeding is to obtain the most economically advantageous varieties. A very important aspect is to shorten the breeding process without losing the quality of the obtained genotypes. Knowledge of the genetic structure and genes actions can contribute to genetic advances resulting in a shorter breeding process. In traditional quantitative genetics, organisms are studied on the basis of only phenotypic observations, and conclusions about the genotype are drawn based on the ways in which quantitative traits are inherited, among other things. This is usually realized by describing the action of genes using genetic parameters, mean and variance functions of phenotypes [1]. Among these parameters is the additive effect of genes and interactions between them [2]. Additive effects are perpetuated in a population as it increases its homozygosity in subsequent generations. Thus, the significant additive effect of genes in the population means that selection starting in early generations offers hope of obtaining transgenic homozygous lines [3].

The introduction of different types of markers into genetic and breeding research has revolutionized these scientific disciplines. Markers, among other molecular markers, are used in selection and for genetic transformation [4]. Molecular markers show genetic differences between individuals with opposite traits in the form of nucleotide sequences and have been developed and used to speed up the breeding process in many crops [5]. Identifying markers significantly associated with plant traits can help develop new varieties with desirable traits. It is very important to localize the genes or groups of genes controlling

a trait (that is, to give the linkage distances of the genes studied to the markers) and determine their effects [6]. The accuracy and correctness of the conclusions depend largely on the precision of determining the location of the gene on the chromosome and the estimation of parameters associated with gene effects [7].

Most important plant traits are quantitative in nature and are influenced by multiple quantitative trait locus (QTLs). The effects of these QTLs on the expression of quantitative trait has been considered in various species such as maize [8], barley [9], rice [10,11], oilseed rape [12–14], wheat [15] and sugar beet [16]. Many loci showing cumulative small effects are often rejected in standard QTL analyses [17,18]. However, by interacting with other loci, they can significantly affect the observed quantitative trait. The effect of epistasis (double interaction) on the observed trait has been considered and reported for, among others: wheat [19], *Triticale* [20], tomato [21], soybean [22], rice [23], pea [24], canola [25], maize [26] and barley [27]. However, higher-order interactions are omitted, although it is difficult to imagine that a quantitative trait is not determined by the interaction of more than two genes. Based on the knowledge of physiological genetics and biochemistry, interactions between gene products are ubiquitous [28].

A previous report [29] described the results obtained for analytical and numerical comparisons of two methods for estimating the parameter associated with the additive-by-additive-by-additive (*aaa*) interaction effect, namely, the phenotypic method, based on extreme groups of homozygous lines, and the genotypic method, based on marker observations. One of the conclusions of these studies was that the estimate of the total additive-by-additive-by-additive interaction effect based on the marker data was in most cases smaller than that based on the phenotype. The explanation for this phenomenon may be simple. Phenotypic data can be used to estimate only the total three-way interaction effect of all hypothetical gene triplets determining a trait. Using marker data, which can be more or less precisely mapped into the genome, individual gene-by-gene-by-gene interaction effects can be estimated. For practical reasons, the number of QTL-QTL-QTL interactions is deliberately low. The sum of the effects obtained is lower than the phenotypic estimate, and this difference may be reinforced by the lack of markers in the regions where the genes are located.

In addition to the above explanation, other possible sources of differences between the calculated estimates should be considered. The above results concern QTL-QTL-QTL interaction effects obtained by the simplest possible methods, i.e., using multiple linear regression on the marker data. Cyplik et al. [30] showed that modifying this regression by using empirical weights provides a better agreement between phenotypic and genotypic estimates.

The purpose of this paper was to compare two methods for estimating the parameter associated with additive-by-additive-by-additive interaction effects: the phenotypic method, traditionally used in quantitative genetics, and the genotypic method, which is based on marker observation and is now routinely used in many species. The comparison was made using a Monte Carlo simulation study. Consideration was also given to modifying the estimation of the additive-by-additive-by-additive of QTL-QTL-QTL interaction effects by using weighted multiple linear regression (in two variants) to determine a bridge between the two compared methodologies.

In the approach described here, the estimation of the parameter associated with the additive-by-additive-by-additive interaction effect was based on extreme groups of homozygous lines and on data for genotypic markers. Weighted multiple linear regression was used to estimate the QTL-QTL-QTL interaction effects, using standard weighted regression with weights based on estimated variances for lines in two variants: using weighted regression to estimate only QTL-QTL-QTL triples interaction effects (unweighted regression for QTL selection and QTL-QTL epistasis) and using weighted regression on all three stages—for QTL selection, epistatic pairs and QTL-QTL-QTL triples. The effectiveness of the proposed method was tested using Monte Carlo simulations for different experimental variants, such as the number and location of chromosomes of QTL-QTL-QTL triples determining the quantitative trait (Table 1), the presence or absence of QTL-QTL

epistasis (Figure 1). Based on the knowledge of the literature, this is the first report of a comparison of methods for estimating a parameter related to the QTL–QTL–QTL triples interaction effect based on simulation studies.

Table 1. The number of QTL–QTL–QTL triples assumed, their designations and locations on the chromosomes, and the coding of individual variants.

Variant Code	The Number of Triples	QTL–QTL–QTL Triples	The Position of Triples on Chromosomes
C01	1	Q6Q7Q8	all on one chromosome
C02	1	Q1Q2Q3	on two different chromosomes
C03	1	Q1Q2Q4	on three different chromosomes
C04	2	Q6Q7Q8, Q9Q10Q11	all on one chromosome
C05	2	Q1Q2Q3, Q4Q5Q6	on two different chromosomes
C06	2	Q1Q2Q4, Q3Q5Q7	on three different chromosomes
C07	3	Q6Q7Q8, Q9Q10Q11, Q12Q13Q14	all on one chromosome
C08	3	Q1Q2Q3, Q4Q5Q6, Q7Q10Q11	on two different chromosomes
C09	3	Q1Q2Q4, Q3Q5Q7, Q4Q6Q9	on three different chromosomes
C10	5	Q6Q7Q8, Q9Q10Q11, Q12Q13Q14, Q1Q2Q3, Q1Q4Q5	three triples on one chromosome and two triples on two chromosomes
C11	5	Q1Q2Q3, Q2Q4Q5, Q4Q6Q7, Q1Q9Q10, Q1Q12Q14	on two different chromosomes
C12	5	Q1Q2Q4, Q3Q5Q6, Q1Q9Q12, Q2Q10Q14, Q5Q10Q12	on three different chromosomes

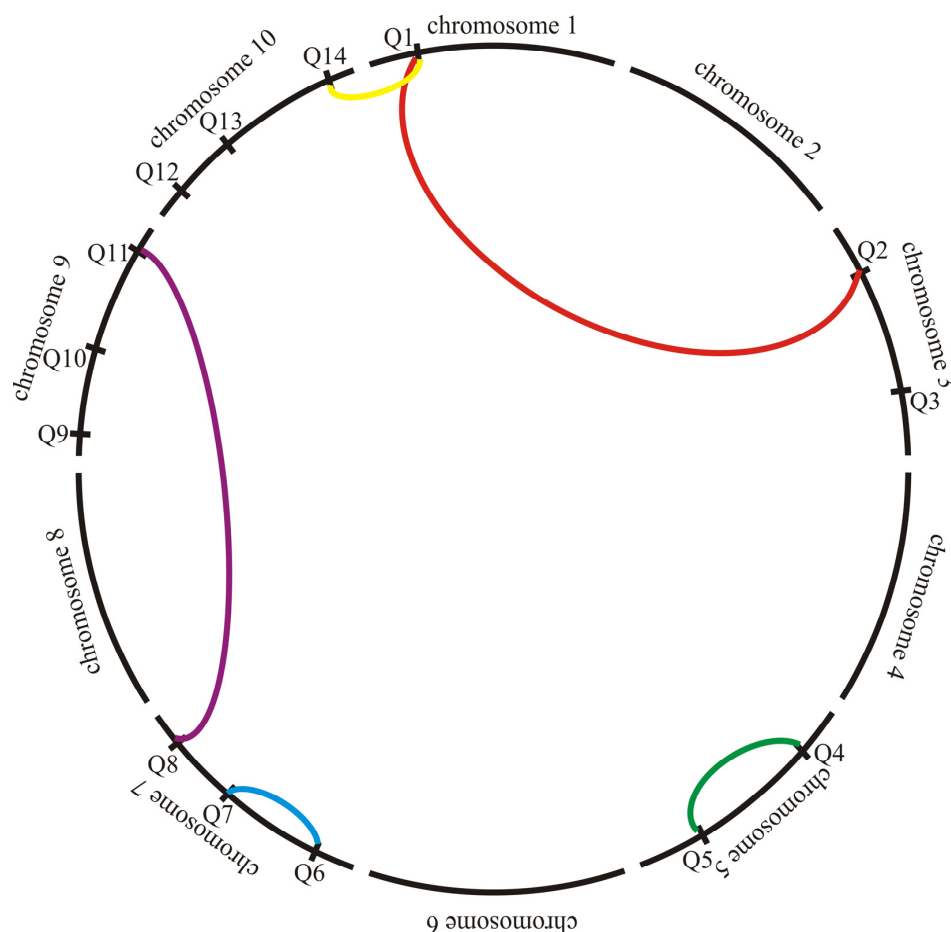


Figure 1. Localization on chromosomes of assumed 14 QTLs and five epistatic pairs.

2. Results

Difficult experimental situations were examined in this large-scale simulation. Tables 2–9 show the results of the simulation performed to compare the estimates of additive-by-additive-by-additive interaction effects obtained by the four methods: the phenotypic methods and three genotypic methods (one unweighted and two weighted). Tables 2–5 show the

results obtained with an assumed error variance of 5, while Tables 6–9 show the results when the error variance was 10. Tables 2, 3, 6 and 7 contain results assuming that the total effect of epistasis was equal to 0, and Tables 4, 5, 8 and 9, assume that the effect of epistasis was significantly different from zero. The results of the simulation studies assuming that all the effects of the *aaa* triples were equal are shown in Tables 2, 4, 6 and 8; while Tables 3, 5, 7 and 9 contain the results obtained assuming different effects of the *aaa* triples.

Table 2. Phenotypic and genotypic estimates of the total additive–additive–additive effect, mean squared errors and mean of coefficients of determination R_j^2 obtained in the simulation study for: the error variance was equal to 5, no epistasis, equal effects for all triples.

Variant Code	Number of QTL–QTL–QTL Interaction Effects	Number of Chromosomes with QTL Triples	Estimate				Mean Squared Error For				R^2
			\widehat{aaa}_p	\widehat{aaa}_{gu}	\widehat{aaa}_{gw1}	\widehat{aaa}_{gw2}	\widehat{aaa}_p	\widehat{aaa}_{gu}	\widehat{aaa}_{gw1}	\widehat{aaa}_{gw2}	
C01	1	1	17.99	17.01	17.63	17.22	0.23	0.31	0.25	0.42	0.89
C02	1	2	18.01	17.09	17.71	17.32	0.27	0.33	0.28	0.41	0.87
C03	1	3	18.32	17.12	18.02	17.59	0.28	0.40	0.33	0.44	0.92
C04	2	1	17.55	16.82	17.30	16.91	0.33	0.42	0.38	0.43	0.84
C05	2	2	17.80	17.09	17.62	17.39	0.37	0.44	0.40	0.50	0.88
C06	2	3	18.27	17.62	17.93	17.52	0.37	0.47	0.42	0.55	0.95
C07	3	1	15.97	15.48	15.82	15.55	0.45	0.53	0.49	0.66	0.94
C08	3	2	16.43	15.66	16.21	16.09	0.47	0.59	0.51	0.68	0.87
C09	3	3	16.91	16.01	16.72	16.49	0.61	0.72	0.68	0.73	0.92
C10	5	1, 2	15.57	14.42	15.02	14.77	0.55	0.71	0.60	0.82	0.90
C11	5	2	16.25	14.93	15.99	15.21	0.56	0.78	0.69	0.88	0.87
C12	5	3	16.66	15.82	16.49	16.31	0.62	0.82	0.77	1.01	0.91

Table 3. Phenotypic and genotypic estimates of the total additive–additive–additive effect, mean squared errors and mean of coefficients of determination R_j^2 obtained in the simulation study for: the error variance was equal to 5, no epistasis, different effects for triples.

Variant Code	Number of QTL–QTL–QTL Interaction Effects	Number of Chromosomes with QTL Triples	Estimate				Mean Squared Error For				R^2
			\widehat{aaa}_p	\widehat{aaa}_{gu}	\widehat{aaa}_{gw1}	\widehat{aaa}_{gw2}	\widehat{aaa}_p	\widehat{aaa}_{gu}	\widehat{aaa}_{gw1}	\widehat{aaa}_{gw2}	
C04	2	1	17.48	16.97	17.42	16.89	0.30	0.44	0.37	0.44	0.84
C05	2	2	17.89	17.32	17.59	17.42	0.41	0.47	0.48	0.52	0.95
C06	2	3	18.33	17.58	17.99	17.99	0.42	0.49	0.45	0.59	0.83
C07	3	1	15.92	15.47	15.89	15.73	0.47	0.51	0.48	0.59	0.90
C08	3	2	16.58	15.60	16.43	15.98	0.51	0.60	0.53	0.66	0.87
C09	3	3	17.02	16.21	16.69	16.55	0.59	0.77	0.63	0.89	0.88
C10	5	1, 2	15.52	14.62	15.17	14.80	0.56	0.73	0.69	0.79	0.88
C11	5	2	16.19	15.23	15.89	15.70	0.59	0.74	0.63	0.90	0.92
C12	5	3	16.70	16.01	16.52	16.25	0.73	1.01	0.88	1.17	0.94

Table 4. Phenotypic and genotypic estimates of the total additive–additive–additive effect, mean squared errors and mean of coefficients of determination R_j^2 obtained in the simulation study for: the error variance was equal to 5, presence of epistasis, equal effects for all triples.

Variant Code	Number of QTL–QTL–QTL Interaction Effects	Number of Chromosomes with QTL Triples	Estimate				Mean Squared Error For				R^2
			\widehat{aaa}_p	\widehat{aaa}_{gu}	\widehat{aaa}_{gw1}	\widehat{aaa}_{gw2}	\widehat{aaa}_p	\widehat{aaa}_{gu}	\widehat{aaa}_{gw1}	\widehat{aaa}_{gw2}	
C01	1	1	17.72	16.53	17.39	16.71	0.19	0.28	0.21	0.39	0.89
C02	1	2	17.84	16.72	17.53	16.82	0.25	0.30	0.27	0.40	0.94
C03	1	3	18.01	16.89	17.72	17.19	0.25	0.37	0.32	0.42	0.90
C04	2	1	17.40	16.66	17.07	16.72	0.30	0.39	0.33	0.40	0.84
C05	2	2	17.71	16.79	17.55	17.01	0.36	0.41	0.40	0.49	0.90
C06	2	3	17.99	16.50	17.62	17.17	0.34	0.40	0.37	0.54	0.86
C07	3	1	15.79	15.03	15.39	15.19	0.40	0.49	0.42	0.58	0.89
C08	3	2	16.29	15.38	16.03	15.57	0.44	0.53	0.49	0.66	0.92
C09	3	3	16.77	15.55	16.59	16.02	0.58	0.69	0.66	0.70	0.89
C10	5	1, 2	15.42	14.29	14.77	14.99	0.50	0.63	0.52	0.77	0.94
C11	5	2	16.09	14.70	15.81	15.33	0.54	0.75	0.62	0.81	0.95
C12	5	3	16.40	14.92	16.22	15.76	0.59	0.81	0.73	0.99	0.88

Table 5. Phenotypic and genotypic estimates of the total additive–additive–additive effect, mean squared errors and mean of coefficients of determination R_j^2 obtained in the simulation study for: the error variance was equal to 5, presence of epistasis, different effects for triples.

Variant Code	Number of QTL–QTL–QTL Interaction Effects	Number of Chromosomes with QTL Triples	Estimate				Mean Squared Error For				R^2
			\widehat{aaa}_p	\widehat{aaa}_{gu}	\widehat{aaa}_{gw1}	\widehat{aaa}_{gw2}	\widehat{aaa}_p	\widehat{aaa}_{gu}	\widehat{aaa}_{gw1}	\widehat{aaa}_{gw2}	
C04	2	1	17.19	16.53	16.98	16.77	0.33	0.40	0.37	0.48	0.95
C05	2	2	17.37	16.80	17.12	16.97	0.36	0.44	0.41	0.53	0.83
C06	2	3	17.42	16.99	17.28	17.17	0.36	0.45	0.39	0.61	0.85
C07	3	1	15.77	15.22	15.59	15.36	0.39	0.52	0.42	0.70	0.91
C08	3	2	15.90	15.39	16.22	15.78	0.42	0.59	0.51	0.68	0.95
C09	3	3	16.82	15.66	16.70	16.32	0.55	0.68	0.60	0.73	0.87
C10	5	1, 2	14.99	14.38	14.73	14.55	0.51	0.63	0.53	0.79	0.89
C11	5	2	15.07	14.42	14.80	14.55	0.54	0.77	0.66	0.93	0.90
C12	5	3	15.18	14.66	14.98	14.81	0.58	0.83	0.69	0.97	0.88

Table 6. Phenotypic and genotypic estimates of the total additive–additive–additive effect, mean squared errors and mean of coefficients of determination R_j^2 obtained in the simulation study for: the error variance was equal to 10, no epistasis, equal effects for all triples.

Variant Code	Number of QTL–QTL–QTL Interaction Effects	Number of Chromosomes with QTL Triples	Estimate				Mean Squared Error For				R^2
			\widehat{aaa}_p	\widehat{aaa}_{gu}	\widehat{aaa}_{gw1}	\widehat{aaa}_{gw2}	\widehat{aaa}_p	\widehat{aaa}_{gu}	\widehat{aaa}_{gw1}	\widehat{aaa}_{gw2}	
C01	1	1	18.03	17.09	17.77	17.39	0.44	0.57	0.49	0.73	0.91
C02	1	2	18.11	17.20	17.79	17.45	0.49	0.60	0.54	0.80	0.87
C03	1	3	18.39	17.22	18.19	17.66	0.54	0.77	0.63	0.91	0.95
C04	2	1	17.61	16.89	17.41	17.01	0.64	0.81	0.75	0.94	0.86
C05	2	2	17.88	17.31	17.77	17.52	0.66	0.79	0.69	0.94	0.87
C06	2	3	18.44	17.63	18.02	17.77	0.68	0.91	0.80	0.99	0.88
C07	3	1	16.02	15.54	15.91	15.62	0.77	0.91	0.84	1.02	0.90
C08	3	2	16.51	15.72	16.43	16.17	0.82	0.93	0.89	1.05	0.92
C09	3	3	16.98	16.29	16.55	16.40	0.83	0.93	0.88	1.11	0.93
C10	5	1, 2	14.99	14.29	14.70	14.42	0.77	0.89	0.81	0.99	0.87
C11	5	2	15.17	14.40	15.01	14.62	0.83	0.99	0.91	1.12	0.84
C12	5	3	15.52	14.80	15.27	14.99	0.90	1.07	0.97	1.23	0.91

Table 7. Phenotypic and genotypic estimates of the total additive–additive–additive effect, mean squared errors and mean of coefficients of determination R_j^2 obtained in the simulation study for: the error variance was equal to 10, no epistasis, different effects for triples.

Variant Code	Number of QTL–QTL–QTL Interaction Effects	Number of Chromosomes with QTL Triples	Estimate				Mean Squared Error For				R^2
			\widehat{aaa}_p	\widehat{aaa}_{gu}	\widehat{aaa}_{gw1}	\widehat{aaa}_{gw2}	\widehat{aaa}_p	\widehat{aaa}_{gu}	\widehat{aaa}_{gw1}	\widehat{aaa}_{gw2}	
C04	2	1	17.51	17.00	17.49	17.23	0.55	0.77	0.61	0.80	0.90
C05	2	2	17.93	17.49	17.60	17.55	0.72	0.91	0.80	0.97	0.87
C06	2	3	18.50	17.69	18.23	18.01	0.79	0.99	0.88	1.07	0.94
C07	3	1	16.09	15.52	15.91	15.79	0.82	0.93	0.87	1.11	0.89
C08	3	2	16.22	15.63	16.09	15.99	0.91	1.21	0.99	1.33	0.86
C09	3	3	16.90	15.99	16.73	16.28	0.92	1.27	1.02	1.44	0.85
C10	5	1, 2	15.29	14.47	15.03	14.71	0.88	0.99	0.92	1.03	0.90
C11	5	2	15.31	14.59	15.28	14.92	0.91	1.08	0.97	1.20	0.87
C12	5	3	15.57	14.66	15.02	15.23	0.97	1.22	1.12	1.40	0.88

Table 8. Phenotypic and genotypic estimates of the total additive–additive–additive effect, mean squared errors and mean of coefficients of determination R_j^2 obtained in the simulation study for: the error variance was equal to 10, presence of epistasis, equal effects for all triples.

Variant Code	Number of QTL–QTL–QTL Interaction Effects	Number of Chromosomes with QTL Triples	Estimate				Mean Squared Error For				R^2
			\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	
C01	1	1	17.77	16.69	17.42	16.88	0.30	0.41	0.37	0.73	0.89
C02	1	2	17.90	16.79	17.66	16.89	0.42	0.55	0.49	0.79	0.94
C03	1	3	18.19	17.17	17.89	17.23	0.44	0.71	0.59	0.79	0.95
C04	2	1	17.42	16.70	17.11	16.99	0.51	0.77	0.57	0.82	0.88
C05	2	2	17.69	16.88	17.58	17.20	0.52	0.71	0.60	0.80	0.97
C06	2	3	18.08	16.91	17.71	17.44	0.58	0.77	0.66	0.84	0.86
C07	3	1	15.80	15.12	15.53	15.27	0.69	0.80	0.77	0.91	0.90
C08	3	2	16.32	15.44	16.42	16.33	0.73	0.92	0.81	0.99	0.92
C09	3	3	16.83	15.61	16.62	16.51	0.74	0.98	0.89	1.02	0.87
C10	5	1, 2	15.21	14.08	14.63	14.49	0.88	1.00	0.93	1.17	0.86
C11	5	2	15.44	14.32	14.77	14.57	0.92	1.19	1.04	1.31	0.91
C12	5	3	15.70	14.79	15.22	15.01	0.98	1.33	1.07	1.40	0.95

Table 9. Phenotypic and genotypic estimates of the total additive–additive–additive effect, mean squared errors and mean of coefficients of determination R_j^2 obtained in the simulation study for: the error variance was equal to 10, presence of epistasis, different effects for triples.

Variant Code	Number of QTL–QTL–QTL Interaction Effects	Number of Chromosomes with QTL Triples	Estimate				Mean Squared Error For				R^2
			\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	\overline{aaa}_p	\overline{aaa}_{gu}	\overline{aaa}_{gw1}	\overline{aaa}_{gw2}	
C04	2	1	17.33	16.55	17.02	16.81	0.49	0.73	0.52	0.80	0.94
C05	2	2	17.49	17.01	17.19	16.99	0.55	0.80	0.58	0.92	0.91
C06	2	3	17.66	17.09	17.31	17.24	0.61	0.92	0.77	0.99	0.87
C07	3	1	15.80	15.49	15.76	15.50	0.66	0.88	0.79	0.92	0.85
C08	3	2	15.93	15.52	15.91	15.80	0.70	0.93	0.82	1.02	0.88
C09	3	3	16.99	15.72	16.82	16.55	0.82	1.09	0.97	1.25	0.90
C10	5	1, 2	14.73	14.21	14.58	14.39	0.79	0.98	0.85	1.07	0.95
C11	5	2	14.82	14.39	14.68	14.40	0.85	1.12	0.93	1.30	0.91
C12	5	3	15.01	14.52	14.95	14.77	0.98	1.21	1.07	1.44	0.92

2.1. Phenotypic Estimates

The results of the simulation showed that the phenotypic estimate was always greater than 15, the true–assumed value, except in four cases for five QTL–QTL–QTL triples: (1) the number of chromosomes containing QTLs was 1, 2, the error variance was equal to 5, the effect of epistasis was different from zero and the effects of *aaa* triples were different (Table 5), (2) the number of chromosomes containing QTLs was 1, 2, the error variance was equal to 10, the effect of epistasis was equal to zero and the effects of all *aaa* triples were equal (Table 6), (3) and (4) the number of chromosomes containing QTLs was 1, 2 (variant code—C10), and 2 (variant code—C11), respectively, the error variance was equal to 10, the effect of epistasis was different from zero, while the effects of *aaa* triples were different (Table 9). The largest values were obtained for one QTL–QTL–QTL triple, also the smallest values for five QTL–QTL–QTL triples (taking into account the analogous distribution of triples in the linkage group). The largest values were obtained for the situation when the QTLs constituting the triple were located in different chromosomes, while the smallest values were for the situation when all three QTLs constituting the triple were located in one linkage group. The phenotypic estimates were larger when assuming no epistasis than when assuming the effect of epistasis on the values of the observed quantitative trait. When considering the effects of triples, equal or different, similar values of phenotypic estimates were obtained. The \overline{aaa}_p values were larger when the error variance was equal to 10 than for an error variance of 5, except when five triples were assumed.

2.2. Genotypic Estimates

The genotypic estimates of *aaa* effects were greater than true-assumed value 15 for 1, 2 and 3 QTL–QTL–QTL triples. For situations where five QTL–QTL–QTL triples were assumed in most cases (58.33%), genotypic estimates were less than 15. With an assumed error variance of 5, in 19 cases (51.78%) genotypic estimates were less than the true-assumed value of *aaa* equal to 15 (Tables 6–9). The largest values of genotypic estimates (unweighted and weighted) were obtained for one QTL–QTL–QTL triple, also the smallest values for five QTL–QTL–QTL triples. The largest values were obtained for situations when QTLs constituting triples were located in different chromosomes, except for the case when five QTL–QTL–QTL triples with different effects were assumed in the absence of epistasis and an error variance of 10 (Table 8). In contrast, the smallest values for cases when all three QTLs constituting the triple were placed in a single coupling group. In most cases, genotypic estimates were larger for different effects than for equal effects of triples. The genotypic estimates were larger with an assumed larger (10) error variance than with a smaller (5) variance in situations where the quantitative trait was determined by 1, 2 or 3 QTL–QTL–QTL triples. In contrast, when five QTL–QTL–QTL triples were assumed, genotypic estimates were larger for an error variance of 5 than for an error variance of 10.

2.3. Phenotypic vs. Genotypic Estimates

The differences between phenotypic estimates and the genotypic estimates for unweighted regression were always positive and were generally smaller when different effects of QTL–QTL–QTL triples were assumed. The relationships $\widehat{aaa}_p > \widehat{aaa}_{gw1} > \widehat{aaa}_{gw2} > \widehat{aaa}_{gu}$ were generally observed, except in three cases: (1) when two QTL–QTL–QTL triples were located in three chromosomes for variance error equal to 5, lack of epistasis and different effects of triples ($\widehat{aaa}_{gw1} = \widehat{aaa}_{gw2} = 17.99$, Table 3), (2) when three QTL–QTL–QTL triples were located in two linkage groups, variance error equal to 5, lack of epistasis and different effects of triples ($\widehat{aaa}_{gw1} > \widehat{aaa}_p$, Table 5), and (3) when three QTL–QTL–QTL triples with equal effects were located in two linkage groups and the effect of epistasis was different from zero: in which case $\widehat{aaa}_{gw1} > \widehat{aaa}_{gw2} > \widehat{aaa}_p$ (Table 8). The weighted estimates were always closer to the phenotypic estimates than the unweighted estimates. The weighted method way (i) showed the best (closest to the true assumed value of 15) ratings of *aaa* triples.

2.4. Mean Squared Errors

In general, a decrease in the estimates was accompanied by an increase in their mean squared error (Tables 2–9). The relationships of mean squared error for $\widehat{aaa}_p < \widehat{aaa}_{gw1} < \widehat{aaa}_{gu} < \widehat{aaa}_{gw2}$ were generally observed, except when two QTL–QTL–QTL triples with different effects were assumed to be present in two linkage groups in case of lack of epistatic effect and variance error was equal to 5, in which case mean squared error for $\widehat{aaa}_{gw1} > \widehat{aaa}_{gw2}$ (Table 3). The mean squared error was larger with an assumed error variance of 10 than with a smaller assumed error variance. The effect values of each triple (equal or different) did not affect the mean squared error values. On the other hand, assuming the presence or absence of the epistasis effect affected the mean squared error values for \widehat{aaa} : large mean squared error values were obtained when assuming no epistasis. The smallest mean squared errors were obtained when the trait was determined by one QTL–QTL–QTL triple, while the largest was obtained with five QTL–QTL–QTL triples. In general, the relationships of mean squared error values for assumed $1QQQ < 2QQQs < 3QQQs < 5QQQs$ were observed. The smallest mean squared error was obtained for QTL–QTL–QTL triples located in one linkage group, while the largest was obtained when QTL–QTL–QTL triples were located across three chromosomes.

2.5. Coefficients of Determination

Variance explained by the QTL–QTL–QTL triple interactions ranged from 83% (Tables 3 and 5) to 97% (Table 8). Coefficients of determination for one QTL–QTL–QTL triple assumed ranged from 0.87 to 0.95, with a mean of 0.91. Coefficients of determination for two QTL–QTL–QTL triples assumed ranged from 0.83 to 0.97, with a mean of 0.89. Coefficients of determination for three and five QTL–QTL–QTL triples assumed were similar and ranged from 0.85 and 0.84 to 0.95 and 0.95, with mean of 0.90 and 0.90, for three and five QTL–QTL–QTL triples assumed, respectively.

3. Discussion

The goal of the breeding process is to obtain new genotypes with traits improved over the parental forms [31]. Parameters related to the additive effect of genes as well as their interactions (such as additive–by–additive epistasis of gene–by–gene interaction effect and additive–by–additive–by–additive of gene–by–gene–by–gene interaction effect) can influence decisions on the suitability of breeding material for this purpose [32–38]. Understanding the genetic architecture of complex traits is a major challenge in the post-genomic era, especially for QTL effects, QTL–by–QTL interactions [39] and QTL–by–QTL–by–QTL interactions.

Monte Carlo simulations are a research tool very widely used to solve various problems by obtaining approximate results. Estimation of epistasis effects using simulation studies has been presented, for example, by Bocianowski [40], Ahsan et al. [41], de los Campos et al. [42], Wang et al. [43] and Sharma et al. [44]. With regards to the comparing methods for estimating additive–by–additive–by–additive of QTL \times QTL \times QTL interaction effects by Monte Carlo simulation studies, there are no publications in the open literature. The paper describes the first simulation study comparing methods for estimating the total additive–by–additive–by–additive effect of gene–by–gene–by–gene interactions. In addition, weighted multiple linear regression was used to estimate QTL–by–QTL–by–QTL interaction effects, with two variants proposed.

The Monte Carlo simulation studies conducted to compare estimation methods could not, of course, take into account all possible experimental situations. It is essential to note that these simulations may not fully capture the complexities and nuances of real-world genetic interactions. The findings should be interpreted with caution until further empirical validation is conducted. However, the parameter combinations used in the analyses presented here correspond to the cases most often encountered in actual QTL studies [45–62]. In simulation studies, the number of lines analyzed is most often assumed to be as high as 500 or 2000 [45,46]. Viana et al. [46] analyzed 400 plants, genotyped for 975 SNPs distributed in 10 chromosomes of 100 cm. One thousand SNP markers were considered in their study by Crawford et al. [47]. The ten chromosomes include *Anthoxanthum odoratum*, *Brassica campestris*, *Sorghum sudanense*, *Zea mays*. Practical experience shows that the number of genes located on individual chromosomes varies greatly [48–51]. The number of QTLs assumed in the present study was 14. The same number of QTLs was presented in their studies by Wu [52] and Balestre et al. [53]. The assumed number of five epistatic pairs was a result observed in many practical experiments [54–57]. Simulation studies based on 10,000 generations were also presented by Goutelle et al. [58], Avery et al. [59], Johnson et al. [60], Sorojsrisom et al. [61], and Wang et al. [62], among others. The parameter combinations assumed in the presented simulation studies represented 84 different experimental situations. The results obtained from the Monte Carlo simulation studies indicated the stability of the properties of the presented methods of estimating additive–by–additive–by–additive of gene–by–gene–by–gene interaction effect over different types of genetic material. The small effect of error variance on the estimation of additive–by–additive–by–additive of gene–by–gene–by–gene interaction effects by all four methods, and on the conclusions regarding the comparison of the proposed estimation methods, indicates good prospects for the applicability of our conclusions for different plant species.

However, it was found that the number of assumed QTL–QTL–QTL triplets, their assumed positions, in one, two or three chromosomes, and their effects (unequal or equal) affected the estimates and their comparison. Phenotypic estimation decreased toward the true value as the number of presumed QTL–QTL–QTL triplets increased (Figure 2). For the largest number of QTL–QTL–QTL triplets, it may have been less than the true value, as genotypic combinations generating extreme lines were not adequately represented in the simulated sample. The phenotypic estimate also tended to the true value when the number of chromosomes with QTL–QTL–QTL triplets decreased, i.e., when there was an accumulation of coupled individual effects. Similarly, phenotypic estimation tended to have a true value when the assumed epistasis effect was different from zero. The phenotypic evaluation was overestimated for a larger assumed error variance. The genotypic estimate decreased its value in a very similar way to the phenotypic estimate. The difference between phenotypic and genotypic estimation was smallest when five QTL–QTL–QTL triplets were assumed, that is when a model close to truly polygenic was used.

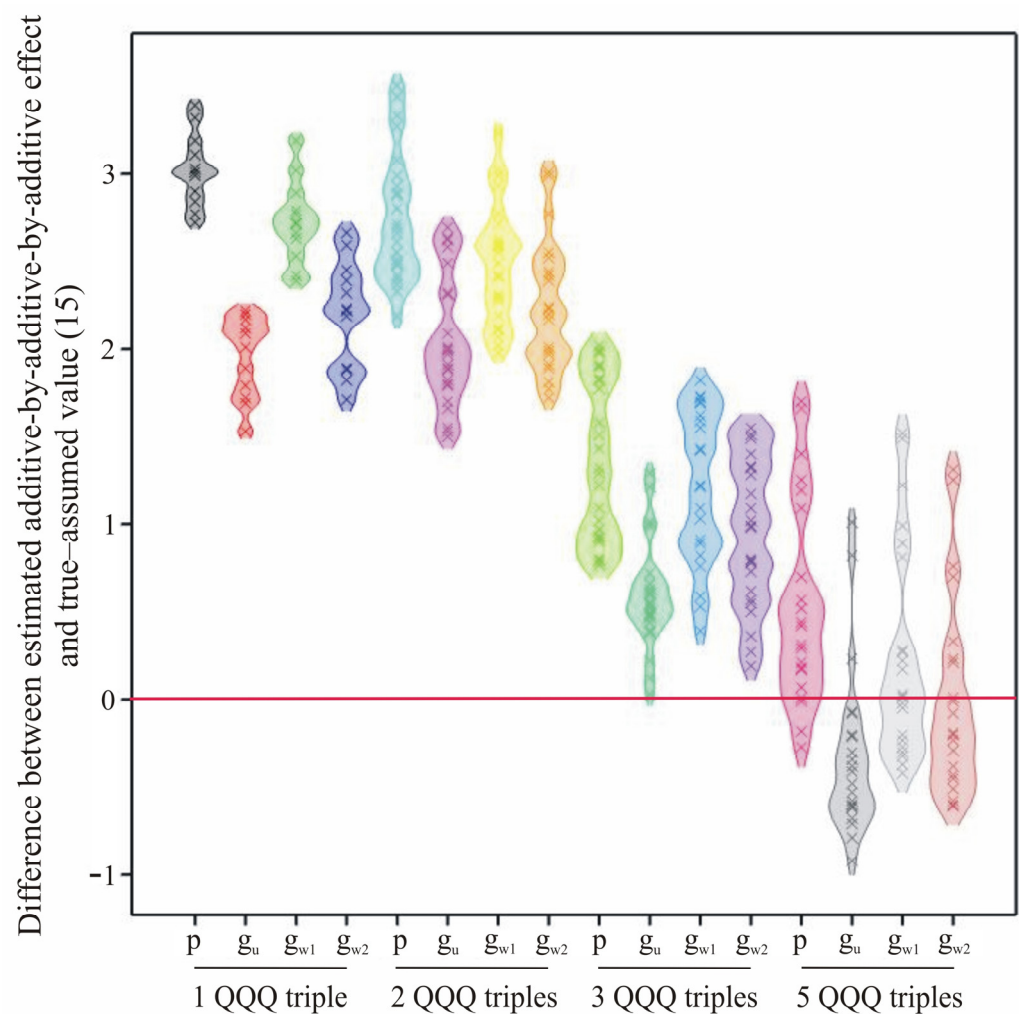


Figure 2. Differences between estimated additive–by–additive–by–additive effect and true–assumed value (15) for particular methods of estimation and different number of QTL–QTL–QTL triplets. p—phenotypic method, g_u —unweighted regression, g_{w1} —weighted regression: weighted triple interaction effects only, g_{w2} —weighted regression: weighted all effects: QTL, epistatic and triple interaction.

The difference between phenotypic and genotypic estimates of the total additive–by–additive–by–additive effect, observed in both numerical [29] and simulation comparisons, is partly due to the fact that genotypic methods do not find all gene–by–gene–by–gene

triplets that determine the trait. However, the use of weighted regression shows that the difference between the phenotypic and genotypic estimation can be reduced, and the total genetic effects will be closer to the true value. Similar results were obtained by estimating QTL–QTL–QTL triple interaction effects for total phenolic content in the mapping population of wheat doubled haploid lines under drought stress conditions [30].

As a direct consequence of interactions, particularly the involvement of QTLs in epistatic [63] and higher-order interactions, the effects of single QTL loci are dependent on the genotypes of other loci. As the results presented here show, consideration of the presence of QTL–QTL epistasis effects had an impact on the evaluation of the effect of QTL–QTL–QTL triple interaction. The lack of epistasis assumed in the simulation studies can be considered in two ways: (1) as the actual absence of epistasis effect or (2) as the summed zero effect of QTL–QTL pairs with epistatic effects different from zero. Attempts to use QTLs in breeding programs must consider not only epistatic effects but also higher-order interactions. Determining the contribution of QTL–QTL–QTL triple interactions is important for understanding the genetic basis of complex traits. Therefore, the genetic models for QTL mapping assuming no QTL–QTL–QTL triple interaction can lead to a biased estimation of QTL parameters and QTL–QTL epistasis [63]. This is indicated by the results obtained by Cyplik et al. [64], where genetic parameters were analyzed for selected traits of maize inbred lines. Although the approach presented in this paper is based on, among other things, a genetic map, the proposed methods to estimate additive–by–additive–by–additive of QTL×QTL×QTL interaction effects can be used in association mapping, as a complement and supplement to the analysis of the effects of major genes and epistasis [65–68].

The total QTLs with additive QTL effects, additive-additive QTL–QTL epistatic effects and additive–additive–additive QTL–QTL–QTL triple interaction effects explained more than 83% of the phenotypic variation. The information obtained in this study will be useful for manipulating the QTLs for plant breeding by marker-assisted selection.

4. Materials and Methods

4.1. Data Set

Consider a population of n significantly differentiated homozygous (recombinant inbred—RI or doubled haploid—DH) plant lines from a cross between two homozygous parents, as is currently practiced in the breeding of selfing species. For this population, an n -vector of phenotypic mean observations $y = [y_1 \ y_2 \ \dots \ y_n]'$ and q n -vectors of marker genotype observations m_l , where $l = 1, 2, \dots, q$, are obtained. The i -th element ($i = 1, 2, \dots, n$) of vector m_l is equal to -1 or 1 , depending on the parental genotype exhibited by the i -th line.

4.2. Estimation of the Additive–by–Additive–by–Additive Interaction Effects of Genes Action

4.2.1. Estimation Based on the Phenotype

The total additive × additive × additive interaction of homozygous loci (three-way interaction) effect on the basis of phenotypic (aaa_p) observations of total phenolic content can be estimated by the formula [29]:

$$\widehat{aaa}_p = \frac{1}{2}(L_{max} + L_{min}) - \bar{L}, \quad (1)$$

where L_{min} and L_{max} are the doubled haploid lines with minimal and maximal mean values, respectively; \bar{L} is the mean of all homozygous lines. The number of genes (number of effective factors, \widehat{K}) obtained, based on phenotypic observations only, was calculated using the formula presented by Kaczmarek et al. [69]:

$$\widehat{K} = \frac{(L_{max} - L_{min})^2}{4V_L}, \quad (2)$$

where V_L is additive variance.

4.2.2. Estimation Based on the Genotype

A multiple linear regression model [70,71] of trait values on observations of markers acting as explanatory variables [72] was used to map QTL. It allows the elimination of markers that reveal an association with QTL when tested individually, but in multiple analyses are not characterized by any effect on the phenotypic trait [73]. The estimation of aaa_g was based on the assumption that genes responsible for the quantitative trait were completely linked to observed molecular markers. After deciding which p markers out of all observed sufficiently well explain the variability of the quantitative trait, we can model phenotypic observation for homozygous lines as

$$y = 1\mu + A\beta + E\gamma + T\delta + e, \tag{3}$$

where y —mean values of the observed trait, 1 —the n -vector of ones, μ —the general mean, A —($n \times p$)-matrix of the form $A = [m_{l_1} \ m_{l_2} \ \dots \ m_{l_p}]$, $l_1, l_2, \dots, l_p \in \{1, 2, \dots, q\}$, β —the p -vector of unknown parameters of the form $\beta' = [a_{l_1} \ a_{l_2} \ \dots \ a_{l_p}]$, E —matrix which columns are products of some columns of matrix A , γ —the vector of unknown parameters of the form $\gamma' = [aa_{l_1l_2} \ aa_{l_1l_3} \ \dots \ aa_{l_{p-1}l_p}]$, T —matrix which columns are three-way products of some columns of matrix A , δ —the vector of unknown parameters of the form $\delta' = [aaa_{l_1l_2l_3} \ aaa_{l_1l_2l_4} \ \dots \ aaa_{l_{p-2}l_{p-1}l_p}]$, e —the n -vector of random variables such that $E(e_i) = 0$, $Cov(e_i, e_j) = 0$ for $i \neq j$, $i, j = 1, 2, \dots, n$. The parameters $a_{l_1}, a_{l_2}, \dots, a_{l_p}$ are the additive effects of the genes controlling the quantitative trait, parameters $aa_{l_1l_2}, aa_{l_1l_3}, \dots, aa_{l_{p-1}l_p}$, are the additive \times additive interaction effects and parameters $aaa_{l_1l_2l_3}, aaa_{l_1l_2l_4}, \dots, aaa_{l_{p-2}l_{p-1}l_p}$ are the additive \times additive \times additive interaction effects. We assume that the epistatic and three-way interaction effects show only loci with significant additive gene action effects. This assumption significantly decreases the number of potentially significant effects and causes the regression model more useful.

Unweighted Regression

The selection of markers chosen for model (3) was made by a stepwise trait selection by Akaike Information Criteria [74]. The procedure consisted of two steps: first we divided markers into groups based on the chromosomes they were located on and performed stepwise trait selection by AIC; after that, we combined the remaining markers into one group and repeated selection as above. All of the remaining markers were combined into the final group and the last feature selection was performed on a model with additive \times additive \times additive interaction effect included. To counteract the multiple comparison problems, we used Bonferroni correction [75].

Denoting by $\alpha' = [\mu \ \beta' \ \gamma' \ \delta']$ and $G = [1 \ A \ E \ T]$ we obtain the model:

$$y = G\alpha + e. \tag{4}$$

If G is of full rank, the estimate of α_u from traditional (unweighted) multiple linear regression model is given by [76]:

$$\widehat{\alpha}_u = (G'G)^{-1}G'y. \tag{5}$$

The total three-way interaction aaa_{gu} effect of genes influencing the total phenolic content from traditional (unweighted) multiple linear regression model can be found as

$$\widehat{aaa}_{gu} = \sum_{k=1}^{p-2} \sum_{\substack{k'=2 \\ k' \neq k}}^{p-1} \sum_{\substack{k''=3 \\ k'' \neq k'}} \widehat{aaa}_{l_k l_{k'} l_{k''}}. \tag{6}$$

Weighted Regression

The modified version of trait regression on marker data in this paper is considered by taking a weighted multiple linear regression, that is, regression with a diagonal matrix W of unknown variances of observations, which, however, may be empirically found by estimation. In this model, the estimate of α_w is:

$$\widehat{\alpha}_w = (G'W^{-1}G)^{-1}G'W^{-1}y, \quad (7)$$

where $W = (w_{ii})$ with w_{ii} being the estimated variance for i -th homozygous line, $i = 1, 2, \dots, n$. Selection of markers for the weighted regression is made in two ways: (i) by the same method as described for the unweighted case (weighted triple interaction effects only) and (ii) using weighted regression to select markers for model (3) (weighted all effects: QTL, epistatic and triple interaction).

The total three-way interaction $\{aaa_{gw1}$ —for (i) way and aaa_{gw2} —for (ii) way} effect of genes influencing the quantitative trait from weighted multiple linear regression model can be found as

$$\widehat{aaa}_{gwx} = \sum_{k=1}^{p-2} \sum_{\substack{k'=2 \\ k' \neq k}}^{p-1} \sum_{\substack{k''=3 \\ k'' \neq k'}} \widehat{aaa}_{l_k l_{k'} l_{k''}}, \quad (8)$$

where x is equal to 1 {for (i) way} or 2 {for (ii) way}.

Coefficient of Determination

The coefficients of determination were used to measure how the model (3) fitted the data and, in this study, was the amount of the phenotypic variance explained by total QTLs with additive effects, QTL-pairs with epistatic effects and QTL-triples with additive-by-additive-by-additive effects (R^2). The R^2 was calculated for unweighted regression and in both cases for weighted regression.

4.2.3. Simulation Studies

In the Monte Carlo simulation studies comparing the “phenotypic” and “genotypic” (unweighted and two ways weighted) estimates of the additive-by-additive-by-additive interaction of QTL effects the following variants of the selected parameters were adopted. The true value of the parameter connected with additive \times additive \times additive interaction effect was assumed to be equal to 15 ($aaa = 15$) and the total mean value of the quantitative trait to 100. In all 500 homozygous lines were analyzed. Each homozygous line was represented by ten plants. We assumed genotyping for 1000 molecular markers distributed in ten chromosomes of 100 cM. The number of molecular markers per chromosome was equal to 100. The distance between markers was equal to 1 cM. The number of QTLs affecting the trait was 14 (each with an additive effect of two). A QTL was marked perfectly so that the marker and QTLs were in complete linkage disequilibrium with no recombination. We allocated one QTL in chromosome 1 (Q1), two QTLs in chromosomes 3 (Q2 and Q3), 5 (Q4 and Q5), and three QTLs in chromosomes 7 (Q6, Q7 and Q8), 9 (Q9, Q10 and Q11), and 10 (Q12, Q13 and Q14) (Table 1). We assumed the action of five epistatic pairs: Q1–Q2, Q4–Q5, Q6–Q7, Q8–Q11, and Q1–Q14 (Figure 1). Thus, we defined two epistatic pairs in the same chromosome: Q4 and Q5 (in chromosome 5) and Q6 and Q7 (in chromosome 7), and three epistatic pairs in distinct chromosomes. Equal epistatic effects were assumed for all five pairs in two variants: two or zero (no epistasis). The number of QTL–QTL–QTL threes with additive-by-additive-by-additive interaction effects that affected the trait was assumed to be one, two, three or five. The QTLs with additive-by-additive-by-additive interaction effects were located in one chromosome (three triples on one chromosome and two triples on two chromosomes), two chromosomes or three chromosomes (each QTL was in a different chromosome) (Table 1). Effects of particular triples of genes were assumed to be: (i) equal for all triples, or (ii) one QTL–QTL–QTL triple effect was much large than the other (for two triples: 10 and 5; for three triples: 9, 3 and 3; for five triples: 7, 2, 2, 2 and 2). The error variance was equal to 5 of 10. A

total of 10,000 data sets containing the vector of phenotypic observations and vectors of marker genotype observations were generated for each combination of the parameters. For each data set the line variances were estimated and the additive-by-additive-by-additive interaction effect was estimated by the phenotypic method (\widehat{aaa}_{jp}), and by unweighted (\widehat{aaa}_{jgu}) and weighted (\widehat{aaa}_{jgw1} , \widehat{aaa}_{jgw2}) ($j = 1, 2, \dots, 10,000$) versions of the genotypic method presented above. Additionally, the coefficients of determination R_j^2 were estimated. Then, the mean values of parameter estimates \widehat{aaa}_p , \widehat{aaa}_{gu} , \widehat{aaa}_{gw1} and \widehat{aaa}_{gw2} for each series were calculated, together with the mean squared errors. Mean values of R^2 were calculated. Statistical analysis was performed using GenStat 22nd edition [77].

5. Conclusions

In practical research, one should expect estimates of the total effect of additive-by-additive-by-additive for triplets of gene-by-gene-by-gene interactions quantitative traits to be smaller than the total phenotype estimate. If a different situation is found, it should be carefully analyzed, checking whether the genetic assumptions of correct segregation and lack of close linkage between markers are met. Weighted multiple linear regression is useful for estimating the parameter associated with additive-by-additive-by-additive of the gene-by-gene-by-gene interaction effect. Regardless of the variant of the approach in the weighted regression, an improvement in the estimation of the parameter under consideration was obtained compared to that obtained using unweighted regression. This indicates the applicability of this estimation method in various experimental situations. The results indicate that the estimation of additive-by-additive-by-additive interaction effects by the weighted regression method can be applied to different plant species.

Author Contributions: Conceptualization, J.B.; methodology, A.C. and J.B.; software, A.C. and J.B.; validation, A.C. and J.B.; formal analysis, A.C. and J.B.; investigation, A.C. and J.B.; resources, A.C. and J.B.; data curation, A.C. and J.B.; writing—original draft preparation, A.C. and J.B.; writing—review and editing, A.C. and J.B.; visualization, J.B.; supervision, J.B.; project administration, J.B.; funding acquisition, J.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Falconer, D.S.; Mackay, T.F.C. *Introduction to Quantitative Genetics*; Longman: Essex, UK, 1996.
2. Bloom, J.S.; Kotenko, I.; Sadhu, M.J.; Treusch, S.; Albert, F.W.; Kruglyak, L. Genetic interactions contribute less than additive effects to quantitative trait variation in yeast. *Nat. Commun.* **2015**, *6*, 8712. [[CrossRef](#)] [[PubMed](#)]
3. Mather, K. *Biometrical Genetics*; Methuen & Co. Ltd.: London, UK, 1949.
4. Young, N.D. A cautiously optimistic vision for marker-assisted breeding. *Mol. Breed.* **1999**, *5*, 505–510. [[CrossRef](#)]
5. Amiteye, S. Basic concepts and methodologies of DNA marker systems in plant molecular breeding. *Heliyon* **2021**, *7*, e08093. [[CrossRef](#)] [[PubMed](#)]
6. Lander, E.S.; Botstein, D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **1989**, *121*, 185–199. [[CrossRef](#)]
7. Melchinger, A.E.; Utz, H.F.; Schön, C.C. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics* **1998**, *149*, 383–403. [[CrossRef](#)]
8. Ren, J.; Zhang, X.; Li, Z.; Wu, P. Genetic Analysis of Maternal Haploid Inducibility for *In Vivo* Haploid Induction in Maize. *Agriculture* **2022**, *12*, 845. [[CrossRef](#)]
9. Sayed, M.A.; Nassar, S.M.; Moustafa, E.S.; Said, M.T.; Börner, A.; Hamada, A. Genetic Mapping Reveals Novel Exotic and Elite QTL Alleles for Salinity Tolerance in Barley. *Agronomy* **2021**, *11*, 1774. [[CrossRef](#)]

10. Kabange, N.R.; Park, S.-Y.; Shin, D.; Lee, S.-M.; Jo, S.-M.; Kwon, Y.; Cha, J.-K.; Song, Y.-C.; Ko, J.-M.; Lee, J.-H. Identification of a Novel QTL for Chlorate Resistance in Rice (*Oryza sativa* L.). *Agriculture* **2020**, *10*, 360. [[CrossRef](#)]
11. Kwon, Y.-H.; Kabange, N.-R.; Lee, J.-Y.; Lee, S.-M.; Cha, J.-K.; Shin, D.-J.; Cho, J.-H.; Kang, J.-W.; Ko, J.-M.; Lee, J.-H. Novel QTL Associated with Shoot Branching Identified in Doubled Haploid Rice (*Oryza sativa* L.) under Low Nitrogen Cultivation. *Genes* **2021**, *12*, 745. [[CrossRef](#)]
12. Botero-Ramírez, A.; Laperche, A.; Guichard, S.; Jubault, M.; Gravot, A.; Strelkov, S.E.; Manzanares-Dauleux, M.J. Clubroot Symptoms and Resting Spore Production in a Doubled Haploid Population of Oilseed Rape (*Brassica napus*) Are Controlled by Four Main QTLs. *Front. Plant Sci.* **2020**, *11*, 604527. [[CrossRef](#)]
13. Bocianowski, J.; Kozak, M.; Liersch, A.; Bartkowiak-Broda, I. A heuristic method of searching for interesting markers in terms of quantitative traits. *Euphytica* **2011**, *181*, 89–100. [[CrossRef](#)]
14. Gacek, K.; Bayer, P.E.; Anderson, R.; Severn-Ellis, A.A.; Wolko, J.; Łopatyńska, A.; Matuszczak, M.; Bocianowski, J.; Edwards, D.; Batley, J. QTL Genetic Mapping Study for Traits Affecting Meal Quality in Winter Oilseed Rape (*Brassica napus* L.). *Genes* **2021**, *12*, 1235. [[CrossRef](#)]
15. Lephuthing, M.C.; Khumalo, T.P.; Tolmay, V.L.; Dube, E.; Tsilo, T.J. Genetic Mapping of Quantitative Trait Loci Associated with Plant Height and Yield Component Traits in a Wheat (*Triticum aestivum* L.) Doubled Haploid Population Derived from Tugela-DN × Elands. *Agronomy* **2022**, *12*, 2283. [[CrossRef](#)]
16. Pegot-Espagnet, P.; Guillaume, O.; Desprez, B.; Devaux, B.; Devaux, P.; Henry, K.; Henry, N.; Willems, G.; Goudemand, E.; Mangin, B. Discovery of interesting new polymorphisms in a sugar beet (elite × exotic) progeny by comparison with an elite panel. *Theor. Appl. Genet.* **2019**, *132*, 3063–3078. [[CrossRef](#)] [[PubMed](#)]
17. Liu, X.; Li, Y.I.; Pritchard, J.K. *Trans* Effects on Gene Expression Can Drive Omnigenic Inheritance. *Cell* **2019**, *177*, 1022–1034. [[CrossRef](#)]
18. O'Connor, L.J.; Schoech, A.P.; Hormozdiari, F.; Gazal, S.; Patterson, N.; Price, A.L. Extreme Polygenicity of Complex Traits Is Explained by Negative Selection. *Am. J. Hum. Genet.* **2019**, *105*, 456–476. [[CrossRef](#)]
19. Pundir, S.; Sharma, R.; Kumar, D.; Singh, V.K.; Chaturvedi, D.; Kanwar, R.S.; Röder, M.S.; Börner, A.; Ganai, W.M.; Gupta, P.K.; et al. QTL mapping for resistance against cereal cyst nematode (*Heterodera avenae* Woll.) in wheat (*Triticum aestivum* L.). *Sci. Rep.* **2022**, *12*, 9586. [[CrossRef](#)]
20. Dhariwal, R.; Fedak, G.; Dion, Y.; Pozniak, C.; Laroche, A.; Eudes, F.; Randhawa, H.S. High Density Single Nucleotide Polymorphism (SNP) Mapping and Quantitative Trait Loci (QTL) Analysis in a Biparental Spring Triticale Population Localized Major and Minor Effect *Fusarium* Head Blight Resistance and Associated Traits QTL. *Genes* **2018**, *9*, 19. [[CrossRef](#)] [[PubMed](#)]
21. Smeda, J.R.; Schillmiller, A.L.; Anderson, T.; Ben-Mahmoud, S.; Ullman, D.E.; Chappell, T.M.; Kessler, A.; Mutschler, M.A. Combination of Acylglucose QTL reveals additive and epistatic genetic interactions and impacts insect oviposition and virus infection. *Mol. Breed.* **2018**, *38*, 3. [[CrossRef](#)]
22. Li, H.; Zhao, T.; Wang, Y.; Yu, D.; Chen, S.; Zhou, R.; Gai, J. Genetic structure composed of additive QTL, epistatic QTL pairs and collective unmapped minor QTL conferring oil content and fatty acid components of soybeans. *Euphytica* **2011**, *182*, 117–132. [[CrossRef](#)]
23. Ali, F.; Chen, W.; Fiaz, S.; Wang, Y.; Wei, X.; Xie, L.; Jiao, G.; Shao, G.; Hu, S.; Tang, S.; et al. QTL Mapping for Grain Appearance Quality Traits Using Doubled Haploid Population of Rice Under Different Environments. *Pak. J. Bot.* **2022**, *54*, 1265–1275. [[CrossRef](#)]
24. Krajewski, P.; Bocianowski, J.; Gawłowska, M.; Kaczmarek, Z.; Pniewski, T.; Świącicki, W.; Wolko, B. QTL for yield components and protein content: A multienvironment study of two pea (*Pisum sativum* L.) populations. *Euphytica* **2012**, *183*, 323–336. [[CrossRef](#)]
25. Yusuf, A.O.; Richter, J.-C.; Möllers, C. Genetic variation and QTL analysis of saturated fatty acids in two doubled haploid populations of oilseed rape (*Brassica napus* L.). *Euphytica* **2022**, *218*, 88. [[CrossRef](#)]
26. Ku, L.X.; Sun, Z.H.; Wang, C.L.; Zhang, J.; Zhao, R.F.; Liu, H.Y.; Tai, G.Q.; Chen, Y.H. QTL mapping and epistasis analysis of brace root traits in maize. *Mol. Breed.* **2012**, *30*, 697–708. [[CrossRef](#)]
27. Beheshtizadeh, H.; Fakheri, B.A.; Aghnoum, R.; Mahdinezhad, N.; Pouredad, S.S.; Masoudi, B. QTL mapping of grain yield and its components under normal and drought stress conditions in barley (*Hordeum vulgare* L.). *Indian J. Genet. Plant Breed.* **2018**, *78*, 69–80. [[CrossRef](#)]
28. Wright, S. Genic and organismic selection. *Evolution* **1980**, *34*, 825–843. [[CrossRef](#)]
29. Cyplik, A.; Bocianowski, J. Analytical and numerical comparisons of two methods of estimation of additive × additive × additive interaction of QTL effects. *J. Appl. Genet.* **2022**, *63*, 213–221. [[CrossRef](#)]
30. Cyplik, A.; Czyczyło-Mysza, I.M.; Jankowicz-Cieslak, J.; Bocianowski, J. QTL × QTL × QTL Interaction Effects for Total Phenolic Content of Wheat Mapping Population of CSDH Lines under Drought Stress by Weighted Multiple Linear Regression. *Agriculture* **2023**, *13*, 850. [[CrossRef](#)]
31. Cullis, B.R.; Smith, A.B.; Cocks, N.A.; Butler, D.G. The Design of Early-Stage Plant Breeding Trials Using Genetic Relatedness. *JABES* **2020**, *25*, 553–578. [[CrossRef](#)]
32. Wang, D.; Salah El-Basyoni, I.; Stephen Baenziger, P.; Cross, J.; Eskridge, K.M.; Dweikat, I. Prediction of genetic values of quantitative traits with epistatic effects in plant breeding populations. *Heredity* **2012**, *109*, 313–319. [[CrossRef](#)] [[PubMed](#)]

33. Bouvet, J.M.; Makouanzi, G.; Cros, D.; Vigneron, P.H. Modeling additive and non-additive effects in a hybrid population using genome-wide genotyping: Prediction accuracy implications. *Heredity* **2016**, *116*, 146–157. [[CrossRef](#)]
34. Voss-Fels, K.P.; Cooper, M.; Hayes, B.J. Accelerating crop genetic gains with genomic selection. *Theor. Appl. Genet.* **2019**, *132*, 669–686. [[CrossRef](#)] [[PubMed](#)]
35. Bocianowski, J.; Warzecha, T.; Nowosad, K.; Bathelt, R. Genotype by environment interaction using AMMI model and estimation of additive and epistasis gene effects for 1000-kernel weight in spring barley (*Hordeum vulgare* L.). *J. Appl. Genet.* **2019**, *60*, 127–135. [[CrossRef](#)] [[PubMed](#)]
36. Ali, M.; Zhang, L.; DeLacy, I.; Arief, V.; Dieters, M.; Pfeiffer, W.H.; Wang, J.; Lu, H. Modeling and simulation of recurrent phenotypic and genomic selections in plant breeding under the presence of epistasis. *Crop. J.* **2020**, *8*, 866–877. [[CrossRef](#)]
37. Labroo, M.R.; Studer, A.J.; Rutkoski, J.E. Heterosis and Hybrid Crop Breeding: A Multidisciplinary Review. *Front. Genet.* **2021**, *12*, 643761. [[CrossRef](#)]
38. Raffo, M.A.; Sarup, P.; Guo, X.; Liu, H.; Andersen, J.R.; Orabi, J.; Jahoor, A.; Jensen, J. Improvement of genomic prediction in advanced wheat breeding lines by including additive-by-additive epistasis. *Theor. Appl. Genet.* **2022**, *135*, 965–978. [[CrossRef](#)]
39. Yang, J.; Zhu, J.; Williams, R.W. Mapping the genetic architecture of complex traits in experimental populations. *Bioinformatics* **2007**, *23*, 1527–1536. [[CrossRef](#)]
40. Bocianowski, J. A comparison of two methods to estimate additive-by-additive interaction of QTL effects by a simulation study. *J. Theor. Biol.* **2012**, *308*, 20–24. [[CrossRef](#)]
41. Ahsan, A.; Monir, M.; Meng, X.; Rahaman, M.; Chen, H.; Chen, M. Identification of epistasis loci underlying rice flowering time by controlling population stratification and polygenic effect. *DNA Res.* **2019**, *26*, 119–130. [[CrossRef](#)]
42. de los Campos, G.; Sorensen, D.A.; Toro, M.A. Imperfect Linkage Disequilibrium Generates Phantom Epistasis (& Perils of Big Data). *G3 Genes | Genomes | Genet.* **2019**, *9*, 1429–1436. [[CrossRef](#)] [[PubMed](#)]
43. Wang, D.; Tang, H.; Liu, J.-F.; Xu, S.; Zhang, Q.; Ning, C. Rapid epistatic mixed-model association studies by controlling multiple polygenic effects. *Bioinformatics* **2020**, *36*, 4833–4837. [[CrossRef](#)] [[PubMed](#)]
44. Sharma, R.; Sadeghian Tehrani, Z.; Kumar, S.; Song, M. Detecting genetic epistasis by differential departure from independence. *Mol. Genet. Genom.* **2022**, *297*, 911–924. [[CrossRef](#)] [[PubMed](#)]
45. Ronin, Y.I.; Korol, A.B.; Nevo, E. Single- and Multiple-Trait Mapping Analysis of Linked Quantitative Trait Loci: Some Asymptotic Analytical Approximations. *Genetics* **1999**, *151*, 387–396. [[CrossRef](#)] [[PubMed](#)]
46. Viana, J.M.S.; Souza, C.A.S. Efficiency of mapping epistatic quantitative trait loci. *Heredity* **2023**. *accepted; online first.* [[CrossRef](#)]
47. Crawford, L.; Zeng, P.; Mukherjee, S.; Zhou, X. Detecting epistasis with the marginal epistasis test in genetic mapping studies of quantitative traits. *PLoS Genet.* **2017**, *13*, e1006869. [[CrossRef](#)]
48. Berke, T.G.; Rocheford, T.R. Quantitative Trait Loci for Tassel Traits in Maize. *Crop. Sci.* **1999**, *39*, 1439–1443. [[CrossRef](#)]
49. Zimnoch-Guzowska, E.; Marczewski, W.; Lebecka, R.; Flis, B.; Schäfer-Pregl, R.; Salamini, F.; Gebhardt, C. QTL Analysis of New Sources of Resistance to *Erwinia carotovora* ssp. *atroseptica* in Potato Done by AFLP, RFLP, and Resistance-Gene-Like Markers. *Crop. Sci.* **2000**, *40*, 1156–1167. [[CrossRef](#)]
50. Burke, J.M.; Tang, S.; Knapp, S.J.; Rieseberg, L.H. Genetic Analysis of Sunflower Domestication. *Genetics* **2002**, *161*, 1257–1267. [[CrossRef](#)]
51. Flint-Garcia, S.A.; Jampatong, C.; Darrah, L.L.; McMullen, M.D. Quantitative Trait Locus Analysis of Stalk Strength in Four Maize Populations. *Crop. Sci.* **2003**, *43*, 13–22. [[CrossRef](#)]
52. Wu, R.L. Mapping Quantitative Trait Loci by Genotyping Haploid Tissues. *Genetics* **1999**, *152*, 1741–1752. [[CrossRef](#)]
53. Balestre, M.; de Souza, C.L. Bayesian reversible-jump for epistasis analysis in genomic studies. *BMC Genom.* **2016**, *17*, 1012. [[CrossRef](#)]
54. Rajurkar, A.B.; Muthukumar, C.; Ayyenar, B.; Thomas, H.B.; Babu, R.C. Mapping consistent additive and epistatic QTLs for plant production traits under drought in target populations of environment using locally adapted landrace in rice (*Oryza sativa* L.). *Plant Prod. Sci.* **2021**, *24*, 388–403. [[CrossRef](#)]
55. Shi, H.; Guan, W.; Shi, Y.; Wang, S.; Fan, H.; Yang, J.; Chen, W.; Zhang, W.; Sun, D.; Jing, R. QTL mapping and candidate gene analysis of seed vigor-related traits during artificial aging in wheat (*Triticum aestivum*). *Sci. Rep.* **2020**, *10*, 22060. [[CrossRef](#)] [[PubMed](#)]
56. Chattopadhyay, K.; Mohanty, S.K.; Vijayan, J.; Marndi, B.C.; Sarkar, A.; Molla, K.A.; Chakraborty, K.; Ray, S.; Ramani Kumar Sarkar, R.K. Genetic Dissection of Component Traits for Salinity Tolerance at Reproductive Stage in Rice. *Plant Mol. Biol. Rep.* **2021**, *39*, 386–402. [[CrossRef](#)]
57. De Leon, T.B.; Pruthi, R.; Jampala, B.; Borjas, A.H.; Subudhi, P.K. Genetic determinants for agronomic and yield-related traits localized on a GBS-SNP linkage map from a *japonica* × *indica* cross in rice. *Plant Gene* **2020**, *24*, 100249. [[CrossRef](#)]
58. Goutelle, S.; Bourguignon, L.; Maire, P.H.; Van Guilder, M.; Conte, J.E., Jr.; Jelliffe, R.W. Population Modeling and Monte Carlo Simulation Study of the Pharmacokinetics and Antituberculosis Pharmacodynamics of Rifampin in Lungs. *Antimicrob. Agents Chemother.* **2009**, *53*, 1520. [[CrossRef](#)]
59. Avery, L.; Rotondi, N.; McKnight, C.; Firestone, M.; Smylie, J.; Rotondi, M. Unweighted regression models perform better than weighted regression techniques for respondent-driven sampling data: Results from a simulation study. *BMC Med. Res. Methodol.* **2019**, *19*, 202. [[CrossRef](#)]

60. Johnson, M.S.; Gopalakrishnan, S.; Goyal, J.; Dillingham, M.E.; Bakerlee, C.W.; Humphrey, P.T.; Jagdish, T.; Jerison, E.R.; Kosheleva, K.; Lawrence, K.R.; et al. Phenotypic and molecular evolution across 10,000 generations in laboratory budding yeast populations. *eLife* **2021**, *10*, e63910. [[CrossRef](#)] [[PubMed](#)]
61. Sorojrisom, E.S.; Haller, B.C.; Ambrose, B.A.; Eaton, D.A.R. Selection on the gametophyte: Modeling alternation of generations in plants. *Appl. Plant Sci.* **2022**, *10*, e11472. [[CrossRef](#)]
62. Wang, X.; Xie, W.; Li, L.-y.; Zhu, J.; Xing, F. Molecular Simulation Study on Mechanical Properties of Microcapsule-Based Self-Healing Cementitious Materials. *Polymers* **2022**, *14*, 611. [[CrossRef](#)]
63. Bocianowski, J. Epistasis interaction of QTL effects as a genetic parameter influencing estimation of the genetic additive effect. *Genet. Mol. Biol.* **2013**, *36*, 93–100. [[CrossRef](#)]
64. Cyplik, A.; Sobiech, A.; Tomkowiak, A.; Bocianowski, J. Genetic Parameters for Selected Traits of Inbred Lines of Maize (*Zea mays* L.). *Appl. Sci.* **2022**, *12*, 6961. [[CrossRef](#)]
65. Rakoczy-Trojanowska, M.; Krajewski, P.; Bocianowski, J.; Schollenberger, M.; Wakuliński, W.; Milczarski, P.; Masojć, P.; Targońska-Karasek, M.; Banaszak, Z.; Banaszak, K.; et al. Identification of Single Nucleotide Polymorphisms Associated with Brown Rust Resistance, α -Amylase Activity and Pre-harvest Sprouting in Rye (*Secale cereale* L.). *Plant Mol. Biol. Rep.* **2017**, *35*, 366–378. [[CrossRef](#)] [[PubMed](#)]
66. Rakoczy-Trojanowska, M.; Orczyk, W.; Krajewski, P.; Bocianowski, J.; Stochmal, A.; Kowalczyk, M. *ScBx* gene based association analysis of hydroxamate content in rye (*Secale cereale* L.). *J. Appl. Genet.* **2017**, *58*, 1–9. [[CrossRef](#)]
67. Tomkowiak, A.; Bocianowski, J.; Wolko, L.; Adamczyk, J.; Mikołajczyk, S.; Kowalczewski, P.L. Identification of Markers Associated with Yield Traits and Morphological Features in Maize (*Zea mays* L.). *Plants* **2019**, *8*, 330. [[CrossRef](#)] [[PubMed](#)]
68. Tomkowiak, A.; Bocianowski, J.; Spychała, J.; Grynia, J.; Sobiech, A.; Kowalczewski, P.L. DArTseq-Based High-Throughput SilicoDArT and SNP Markers Applied for Association Mapping of Genes Related to Maize Morphology. *Int. J. Mol. Sci.* **2021**, *22*, 5840. [[CrossRef](#)]
69. Kaczmarek, Z.; Surma, M.; Adamski, T. Epistatic effects in estimation of the number of genes on the basis of doubled haploid lines. *Genet. Pol.* **1988**, *29*, 353–359.
70. Jansen, R.C. Interval mapping of multiple quantitative traits. *Genetics* **1993**, *135*, 205–214. [[CrossRef](#)] [[PubMed](#)]
71. Haley, C.S.; Knott, S.A.; Elsen, J.M. Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics* **1994**, *136*, 1195–1207. [[CrossRef](#)] [[PubMed](#)]
72. Jansen, R.C. A general Monte Carlo method for mapping multiple quantitative trait loci. *Genetics* **1996**, *142*, 305–311. [[CrossRef](#)]
73. Bocianowski, J.; Krajewski, P. Comparison of the genetic additive effect estimators based on phenotypic observations and on molecular marker data. *Euphytica* **2009**, *165*, 113–122. [[CrossRef](#)]
74. Akaike, H. Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*; Parzen, E., Tanabe, K., Kitagawa, G., Eds.; Springer Series in Statistics (Perspectives in Statistics); Springer: New York, NY, USA, 1998; pp. 199–213. [[CrossRef](#)]
75. Province, M.A. 30 Sequential methods of analysis for genome scan. *Adv. Genet.* **2001**, *42*, 499–514. [[CrossRef](#)] [[PubMed](#)]
76. Searle, S.R. *Matrix Algebra Useful for Statistics*; Wiley: New York, NY, USA, 1982.
77. VSN International. *Genstat Reference Manual (Release 22), Part 2 Directives*; VSN International: Hemel Hempstead, UK, 2022.

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.

ROZDZIAŁ III

Oświadczenia autorów

mgr inż. Adrian Cyplik

PESEL: 94010706036

Katedra Metod Matematycznych i Statystycznych, Uniwersytet Przyrodniczy w Poznaniu

ul. Wojska Polskiego 28, 60-637 Poznań

Nr telefonu: +48 603 843 080

E-mail: adrian.cyplik@up.poznan.pl

Oświadczenie o współautorstwie

Niniejszym oświadczam, że w pracy A. Cyplik, J. Bocianowski, 2022, **Analytical and numerical comparisons of two methods of estimation of additive \times additive \times additive interaction of QTL effects**, *Journal of Applied Genetics*, 63(2), 213–221, DOI:10.1007/s13353-021-00676-7 mój indywidualny udział w jej powstaniu polegał na: przygotowaniu projektu badań, opracowaniu krytycznego przeglądu literatury, opracowaniu założeń i metod, przeprowadzeniu badań, w tym zbieraniu danych, analizie wyników, interpretacji wyników i sformułowaniu wniosków, obliczeniach statystycznych, opracowaniu grafik, opracowaniu manuskryptu, obsłudze oprogramowania, walidacji oraz odpowiedziach na uwagi recenzentów.

Data 30.08.2023

Podpis

Adrian Cyplik

Prof. UPP dr hab. Jan Bocianowski

PESEL: 73101001793

Katedra Metod Matematycznych i Statystycznych, Uniwersytet Przyrodniczy w Poznaniu

ul. Wojska Polskiego 28, 60-637 Poznań

Nr telefonu: +48 61 8487143

E-mail: jan.bocianowski@up.poznan.pl

Oświadczenie o współautorstwie

Niniejszym oświadczam, że w pracy A. Cyplik, J. Bocianowski, 2022, **Analytical and numerical comparisons of two methods of estimation of additive \times additive \times additive interaction of QTL effects**, *Journal of Applied Genetics*, 63(2), 213–221, DOI:10.1007/s13353-021-00676-7 mój indywidualny udział w jej powstaniu polegał na: autorstwie koncepcji, przygotowaniu projektu badań, opracowaniu krytycznego przeglądu literatury, opracowaniu założeń i metod, przeprowadzeniu badań, w tym zbieraniu danych, analizie wyników, interpretacji wyników i sformułowaniu wniosków, obliczeniach statystycznych, opracowaniu manuskryptu, nadzorze i administracji projektu, pozyskiwaniu finansowania, walidacji oraz odpowiedziach na uwagi recenzentów.

Data 30.09.2023

Podpis

Jan Bocianowski

mgr inż. Adrian Cyplik

PESEL: 94010706036

Katedra Metod Matematycznych i Statystycznych, Uniwersytet Przyrodniczy w Poznaniu

ul. Wojska Polskiego 28, 60-637 Poznań

Nr telefonu: +48 603 843 080

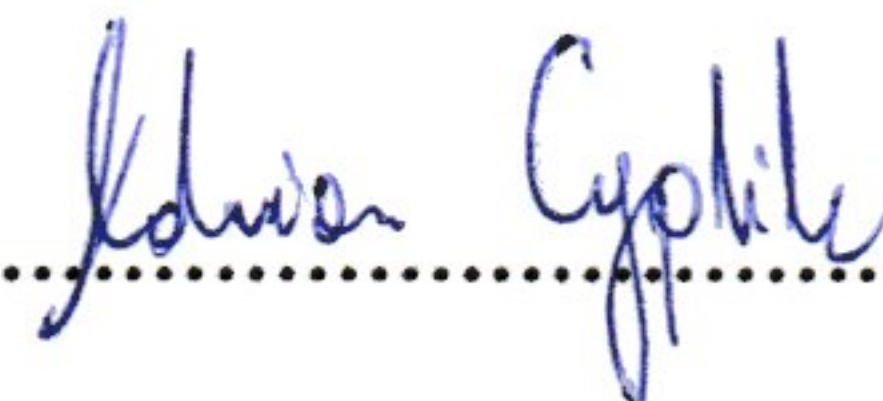
E-mail: adrian.cyplik@up.poznan.pl

Oświadczenie o współautorstwie

Niniejszym oświadczam, że w pracy A. Cyplik, A. Sobiech, A. Tomkowiak, J. Bocianowski, 2022, **Genetic Parameters for Selected Traits of Inbred Lines of Maize (*Zea mays L.*)**, *Applied Sciences*, 12(14), 6961, DOI:10.3390/app12146961 mój indywidualny udział w jej powstaniu polegał na autorstwie koncepcji, opracowaniu założeń i metod, obsłudze oprogramowania, walidacji, analizie formalnej, przeprowadzeniu badań, opracowaniu manuskryptu oraz odpowiedziach na uwagi recenzentów.

Data30.08.2023.....

Podpis

..........

mgr inż. Aleksandra Sobiech

PESEL: 96041407544

Katedra Genetyki i Hodowli Roślin, Uniwersytet Przyrodniczy w Poznaniu

ul. Dojazd 11, 60-632 Poznań

Nr telefonu: 61 848 76 22

E-mail: aleksandra.sobiech@up.poznan.pl

Oświadczenie o współautorstwie

Niniejszym oświadczam, że w pracy A. Cyplik, A. Sobiech, A. Tomkowiak, J. Bocianowski, 2022, **Genetic Parameters for Selected Traits of Inbred Lines of Maize (*Zea mays L.*)**, *Applied Sciences*, 12(14), 6961, DOI:10.3390/app12146961 mój indywidualny udział w jej powstaniu polegał na walidacji, przeprowadzeniu badań, zarządzaniu zasobami, przygotowaniu danych, opracowaniu manuskryptu oraz odpowiedziach na uwagi recenzentów.

Data 30.06.2023r.

Podpis

Aleksandra Sobiech

Prof. UPP dr hab. Agnieszka Tomkowiak

PESEL: 78031603662

Katedra Genetyki i Hodowli Roślin, Uniwersytet Przyrodniczy w Poznaniu

ul. Dojazd 11, 60-632 Poznań

Nr telefonu: 61 848 76 80

E-mail: agnieszka.tomkowiak@up.poznan.pl

Oświadczenie o współautorstwie

Niniejszym oświadczam, że w pracy A. Cyplik, A. Sobiech, A. Tomkowiak, J. Bocianowski, 2022, **Genetic Parameters for Selected Traits of Inbred Lines of Maize (*Zea mays L.*)**, *Applied Sciences*, 12(14), 6961, DOI:10.3390/app12146961 mój indywidualny udział w jej powstaniu polegał na walidacji, przeprowadzeniu badań, zarządzaniu zasobami, przygotowaniu danych, opracowaniu manuskryptu oraz odpowiedziach na uwagi recenzentów.

Data 18.08.2023

Podpis

Agnieszka Tomkowiak

Prof. UPP dr hab. Jan Bocianowski

PESEL: 73101001793

Katedra Metod Matematycznych i Statystycznych, Uniwersytet Przyrodniczy w Poznaniu

ul. Wojska Polskiego 28, 60-637 Poznań

Nr telefonu: +48 61 8487143

E-mail: jan.bocianowski@up.poznan.pl

Oświadczenie o współautorstwie

Niniejszym oświadczam, że w pracy A. Cyplik, A. Sobiech, A. Tomkowiak, J. Bocianowski, 2022, **Genetic Parameters for Selected Traits of Inbred Lines of Maize (*Zea mays L.*)**, *Applied Sciences*, 12(14), 6961, DOI:10.3390/app12146961 mój indywidualny udział w jej powstaniu polegał na autorstwie koncepcji, opracowaniu założeń i metod, obsłudze oprogramowania, walidacji, analizie formalnej, przeprowadzeniu badań, opracowaniu manuskryptu, odpowiedziach na uwagi recenzentów, nadzorze i administracji projektu oraz pozyskiwaniu finansowania.

Data 30.08.2023

Podpis

Jan Bocianowski

mgr inż. Adrian Cyplik

PESEL: 94010706036

Katedra Metod Matematycznych i Statystycznych, Uniwersytet Przyrodniczy w Poznaniu

ul. Wojska Polskiego 28, 60-637 Poznań

Nr telefonu: +48 603 843 080

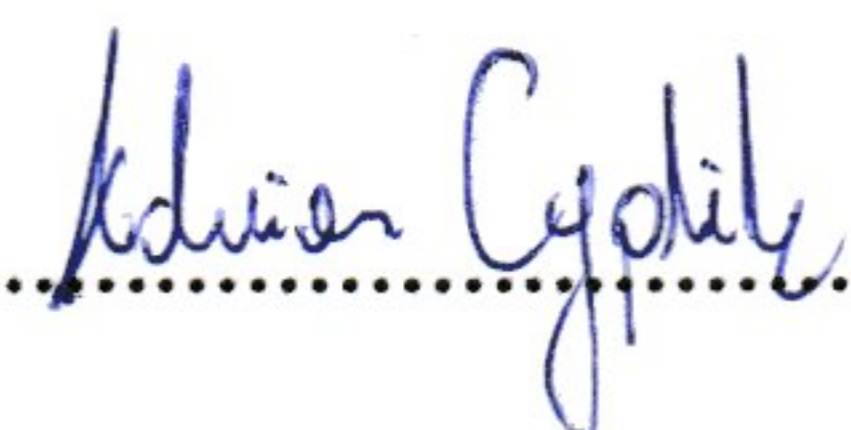
E-mail: adrian.cyplik@up.poznan.pl

Oświadczenie o współautorstwie

Niniejszym oświadczam, że w pracy A. Cyplik, I. M. Czyczyło-Mysza, J. Jankowicz-Cieślak, J. Bocianowski, 2023, **QTL×QTL×QTL Interaction Effects for Total Phenolic Content of Wheat Mapping Population of CSDH Lines under Drought Stress by Weighted Multiple Linear Regression**, *Agriculture*, 13(4), 850, DOI:10.3390/agriculture13040850 mój indywidualny udział w jej powstaniu polegał na opracowaniu założeń i metod, obsłudze oprogramowania, walidacji, analizie formalnej, przeprowadzeniu badań, opracowaniu manuskryptu oraz odpowiedziach na uwagi recenzentów.

Data30.08.2023.....

Podpis

..........

dr hab. inż. Ilona Mieczysława Czyczyło-Mysza, prof. IFR PAN

PESEL: 72010114644

Instytut Fizjologii Roślin im. Franciszka Górskiego Polskiej Akademii Nauk

Niezapominajek 21, 30-239 Kraków

Nr telefonu: +48 12 4251833 (wew./ext. 109)

E-mail: i.czyczylo@ifr-pan.edu.pl

Oświadczenie o współautorstwie

Niniejszym oświadczam, że w pracy A. Cyplik, I. M. Czyczyło-Mysza, J. Jankowicz-Cieslak, J. Bocianowski, 2023, **QTL×QTL×QTL Interaction Effects for Total Phenolic Content of Wheat Mapping Population of CSDH Lines under Drought Stress by Weighted Multiple Linear Regression**, *Agriculture*, 13(4), 850, DOI:10.3390/agriculture13040850 mój indywidualny udział w jej powstaniu polegał na opracowaniu założeń i metod, walidacji, przeprowadzeniu badań, zarządzaniu zasobami, przygotowaniu danych, opracowaniu manuskryptu, odpowiedziach na uwagi recenzentów oraz pozyskiwaniu finansowania.

Data 29.06.2023r.

Podpis

Czyczyło-Mysza

dr hab. Joanna Jankowicz-Cieslak

PESEL: 76012303008

Plant Breeding and Genetics Laboratory, Joint FAO/IAEA Centre of Nuclear Techniques in Food and Agriculture, IAEA Laboratories Seibersdorf, International Atomic Energy Agency, Vienna International Centre

P.O. Box 100, A-1400 Vienna, Austria

Nr telefonu: +43 6648709441

E-mail: J.Jankowicz@iaea.org

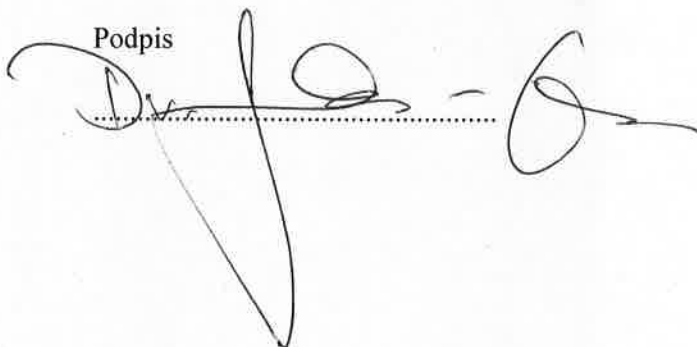
Oświadczenie o współautorstwie

Niniejszym oświadczam, że w pracy A. Cyplik, I. M. Czyczyło-Mysza, J. Jankowicz-Cieslak, J. Bocianowski, 2023, **QTL×QTL×QTL Interaction Effects for Total Phenolic Content of Wheat Mapping Population of CSDH Lines under Drought Stress by Weighted Multiple Linear Regression**, *Agriculture*, 13(4), 850, DOI:10.3390/agriculture13040850 mój indywidualny udział w jej powstaniu polegał na walidacji, opracowaniu manuskryptu, odpowiedziach na uwagi recenzentów.

Data

05.07.2023

Podpis



Prof. UPP dr hab. Jan Bocianowski

PESEL: 73101001793

Katedra Metod Matematycznych i Statystycznych, Uniwersytet Przyrodniczy w Poznaniu

ul. Wojska Polskiego 28, 60-637 Poznań

Nr telefonu: +48 61 8487143

E-mail: jan.bocianowski@up.poznan.pl

Oświadczenie o współautorstwie

Niniejszym oświadczam, że w pracy A. Cyplik, I. M. Czyczyło-Mysza, J. Jankowicz-Cieslak, J. Bocianowski, 2023, **QTL×QTL×QTL Interaction Effects for Total Phenolic Content of Wheat Mapping Population of CSDH Lines under Drought Stress by Weighted Multiple Linear Regression**, *Agriculture*, 13(4), 850, DOI:10.3390/agriculture13040850 mój indywidualny udział w jej powstaniu polegał na autorstwie koncepcji, opracowaniu założeń i metod, obsłudze oprogramowania, walidacji, analizie formalnej, przeprowadzeniu badań, zarządzaniu zasobami, opracowaniu manuskryptu, odpowiedziach na uwagi recenzentów, nadzorze i administracji projektu, pozyskiwaniu finansowania.

Data 30.03.2023

Podpis

Jan Bocianowski

Imię i nazwisko doktoranta: Adrian Cyplik

Nr PESEL: 94010706036

Afiliacja: Katedra Metod Matematycznych i Statystycznych, Uniwersytet Przyrodniczy w Poznaniu

Adres do korespondencji: ul. Wojska Polskiego 28, 60-637 Poznań

Nr telefonu: +48 603 843 080

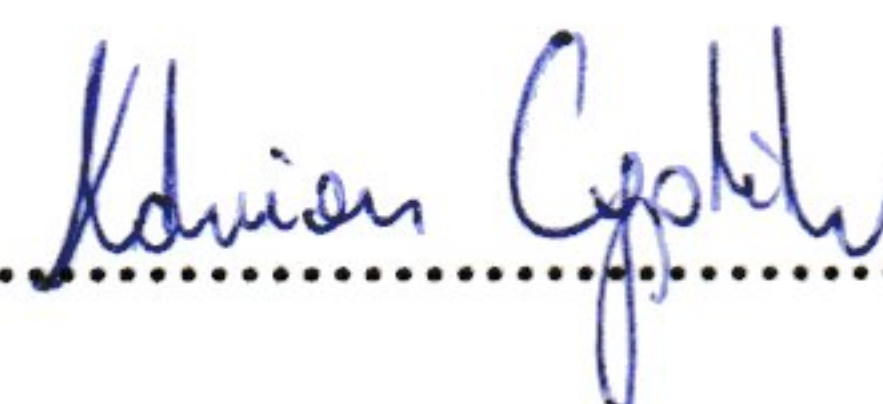
E-mail: adrian.cyplik@up.poznan.pl

Oświadczenie o współautorstwie

Niniejszym oświadczam, że w pracy A. Cyplik, J. Bocianowski, 2023, **A Comparison of Methods to Estimate Additive-by-Additive-by-Additive of QTL×QTL×QTL Interaction Effects by Monte Carlo Simulation Studies**, International Journal of Molecular Sciences, 24(12), 10043 mój indywidualny udział w jej powstaniu polegał na: przygotowanie projektu badań, opracowanie krytycznego przeglądu literatury, opracowanie założeń i metod, przeprowadzenie badań, w tym zbieranie danych, analiza wyników, interpretacja wyników i sformułowanie wniosków, obliczenia statystyczne, opracowanie manuskryptu.

Data30.08.2023.....

Podpis

.....

Prof. UPP dr hab. Jan Bocianowski

PESEL: 73101001793

Katedra Metod Matematycznych i Statystycznych, Uniwersytet Przyrodniczy w Poznaniu

ul. Wojska Polskiego 28, 60-637 Poznań

Nr telefonu: +48 61 8487143

E-mail: jan.bocianowski@up.poznan.pl

Oświadczenie o współautorstwie

Niniejszym oświadczam, że w pracy A. Cyplik, J. Bocianowski, 2023, **A Comparison of Methods to Estimate Additive-by-Additive-by-Additive of QTL×QTL×QTL Interaction Effects by Monte Carlo Simulation Studies**, *International Journal of Molecular Sciences*, 24(12), 10043, <https://doi.org/10.3390/ijms241210043> mój indywidualny udział w jej powstaniu polegał na: autorstwie koncepcji, przygotowaniu projektu badań, opracowaniu krytycznego przeglądu literatury, opracowaniu założeń i metod, przeprowadzeniu badań, w tym zbieraniu danych, analizie wyników, interpretacji wyników i sformułowaniu wniosków, obliczeniach statystycznych, opracowaniu grafik, opracowaniu manuskryptu, nadzorze i administracji projektu, pozyskiwaniu finansowania oraz odpowiedziach na uwagi recenzentów.

Data 30.09.2023

Podpis

Jan Bocianowski