International Journal of
Molecular Sciences

MDPI

Article

Multi-Omics Analysis of the Transcriptome and Metabolome as a Key to Understanding the Mechanisms Underlying Maize Yield

Maciej Lenort [1], Agnieszka Tomkowiak [1,*], Jan Bocianowski [2,*], Anna Przybylska-Balcerek [3], Karolina Jarzyniak [4], Przemysław Olejnik [4] and Przemysław Gawrysiak [1]

[1] Department of Genetics and Plant Breeding, Poznań University of Life Sciences, Dojazd 11, 60-632 Poznań, Poland
[2] Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Wojska Polskiego 28, 60-637 Poznań, Poland
[3] Department of Chemistry, Poznań University of Life Sciences, Wojska Polskiego 75, 60-625 Poznań, Poland
[4] Department of Biochemistry and Biotechnology, Poznań University of Life Sciences, Dojazd 11, 60-632 Poznań, Poland
* Correspondence: agnieszka.tomkowiak@up.poznan.pl (A.T.); jan.bocianowski@up.poznan.pl (J.B.)

Abstract

Maize (*Zea mays* L.) grain yield is a complex trait shaped by genetic, physiological, and environmental variation. We integrated a three-year field assessment of 20 genotypes with quantitative reverse-transcription PCR analysis of 33 candidate genes and profiling of phenolic acids, flavonoids, carotenoids, and oxidative enzyme activities to identify molecular and biochemical patterns associated with yield potential. Expression profiles were compared between yield groups and gene–biochemical relationships assessed by linear regression. Three-year mean yield ranged from 3.46 to 14.01 t ha^{-1}. Fourteen genes showed yield-associated expression patterns: five with high yield potential and nine with low yield potential. High-yielding SMH lines consistently had higher contents of all four phenolic acids, five flavonoids, and four carotenoids, and higher activities of polyphenol oxidase (PPO), catalase (CAT), and peroxidase (POD), than traditional cultivars. G32 expression was negatively associated with all 16 biochemical and enzymatic parameters, explaining 42.8–48.7% of their variance. Principal component analysis showed substantial transcriptional and metabolomic variation but no complete separation by yield potential. These findings identify coordinated transcriptional and biochemical patterns and prioritize G32 for functional validation; however, they demonstrate associations, not causality.

Keywords: *Zea mays*; grain yield; candidate genes; qRT-PCR; gene expression; gene–metabolite associations; phenolic acids; flavonoids; carotenoids; oxidoreductive enzymes

Check for updates

Academic Editor: Fabrizio Carbone

Received: 4 November 2026
Revised: 16 November 2026
Accepted: 18 November 2026
Published: 20 November 2026

Copyright: © 2026 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.

1. Introduction

Maize (*Zea mays* L.) is one of the most economically important crop species worldwide, serving as a source of food and feed and as a raw material for the food and energy industries. Contemporary breeding programs focus on improving resistance to biotic stresses, tolerance to water deficit and high temperatures, nutrient-use efficiency, grain quality, and the suitability of cultivars for cultivation at higher plant densities. Long-term progress in maize breeding has been associated, among other factors, with increased efficiency of solar radiation interception and utilization, as well as more favorable biomass allocation to reproductive organs [1,2]. However, sustaining this progress under increasing environmental variability requires a better understanding of the mechanisms underlying the stability of

agronomic traits [3]. Among the current priorities in maize improvement, increasing and stabilizing grain yield remains particularly important.

Grain yield is a complex quantitative trait determined by numerous phenotypic components and interactions among reproductive organ development, assimilate transport, and the dynamics of grain filling. Its final value is influenced by several factors, including the number of ears, the number of kernel rows and kernels per ear, individual kernel weight, and the duration and rate of reserve accumulation. The polygenic basis of these traits has been confirmed by linkage mapping, association studies, and QTL meta-analyses. Integrating QTL mapping with genome-wide association studies (GWAS) has enabled the identification of numerous genomic regions associated with maize yield-related traits and the functional validation of a candidate gene associated with kernel row number [4]. In turn, a QTL meta-analysis identified regions shared by several yield components, encompassing genes involved in reproductive development, hormone signaling, and sucrose and starch metabolism [5]. These findings confirm the complexity of the genetic architecture of grain yield; however, they do not yet allow the functional mechanisms underlying individual associations to be unequivocally determined.

The interpretation of association study results remains constrained by the structure of the populations analyzed, allele frequencies, phenotyping accuracy, the statistical model employed, and the effects of the environment and pleiotropy [6]. Identifying a marker located within or near a gene does not necessarily demonstrate that the gene is functionally involved in determining the trait of interest. Integrating transcriptomic data may reduce the number of candidate genes and identify genes whose activity is consistent with the observed phenotypic variation. Nevertheless, this approach does not establish whether changes in transcript abundance lead to modifications in specific metabolic pathways or whether such effects translate into differences in individual yield components. Therefore, analyses conducted on the same plant material are needed to simultaneously investigate genotype, gene expression, metabolic profiles, and phenotype.

Metabolites constitute a functional link between gene activity and the final phenotypic response of a plant. Previous studies of maize productivity have focused primarily on compounds involved in primary carbon and nitrogen metabolism. Under nitrogen deficiency, reductions in kernel number and kernel weight were accompanied by starch accumulation in leaves and sugar accumulation in developing kernels, suggesting a reduced capacity of sink organs to utilize available assimilates [7]. Relationships among nitrogen supply, sucrose and starch contents, the activity of enzymes involved in their metabolism, phytohormone levels, and grain-filling dynamics were also described by Yue et al. [8]. Furthermore, a multi-omics approach enabled the identification of ZmAVT1A-1 as a regulator of amino acid accumulation and nitrogen distribution between vegetative tissues and kernels, while also demonstrating the pleiotropic effects of this gene on agronomic traits [9]. Available findings confirm the importance of primary metabolites and hormonal regulators but provide limited insight into the potential contributions of secondary metabolites and the enzymatic antioxidant system to variation in grain yield.

Phenylpropanoids, flavonoids, and carotenoids are among the groups of compounds that require further characterization. Profiling polyphenolic metabolites in different kernel tissues and at successive stages of germination revealed pronounced tissue- and developmental stage-dependent patterns of their distribution [4]. Substantial variation in ferulic acid, p-coumaric acid, and diferulic acid contents was also observed among maize materials of different genetic origins [10]. Moreover, water deficit altered phenolic acid contents in leaves, accompanied by changes in the expression of genes involved in the phenylpropanoid pathway [11]. Carotenoids constitute another distinct group of compounds. An association analysis involving 244 inbred lines revealed numerous loci associated with lutein, zeax-

anthin, β-cryptoxanthin, and β-carotene contents and identified stable genomic regions suitable for further candidate gene identification [12]. These studies document the genetic and environmental variation in individual metabolite classes but do not establish whether their overall profiles are associated with the mechanisms determining grain yield.

The relationships between these groups of compounds and the activity of enzymes involved in regulating cellular redox status also remain insufficiently understood. Peroxidases, polyphenol oxidases, and catalase participate in phenolic compound transformations and the regulation of reactive oxygen species levels; however, their activities depend strongly on tissue type, developmental stage, and environmental conditions. To date, no comprehensive assessment has been conducted in which the contents of ferulic, p-coumaric, caffeic, and diferulic acids; luteolin and its derivatives; lutein; zeaxanthin; α-carotene; and β-carotene, together with POD, PPO, and CAT activities, were analyzed jointly with candidate gene expression and yield components in the same maize material. The absence of such an integrated approach limits the ability to determine whether the observed metabolic differences constitute part of a mechanism associated with productivity or merely represent an independent manifestation of variation among the genotypes investigated.

Establishing functional links between genetic variants and specific metabolic processes is also important for the practical application of research findings. In maize, CRISPR/Cas9-mediated editing of *CLE* gene promoters has enabled the generation of quantitative variation in traits associated with meristem development and grain yield, demonstrating that modifying gene expression levels may be more advantageous than completely disrupting gene function [13]. The ability to precisely modulate gene expression underpins the concept of cis-regulatory element editing, proposed as a strategy for generating novel breeding alleles while limiting undesirable pleiotropic effects [14]. The importance of this regulatory level is further supported by maps of cell type-specific cis-regulatory variation in maize [15]. However, the application of these tools requires the prior identification of genes for which the relationships among expression, metabolic changes, and phenotype have been experimentally established.

In view of the above, the aim of this study was to evaluate the relationships among the expression of candidate genes associated with maize grain yield, the profiles of selected phenolic acids, flavonoids, and carotenoids, the activities of peroxidases, polyphenol oxidases, and catalase, and variation in yield components. The study was designed to determine whether previously identified candidate genes could be linked to specific metabolic processes and, subsequently, to identify candidate genes requiring further functional validation and potentially useful for marker-assisted selection or precise genome editing.

2. Results

2.1. Field Experiment

Grain yield data (t·ha^{-1}) collected over the 2020–2022 period, with three replicates per year, were subjected to a two-way analysis of variance (ANOVA), with genotype (df = 19), year (df = 2), and the genotype × year interaction (df = 38) as the main factors. Tukey's honestly significant difference (HSD) post hoc test was used to compare differences between means at a significance level of $\alpha = 0.05$. The calculations accounted for standard deviations (SD).

The analysis revealed a highly significant effect of genotype on grain yield ($F_{19;120} = 1474.51$, $p < 0.001$), confirming substantial genetic variation among the tested entries, including SMH breeding lines and conventional cultivars. In contrast, the main effect of year was not statistically significant ($F_{2;120} = 2.37$, $p = 0.098$), indicating that mean grain yield, averaged across genotypes, did not differ significantly among years. The genotype × year interaction was statistically significant ($F_{38;120} = 3.14$, $p < 0.001$), demonstrating that

the tested genotypes responded differently to the environmental conditions prevailing in individual years.

The SMH breeding lines exhibited significantly higher yield potential than the conventional cultivars. The highest mean grain yield was recorded for line SMH 58024, at 14.01 t·ha^{-1}, followed by SMH 57924, which yielded 13.41 t·ha^{-1}. The low standard deviation values (SD < 0.60 t·ha^{-1}) observed for most of the tested entries indicate limited year-to-year variation within the three-year experiment. A statistically significant genotype × year interaction was also confirmed ($p < 0.001$) (Tables 1 and S1).

Table 1. Average grain yield (t·ha^{-1}) in individual years, multi-year average (Mean), and standard deviation (SD).

Variety/Line	2020 (t·ha^{-1})	2021 (t·ha^{-1})	2022 (t·ha^{-1})	Mean	SD
SMH 58024	14.37	13.47	14.20	14.01	0.53
SMH 57924	13.93	12.90	13.40	13.41	0.52
SMH 57624	13.00	12.43	12.27	12.57	0.55
SMH 58224	12.70	12.13	12.77	12.53	0.39
SMH 57324	11.87	12.57	12.87	12.43	0.60
SMH 57023	12.30	11.57	12.23	12.03	0.47
SMH 57524	12.10	11.77	12.20	12.02	0.35
SMH 57724	11.80	12.40	11.47	11.89	0.49
SMH 57824	11.70	11.30	11.53	11.51	0.33
SMH 58324	10.60	11.00	11.23	10.94	0.40
Wigor	6.37	6.33	6.40	6.37	0.41
Lancaster	5.23	5.87	5.43	5.51	0.37
Szyldecka	4.70	4.57	4.37	4.54	0.31
Małopolanka	4.17	4.73	4.67	4.52	0.35
Wielkopolanka	3.97	3.87	3.80	3.88	0.17
Złota Górecka	3.97	3.83	3.73	3.84	0.19
Bydgoska wczesna	4.20	3.47	3.83	3.83	0.37
Dar Północy	3.23	3.73	3.83	3.60	0.38
Wawrzyńczycka	3.37	3.70	3.33	3.47	0.32
Muszyński	3.83	3.27	3.27	3.46	0.36

2.2. Analysis of the Transcriptional Activity of Candidate Genes

To investigate the relationships between the transcription levels of the 33 identified candidate genes and the yield potential of *Zea mays* L., the relative expression profiles of these genes were analyzed using qRT-PCR across a panel of 20 extreme maize genotypes.

Based on the observed relative transcript abundance profiles across the 20 maize genotypes, the candidate genes were descriptively categorized into four expression-pattern groups. This classification was intended to facilitate the presentation and comparison of transcriptional profiles and did not represent a statistically validated grouping or evidence of direct functional relationships with grain yield. The categories reflected predominant expression tendencies, the absence of a consistent yield-group-associated pattern, or pronounced genotype-specific variation. The functional-transcriptional groups were established based on differences in gene expression among maize genotypes, as determined by one-way ANOVA followed by Tukey's honestly significant difference (HSD) post hoc test ($p < 0.05$).

2.2.1. Group I—Genes Showing Higher or Distinct Expression Patterns Predominantly in Modern SMH Breeding Hybrids: [G2 (*LOC100285609*)—RNA Polymerase II Transcriptional Coactivator *KELP*; G5 (*LOC103635969*)—Putative Protein Phosphatase 2C; G6 (*LOC100276990*)—Uncharacterized Protein; G10 (*LOC100542095*)—Isoamylase-Type Starch Debranching Enzyme 3 Isoform; and G13 (*LOC103643269*)—Probable Arabinosyltransferase *ARAD1*] (Figure 1)

Analysis of the expression profiles of the five candidate genes (G2, G5, G6, G10, and G13) across the complete panel of 20 maize genotypes revealed clear transcriptional divergence between the low-yielding cultivars (entries 1–10, from Szyldecka to Wielkopolanka) and the modern, high-yielding hybrids (entries 11–20, from SMH 57023 to SMH 58324). Although most of the genes examined exhibited low basal transcript abundance across the population as a whole, a coordinated pattern of transcriptional activation was observed in the high-yielding group, with deviations attributable to genotype-specific increases in expression in selected cultivars. One particularly noteworthy pattern was the pronounced upregulation of G5, G10, and G13 in the high-yielding genotype SMH 58024. This pattern may reflect regulation by a shared upstream transcriptional regulator or another trans-acting factor in this line. The genotype combines high expression of a starch biosynthesis-related enzyme (G10), a signaling regulator (G5), and a cell wall biosynthesis-related enzyme (G13), making it a candidate for further investigation of sink-related and metabolic processes (Figure 1).

Furthermore, a distinct transcriptional shift in G5 and G10 expression was observed between entry 10 (Wielkopolanka) and entry 11 (SMH 57023). In the first 10 low-yielding cultivars, expression remained close to zero, whereas the transition to SMH 57023 marked an abrupt shift toward higher transcript abundance. The activation of G5 and G10 may therefore represent a characteristic feature of the modern breeding program from which the SMH genotypes originated. These results suggest that the SilicoDArT 28630 and SNP 28634 markers identified through association mapping are consistently associated with the high-yielding genotypes.

In addition, Szyldecka was the only genotype in the low-yielding group that exhibited very high expression levels of G5, G10, and G13. This observation is consistent with the polygenic nature of grain yield, as the transcriptional potential of these loci in Szyldecka may be counteracted by other genetic factors, such as lower photosynthetic capacity or greater sensitivity to environmental stresses. Gene G2 maintained low expression levels in 19 of the 20 genotypes, with a single distinct expression peak observed in the cultivar Wigor. This pattern may reflect genotype-specific regulation. However, it limits the suitability of G2 as a universal selection marker while making it an interesting candidate for investigating cultivar-specific regulatory mechanisms (Figure 1).

The most promising candidate indicator genes potentially associated with increased grain yield were G5 (*LOC103635969*)—putative protein phosphatase 2C; G10 (*LOC100542095*)—isoamylase-type starch debranching enzyme 3 isoform; and G13 (*LOC103643269*)—probable arabinosyltransferase *ARAD1*.

Protein phosphatases of the PP2C family are key cellular regulators involved, among other processes, in environmental stress responses and plant hormone signaling. Higher expression of this gene may help plants maintain metabolic efficiency and delay senescence of the photosynthetic apparatus, potentially extending photosynthetic activity and contributing to increased grain yield.

Isoamylases (ISO3) are enzymes directly involved in starch biosynthesis and the proper formation of starch structures in the developing kernel endosperm. Elevated expression of this gene may enhance sink strength, defined as the capacity of developing kernels

to efficiently accumulate sugars and convert them into starch, which is a fundamental biological determinant of kernel filling and final kernel weight.

Arabinosyltransferases participate in cell wall biosynthesis and modification, including processes associated with pectin biosynthesis. Their activity in plant tissues may promote dynamic cell growth and improved development of vascular tissues, thereby facilitating the transport of nutrients to the ears (Figure 1).

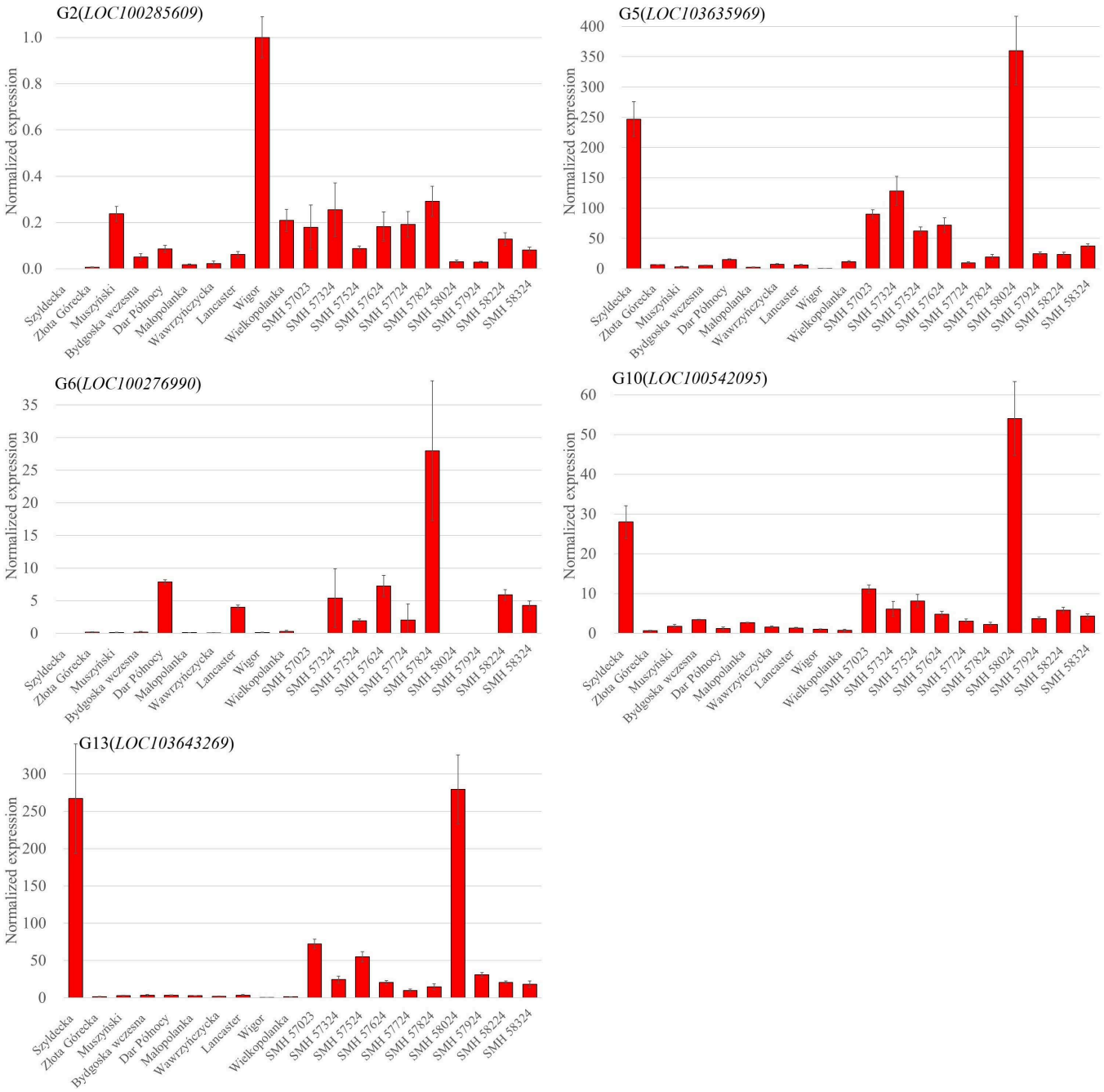


Figure 1. Graphs showing the transcriptional activity of candidate genes associated with maize yield.

2.2.2. Group II—Genes Showing Higher Expression Patterns Predominantly in Traditional Open-Pollinated Cultivars: [G1 (*LOC100282883*)—Uncharacterized Protein; G3 (*LOC100282625*)—Aspartate Aminotransferase; G15 (*LOC100502264*)—Uncharacterized Protein; G26 (*LOC100282937*)—RAF-Type Zinc Finger Domain-Containing Protein 1; G27 (*LOC100279888*)—Uncharacterized Protein; G29 (*LOC103637325*)—Uncharacterized Protein; G30 (*LOC103636090*)—Sphingoid Long-Chain Bases Kinase 1; G31 (*LOC100191174*)—Uncharacterized Protein; and G32 (*LOC103636139*)—Uncharacterized Protein] (Figure 2)

Analysis of the transcription levels of the genes examined by RT-qPCR across a panel of 20 extreme genotypes (10 low-yielding and 10 high-yielding) revealed diverse expression patterns, with several loci showing higher transcript abundance predominantly in the low-yielding group. The predominant trend observed was elevated expression in low-yielding genotypes (Figure 2). This pattern was particularly pronounced for G29, G30, and G32, whose expression was almost completely suppressed in most high-yielding lines. The suppression of G30, which encodes sphingoid long-chain bases kinase 1 (*LCBK1*), is of particular biological interest. Sphingolipids are key signaling molecules involved in the induction of abiotic stress responses and programmed cell death (PCD). Higher G30 expression in low-yielding cultivars may be associated with differences in stress-related signaling; however, the present data do not establish a trade-off between defense and biomass accumulation. A similar expression pattern, potentially indicative of a negative regulatory role, was observed for G26, which encodes a RAF-type zinc finger domain-containing protein. Given the regulatory functions of members of this protein family, reduced expression in high-yielding lines may be associated with broader differences in growth-related regulation (Figure 2).

In contrast, genes involved in fundamental metabolic pathways, such as G3, which encodes aspartate aminotransferase, and G31, exhibited expression patterns dependent on the individual genetic background, without a clear distinction between yield groups. Isolated expression peaks in high-yielding lines, such as those observed for G26 and G27 in SMH 57924, suggest the possible involvement of genotype-specific compensatory epistatic mechanisms. Meanwhile, G1 and G15, which encode uncharacterized proteins, may reflect polymorphism in regulatory regions, with their expression patterns more closely resembling those of genes exhibiting higher activity in low-yielding genotypes (Figure 2).

The absence of overlap between the confidence intervals of the two extreme genotype groups is consistent with differences in transcript levels for selected genes. These findings suggest that the identified SNP/SilicoDArT polymorphisms may be associated with cis-regulatory variation affecting transcription. The candidate genes analyzed, particularly G26 (LOC100282937), encoding a RAF-type zinc finger domain-containing protein, and the uncharacterized proteins G29 (LOC103637325) and G32 (LOC103636139), represent interesting candidates for further functional investigation. Their expression patterns suggest potential relationships with grain yield-related traits and biochemical variation. However, because the molecular functions of G29 and G32 have not yet been established, their suitability as targets for marker-assisted selection or genome editing remains to be determined. Further functional characterization is required before considering targeted modifications of their expression (Figure 2).

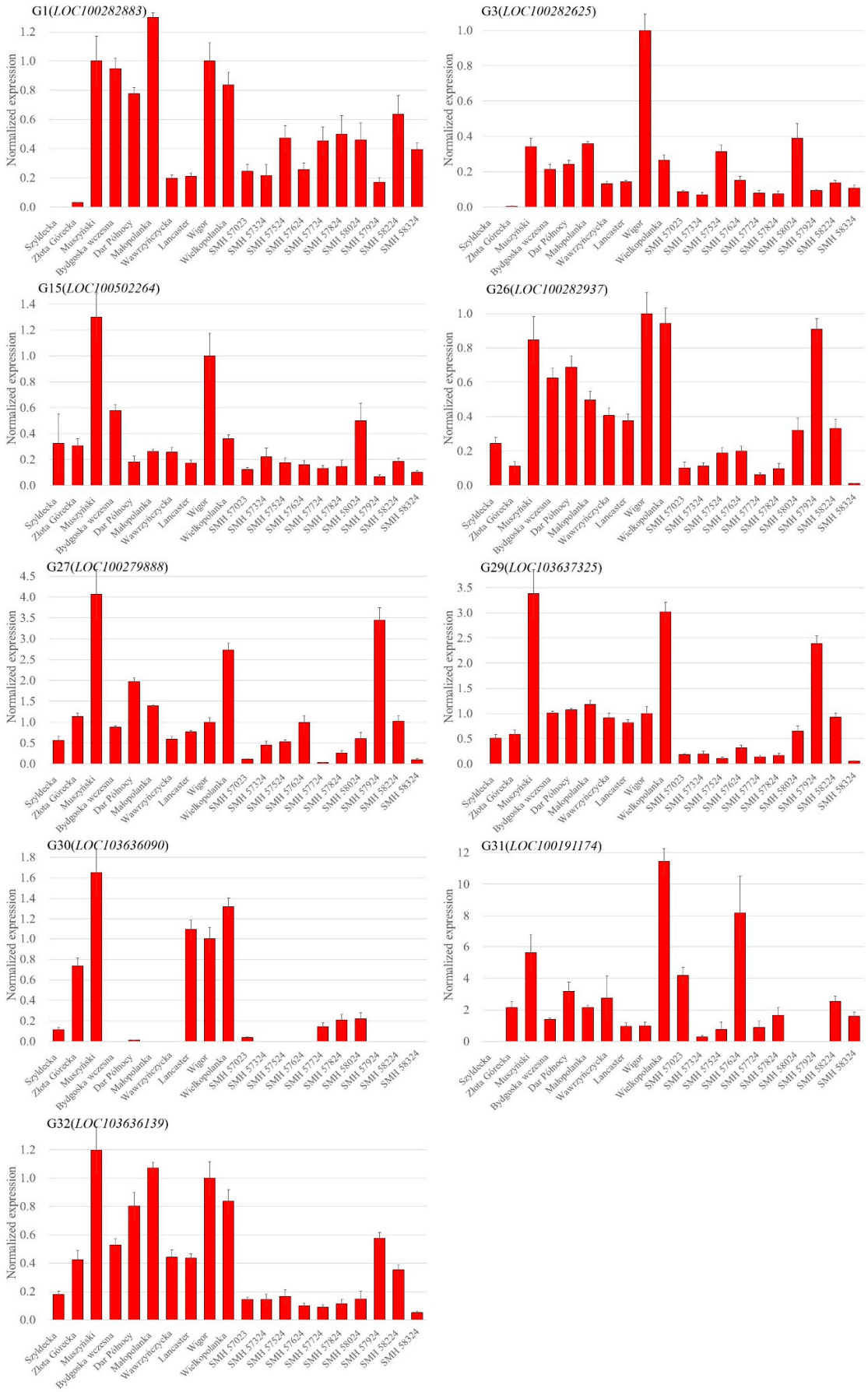


Figure 2. Graphs showing the transcriptional activity of candidate genes associated with maize yield.

2.2.3. Group III—Genes Exhibiting Stable, Constitutive Expression with Similar Transcript Levels Irrespective of the Yield Potential of the Genotype Examined (No Substantial Differences Between the Low- and High-Yielding Groups): [G4 (*LOC100541615*)—Sucrose Transporter 1; G7 (*LOC100276990*)—Uncharacterized Protein; G8 (*LOC103633409*)—Uncharacterized ATP-Dependent Helicase; G9 (*LOC100384172*)—Peroxisomal Ion Protease Homolog 2; G11 (*LOC100194289*)—Ribosomal Protein L32-like; G12 (*LOC103655564*)—Uncharacterized Protein; G14 (*LOC103638070*)—Probable Sugar-Phosphate/Phosphate Translocator *AT1G06470*; G16 (*LOC100382335*)—Uncharacterized Protein; G19 (*LOC103625798*)—Hydroxyproline O-Galactosyltransferase *GALT*; G20 (*LOC100383550*)—Uncharacterized Protein; G22 (*LOC103639124*)—Alpha-N-Acetylglucosaminidase; G23 (*LOC100383602*)—Putative Calcium-Dependent Protein Kinase Family Protein; G24 (*LOC100383055*)—Uncharacterized Protein; G25 (*LOC100383055*)—Uncharacterized Protein; G28 (*LOC103655222*)—Glycerol-3-Phosphate 2-O-Acyltransferase 6; and G33 (*LOC100383455*)—U-Box Domain-Containing Protein 7] (Figures 3 and 4)

Comparison of mean transcript expression levels between the low- and high-yielding genotype groups revealed no consistent, universal trend toward either increased or decreased expression across the entire set of candidate genes analyzed (Figures 3 and 4). Most of the loci examined, including G8, G11, G19, G23, G24, and G25, exhibited high within-group variability, indicating that their expression was strongly dependent on the specific genetic background (genotype effect) rather than on grain yield phenotype as a complex trait per se.

Nevertheless, specific patterns of transcriptional activation, characterized by expression peaks in selected high-yielding genotypes, were observed for some genes. G7, G14, G16, and G20 exhibited highly genotype-specific and pronounced overexpression in the high-yielding line SMH 58024 (Figures 3 and 4). In the remaining genotypes, the expression of these genes remained at low or moderate levels.

For selected genes, the highest levels of transcriptional activation were observed in low-yielding lines. G28 was overexpressed in the cultivar Szyldecka, whereas transcript abundance remained relatively low in most of the other genotypes, with the exception of SMH 58024. A similar pattern was observed for G9, whose expression in the cultivar Muszyński reached the highest value among all genotype–gene combinations analyzed (Figures 3 and 4).

Among the genes involved in carbohydrate metabolism, G4, which encodes sucrose transporter 1, exhibited the highest expression levels in the low-yielding cultivar Małopolanka and the high-yielding line SMH 58224, while transcript abundance remained low in the other genotypes examined.

The presence of statistically significant SNP and SilicoDArT markers mapped within the genes investigated supports their genomic association with grain yield. The absence of a simple, direct correlation between transcriptional activation levels (mRNA abundance) and grain yield phenotype in most genotypes suggests that the effects of the identified allelic variants may operate through non-transcriptional mechanisms. For example, structural mutations may be involved: SNPs located in coding regions may cause amino acid substitutions that alter protein structure, enzymatic activity, substrate affinity, or conformational stability without affecting the transcriptional output of the gene itself. Alternatively, these associations may reflect linkage disequilibrium (LD), whereby the mapped markers are in close genetic linkage with the causal variant underlying the quantitative trait locus (QTL), located in adjacent regulatory or genomic regions.

Furthermore, analysis of the standard deviation values presented in the plots indicates pronounced differences among selected genotypes; however, overlap or non-overlap of

error bars alone does not establish statistical significance. However, at the group level, the difference in mean transcript expression between the low-yielding and high-yielding populations does not reach statistical significance as a universal phenotypic indicator. Overall, the transcriptional activation profiles are highly individual and genotype-specific (Figures 3 and 4).

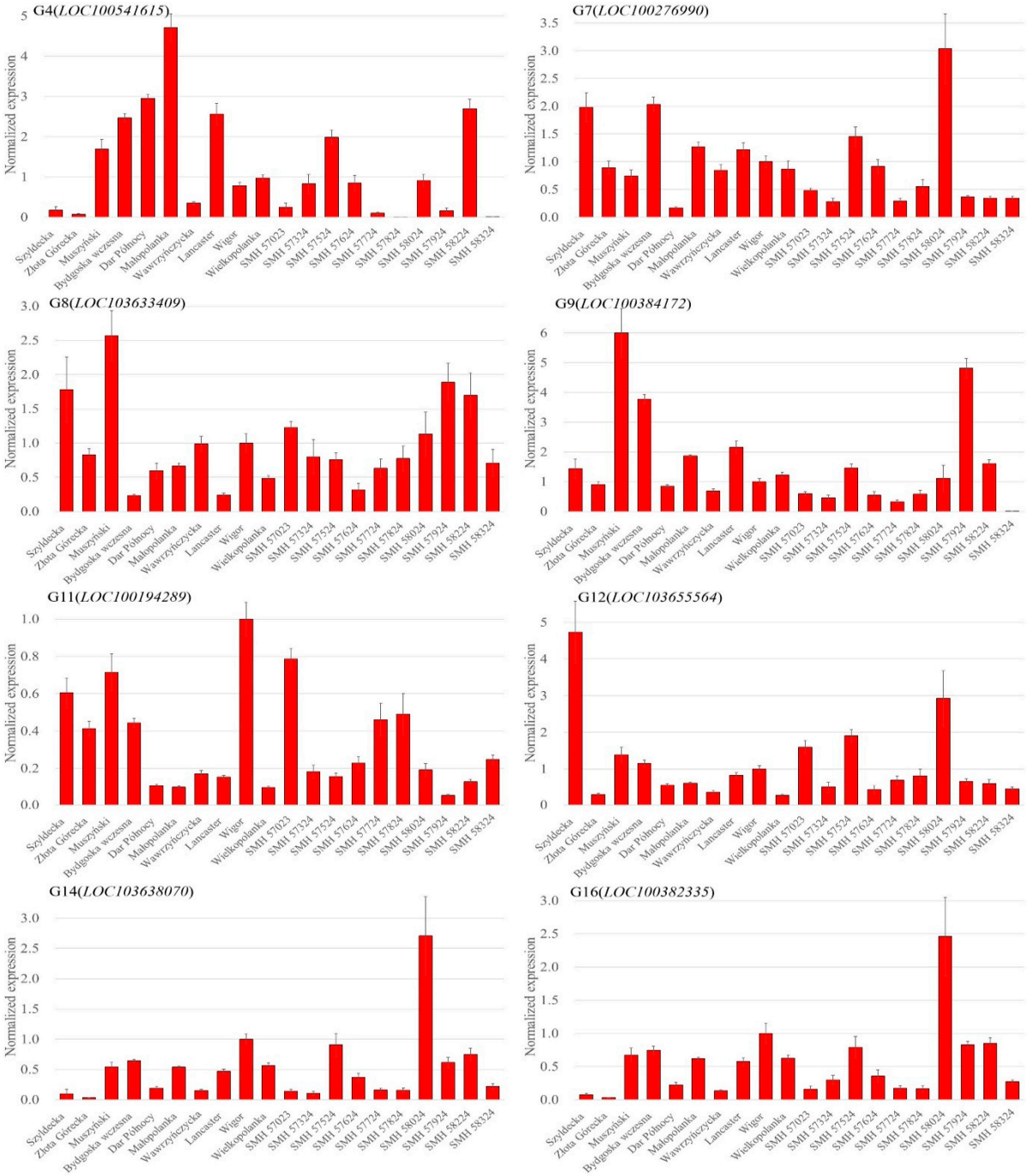


Figure 3. Graphs showing the transcriptional activity of candidate genes associated with maize yield.

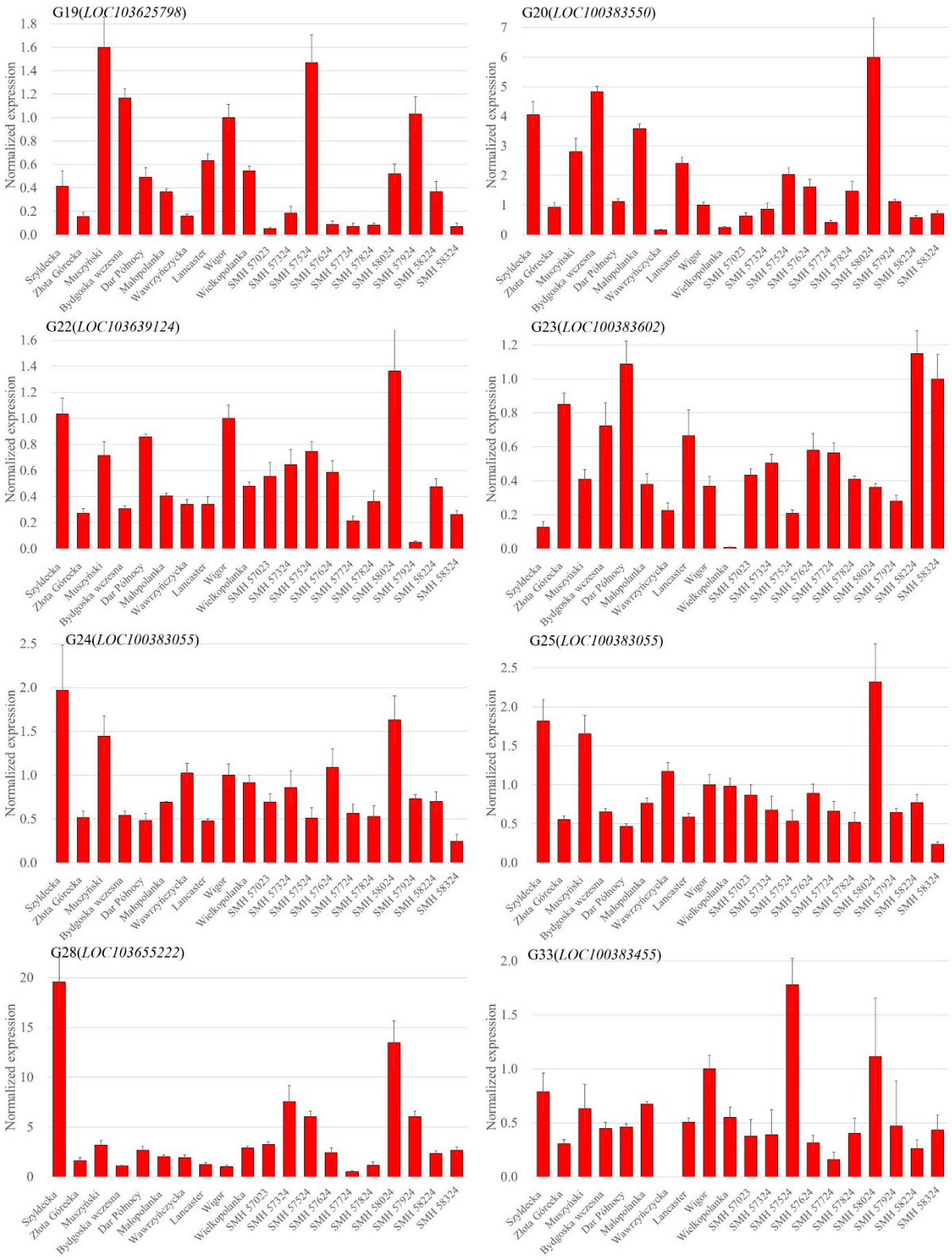


Figure 4. Graphs showing the transcriptional activity of candidate genes associated with maize yield.

2.2.4. Group IV—Genes Exhibiting a Predominantly Genotype-Specific Expression Pattern. Loci Characterized by Generally Low Transcript Abundance in Most of the Genotypes Examined but Displaying Pronounced, Isolated Expression Peaks in Individual Genotypes: [G17 (*LOC103654750*)—Ubiquitin Carboxyl-Terminal Hydrolase 15 Isoform X2; G18 (*LOC100193686*)—Uncharacterized Protein; and G21 (*LOC103650272*)—Uncharacterized Protein] (Figure 5)

The expression profile of G17 revealed no clear difference between the low-yielding and high-yielding genotype groups (Figure 5). Instead, highly genotype-specific transcript accumulation was observed. A markedly elevated expression level was recorded in the low-yielding cultivar Muszyński, whereas transcript abundance remained at basal or near-zero levels in most of the genotypes examined, including Małopolanka, Wawrzyńczycka, and lines SMH 57023–SMH 57824. Only a slight increase in transcript abundance was detected in the cultivars Szyldecka and SMH 57924.

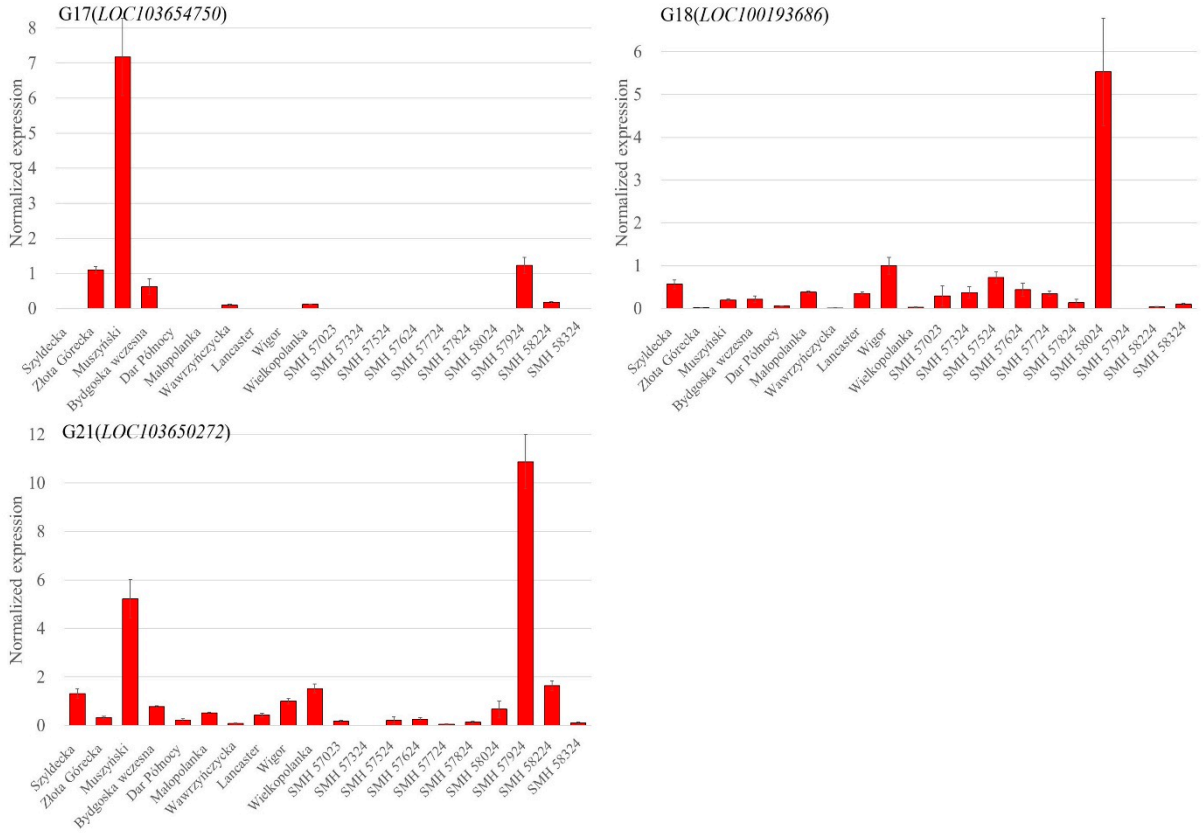


Figure 5. Graphs showing the transcriptional activity of candidate genes associated with maize yield.

Analysis of transcript levels at locus *LOC100193686* (G18) revealed generally low and relatively uniform expression across most of the genotypes analyzed (Figure 5). The exception was the high-yielding line SMH 58024, in which a distinct peak of transcript accumulation was observed. G21 expression exhibited substantial variation among genotypes, without a clear trend associated with yield potential. The highest transcript accumulation was recorded in the high-yielding genotype SMH 57924, while a second pronounced expression peak was observed in the low-yielding cultivar Muszyński. The remaining genotypes exhibited low to moderate transcript levels (Figure 5).

Although SNP and SilicoDArT markers mapped within loci *LOC103654750* (G17), *LOC100193686* (G18), and *LOC103650272* (G21) showed statistically significant associations with high grain yield at the DNA level, the transcriptomic profiles of the 20 extreme

genotypes did not confirm statistically significant differences between the low-yielding and high-yielding groups as a whole. The genetic associations identified by GWAS may therefore reflect structural or coding mutations, such as nonsynonymous SNPs that alter protein stability or enzymatic activity, rather than changes in promoter strength or mRNA expression levels.

In the case of G17, which encodes ubiquitin carboxyl-terminal hydrolase 15, cellular regulation may occur predominantly at the post-translational level through deubiquitination processes rather than through direct upregulation of mRNA expression. The pronounced expression peaks observed in individual lines, such as Muszyński and SMH 57924, suggest the involvement of unique trans-regulatory networks or genotype-specific responses to environmental conditions, rather than a universal transcriptional mechanism determining grain yield.

2.2.5. Principal Component Analysis (PCA) of Candidate Gene Expression Profiles

To characterize the variation in the expression profiles of candidate genes associated with the yield potential of maize (*Zea mays* L.), principal component analysis (PCA) was performed on 20 genotypes, comprising 10 low-yielding and 10 high-yielding genotypes. The analysis enabled the assessment of similarities among multivariate transcriptional profiles and the identification of genotypes exhibiting the most distinct expression patterns. The first two principal components jointly explained 99.50% of the total data variance, with PC_1 accounting for 97.04% and PC_2 for 2.46% (Figure 6). This indicates that nearly all the variation captured by the analysis was represented in the two-dimensional PCA space, with a clear predominance of the first principal component.

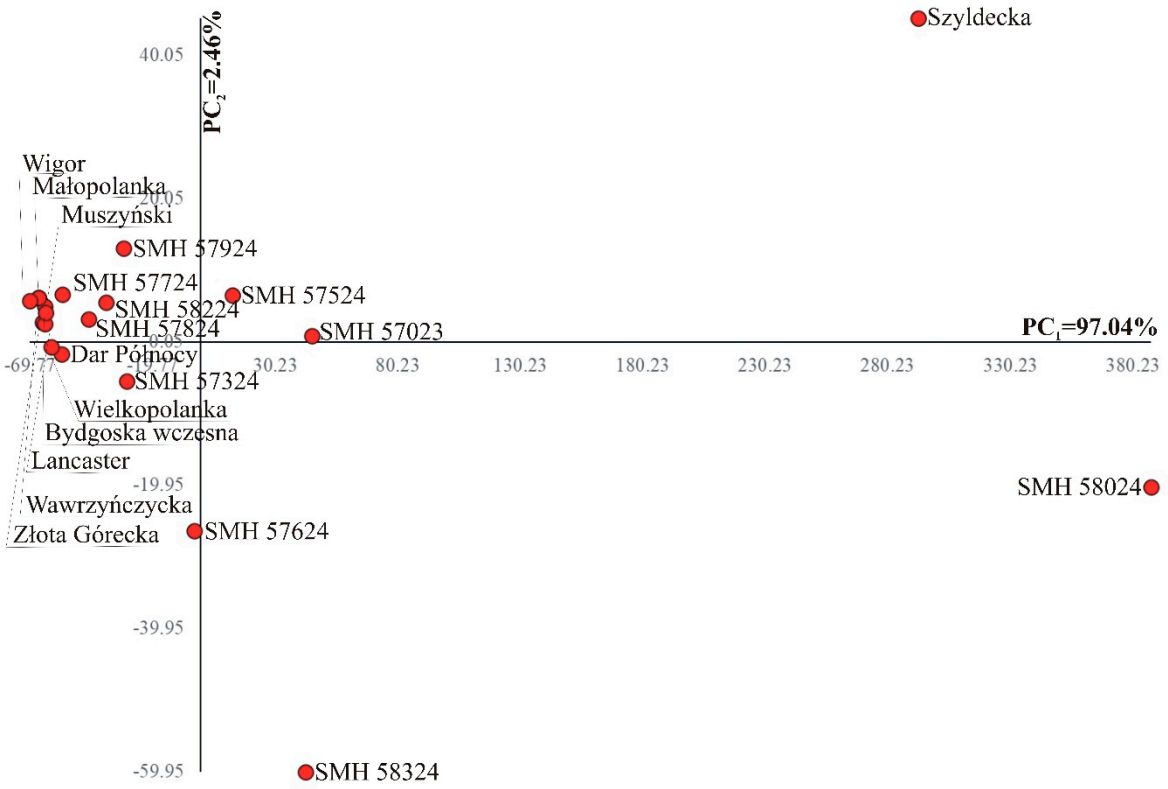


Figure 6. Distribution of genotypes in the space of the first two principal components (PCA) calculated based on the expression profiles of candidate genes.

The distribution of genotypes in the PCA plot indicated substantial variation in expression profiles; however, no clear separation of all low-yielding and high-yielding geno-

types was observed. Most traditional cultivars, including Wigor, Małopolanka, Muszyński, Wielkopolanka, Bydgoska Wczesna, Lancaster, Wawrzyńczycka, and Złota Górecka, clustered near the origin of the coordinate system. The similar positioning of some SMH lines suggests that their expression profiles were relatively similar to those observed in traditional cultivars. At the same time, the distribution of the remaining SMH lines indicates substantial transcriptional heterogeneity within the high-yielding group (Figure 6).

The most pronounced deviations along the first principal component were observed for the cultivar Szyldecka and the line SMH 58024. Both genotypes were among the most distant from the main cluster but occupied opposite positions along the second principal component. Szyldecka exhibited a high positive PC_2 score, whereas SMH 58024 had a negative score. This distribution indicates that, despite their substantial divergence along PC_1, the expression profiles of these two genotypes also differed along the dimension of variation represented by PC_2. This finding is particularly noteworthy in light of the RT-qPCR results, which revealed elevated expression of G5, G10, and G13 in both Szyldecka and SMH 58024. In SMH 58024, elevated expression of G7, G14, G16, and G20, together with a marked increase in G18 expression, was also observed. The combination of a distinct PCA position and the activation of several genes indicates that SMH 58024 is characterized by a specific transcriptional profile. In contrast, the different position of Szyldecka along PC_2 suggests that similarities in the expression of selected loci do not necessarily imply similarity in the overall transcriptional profile (Figure 6).

Of particular interest is the fact that Szyldecka belongs to the low-yielding group despite exhibiting elevated expression of several genes classified as showing a tendency toward an association with high grain yield. This observation indicates that the activation of individual candidate genes is insufficient to explain differences in yield potential. In SMH 58024, elevated expression of a larger number of genes coincided with its classification as a high-yielding genotype; however, the results presented do not allow a definitive conclusion as to whether these changes directly contribute to its higher yield. The differences between these two genotypes may result from distinct configurations of expression across the remaining loci, interactions among these loci, or other regulatory mechanisms whose contributions require further investigation (Figure 6). The contrasting phenotypic performance of Szyldecka and SMH 58024, despite their similarly elevated expression of G5, G10, and G13, raises an important question regarding the biological significance of transcriptional activation in genotypes with different yield potential. One possible explanation is that increased transcript abundance does not necessarily translate into a corresponding increase in protein abundance or enzymatic activity. Post-transcriptional and post-translational regulatory mechanisms, including differences in protein stability, phosphorylation status, and protein turnover, may modify the functional consequences of gene expression. This possibility is particularly relevant for G5, which encodes a putative protein phosphatase 2C (PP2C), a protein family involved in the regulation of diverse cellular signaling pathways. Similarly, the biological effects of increased G10 expression, encoding an isoamylase-type starch debranching enzyme, may depend on the metabolic context and the availability of substrates, whereas the expression of G13, encoding a probable arabinosyltransferase ARAD1, may reflect changes in cell wall metabolism that do not necessarily translate into improved grain productivity. Another possible explanation is that elevated expression of these genes in Szyldecka represents a compensatory or stress-responsive mechanism rather than a process directly promoting yield formation. Regulatory feedback mechanisms may induce transcriptional activation in response to physiological or metabolic limitations, without necessarily restoring the efficiency of the affected processes. Furthermore, differences in carbon assimilation, assimilate partitioning, and source–sink relationships may influence the extent to which molecular and metabolic activities in vegetative tissues are

translated into reproductive productivity. Consequently, even pronounced activation of selected candidate genes may coexist with limited grain yield if other physiological processes constrain assimilate production, transport, or utilization during kernel development. These mechanisms remain hypothetical in the present study, as protein abundance, enzyme activity specifically associated with the candidate genes, photosynthetic performance, and assimilate transport were not directly investigated. The distinct position of Szyldecka in the PCA space further supports the interpretation that its transcriptional profile differs substantially from those of the other genotypes. Importantly, PCA reflects the combined variation across all 33 candidate genes, rather than the expression of individual loci or grain yield itself. Therefore, the extreme coordinate displacement of Szyldecka, despite its elevated expression of G5, G10, and G13, may reflect a distinct overall transcriptional configuration involving additional genes and regulatory interactions. This observation emphasizes that the relationship between gene expression and grain yield depends not only on the activity of individual candidate genes but also on their interactions within the broader genetic and physiological background.

The line SMH 58324 was another clearly distinct genotype, exhibiting an extremely negative PC_2 score while showing a substantially smaller deviation along PC_1 than Szyldecka and SMH 58024. This position indicates that its expression profile differed primarily along the dimension of variation represented by the second principal component. Since the RT-qPCR results presented do not identify a characteristic gene activation pattern for SMH 58324, its PCA position cannot currently be attributed to specific loci (Figure 6).

Variation in PCA positioning was also observed among the remaining SMH lines. The genotypes SMH 57524 and SMH 57023 were located on the positive side of PC_1, whereas SMH 57624 exhibited a distinctly negative PC_2 score. SMH 57924 was positioned closer to the main cluster, despite showing elevated expression of several genes in the RT-qPCR analysis, including G21, G26, and G27. In contrast, SMH 57324 was located near the main cluster, with a slight deviation toward negative PC_2 values. Taken together, these observations indicate that the specific expression of individual genes does not necessarily result in an equally pronounced separation of a genotype in multivariate analysis (Figure 6).

In summary, PCA revealed substantial variation in candidate gene expression profiles across the maize panel, but no clear grouping of genotypes according to their yield potential was observed. These findings indicate that the relationships between gene expression and grain yield are complex and cannot be explained solely by the activity of individual loci (Figure 6).

2.3. Metabolomic Analysis and Metabolite Profiling

This study investigated compounds representing several important aspects of plant metabolism: phenolic acids, flavonoids, carotenoids, and the activities of selected oxidoreductive enzymes. The selection of these parameters enabled the characterization of genotypic variation in phenolic compound metabolism, antioxidant defense mechanisms, and the functioning of systems associated with photosynthetic pigments. These processes may indirectly affect grain yield and the maintenance of plant physiological performance, their capacity to adapt to changing environmental conditions, and the efficiency with which resources are utilized during growth and development.

2.3.1. Variation in Phenolic Acid Contents Among Maize Genotypes with Contrasting Yield Potential

Analysis of caffeic, diferulic, ferulic, and p-coumaric acid contents revealed pronounced differences in phenolic acid profiles between traditional low-yielding maize cultivars and modern, high-yielding SMH breeding lines. In all four cases, consistently higher contents of

the compounds analyzed were observed in the SMH group. Among the traditional cultivars, Muszyński exhibited relatively low values, whereas among the high-yielding genotypes, the highest values for all four parameters were recorded in the line SMH 58324. A similar pattern of variation was observed in the other SMH lines, although quantitative differences in metabolite accumulation occurred among individual genotypes (Figure 7).

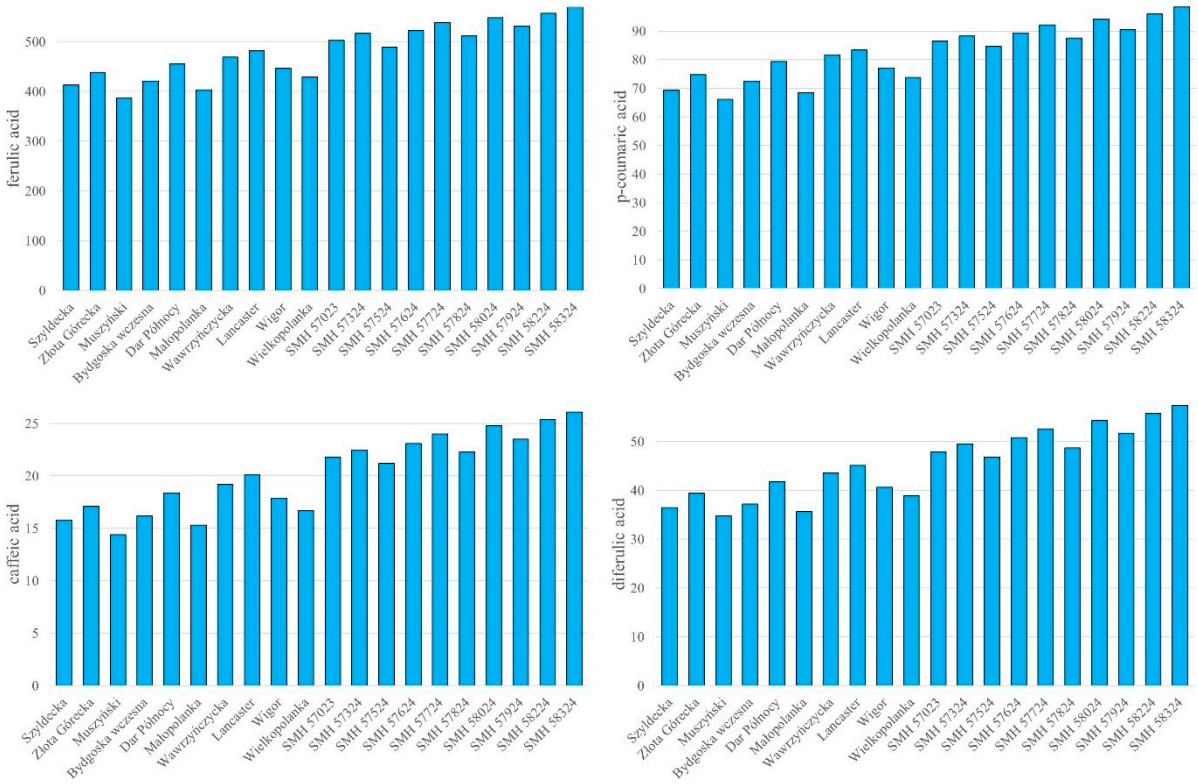


Figure 7. Graphs illustrating changes in phenolic acid contents ($\mu g \cdot g^{-1}$) in low- and high-yielding maize genotypes.

This pattern is particularly noteworthy in the context of the field experiment results, in which modern SMH lines exhibited substantially higher grain yields than traditional cultivars. Mean grain yields among the SMH lines ranged from approximately 10.94 to 14.01 $t \cdot ha^{-1}$, whereas those of the traditional cultivars ranged from approximately 3.46 to 6.37 $t \cdot ha^{-1}$. Thus, the comparison of grain yield and phenolic acid profiles indicates a co-occurrence of high yield potential and elevated contents of the compounds analyzed across the genotype panel. The consistency of this pattern across four different phenolic acids suggests that the observed differences may reflect broader variation in phenylpropanoid metabolism rather than changes in a single metabolite alone (Figure 7).

From a physiological perspective, phenolic acids participate in processes associated with protection against environmental stressors and the metabolism of cell wall components. Ferulic and diferulic acids are of particular importance because they can contribute to cross-linking cell wall constituents, thereby influencing cell wall structural properties. Caffeic and p-coumaric acids, in turn, belong to the phenylpropanoid pathway, which encompasses numerous compounds with structural and defensive functions.

An important aspect of the results is the variation observed within the high-yielding group. Although all SMH lines exhibited higher contents of the phenolic acids analyzed than the traditional cultivars, the individual lines differed in the levels of compound accumulation. The line SMH 58324, which exhibited the highest values for all four parameters, also represented a genotype with high yield potential. This observation suggests a possible

relationship between metabolite profiles and genetic variation among the lines; however, it does not establish that elevated phenolic acid contents directly cause higher grain yield (Figure 7).

In the context of the transcriptomic analysis, these findings provide a rationale for investigating relationships between candidate gene expression and phenolic acid contents. Integrating grain yield, gene expression, and metabolite profile data may help determine whether differences in the activity of selected genes coincide with changes in phenylpropanoid metabolism and whether these relationships are consistent across genotypes with contrasting yield potential. Furthermore, the higher contents of caffeic, diferulic, ferulic, and p-coumaric acids observed in the SMH hybrids represent a consistent pattern of biochemical differences between the two genotype groups. However, because the groups also differ in breeding history and genetic background, these differences cannot be attributed specifically to grain yield potential. They may reflect broader genetic and metabolic divergence associated with historical breeding and selection. These findings indicate that phenolic acid profiles may represent a valuable component of the multidimensional characterization of genotypes differing in grain yield. At this stage, however, this relationship should be interpreted as the co-occurrence of a specific metabolic profile with high yield potential rather than as evidence of a direct effect of the compounds analyzed on grain yield (Figure 7).

2.3.2. Variation in Flavonoid Contents Among Maize Genotypes with Contrasting Yield Potential

Analysis of flavonoid contents revealed pronounced differences between traditional maize cultivars with lower grain yields and modern, high-yielding SMH breeding lines. Across all five compounds analyzed, namely luteolin, luteolin-7-O-glucoside, luteolin-7-O-rutinoside, luteolin-7-O-glucuronide, and luteolin-7-O-(6″-malonyl)-glucoside, an overall pattern of higher contents was observed in the SMH group. These differences were particularly evident for luteolin derivatives, whose contents were generally higher in the SMH lines than in the traditional cultivars. At the same time, variation within both groups indicated genotype-dependent differences in the accumulation of individual flavonoids (Figure 8).

The most consistent pattern was observed for luteolin-7-O-rutinoside, luteolin-7-O-glucuronide, and luteolin-7-O-(6″-malonyl)-glucoside. For these compounds, all SMH lines exhibited higher values than the traditional cultivars, with the highest levels recorded in SMH 58324. A similar trend was observed for luteolin and luteolin-7-O-glucoside. Among the traditional cultivars, Muszyński was characterized by relatively low contents of the compounds analyzed, whereas the remaining cultivars exhibited varying intermediate values. The consistency of this pattern across several derivatives of the same flavonoid suggests that the observed differences may reflect broader variation in luteolin metabolism and its biotransformation rather than changes in a single compound alone (Figure 8).

These findings are particularly relevant in the context of maize grain yield, as the SMH lines also exhibited higher yields than the traditional cultivars. Previous field experiment results indicated grain yields ranging from approximately 10.94 to 14.01 t·ha^{-1} for the SMH lines, compared with approximately 3.46 to 6.37 t·ha^{-1} for the traditional cultivars. Taken together, these observations indicate a co-occurrence of high yield potential and elevated contents of the flavonoids analyzed in the plant material examined. The consistency of the direction of variation across several luteolin derivatives reinforces the relevance of this pattern as a component of the biochemical characterization of genotype groups differing in grain yield.

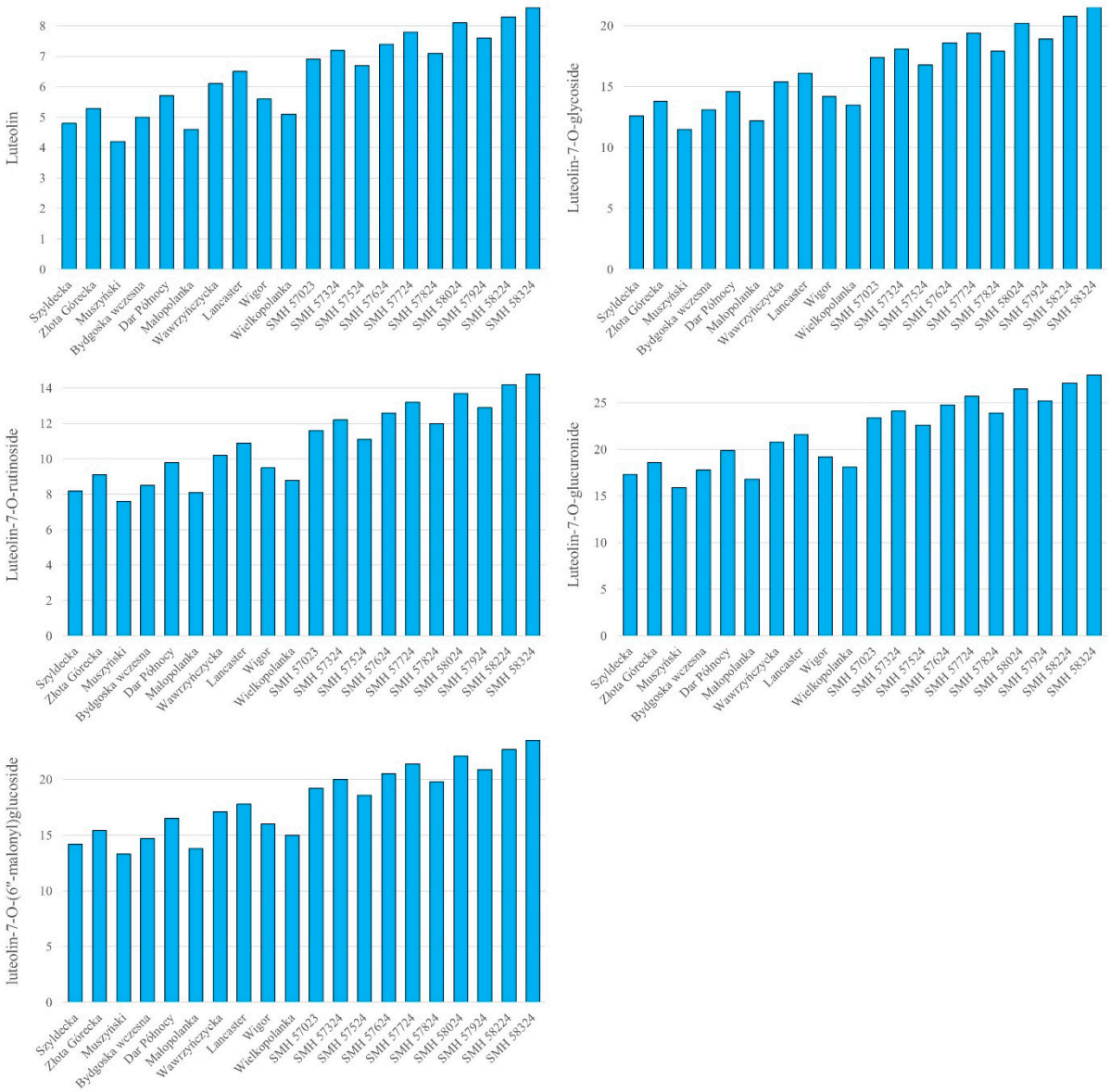


Figure 8. Graphs illustrating changes in flavonoid contents ($\mu g \cdot g^{-1}$) in low- and high-yielding maize genotypes.

Luteolin and its derivatives belong to a class of flavonoids involved in plant responses to environmental factors, protection against radiation, and the regulation of selected physiological processes. Differences in the contents of individual derivatives may result not only from changes in the overall activity of the flavonoid biosynthetic pathway but also from variation in their modification through glycosylation, glucuronidation, and malonylation. The observed increase in the contents of several luteolin derivatives in the SMH lines may therefore indicate distinct metabolic characteristics of these genotypes.

Elevated flavonoid contents may be considered a component of the biochemical variation among genotypes with contrasting yield potential; however, at this stage, they do not provide direct evidence of a causal contribution to grain yield determination. In particular, it should not be assumed that higher contents of each individual flavonoid automatically translate into greater plant productivity. This relationship may reflect broader genetic differences between modern breeding lines and traditional cultivars, encompassing multiple physiological and metabolic processes (Figure 8).

In the context of the transcriptomic analysis, these findings provide a rationale for further integration of candidate gene expression data with flavonoid contents. Such an approach may help identify relationships among genetic variation, the regulation of metabolic processes, and the grain yield phenotype.

In summary, the profiles of the five flavonoids analyzed exhibited a consistent pattern of variation between traditional maize cultivars and high-yielding SMH lines. The higher contents of luteolin and its derivatives in the SMH group, particularly in SMH 58324, represent a valuable component of the metabolomic characterization of genotypes with contrasting yield potential. These findings support the hypothesis that high grain yield co-occurs with a distinct flavonoid metabolic profile; however, further analyses are required to establish the strength and nature of this relationship (Figure 8).

2.3.3. Variation in Carotenoid Contents Among Maize Genotypes with Contrasting Yield Potential

Analysis of carotenoid contents in the leaves of the maize genotypes examined revealed pronounced differences between traditional, lower-yielding cultivars and high-yielding SMH breeding lines. For all four compounds analyzed—zeaxanthin, lutein, β-carotene, and α-carotene—higher values were observed in the SMH group than in the traditional cultivar group. Among the traditional cultivars, Muszyński exhibited relatively low contents of the carotenoids analyzed, whereas the highest values for all four parameters among the high-yielding genotypes were recorded in the line SMH 58324. Despite variation among individual SMH lines, the consistent direction of change across all compounds examined indicates distinct characteristics of leaf carotenoid metabolism in genotypes differing in yield potential (Figure 9).

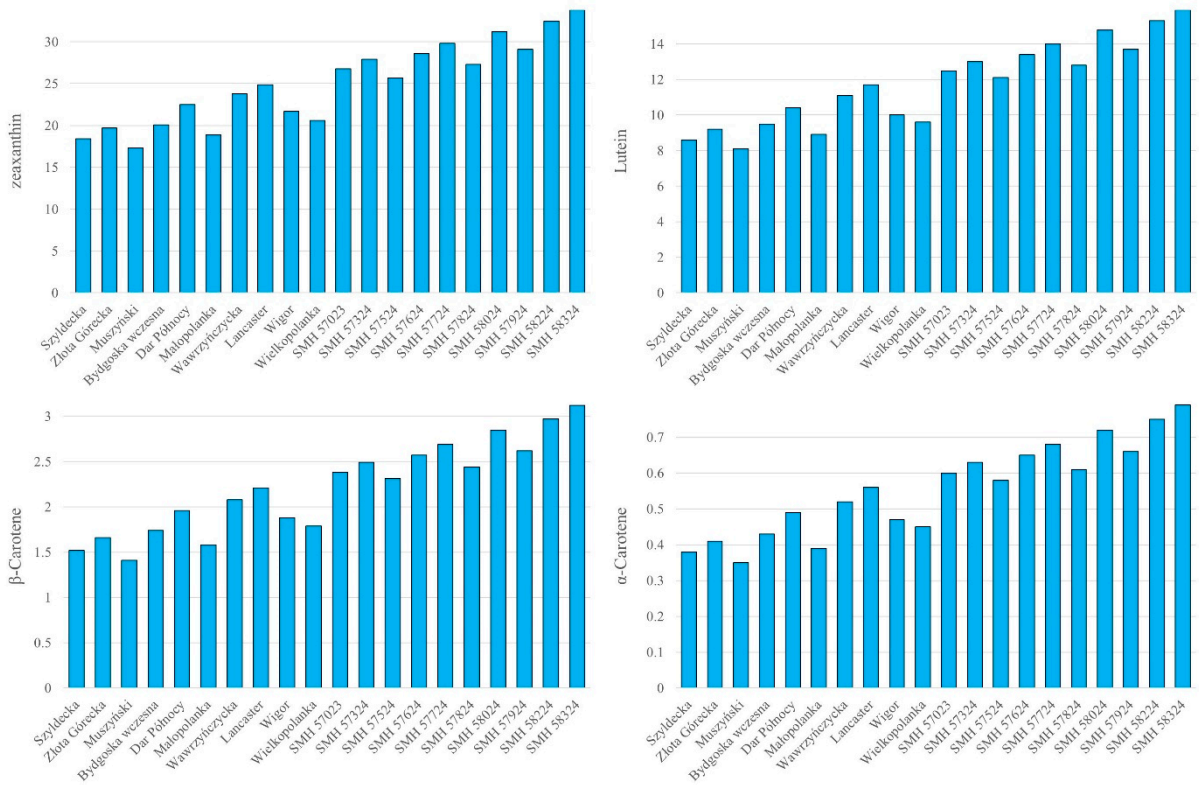


Figure 9. Graphs illustrating changes in carotenoid contents (mg·kg^{-1}) in low- and high-yielding maize genotypes.

The observed differences are physiologically relevant because carotenoids perform essential functions in the organization and protection of the photosynthetic apparatus. Lutein and β-carotene are components of photosynthetic complexes, contributing to their structural stabilization and the regulation of excitation energy transfer. Zeaxanthin plays a key role in the xanthophyll cycle-dependent dissipation of excess light energy as heat, thereby reducing the risk of photodamage to the photosynthetic apparatus. β-Carotene also participates in quenching reactive oxygen species generated under conditions of excessive excitation of the photosystems. Consequently, differences in the contents of individual carotenoids may reflect variation in the photoprotective capacity of leaves and the biochemical properties of the photosynthetic apparatus.

In the context of grain yield, the co-occurrence of higher leaf carotenoid contents in SMH lines and their greater yield potential is particularly noteworthy. Previous field experiment results indicated grain yields of approximately 10.94–14.01 t·ha^{-1} in the SMH group, compared with approximately 3.46–6.37 t·ha^{-1} in the traditional cultivar group. Maintenance of photosynthetic apparatus functionality, particularly under excessive irradiance or other stress conditions, may help preserve the capacity of leaves to assimilate carbon dioxide and sustain the production of assimilates required for plant growth and grain filling. From this perspective, higher carotenoid contents in the leaves of SMH lines may represent a component of a physiological profile associated with high yield potential (Figure 9).

However, it should be emphasized that the total content of individual carotenoids is not a direct measure of photosynthetic rate or radiation-use efficiency. The functional significance of these compounds depends, among other factors, on their distribution within photosynthetic complexes, the relative proportions of individual pigments, the physiological status of the plants, and environmental conditions. In the case of zeaxanthin, in particular, its contribution to photoprotection depends on dynamic transformations within the xanthophyll cycle. Therefore, a single measurement of zeaxanthin content cannot independently determine the extent of excess energy dissipation. Similarly, higher lutein and carotenoid contents do not, in themselves, demonstrate greater photosynthetic efficiency.

The consistent direction of variation across the four carotenoids indicates that the differences between the groups examined involve several components of the leaf pigment system. The highest values of all compounds analyzed in SMH 58324 further highlight its distinct biochemical profile. Combined with the observations concerning phenolic acids and flavonoids, these results suggest that the high-yielding SMH lines exhibit broader metabolomic differentiation compared with traditional cultivars. Nevertheless, these findings do not establish that changes in carotenoid contents directly determine increased grain yield (Figure 9).

In summary, the higher contents of zeaxanthin, lutein, β-carotene, and α-carotene in the leaves of high-yielding SMH lines indicate consistent differences in carotenoid profiles among genotypes with contrasting yield potential. These findings support the hypothesis that variation in photosynthetic pigment composition constitutes one component of the broader physiological and metabolic differentiation among the genotypes examined. However, the potential contribution of these differences to grain yield requires further validation, as carotenoid content alone does not allow direct conclusions to be drawn regarding photosynthetic efficiency or a causal effect on grain yield (Figure 9).

2.3.4. Variation in Oxidoreductase Activities Among Maize Genotypes with Contrasting Yield Potential

Analysis of polyphenol oxidase (PPO), catalase (CAT), and peroxidase (POD) activities in the leaves of the maize genotypes examined revealed pronounced differences between traditional cultivars and high-yielding SMH breeding lines. For all three parameters ana-

lyzed, higher values were recorded in the SMH lines than in the traditional cultivar group. PPO activity in the traditional cultivars ranged approximately from 17 to 24 U·g^{-1}, whereas in the SMH lines it ranged from approximately 26 to 35 U·g^{-1}. CAT activity values were approximately 73–96 U·g^{-1} and 101–131 U·g^{-1}, respectively, while POD activity ranged from approximately 40 to 54 U·g^{-1} in the traditional cultivars and from approximately 57 to 74 U·g^{-1} in the SMH lines. The separation of the value ranges between the two groups indicates consistent differences in the activities of the enzymes examined across the genetic material analyzed. The elevated polyphenol oxidase activity observed in the leaves of SMH lines indicates differences in the activity of oxidative transformations of phenolic compounds (Figure 10).

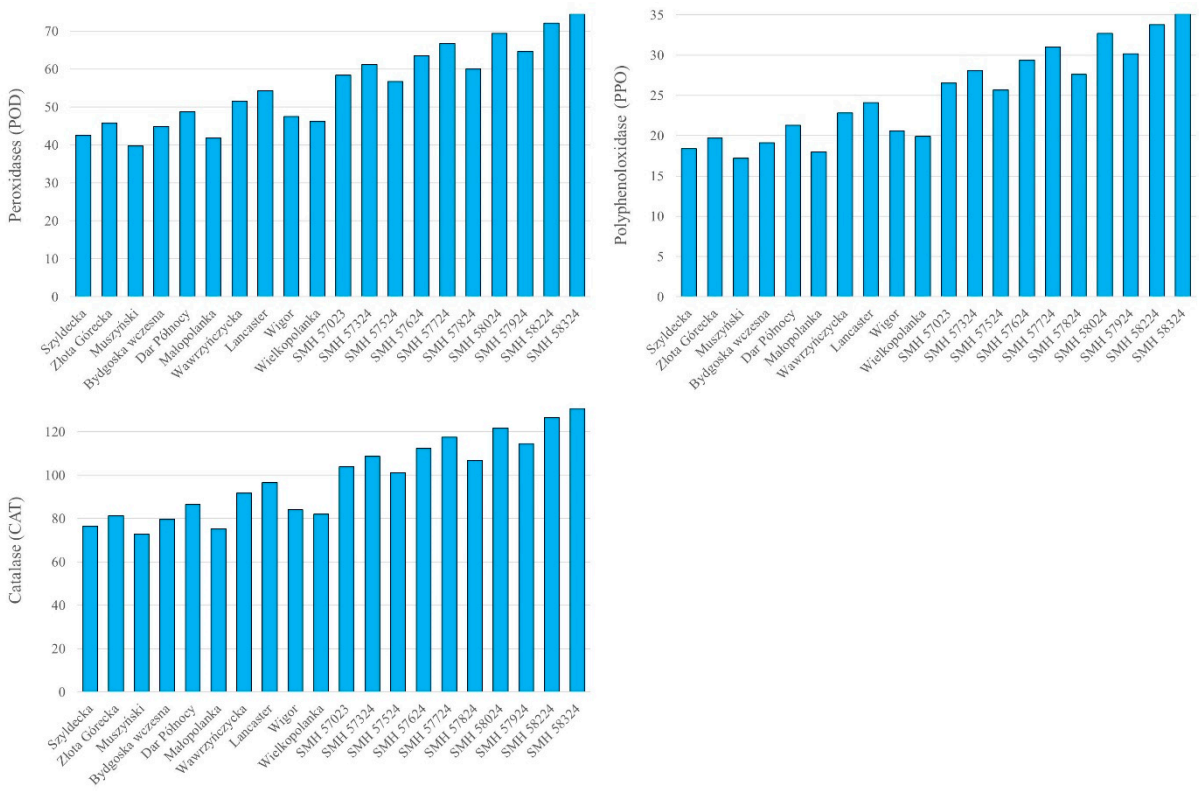


Figure 10. Graphs illustrating changes in oxidoreductase activities (U·g^{-1}) in low- and high-yielding maize genotypes.

PPO catalyzes the oxygen-dependent oxidation of specific mono- and o-diphenols to the corresponding quinones, which may subsequently participate in reactions leading to the formation of products with defensive properties. This enzyme is involved, among other processes, in plant responses to tissue damage and certain biotic factors. Higher PPO activity in SMH lines may therefore reflect distinct regulation of phenolic metabolism and a potentially greater capacity to activate selected defense-related reactions. However, it does not independently demonstrate increased plant resistance, since the activity of this enzyme also depends on substrate availability, cellular localization, and physiological conditions (Figure 10).

The differences in catalase and peroxidase activities are particularly relevant because these enzymes participate in the regulation of reactive oxygen species levels. Catalase catalyzes the decomposition of hydrogen peroxide into water and oxygen, thereby limiting the risk of its excessive accumulation in cells. Peroxidases, in turn, use hydrogen peroxide to oxidize specific substrates and participate in the regulation of redox balance, phenolic compound transformations, cell wall remodeling, and defense responses. These enzymes

differ in their localization, substrate specificity, and physiological functions; therefore, their activities should be interpreted separately and subsequently considered in the context of the antioxidant system as a whole. The higher CAT and POD activities observed in the leaves of SMH lines suggest that these genotypes exhibit distinct regulation of processes associated with hydrogen peroxide metabolism (Figure 10). This may indicate a greater capacity for enzymatic control of oxidative homeostasis, which is important for maintaining cellular structural integrity and chloroplast function. Maintaining a balance between the production and scavenging of reactive oxygen species is particularly important in photosynthetic tissues, where electron transport can generate these species, especially under excessive irradiance or other stress conditions. Nevertheless, elevated enzyme activity may reflect either more effective redox regulation or an increased demand for defense mechanisms (Figure 10).

Previous grain yield data indicated that the SMH lines achieved higher yields than the traditional cultivars. The simultaneous occurrence of higher PPO, CAT, and POD activities in the leaves of SMH lines therefore indicates an association between high yield potential and a distinct profile of enzyme activities related to phenolic metabolism and the regulation of oxidative balance. From a biological perspective, effective control of reactive oxygen species may help preserve cellular functions and the photosynthetic apparatus, thereby supporting the continued production of assimilates required for kernel growth and development (Figure 10). However, the results presented do not establish that higher activities of these enzymes directly cause increased grain yield.

The consistency between the observed changes in enzyme activities and the previously identified differences in phenolic acid, flavonoid, and carotenoid contents further increases the significance of these findings. The concurrent occurrence of these differences suggests that high-yielding SMH lines may exhibit distinct coordination of several interconnected components of secondary metabolism and cellular protection. Phenolic compounds may serve as substrates for reactions catalyzed by PPO and certain peroxidases, whereas carotenoids contribute to the protection of the photosynthetic apparatus against excessive excitation and oxidative stress. The co-occurrence of these characteristics supports the need for an integrated metabolomic and enzymatic analysis, although it does not establish direct causal relationships among metabolite contents, enzyme activities, and grain yield (Figure 10).

In summary, the high-yielding SMH lines exhibited higher leaf activities of polyphenol oxidase, catalase, and peroxidases than the traditional cultivars. These findings indicate consistent differences in enzymatic mechanisms associated with phenolic metabolism and the regulation of oxidative homeostasis. Together with the distinct metabolite profiles observed in the leaves, these results provide a basis for further investigations into the physiological and biochemical determinants of high grain yield in maize.

2.3.5. Principal Component Analysis (PCA) of Selected Metabolite Profiles

Principal component analysis (PCA) revealed pronounced variation in the leaf metabolomic profiles of the maize genotypes examined. The first principal component (PC_1) explained 99.79% of the total variance, whereas the second principal component (PC_2) accounted for 0.19%. Together, the two components represented 99.98% of the variation captured in the displayed projection. The distribution of genotypes indicated partial separation of traditional cultivars from SMH breeding lines, with most traditional cultivars located on the negative side of PC_1 and most SMH lines on the positive side. This arrangement suggests that the principal differences in the biochemical profiles analyzed were associated with the PC_1 axis. At the same time, variation in genotype positioning along PC_2 indicated additional differences among individual lines and cultivars (Figure 11).

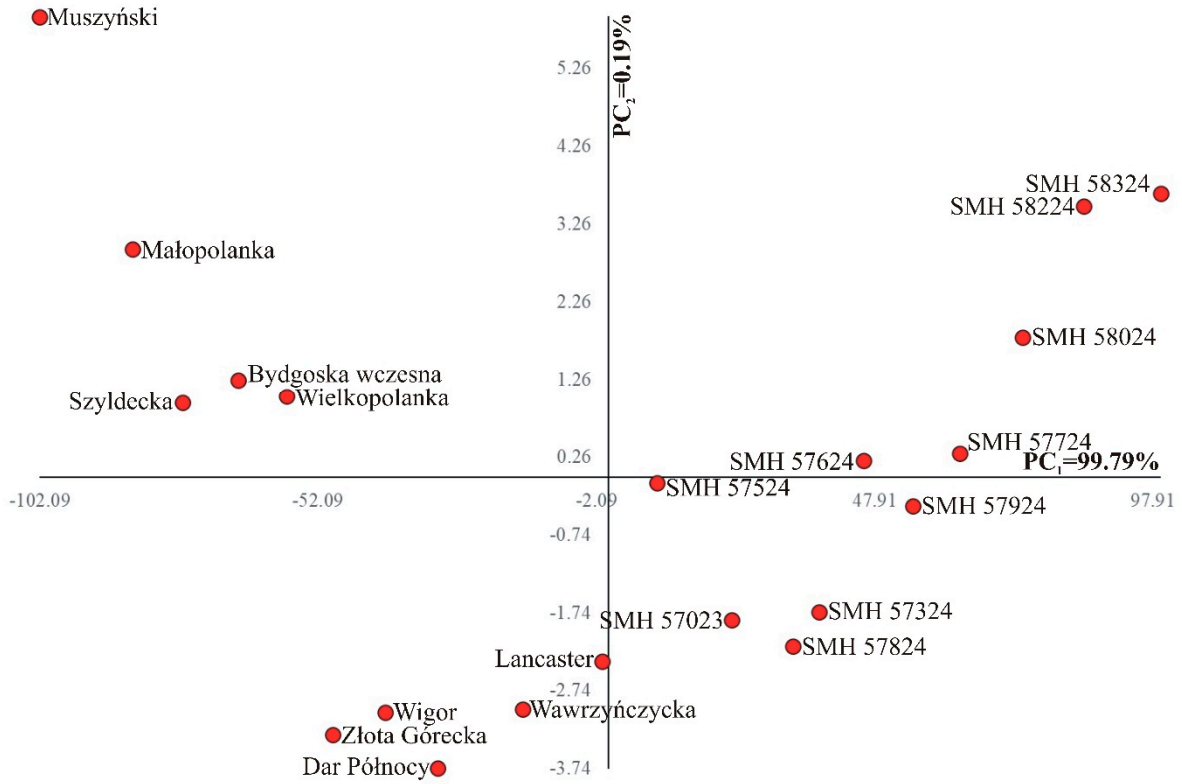


Figure 11. Distribution of genotypes (cultivars and lines) in the space of the first two principal components (PCA) calculated based on metabolic profiles.

The most distant positions in the positive regions of both principal components were occupied by the lines SMH 58224 and SMH 58324, whose metabolomic profiles were clearly distinct from those of most traditional cultivars. The line SMH 58024 also exhibited a high positive PC_1 score but occupied an intermediate position along PC_2. A different coordinate pattern was observed for SMH 57023, SMH 57324, and SMH 57824, which were located in the lower part of the plot. In contrast, SMH 57524 and SMH 57624 were positioned close to the origin. The variation in the distribution of the SMH lines indicates that, despite belonging to the high-yielding group, they did not exhibit uniform metabolomic profiles (Figure 11).

Pronounced variation was also observed among the traditional maize cultivars. Małopolanka occupied a position characterized by a relatively high positive PC_2 score and a clearly negative PC_1 score, whereas Muszyński was located in the upper part of the plot. Szyldecka, Bydgoska Wczesna, and Wielkopolanka formed a region characterized by positive PC_2 values, while Wigor, Złota Górecka, Dar Północy, Lancaster, and Wawrzyńczycka were positioned below the origin. This distribution indicates substantial variation in biochemical profiles within the traditional cultivar group as well, highlighting the importance of genotype-specific characterization in metabolomic analyses (Figure 11).

In the context of the previously reported results on carotenoid and flavonoid contents, phenolic acid profiles, and the activities of enzymes involved in oxidative metabolism, the observed separation of genotypes indicates broader biochemical differentiation between the groups examined. The partial separation of most SMH lines from the traditional cultivars is consistent with the hypothesis that high yield potential co-occurs with distinct leaf metabolomic profiles. However, this does not imply that all SMH lines exhibit identical patterns of change in individual metabolites or that these differences directly determine grain yield. PCA reveals similarities and differences among multivariate profiles but does

not, in itself, establish causal relationships between biochemical composition and plant productivity (Figure 11).

In summary, PCA revealed pronounced variation in the metabolomic profiles of the maize genotypes examined and partial separation of traditional cultivars from the SMH lines. At the same time, variation in the positioning of genotypes within both groups indicates distinct biochemical characteristics of individual genotypes.

2.4. Integration of Transcriptomic and Metabolomic Data

To examine molecular and biochemical patterns associated with differences in yield potential among the maize (*Zea mays* L.) genotypes examined, an integrated multivariate analysis was performed, combining transcriptional activity profiles of 33 candidate genes (classified into four distinct functional groups) with the accumulation levels of 16 metabolites. The dataset comprised phenolic acids (ferulic, p-coumaric, caffeic, and diferulic acids), flavonoids (luteolin, luteolin-7-O-glucoside, luteolin-7-O-rutinoside, luteolin-7-O-glucuronide, and luteolin-7-O-(6″-malonyl)-glucoside), carotenoids (zeaxanthin, lutein, β-carotene, and α-carotene), and the activities of oxidoreductive enzymes (peroxidase—POD, polyphenol oxidase—PPO, and catalase—CAT).

Linear regression analysis revealed highly statistically significant relationships between the expression levels of selected candidate genes and the concentrations of the metabolites analyzed, as well as enzymatic activity. The most prominent pattern identified by the regression analysis was a strong negative association between the expression of Group II genes, which exhibited higher transcript levels in low-yielding genotypes, and the entire set of metabolic and enzymatic parameters analyzed. In particular, genes G1 (*LOC100282883*), G15 (*LOC100502264*), G26 (*LOC100282937*, encoding a RAF-type zinc finger domain protein), G29 (*LOC103637325*), and G32 (*LOC103636139)* exhibited significant negative regression coefficients for all 16 metabolic and enzymatic parameters analyzed ($p < 0.05$).

Among these genes, G32 (*LOC103636139*) showed the strongest negative associations with the biochemical parameters, with its transcript level accounting for 42.8–48.7% of the total variance in the metabolites examined (Pr). Elevated G32 expression had a pronounced negative association with the accumulation of ferulic acid ($\beta = -109.00$, $p < 0.001$, Pr = 48.7%), p-coumaric acid ($\beta = -18.66$, $p < 0.001$), caffeic acid ($\beta = -7.19$, $p < 0.001$), diferulic acid ($\beta = -13.60$, $p < 0.001$), and luteolin derivatives (e.g., luteolin-7-O-glucuronide: $\beta = -7.41$, $p < 0.001$, Pr = 48.6%). Furthermore, G32 expression was negatively associated with carotenoid content (α-carotene: $\beta = -20.48$, $p < 0.001$) and the activities of polyphenol oxidase ($\beta = -34.99$, $p < 0.001$, Pr = 45.2%) and peroxidases ($\beta = -10.81$, $p < 0.001$, Pr = 44.2%) (Table 2).

A similar broad negative association pattern was observed for G15, whose expression accounted for 22.3–28.3% of the variance in metabolite levels (including ferulic acid: $\beta = -100.10$, $p = 0.009$; PPO: $\beta = -30.70$, $p = 0.018$), G1 (Pr = 21.5–30.6%; ferulic acid: $\beta = -90.20$, $p = 0.008$; PPO: $\beta = -26.80$, $p = 0.021$), and G26 (Pr = 17.4–22.0%; ferulic acid: $\beta = -86.30$, $p = 0.023$). In the case of G29, negative associations were observed with phenolic acids, flavonoids, carotenoids, and PPO activity (Pr = 15.2–20.1%, $p < 0.05$) (Table 2).

These observations show that higher transcriptional activity of Group II genes, characteristic of traditional open-pollinated cultivars with low grain yield, is associated with lower levels of the measured phenolic compounds and carotenoids and with lower oxidoreductive enzyme activities. Consequently, transcriptional downregulation of Group II genes in modern SMH breeding lines may co-occur with higher secondary metabolite levels and oxidoreductive enzyme activities.

Table 2. The influence of gene expression on the values of individual metabolites. Red shading indicates significant negative associations.

	Ferulic Acid			p-Coumaric Acid			Caffeic Acid			Diferulic Acid			Luteolin			Luteolin-7-O-Glucoside			Luteolin 7-O-Rutinoside			Luteolin-7-O-Glucuronide		
	Effect [1]	p^2	Pr [3]	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr
G1	−90.2	0.008	30.6	−15.55	0.008	30.4	−5.77	0.011	28.1	−11.07	0.012	27.5	−2.129	0.01	28.8	−4.73	0.013	26.9	−3.35	0.016	25.6	−5.93	0.011	28.1
G2	−29.9	0.62		−5.1	0.621		−1.96	0.623		−4.14	0.592		−0.74	0.613		−1.82	0.585		−1.26	0.602		−2.25	0.584	
G3	−87.6	0.137	7.4	−14.59	0.153	6.4	−5.59	0.153	6.4	−10.9	0.151	6.6	−2.02	0.158	6.2	−4.78	0.144	7	−3.36	0.155	6.3	−5.96	0.138	7.3
G4	−15.53	0.117	8.3	−2.57	0.139	6.8	−0.989	0.133	7.2	−1.9	0.137	7	−0.359	0.136	7	−0.823	0.134	7.1	−0.572	0.15	6.2	−1.019	0.131	7.3
G5	0.134	0.341		0.021	0.393		0.00936	0.313	0.4	0.0175	0.331		0.00337	0.32	0.2	0.00784	0.312	0.4	0.00536	0.339		0.00964	0.312	0.4
G6	3.23	0.12	10.4	0.543	0.134	9.3	0.214	0.12	10.4	0.389	0.147	8.3	0.0749	0.138	9	0.172	0.135	9.2	0.1226	0.141	8.7	0.219	0.122	10.2
G7	−18.4	0.303	0.6	−3.28	0.295	0.9	−1.17	0.326	0.1	−2.35	0.308	0.5	−0.403	0.353		−0.942	0.343		−0.724	0.312	0.4	−1.2	0.327	0.1
G8	−9.1	0.675		−1.94	0.609		−0.48	0.741		−0.72	0.798		−0.204	0.698		−0.38	0.75		−0.244	0.78		−0.47	0.751	
G9	−14.65	0.07	12.4	−2.61	0.064	13.2	−0.936	0.082	11.2	−1.751	0.095	10	−0.343	0.081	11.3	−0.772	0.086	10.8	−0.54	0.098	9.7	−0.947	0.087	10.7
G10	0.96	0.357		0.155	0.394		0.0675	0.325	0.1	0.128	0.335		0.0247	0.324	0.1	0.0575	0.315	0.3	0.0396	0.339		0.0697	0.323	0.2
G11	−62.1	0.196	4	−11.13	0.185	4.5	−4.01	0.208	3.6	−8.1	0.189	4.3	−1.49	0.199	3.9	−3.43	0.197	4	−2.49	0.195	4.1	−4.16	0.203	3.8
G12	−9	0.461		−1.76	0.406		−0.52	0.519		−1.1	0.483		−0.193	0.512		−0.448	0.506		−0.349	0.473		−0.537	0.517	
G13	0.067	0.679		0.0089	0.754		0.0051	0.635		0.0091	0.66		0.00184	0.638		0.0043	0.631		0.00277	0.669		0.0053	0.632	
G14	17.8	0.425		3.22	0.407		1.28	0.385		2.53	0.376		0.477	0.374		1.09	0.373		0.779	0.379		1.28	0.398	
G15	−100.1	0.009	28.3	−17.27	0.01	27.4	−6.47	0.011	26.8	−12.21	0.015	24.9	−2.371	0.011	26.9	−5.33	0.013	25.8	−3.78	0.015	24.7	−6.67	0.011	26.9
G16	18.6	0.445		3.35	0.43		1.34	0.406		2.68	0.391		0.501	0.393		1.15	0.39		0.836	0.388		1.34	0.417	
G17	−12.55	0.234	12.2	−2.26	0.212	14.9	−0.764	0.278	7.4	−1.4	0.306	4.7	−0.294	0.252	10.1	−0.642	0.27	8.2	−0.435	0.309	4.4	−0.774	0.282	7
G18	12.3	0.259	2	2.11	0.27	1.6	0.848	0.24	2.6	1.63	0.244	2.5	0.314	0.233	2.8	0.712	0.238	2.7	0.507	0.244	2.5	0.862	0.244	2.5
G19	−48.7	0.062	13.5	−8.24	0.071	12.3	−3.05	0.078	11.6	−5.85	0.083	11.1	−1.122	0.076	11.8	−2.59	0.073	12.1	−1.864	0.075	12	−3.18	0.074	12.1
G20	−9.45	0.233	2.7	−1.73	0.21	3.5	−0.587	0.264	1.7	−1.173	0.25	2.1	−0.209	0.276	1.4	−0.474	0.281	1.2	−0.346	0.277	1.3	−0.59	0.276	1.4
G21	−0.14	0.98		−0.118	0.899		−0.002	0.995		0.032	0.963		−0.004	0.976		−0.004	0.99		0.01	0.963		0.003	0.994	
G22	−24.5	0.545		−4.16	0.557		−1.41	0.6		−2.85	0.584		−0.551	0.574		−1.26	0.574		−0.95	0.559		−1.55	0.574	
G23	56.9	0.164	5.5	10.55	0.138	6.9	3.63	0.181	4.7	7.35	0.162	5.6	1.36	0.169	5.3	3.19	0.157	5.9	2.43	0.135	7.1	3.81	0.171	5.2
G24	−36.3	0.225	3	−6.92	0.184	4.6	−2.28	0.251	2.1	−4.34	0.261	1.8	−0.85	0.241	2.4	−1.89	0.255	2	−1.38	0.249	2.1	−2.33	0.253	2
G25	−24	0.35		−4.46	0.319	0.3	−1.47	0.387		−2.74	0.408		−0.543	0.382		−1.19	0.402		−0.88	0.392		−1.5	0.391	
G26	−86.3	0.023	21.3	−14.64	0.028	19.8	−5.73	0.023	21.5	−10.81	0.027	20	−2.084	0.023	21.3	−4.78	0.023	21.5	−3.36	0.027	20	−5.94	0.021	22
G27	−21.4	0.058	14	−3.78	0.056	14.4	−1.399	0.062	13.4	−2.55	0.082	11.2	−0.523	0.056	14.3	−1.155	0.066	13	−0.799	0.079	11.5	−1.43	0.064	13.2
G28	−0.19	0.945		−0.097	0.843		−0.002	0.99		−0.014	0.969		−0.002	0.977		−0.003	0.987		−0.007	0.953		−0.002	0.992	

Table 2. *Cont.*

	Ferulic Acid			p-Coumaric Acid			Caffeic Acid			Diferulic Acid			Luteolin			Luteolin-7-O-Glucoside			Luteolin 7-O-Rutinoside			Luteolin-7-O-Glucuronide		
	Effect [1]	p^2	Pr [3]	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr
G29	−28.7	0.027	20.1	−5.01	0.028	19.9	−1.887	0.029	19.7	−3.45	0.041	16.8	−0.695	0.027	20.1	−1.539	0.033	18.5	−1.08	0.039	17.2	−1.931	0.029	19.5
G30	−53.1	0.053	28.5	−9.01	0.062	26	−3.49	0.054	28.2	−6.29	0.068	24.8	−1.255	0.06	26.5	−2.82	0.061	26.5	−1.966	0.067	24.9	−3.61	0.051	28.8
G31	−4.4	0.349		−0.781	0.34		−0.285	0.36		−0.507	0.402		−0.112	0.325	0.2	−0.228	0.383		−0.168	0.37		−0.291	0.365	
G32	−109	<0.001	48.7	−18.66	<0.001	46.5	−7.19	<0.001	48.4	−13.6	<0.001	45.6	−2.627	<0.001	48.3	−5.98	<0.001	47.8	−4.19	<0.001	44.6	−7.41	<0.001	48.6
G33	−25.5	0.488		−4.06	0.528		−1.42	0.559		−2.96	0.531		−0.534	0.548		−1.34	0.511		−1.05	0.475		−1.6	0.522	

	luteolin-7-O-(6″malonyl)glucoside			zeaxanthin			lutein			β-carotene			α-carotene			peroxidase			Polyphenol oxidase			catalase		
	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr	Effect	p	Pr
G1	−4.87	0.014	26.3	−7.1	0.024	22.3	−3.35	0.025	21.7	−0.712	0.026	21.5	−0.185	0.025	21.8	−15.6	0.022	22.8	−8.18	0.025	21.9	−26.8	0.021	23.2
G2	−1.76	0.612		−2.89	0.593		−1.57	0.541		−0.295	0.59		−0.075	0.599		−7	0.553		−3.8	0.544		−12.1	0.544	
G3	−4.88	0.152	6.5	−7.27	0.171	5.5	−3.52	0.164	5.8	−0.723	0.181	5	−0.186	0.184	4.9	−16.6	0.151	6.6	−8.77	0.154	6.4	−28.8	0.143	7
G4	−0.857	0.134	7.2	−1.163	0.198	4	−0.538	0.211	3.5	−0.1157	0.21	3.5	−0.0301	0.207	3.6	−2.73	0.159	5.7	−1.454	0.159	5.8	−4.68	0.158	5.8
G5	0.008	0.322	0.2	0.0116	0.359		0.00576	0.339		0.00117	0.368		0.000307	0.358		0.0276	0.312	0.4	0.0155	0.283	1.2	0.0471	0.311	0.4
G6	0.182	0.128	9.7	0.271	0.149	8.1	0.1273	0.153	7.9	0.0272	0.155	7.8	0.00692	0.163	7.2	0.565	0.166	7	0.305	0.157	7.6	0.989	0.155	7.7
G7	−1.03	0.318	0.3	−1.58	0.328	0.1	−0.688	0.372		−0.155	0.349		−0.04	0.348		−3.15	0.367		−1.64	0.379		−5.53	0.354	
G8	−0.34	0.786		−0.72	0.713		−0.332	0.723		−0.087	0.665		−0.0222	0.668		−1.04	0.807		−0.3	0.894		−1.35	0.852	
G9	−0.784	0.095	10	−1.218	0.098	9.7	−0.558	0.112	8.6	−0.1242	0.098	9.7	−0.0321	0.098	9.8	−2.51	0.115	8.4	−1.271	0.135	7.1	−4.15	0.128	7.5
G10	0.0585	0.327	0.1	0.0872	0.352		0.0439	0.323	0.2	0.00884	0.354		0.00231	0.35		0.207	0.304	0.6	0.117	0.275	1.4	0.351	0.307	0.5
G11	−3.47	0.211	3.5	−5.91	0.172	5.1	−2.89	0.16	5.7	−0.617	0.161	5.6	−0.159	0.162	5.6	−12.4	0.184	4.6	−6.55	0.187	4.4	−21	0.187	4.4
G12	−0.471	0.502		−0.83	0.45		−0.372	0.477		−0.086	0.443		−0.0217	0.453		−1.53	0.518		−0.74	0.56		−2.54	0.53	
G13	0.00427	0.647		0.0055	0.706		0.00294	0.672		0.00054	0.718		0.000145	0.706		0.0155	0.622		0.0093	0.578		0.0265	0.622	
G14	1.12	0.38		1.86	0.352		0.944	0.32	0.2	0.199	0.329	0	0.0511	0.331		4.12	0.339		2.28	0.319	0.3	6.88	0.35	
G15	−5.44	0.015	24.7	−8.55	0.015	24.8	−4.05	0.015	24.6	−0.87	0.015	24.7	−0.2243	0.015	24.6	−18.05	0.017	23.6	−9.39	0.021	22.3	−30.7	0.018	23.3
G16	1.19	0.394		2.01	0.357		1.02	0.324	0.1	0.215	0.334		0.0554	0.335		4.44	0.345		2.47	0.324	0.2	7.44	0.355	
G17	−0.626	0.304	4.9	−1.043	0.278	7.4	−0.49	0.283	6.9	−0.1096	0.258	9.5	−0.028	0.263	9	−2.12	0.303	5.1	−1.03	0.355	0.6	−3.38	0.339	1.9
G18	0.736	0.24	2.6	1.132	0.251	2.2	0.563	0.229	3	0.1169	0.245	2.4	0.0302	0.245	2.4	2.54	0.231	2.9	1.41	0.21	3.7	4.28	0.237	2.7
G19	−2.64	0.08	11.4	−4.07	0.086	10.8	−1.89	0.095	10	−0.402	0.097	9.8	−0.1039	0.097	9.8	−8.72	0.088	10.6	−4.53	0.097	9.8	−14.59	0.095	10
G20	−0.498	0.276	1.4	−0.764	0.287	1.1	−0.331	0.334		−0.0763	0.297	0.8	−0.0199	0.294	0.9	−1.53	0.323	0.2	−0.765	0.355		−2.61	0.324	0.2

Table 2. *Cont.*

	luteolin-7-O-(6″malonyl)glucoside			zeaxanthin			lutein			β-carotene			α-carotene			peroxidase			Polyphenol oxidase			catalase		
	Effect	***p***	**Pr**	**Effect**	***p***	**Pr**	**Effect**	***p***	**Pr**	**Effect**	***p***	**Pr**	**Effect**	***p***	**Pr**	**Effect**	***p***	**Pr**	**Effect**	***p***	**Pr**	**Effect**	***p***	**Pr**
G21	0.01	0.974		−0.009	0.985		0	0.999		−0.0022	0.964		−0.0005	0.967		0.06	0.953		0.076	0.89		0.18	0.918	
G22	−1.27	0.585		−2.1	0.565		−1.02	0.559		−0.213	0.568		−0.0532	0.581		−4.6	0.559		−2.25	0.591		−7.9	0.556	
G23	3.37	0.151	6.1	5.29	0.151	6.2	2.49	0.155	5.9	0.539	0.151	6.2	0.1374	0.157	5.9	11.36	0.152	6.1	5.82	0.169	5.2	18.8	0.165	5.4
G24	−1.96	0.257	1.9	−3.25	0.229	2.8	−1.52	0.238	2.5	−0.343	0.213	3.4	−0.0865	0.224	3	−6.31	0.28	1.3	−3.07	0.324	0.1	−10.58	0.289	1
G25	−1.25	0.4		−2.03	0.381		−0.92	0.404		−0.214	0.366		−0.054	0.377		−3.86	0.441		−1.83	0.493		−6.51	0.447	
G26	−4.91	0.025	20.7	−7.34	0.034	18.4	−3.51	0.033	18.6	−0.736	0.037	17.6	−0.1896	0.038	17.4	−16.26	0.029	19.7	−8.61	0.03	19.4	−28	0.027	20.1
G27	−1.172	0.074	12.1	−1.837	0.074	12.1	−0.868	0.076	11.8	−0.1876	0.073	12.1	−0.0482	0.075	11.9	−3.83	0.084	10.9	−1.95	0.1	9.6	−6.41	0.091	10.3
G28	−0.005	0.977		−0.03	0.905		−0.01	0.931		−0.0037	0.886		−0.0008	0.905		0	1		0.021	0.942		0.014	0.988	
G29	−1.585	0.035	18	−2.41	0.042	16.6	−1.136	0.044	16.2	−0.244	0.044	16.3	−0.0627	0.045	16	−5.07	0.048	15.6	−2.62	0.055	14.4	−8.58	0.05	15.2
G30	−2.94	0.058	27	−4.4	0.077	23	−2.08	0.082	22	−0.441	0.085	21.6	−0.1122	0.089	20.9	−9.32	0.076	23.1	−4.93	0.077	22.9	−15.7	0.08	22.5
G31	−0.244	0.368		−0.364	0.391		−0.174	0.384		−0.0365	0.396		−0.0091	0.414		−0.734	0.424		−0.378	0.437		−1.25	0.425	
G32	−6.15	<0.001	46.6	−9.32	<0.001	43.1	−4.45	<0.001	43.4	−0.947	0.001	42.8	−0.2448	0.001	42.9	−20.48	<0.001	45.1	−10.81	<0.001	44.2	−34.99	<0.001	45.2
G33	−1.39	0.509		−2.12	0.522		−0.96	0.541		−0.199	0.557		−0.0514	0.557		−4.86	0.496		−2.46	0.516		−8	0.509	

[1]—The slope of the regression line (β). It determines the effect of gene expression on the value of a particular metabolite.; [2]—*p*-value. The probability of making a Type I error.; [3]—Percentage of the total variability of the metabolite level explained by the variability in gene expression.

In contrast to Group II, candidate genes belonging to Group I (genes co-activated in high-yielding lines, e.g., G5—PP2C, G10—ISO3, and G13—*ARAD1*), Group III (genes with constitutive or background expression), and Group IV (genotype-specific genes, e.g., G17, G18, and G21) did not exhibit universal, statistically significant regression relationships across the entire population ($p > 0.05$). Although Group I genes showed coordinated activation in high-yielding genotypes, including the highest-yielding line SMH 58024, their relationships with metabolite levels may depend on genetic background or processes not assessed in the present analysis—for example, starch accumulation in kernels (G10) or cell wall biosynthesis (G13)—rather than directly affecting the concentrations of soluble secondary metabolites in leaves.

In summary, the integrated transcriptomic–metabolomic analysis revealed several statistically significant associations between candidate gene expression and selected biochemical parameters. In particular, genes belonging to Group II, including G32, G15, G26, and G1, showed negative associations with several phenolic acids, flavonoids, carotenoids, and oxidoreductive enzyme activities. These relationships describe coordinated variation in transcript abundance and leaf biochemical profiles across the investigated genotype panel. However, they do not demonstrate that these genes directly regulate the corresponding metabolic pathways or that the observed biochemical differences contribute causally to increased grain yield. Given the distinct genetic backgrounds and breeding histories of the two genotype groups, these associations may also reflect broader differences between modern SMH hybrids and traditional open-pollinated cultivars. Further studies using genetically diverse and appropriately balanced materials are required to establish the biological significance and direction of these relationships.

3. Discussion

Maize (*Zea mays* L.) is of critical importance to global food and feed production and serves as a source of industrial raw materials, while its cultivation underpins the livelihoods of a very large number of farms [16,17]. Contemporary breeding programs focus on improving drought and heat tolerance, resistance to pathogens, water- and nutrient-use efficiency, grain quality, and production stability under conditions of increasing climate variability [18]. All these objectives are important for reducing losses and maintaining productivity; however, increasing grain yield remains one of the most important goals. This is a particularly challenging task because yield depends on kernel number and weight, ear development, assimilate partitioning, and the coordinated action of numerous regulatory pathways [19,20].

Maize grain yield is a quantitative trait with a complex genetic architecture that is strongly modified by the environment. In our study, a field experiment conducted over the 2020–2022 period revealed a very strong genotype effect ($p < 0.001$), no significant main effect of year ($p = 0.098$), and a significant genotype × year interaction ($p < 0.001$). The highest mean grain yield was obtained by the line SMH 58024, at 14.01 t·ha^{-1}, followed by SMH 57924, at 13.41 t·ha^{-1}. The lowest values were recorded for the cultivars Muszyński (3.46 t·ha^{-1}) and Wawrzyńczycka (3.47 t·ha^{-1}). The difference between SMH 58024 and Muszyński was 10.55 t·ha^{-1}, indicating substantial variation within the material examined. The absence of a significant main effect of year does not mean that environmental conditions had no influence on yield. It indicates only that no significant difference in mean yield between years was detected after averaging across genotypes. The significant interaction, in turn, indicates that individual genotypes responded differently to the conditions prevailing in each growing season. The three-year results confirm substantial variation in yield potential but do not yet justify claims of broad genotype stability across different environments.

Studies by other authors have demonstrated the potential for genomic prediction of grain yield under diverse European conditions, while also confirming the need to account for genotype responses to specific environments [21]. QTL mapping and genome-wide association studies (GWAS) make it possible to identify genomic regions associated with complex traits [22], whereas combining crop growth models with genome analysis facilitates the biological interpretation of genotype–environment interactions [23–25]. It is also increasingly emphasized that transcriptome analysis can complement GWAS [26], and that the use of gene expression and metabolic data can improve the predictive value of models [27]. This means that a marker statistically associated with a trait should be regarded as a starting point for identifying the underlying gene and mechanism rather than as definitive evidence of a causal relationship. A recent review of maize seed development emphasizes that the final size, weight, composition, and vigor of the grain result from the coordinated development of the embryo, endosperm, and other tissues, as well as the activity of regulatory networks controlling nutrient transport and the deposition of starch and storage proteins [28].

In previous studies conducted by our research team, the application of DArTseq technology enabled the identification of SilicoDArT and SNP markers associated with ear length and diameter, row number and kernel number, thousand-kernel weight, and grain yield [29]. Among the markers associated with the same traits at two locations, some were located within genes, enabling the team to identify candidate genes for further characterization [29]. A subsequent study extended this strategy by developing diagnostic assays and evaluating markers in materials differing in productivity [30]. Subsequently, expression analysis of six candidate genes identified *CCR1*, *AAT*, and *SUT1* as particularly relevant to yield formation [31], while a further GWAS identified additional candidate genes, including genes encoding a U-box domain-containing protein and a WUSCHEL-related homeobox family protein [32].

The *CCR1* gene encodes cinnamoyl-CoA reductase, a component of the phenylpropanoid pathway leading to the biosynthesis of monolignols and lignin. Characterization of a maize *CCR1* mutant demonstrated that altered activity of this gene affects lignification, fiber development, and the global expression profile [33]. López-Malvar et al. [34] demonstrated an association between cell wall composition, structural characteristics of the stem, and biomass yield. In turn, the aspartate aminotransferase encoded by *AAT* participates in amino acid metabolism and links carbon metabolism with nitrogen metabolism. Alanine and aspartate aminotransferases perform important functions in the metabolism of C4 plants [35], whereas overexpression of *AAT* genes in rice has been shown to alter nitrogen metabolism and increase amino acid content in seeds [36]. The potential effect of *AAT* on yield may therefore result from changes in the availability of nitrogenous compounds, the supply of nutrients to developing kernels, and the relationship between photosynthetic carbon assimilation and amino acid synthesis. *SUT1* represents a different mechanism, but one directly associated with source–sink relationships. Studies of the expression and localization of *ZmSUT1* have indicated its involvement in sucrose loading into the phloem [37], while mutant experiments have confirmed the function of this transporter in maize leaves [38]. *ZmSUT1* may mediate both sucrose uptake into the phloem in mature leaves and sugar release from phloem vessels into other tissues [39]. Efficient transport of sugars from leaves to developing kernels may affect kernel number and weight and, ultimately, grain yield.

Among the genes examined in the present study, one of the strongest candidates for functional validation is the gene encoding isoamylase-type starch debranching enzyme isoform 3, *LOC100542095* (G10). In the material examined, high G10 expression co-occurred with high grain yield in the line SMH 58024. Furthermore, isoamylases participate in the metabolism of branched glucans and the proper formation of starch structure, which

constitutes the main storage material in maize kernels. Differences in starch metabolism may affect grain composition and weight. Another important gene is putative protein phosphatase 2C *LOC103635969* (G5), a candidate associated with physiological regulation. Strong transcriptional activation of this gene was observed in the high-yielding line SMH 58024. PP2C phosphatases comprise a large protein family with diverse regulatory functions. Some members participate in stress-response signaling pathways and hormonal regulation. Possible mechanisms may involve the regulation of cellular signaling, the maintenance of leaf physiological activity, or responses to stress. However, each of these hypotheses requires further investigation to establish the PP2C subfamily to which *LOC103635969* (G5) belongs and to identify the proteins and pathways it regulates. Another potentially important gene is probable arabinosyltransferase *ARAD1 LOC103643269* (G13). High expression was observed in both the high-yielding line SMH 58024 and the low-yielding cultivar Szyldecka. This observation is particularly important because it indicates that activation of this gene is not sufficient to ensure high grain yield. The gene may be involved in the modification of cell wall polysaccharides. Potential consequences include changes in the mechanical properties of tissues, cell growth, or organ function. However, no direct evidence currently indicates that G13 improves assimilate transport to the ears or increases kernel weight. The contrasting expression patterns observed in Szyldecka and SMH 58024 suggest that the potential effects of these genes may be strongly dependent on the genetic background and the broader regulatory context. In particular, elevated G5 and G13 expression in Szyldecka does not necessarily indicate enhanced activity of the corresponding proteins or more efficient physiological processes related to grain yield formation. Another important candidate is the uncharacterized protein *LOC103636139* (G32). This gene is one of the most interesting candidates identified through regression analysis. In the present study, negative relationships were reported between its expression and all 16 biochemical and enzymatic parameters examined. For ferulic acid, the regression coefficient was $\beta = -109.00$, $p < 0.001$, and Pr = 48.7%; for PPO activity, $\beta = -34.99$ and Pr = 45.2%. Such a large proportion of explained variance is noteworthy because it may indicate an association between G32 and a broad physiological profile of the plant. However, this does not mean that G32 is a direct repressor of all these pathways. Possible explanations include a regulatory function of the protein under investigation, an indirect relationship resulting from genetic differences between the two groups, the influence of a common regulatory factor affecting both G32 expression and metabolism, or a statistical effect arising from the small number of genotypes or strong collinearity. An additional limitation is that the molecular function of this protein is unknown. It is therefore not clear whether G32 encodes a transcription factor, a signaling protein, an enzyme, or a component of another cellular process. Nevertheless, G32 warrants high priority for functional studies because of the consistency of its statistical associations.

Other candidate genes for investigating the mechanisms underlying the negative relationships with the biochemical profile include uncharacterized protein *LOC100502264* (G15), RAF-type zinc finger domain-containing protein 1 *LOC100282937* (G26), and uncharacterized protein *LOC103637325* (G29). An important limitation is that several of these genes are annotated as encoding uncharacterized proteins. Although these loci exhibited significant associations with the biochemical parameters analyzed, their precise molecular functions and potential roles in the regulation of metabolic processes remain unknown. Therefore, the negative regression relationships observed for these genes should not be interpreted as evidence of direct inhibitory or regulatory activity. In particular, the strong associations identified for G32 indicate its potential relevance to the biochemical variation observed among the genotypes, but do not establish whether the encoded protein acts as a transcriptional regulator, a signaling component, an enzyme, or an indirect participant in these processes. Further characterization of these loci will require complementary bioin-

formatic and experimental approaches. In silico analyses, including sequence homology searches against annotated plant proteins, identification of conserved protein domains and motifs, and protein structure and subcellular localization predictions, could provide initial clues regarding their molecular functions. Comparative analyses with homologous proteins from other plant species may also help establish their evolutionary conservation and identify potentially related biological processes. However, such predictions would need to be supported by experimental evidence, including functional characterization of the encoded proteins, assessment of their effects on relevant metabolic pathways, and validation of their roles in maize plants. Accordingly, G15, G29, and G32 should currently be regarded as candidates for further functional investigation rather than validated targets for direct genome editing. Their potential application in marker-assisted selection also requires confirmation of marker–trait associations in independent and genetically diverse breeding populations. In the case of genome editing, functional characterization should precede the design of targeted modifications, followed by the assessment of their effects on yield-related traits and possible pleiotropic consequences. This stepwise approach is particularly important for genes with unknown molecular functions, for which the direction and biological consequences of modifying their expression cannot yet be reliably predicted.

PCA of the expression profiles of 33 genes showed that the first two principal components jointly explained 99.50% of the variation: PC_1 accounted for 97.04% and PC_2 for 2.46%. However, the analysis did not yield an unambiguous separation of all low- and high-yielding genotypes. Szyldecka, SMH 58024, and SMH 58324 were particularly distinctive. The latter two high-yielding lines were characterized by the highest mean grain yield in the field experiment, high expression of several Group I genes, including G5, G10, and G13, and expression profiles that differed from those of most other genotypes. SMH 58024 is therefore a valuable material for investigating combinations of molecular traits associated with high grain yield. However, it cannot be concluded that its performance is attributable to a single specific gene. High yield may result from multiple alleles, their interactions, and physiological traits. The observed relationships between candidate gene expression and grain yield should be interpreted as associations rather than evidence of direct causal effects. The elevated expression of G5, G10, and G13 in both the high-yielding line SMH 58024 and the low-yielding cultivar Szyldecka demonstrates that the transcriptional activation of individual genes is not sufficient to explain differences in grain yield. These observations suggest that the contribution of candidate genes may depend on the broader genetic background, interactions among loci, and additional physiological and regulatory processes. Therefore, the genes identified in this study should be considered candidates potentially associated with yield-related processes, rather than confirmed determinants of high grain yield. A further limitation of the present study concerns the differences in developmental stage and experimental conditions between molecular and biochemical analyses and grain yield assessment. Gene expression and biochemical parameters were evaluated in seedling leaf material obtained from plants grown under controlled phytotron conditions, whereas grain yield was determined in separate field experiments conducted over three growing seasons and assessed at plant maturity. Consequently, the molecular and biochemical measurements represent the physiological and metabolic status of the investigated genotypes during early vegetative development, rather than during the reproductive stages directly associated with ear development, kernel set, and grain filling. Nevertheless, the analysis of early vegetative material provides information on genotype-dependent molecular and biochemical variation under standardized environmental conditions. Such information may help characterize differences among genetic materials at an early developmental stage and identify candidate genes and biochemical traits for further investigation. However, establishing whether these early-stage patterns are maintained during reproductive devel-

opment and whether they are functionally related to final grain yield will require additional experiments involving repeated sampling at different developmental stages, particularly around flowering and during grain filling.

In earlier studies conducted by our research team, candidate genes associated with grain yield included a gene encoding a U-box domain-containing protein and a member of the WOX protein family [32]. U-box E3 ligases participate in the selective targeting of proteins for degradation and regulate growth, development, and stress responses [40,41]. Selected PUB ligases have also been shown to participate in the control of organ growth [42]. WOX proteins, in turn, regulate meristem development and tissue differentiation. Tissue-specific functions of *WOX3* have also been described in maize [43], while *WOX11* participates in adventitious root development in rice [44]. Further evidence linking developmental regulation with yield-related traits was provided by Li et al. [45]. Disruption of *ZmDLR9* affected pre-rRNA processing and ribosome assembly, resulting in reduced protein translation, impaired cell proliferation, and delayed lateral root development. The *dlr9-1* mutant also exhibited reduced kernel length, kernel width, and 100-kernel weight, indicating that disturbances in fundamental cellular processes may have broader consequences for plant development and kernel-related traits. Together, these findings illustrate that genes controlling developmental pathways and fundamental cellular processes may influence yield-related traits through changes in root architecture, organ development, and overall plant growth.

Genetic control of grain yield is closely linked to metabolism. The integration of transcriptomic, proteomic, and metabolic data during maize development has demonstrated that different regulatory layers form complex networks that depend on tissue type and developmental stage [46]. Studies of the genetic architecture of grain composition have also confirmed that the contents of major compound groups in the grain are controlled by multiple loci [47]. Other studies have mapped loci affecting carotenoid content [48], while manipulation of raffinose-family oligosaccharides has been shown to affect maize seed vigor [49]. These findings indicate that sugars, amino acids, organic acids, lipids, carotenoids, phenolic compounds, and metabolites associated with the antioxidant system may be related to grain yield or to traits that support yield formation.

Maize grain yield may be associated with variation in multiple groups of metabolites involved in carbon and nitrogen metabolism, grain filling, and plant responses to adverse environmental conditions [50]. Studies by other authors have placed particular emphasis on soluble sugars, starch, amino acids, organic acids, and phytohormones. Nitrogen deficiency may restrict assimilate transport and utilization, leading to starch accumulation in leaves and sugar accumulation in developing kernels, while simultaneously reducing kernel number and weight. These findings highlight the importance of the relationship between metabolite availability and the capacity of the grain as a sink organ [7]. Adequate nitrogen supply to maize has also been shown to affect sucrose and starch contents, the activity of enzymes involved in starch metabolism, endogenous hormone levels, and the course of grain filling [8]. The importance of amino acid metabolism was further confirmed by multi-omics studies in which *ZmAVT1A-1* was identified as a regulator of amino acid accumulation and nitrogen distribution between vegetative tissues and kernels; modifications of this process had pleiotropic effects on maize agronomic traits [9].

The present study did not examine the aforementioned groups of primary metabolites or phytohormones. Instead, it focused on phenolic acids, flavonoids, carotenoids, and the activity of selected oxidative and antioxidant enzymes. Ferulic, p-coumaric, caffeic, and diferulic acids are metabolites of the phenylpropanoid pathway, and ferulic and p-coumaric acids, as well as diferulates, occur in maize grain in both soluble and cell wall-bound forms. Their contents vary considerably among maize genotypes, suggesting that the profiles of these compounds may be genetically determined [10]. Changes in ferulic, p-coumaric, and

caffeic acid contents have also been observed in maize leaves subjected to water deficit, alongside changes in the expression of phenylpropanoid pathway genes, linking this group of metabolites to plant responses to environmental stress [11].

Luteolin and its derivatives measured in this study—luteolin-7-O-glucoside, luteolin-7-O-rutinoside, luteolin-7-O-glucuronide, and luteolin-7-O-(6″-malonyl)glucoside—belong to the flavone group, whose functions are associated with antioxidant protection and plant stress responses. Increased flavonoid accumulation in maize seedlings may limit oxidative damage during drought [42], whereas flavones secreted by maize roots may affect the composition of the rhizosphere microbiome and support plant development under limited nitrogen availability [51].

The third group analyzed comprised the carotenoids lutein, zeaxanthin, β-carotene, and α-carotene. These compounds function as photosynthetic and photoprotective pigments. Their content in grain is controlled by the activity of genes involved in carotenoid biosynthesis and degradation and varies markedly among maize genotypes [52]. Substantial variation in lutein, zeaxanthin, and β-carotene contents has also been reported among commercial maize hybrids [53].

Analysis of peroxidase (POD), polyphenol oxidase (PPO), and catalase (CAT) activities complements the metabolite measurements by providing information on the enzymatic oxidative response. POD and CAT participate in regulating reactive oxygen species levels, and their activities may change in response to environmental conditions and fertilization. In maize experiments, nitrogen supply has been shown to modify antioxidant enzyme activity, photosynthetic capacity, and grain yield [8]. PPO, in turn, participates in the oxidation of phenolic compounds, and variation in PPO and POD activities and their isoforms has been reported among maize genotypes differing in the phenolic content of ear silks [54].

In our study, metabolomic analysis revealed clear differences in biochemical profiles between traditional maize cultivars and modern, high-yielding SMH lines. These differences encompassed several groups of compounds and physiological parameters. Overall, the SMH lines exhibited higher contents of the metabolites analyzed and higher activities of polyphenol oxidase (PPO), catalase (CAT), and peroxidases (POD) than the traditional cultivars. The consistent direction of changes across multiple parameters indicates that the differences between the groups were not restricted to a single compound but reflected broader metabolic variation.

For phenolic acids, higher values were observed for caffeic, diferulic, ferulic, and p-coumaric acids. A similar pattern was found for the five flavonoids analyzed, comprising luteolin and its derivatives, as well as for the four carotenoids zeaxanthin, lutein, β-carotene, and α-carotene. At the same time, the SMH lines exhibited higher activities of enzymes involved in phenolic compound transformations and the regulation of oxidative balance. Taken together, these results indicate a distinct biochemical profile in the lines examined, encompassing both secondary metabolism and mechanisms associated with protection of the photosynthetic apparatus and control of oxidative processes. However, it should be emphasized that metabolite content or enzyme activity alone does not directly measure photosynthetic efficiency or stress tolerance.

The comparison between SMH 58324 and SMH 58024 is particularly interesting because these lines represent different combinations of biochemical traits and yield performance. SMH 58324 exhibited the highest values for most of the metabolites analyzed, including all four phenolic acids and the carotenoids, as well as high contents of the flavonoids examined. At the same time, its mean grain yield was 10.94 t·ha^{-1}, placing it last among the ten high-yielding SMH lines. In contrast, SMH 58024 achieved the highest mean grain yield across the entire panel, at 14.01 t·ha^{-1}, even though its metabolomic profile ranked third among the SMH lines. The discrepancy between metabolite accumulation and grain yield is important

for interpreting these results. It indicates that intensive accumulation of the compounds analyzed may be part of the characteristic metabolic profile of a particular genotype but does not necessarily translate directly into maximum productivity. The high yield of SMH 58024 may have been associated with a favorable combination of multiple physiological processes that are not fully reflected by a single measurement of leaf metabolite content. In contrast, the profile of SMH 58324 indicates substantial variation in the metabolism of the compounds examined. This discrepancy suggests that the metabolic profiles examined represent only selected components of the physiological characteristics of individual genotypes. Their relationships with grain yield may depend on other metabolic pathways, physiological processes, and interactions that were not directly assessed in the present study.

Multivariate PCA also revealed variation in the metabolomic profiles of the genotypes examined and partial separation of the SMH lines from the traditional cultivars. The particularly distant positions of SMH 58224 and SMH 58324 in the positive region of the principal component space indicated their biochemical distinctiveness. At the same time, the first principal component explained 99.79% of the total variation, whereas the second explained 0.19%. Biological interpretation is strengthened when evidence is consistent across several data levels. If a favorable marker allele co-occurs with a particular level of gene expression, a change in a metabolite belonging to the same pathway, and a favorable value for a yield component, this provides a sound basis for hypothesizing a role for the gene in question.

The results may have practical significance if the markers identified prove to be reproducible, easy to assay, and associated with effects sufficiently large to be useful in breeding. Molecular markers are used in commercial breeding programs [55], and the potential applications of marker-assisted selection (MAS) in maize improvement include the selection of quality and resistance traits [56]. For highly polygenic traits such as grain yield, however, genomic selection may outperform selection based on individual markers, as demonstrated for grain yield and physiological traits under different water-availability conditions [57]. The practical application of genomic knowledge was also demonstrated by Yao et al. [58], who used genomically assisted pyramiding of favorable alleles of genes regulating plant architecture to develop maize hybrids better adapted to high-density cultivation. Across experiments conducted in eight environments, the new hybrids achieved plot yields 4.1–9.2% higher than that of the original hybrid.

Gene identification and functional characterization may also provide targets for genome editing. Precision genome engineering enables modifications to coding and regulatory sequences [59], and CRISPR/Cas9 has been used in maize both to investigate gene function and to develop materials with altered architecture, yield, quality, or stress tolerance [60]. Newer systems, including base editing and prime editing, expand the range of possible modifications, although off-target effects, delivery methods, and regulatory considerations remain important [61].

For candidate genes previously identified by our research team (*CCR1*, *AAT*, *SUT1*, U-box regulators, and *WOX* genes), as well as the important candidate genes presented in the current study [*LOC103635969* (G5), *LOC100542095* (G10), *LOC103643269* (G13), *LOC103636139* (G32), *LOC100502264* (G15), *LOC100282937* (G26), and *LOC103637325* (G29)], functional validation should be the priority before designing modifications. Particular attention should be paid to pleiotropy, because modifying genes that control assimilate transport, nitrogen metabolism, or meristem development may improve one yield component at the expense of another trait.

A particularly relevant example is the editing of promoters of maize *CLE* genes. Liu et al. [13] used CRISPR–Cas9 to generate weak alleles of the *ZmCLE7* and *ZmFCP1* promoters, as well as alleles of the *ZmCLE1E5* gene, resulting in quantitative changes in meristem

size and increases in several yield-related traits. This study demonstrates that precisely tuning the activity of a regulatory gene may be more advantageous than completely eliminating its function. The value of identifying genes that regulate phenylpropanoid metabolism is not necessarily limited to the use of molecular markers. Oliveira et al. [62] used CRISPR/Cas9 to modify the maize *ZmPMT1* gene, which encodes p-coumaroyl-CoA transferase 1. The resulting mutation reduced lignin p-coumaroylation and altered its composition without reducing the total lignin content. This study demonstrates that precise editing of a single gene in the phenylpropanoid pathway can produce predictable changes in metabolite profiles and biomass properties. However, no direct effect of this modification on grain yield was demonstrated. For genes involved in multiple physiological processes, complete loss of function may lead to undesirable pleiotropic effects. An alternative is to edit cis-regulatory elements that control the level, timing, or tissue specificity of gene expression. Swinnen et al. [14] indicated that modifying such elements may enable the creation of new breeding alleles without altering the protein-coding sequence and may limit the consequences arising from the multiple effects of a gene. This strategy may be particularly useful for candidate genes whose activity needs to be quantitatively adjusted rather than completely eliminated.

4. Materials and Methods

4.1. Plant Material

The plant material comprised 20 reference genotypes characterized by different grain yield levels. Of these, ten genotypes (SMH 57023, SMH 57324, SMH 57524, SMH 57624, SMH 57724, SMH 57824, SMH 58024, SMH 57924, SMH 58224, and SMH 58324) were characterized by high grain yields, whereas the remaining ten (Szyldecka, Złota Górecka, Muszyński, Bydgoska Wczesna, Dar Północy, Małopolanka, Wawrzyńczycka, Lancaster, Wigor, and Wielkopolanka) exhibited low grain yields. The high-yielding genotypes were modern, selected maize hybrids, whereas the low-yielding genotypes represented traditional Polish open-pollinated maize cultivars cultivated in the mid-twentieth century. The plant material was obtained from Plant Breeding Smolice Ltd., IHAR Group (Smolice, Poland) (high-yielding genotypes), and the Plant Breeding and Acclimatization Institute—National Research Institute (Radzików, Poland) (low-yielding genotypes). The two genotype groups differed not only in grain yield but also in their breeding history and genetic background. Specifically, the high-yielding group consisted of modern breeding hybrids, whereas the low-yielding group comprised traditional open-pollinated cultivars developed and cultivated in the mid-twentieth century. Therefore, breeding status and yield group were not independently represented in the experimental material. This aspect of the experimental design was considered when interpreting the transcriptomic and metabolomic differences observed between the groups.

The genotypes were selected based on three years of observations conducted in experimental plots at HR Smolice. They were sown in three replicates on plots with an area of 10 m^2 and represented contrasting levels of grain yield across the three-year field experiment.

4.2. Methodology

4.2.1. Candidate Genes Associated with Maize Grain Yield Subjected to Functional Analysis

The candidate genes presented in Table 3 were identified by our research team between 2022 and 2025 using next-generation sequencing (NGS), association mapping, and physical mapping. The methods and results concerning candidate gene identification have been described in detail in the publications by Tomkowiak et al. [29], Nowak et al. [30], Tomkowiak et al. [31], and Lenort et al. [32]. Table 3 of the present study also includes genes that were not reported in the aforementioned publications but were found to be of interest during the functional analysis.

Table 3. Candidate genes selected by the Team in 2022–2025.

	Candidate Genes	Locus	Chromosomal Localization	Molecular Marker	Marker Type	Publication
1	Uncharacterized protein	*LOC100282883*	Chr.2	16,703	SilicoDArT	[29]
2	RNA polymerase II transcriptional coactivator *KELP*	*LOC100285609*	Chr.3	3233	SilicoDArT	[31]
3	Aspartate aminotransferase	*LOC100282625*	Chr.5	11,657	SilicoDArT	[31]
4	Sucrose transporter 1	*LOC100541615*	Chr.1	12,812	SilicoDArT	[31]
5	Putative protein phosphatase 2C	*LOC103635969*	Chr.8	28,630	SilicoDArT	[30]
6	Uncharacterized protein	*LOC100276990*	Chr.4	28,631	SilicoDArT	this study
7	Uncharacterized protein	*LOC100276990*	Chr.4	28,631	SNP	[30]
8	Uncharacterized ATP-dependent helicase *YprA*	*LOC103633409*	Chr.7	28,632	SNP	[30]
9	Ion protease homolog 2, peroxisomal	*LOC100384172*	Chr.7	28,633	SNP	[30]
10	Isoamylase-type starch debranching enzyme *ISO3* isoform	*LOC100542095*	Chr.7	28,634	SNP	[30]
11	Ribosomal protein L32-like	*LOC100194289*	Chr.1	31,977	SNP	[30]
12	Uncharacterized protein	*LOC103655564*	Chr.4	29,503	SNP	[30]
13	Probable arabinosyltransferase *ARAD1*	*LOC103643269*	Chr.1	28,625	SNP	[30]
14	Probable sugar phosphate/phosphate translocator *AT1G06470*	*LOC103638070*	Chr.9	28,640	SNP	[30]
15	Uncharacterized protein	*LOC100502264*	Chr.1	28,648	SNP	[30]
16	Uncharacterized protein	*LOC100382335*	Chr.8	28,639	SNP	[30]
17	Ubiquitin carboxyl-terminal hydrolase 15 isoform X2	*LOC103654750*	Chr.4	28,649	SNP	[30]
18	Uncharacterized protein	*LOC100193686*	Chr.4	30,773	SNP	[30]
19	Hydroxyproline O-galactosyltransferase *GALT*	*LOC103625798*	Chr.5	29,294	SNP	[30]
20	Uncharacterized protein	*LOC100383550*	Chr.9	28,262	SNP	[30]
21	Uncharacterized protein	*LOC103650272*	Chr.3	28,263	SNP	[30]
22	Alpha-N-acetylglucosaminidase	*LOC103639124*	Chr.9	19,463	SilicoDArT	this study
23	Putative calcium-dependent protein kinase family protein	*LOC100383602*	Chr.3	25,386	SilicoDArT	this study
24	Uncharacterized protein	*LOC100383055*	Chr.4	25,872	SilicoDArT	this study
25	Uncharacterized protein	*LOC100383055*	Chr.4	42,662	SNP	this study
26	RAF-type zinc finger domain-containing protein 1	*LOC100282937*	Chr.9	43,826	SNP	this study
27	Uncharacterized protein	*LOC100279888*	Chr.4	52,996	SNP	this study
28	Glycerol-3-phosphate 2-O-acyltransferase 6	*LOC103655222*	Chr.4	56,891	SNP	this study
29	Uncharacterized protein	*LOC103637325*	Chr.8	9,692,004	SNP	[32]
30	Sphingoid long-chain base kinase 1	*LOC103636090*	Chr.8	4,768,759	SilicoDArT	[32]

Table 3. *Cont.*

	Candidate Genes	Locus	Chromosomal Localization	Molecular Marker	Marker Type	Publication
31	Uncharacterized protein	*LOC100191174*	Chr.10	2,448,946	SilicoDArT	[32]
32	Uncharacterized protein	*LOC103636139*	Chr.8	2,492,509	SilicoDArT	[32]
33	U-box domain-containing protein 7	*LOC100383455*	Chr.8	5,587,791	SNP	[32]

4.2.2. Growth Conditions and Experimental Setup in the Phytotron

The experiment involving 20 maize genotypes subjected to gene expression analysis was conducted under controlled phytotron conditions, with three biological replicates per genotype (60 pots in total). Seeds were sown in disposable pots with a volume of 2 dm^3, filled with sterile peat substrate with a defined water-holding capacity and an adjusted pH of 6.5. The photoperiod was set to 16 h of light and 8 h of darkness. The temperature was maintained at 24 °C during the day and 18 °C at night. Relative air humidity was maintained at 60–70%. The light intensity was 600 $\mu mol \cdot m^{-2} \cdot s^{-1}$. Watering was performed regularly to maintain soil moisture at approximately 70% of the soil's field capacity.

4.2.3. Field Experimental Procedures

Field experiments involving 20 maize (*Zea mays* L.) genotypes grown for grain production were conducted in 2020–2022 in Smolice (51°42′36″ N, 17°09′36″ E), western Poland. The experiment was established using a randomized complete block design (RCBD) with three replicates. Each genotype constituted a separate experimental treatment, resulting in a total of 60 plots. Each plot covered an area of 10 m^2. Within each replicate, the treatments were arranged randomly. Sowing was performed in spring, at dates adjusted to soil temperature and moisture conditions (between 15 and 25 April). Before sowing, seedbed preparation was carried out, and mineral fertilizers were applied at rates determined on the basis of agronomic records and soil fertility. During the growing season, crop maintenance and herbicide treatments were performed in accordance with integrated plant protection principles. Grain was harvested in autumn, after the plants had reached maturity suitable for mechanical harvesting (between 10 and 20 November). Assessments and measurements were performed separately for each plot, which was treated as an experimental unit. Statistical analyses were conducted with genotype as the factor under investigation and replicate as the blocking effect.

4.2.4. RNA Isolation

Total RNA was isolated using the Maxwell® RSC automated instrument (Promega, Madison, WI, USA) and the Maxwell® RSC simplyRNA Tissue Kit (cat. No. AS1340), according to the manufacturer's instructions. Samples were homogenized in lysis buffer supplemented with 1-thioglycerol. The automated purification process included digestion of genomic DNA with DNase I directly in the cartridge. The isolated RNA was eluted in 30 µL of nuclease-free water. RNA concentration and quality were assessed using a NanoDrop spectrophotometer.

4.2.5. Reverse Transcription

Complementary DNA (cDNA) was synthesized using reagents from Bio-Rad (Hercules, CA, USA). A total of 500 ng of RNA was used for each reaction. Reverse transcription was performed in a final volume of 20 µL using the iScript Reverse Transcription Supermix kit (Bio-Rad, cat. No. 1708841), according to the manufacturer's protocol. The reaction was

carried out in a thermal cycler using the following temperature profile: initial priming for 5 min at 25 °C, reverse transcription for 35 min at 45 °C, enzyme inactivation for 5 min at 85 °C, followed by cooling and maintenance at 4 °C. The resulting cDNA products were stored at −20 °C until further analysis by RT-qPCR.

4.2.6. RT-qPCR Analysis

The reactions were performed according to the manufacturer's instructions, using a constant amount of cDNA template equal to 200 ng in a total reaction volume of 20 µL. cDNA amplification was carried out using the CFX Opus 96 Real-Time PCR System (Bio-Rad Laboratories; cat. No. 12011319). The reactions were performed using iTaq Universal SYBR® Green Supermix (Bio-Rad Laboratories; cat. No. 1725120). Reactions were conducted in 96-well Hard-Shell PCR Plates (Bio-Rad Laboratories; cat. No. HSP9655), sealed with Microseal 'B' PCR Plate Sealing Film (Bio-Rad Laboratories; cat. No. MSB1001). Each measurement was performed in three technical replicates. The thermal cycling conditions consisted of an initial denaturation step of 30 s at 95 °C, followed by 40 amplification cycles comprising 10 s at 95 °C and 30 s at 60 °C. After amplification, melting curve analysis was performed over the range of 65–95 °C, with increments of 0.5 °C, to confirm the specificity of the reaction products. Raw fluorescence data were analyzed, and quantification cycle (Cq) values were determined using CFX Maestro Software (Bio-Rad Laboratories; cat. No. 12013758). The stability of reference gene expression was assessed using the Reference Gene Selector module integrated into CFX Maestro software v. 2.3. Changes in the expression levels of target genes relative to reference genes were calculated using the comparative $2^{-\Delta\Delta Cq}$ method, with reaction efficiencies determined from standard curves for each gene analyzed. The primers presented in Table 4 were designed to span exon–exon junctions.

Table 4. Sequences of designed primers used for RT-qPCR analyses.

	Candidate Genes	Locus	Forward	Reverse	Product Length	Tm
1	Uncharacterized protein	*LOC100282883*	CTCGAAC-GAAAACAGGGGCA	AGCTGGTGAC-CATTTGGCAT	102	59
2	RNA polymerase II transcriptional coactivator *KELP*	*LOC100285609*	GCCCTCCGC-CAAAGGTATTA	CCGCTGGAACAT-CACCAAAG	220	59
3	Aspartate aminotransferase	*LOC100282625*	GCCAACGTTGTTG-GTGATCC	TCATGTTGT-CACTCTGCGCT	205	59
4	Sucrose transporter 1	*LOC100541615*	CCAGAACCTCCC-CACAAAGG	GGAACCAAGA-GAGCCAGGTG	143	60
5	Putative protein phosphatase 2C	*LOC103635969*	CGCCAACCAAGC-CGGAAT	TTCCC-CTCCAGGGCTTGTT	248	60
6	Uncharacterized protein	*LOC100276990*	CTCGAAC-GAAAACAGGGGCA	AGCTGGTGAC-CATTTGGCAT	115	60
7	Uncharacterized protein	*LOC100276990*	GCTTCTTGGC-CAAATGCTGT	TCATGCGCA-CAATTCGTTGG	109	59
8	Uncharacterized ATP-dependent helicase *YprA*	*LOC103633409*	TCACAGGCC-CAAATGCTCTT	GCACGTCAGT-GATTGGATGC	114	59
9	Ion protease homolog 2, peroxisomal	*LOC100384172*	TTCTCTCTGGCATG-GAGATTC	AGCCTAA-CAAAAATGGGCGT	235	57
10	Isoamylase-type starch debranching enzyme *ISO3* isoform	*LOC100542095*	GCCCATCTTGC-TATCCTCCC	GAGTGTGATCG-GACGCTTGG	135	59

Table 4. *Cont.*

	Candidate Genes	Locus	Forward	Reverse	Product Length	Tm
11	Ribosomal protein L32-like	*LOC100194289*	GCTCATGATGCA-CAACAGAACC	ATGGCAGGATC-TATTTGGCCG	187	60
12	Uncharacterized protein	*LOC103655564*	TTGTTTATC-CACGGGGGCAG	AGACGATGGCTCA-CAAAGCA	112	60
13	Probable arabinosyltransferase *ARAD1*	*LOC103643269*	CGTGTTGTCT-GACTTCGGGA	GAACCTTCC-CGCCAGCTTTC	179	60
14	Probable sugar phosphate/phosphate translocator *AT1G06470*	*LOC103638070*	CAACTGC-CACAGGCTTGTTC	CCGATGGAGT-GCTTGACAGT	231	60
15	Uncharacterized protein	*LOC100502264*	CATAGCTCTTTGTG-GCTGGC	GGGGCAAATG-GAACTTCAGC	167	59
16	Uncharacterized protein	*LOC100382335*	CAACTGC-CACAGGCTTGTTC	CCGATGGAGT-GCTTGACAGT	231	60
17	Ubiquitin carboxyl-terminal hydrolase 15 isoform X2	*LOC103654750*	CCAGCACCAGCG-GTGAAAT	GTGTGCACCTC-CATACGTCA	197	60
18	Uncharacterized protein	*LOC100193686*	CGTCT-GCTCAGGAACCGC	CGCGTG-TACGCGTTTAGAAT	143	60
19	Hydroxyproline O-galactosyltransferase *GALT*	*LOC103625798*	TGCGAG-TATGGGGTCCATGT	CACGTAATG-GCTTGTGGTGA	160	58
20	Uncharacterized protein	*LOC100383550*	CTCAGCCCAT-ACAGGCTTCT	AGACAACAGACAT-GCCCCAG	118	59
21	Uncharacterized protein	*LOC103650272*	TCTC-CGTTGTCAGC-GAATCA	GTTCTTACTTC-CGCCAGGCT	193	60
22	Alpha-N-acetylglucosaminidase	*LOC103639124*	GGGTGGACCACT-GTCACAAA	ACGATGGAA-GAACAGGGACC	104	60
23	Putative calcium-dependent protein kinase family protein	*LOC100383602*	AAGGAC-CATAGCGGGTACA	CATGACAGTTA-GAACATGCGTCTC	209	57
24	Uncharacterized protein (SilicoDArT)	*LOC100383055*	AGAGTGTCAAC-CTGGGGGTA	CCAACCGGAAT-GTCCACGTA	110	60
25	Uncharacterized protein (SNP)	*LOC100383055*	AGAGTGTCAAC-CTGGGGGTA	CCAACCGGAAT-GTCCACGTA	110	60
26	RAF-type zinc finger domain-containing protein 1	*LOC100282937*	GTGCCCACAAA-GAATCGTCG	GGAATTGCTCTG-GCACTGGA	210	60
27	Uncharacterized protein	*LOC100279888*	TGCTGGTGTAGC-TAAGGTTGAT	AAGCAT-CAAGGGGATCAA-GAGG	200	60
28	Glycerol-3-phosphate 2-O-acyltransferase 6	*LOC103655222*	ACAAGTACCGGAT-GATCGCC	TTCTGTGCAT-GAGTCCCCAC	113	60
29	Uncharacterized protein	*LOC103637325*	CCATTTGATCAAG-GTGGTGCG	TTGCGCCGCTGGA-GATG	240	60
30	Sphingoid long-chain base kinase 1	*LOC103636090*	CACCCCATG-GCATCTAGCAA	GTCCTGTTGTGCA-GAACCTTAAA	191	60
31	Uncharacterized protein	*LOC100191174*	GAAGTTCCT-GAGCAGTTCCCA	CCTTTTCCATTTGC-CTCGACC	158	60
32	Uncharacterized protein	*LOC103636139*	GGCTCCTCTGC-CAAGATACT	GAACTGTTGGCT-TAACTGGCA	237	58
33	U-box domain-containing protein 7	*LOC100383455*	TCTGACTGGCTCT-GAAGACG	TACCTGAGCTC-CAACATCCAG	223	59

4.2.7. Selection and Analysis of Reference Genes

The selection of reference genes began with a literature review aimed at identifying genes exhibiting high expression stability in the biological material under investigation. In the present study, an initial group of candidate genes was selected, taking into account genes commonly used in maize research. Ultimately, three genes were selected: elongation factor 1-alpha (*EF1α*), beta-tubulin (*β-TUB*), and cyclophilin (*CYP*). For each candidate reference gene, standard curve analysis was performed to determine the reaction efficiency (E) and correlation coefficient (R^2). These parameters are essential for assessing the technical performance of the qRT-PCR assay before proceeding with data normalization. Based on the resulting standard curves, genes exhibiting the most favorable reaction parameters were selected. The key steps of the procedure included qRT-PCR analysis of the selected candidates using test cDNA as the template, followed by comparison of the stability of the selected reference genes across the test samples using the dedicated Gene Study software module in the CFX Maestro system. Finally, the selected reference genes, beta-tubulin (*β-TUB*) and cyclophilin (*CYP*), were validated to assess their suitability for data normalization under the specific experimental conditions.

4.2.8. Metabolite Analyses

Determination of Phenolic Compounds

Phenolic compounds in samples were analyzed after alkaline and acidic hydrolysis [63]. Analysis was performed using an Acquity H class UPLC system equipped with a Waters Acquity PDA detector (Waters, Milford, MA, USA). Chromatographic separation was performed on an Acquity UPLC® BEH C18 column (100 mm × 2.1 mm, particle size 1.7 µm) (Waters, Ireland). Gradient elution was performed using the following mobile phases: A: acetonitrile with 0.1% formic acid, B: 1% aqueous formic acid mixture (pH = 2). Concentrations of phenolic compounds were determined using an internal standard at wavelengths λ = 320 nm and 280 nm. Compounds were identified by comparing the retention time of each analyzed peak with that of the corresponding standard and by adding a specific amount of the standard to the analyzed samples and a repeated analysis. The detection limit was 1 $\mu g \cdot g^{-1}$. Flavonoids (luteolin-7-O-glucuronide, luteolin-7-O-(6″-malonyl)-glucoside, luteolin-7-O-glycoside, luteolin-7-O-rutinoside). Phenolic acids (p-coumaric acid, caffeic acid, diferulic acid) were analyzed [64].

Determination of Carotenoids

Carotenoid isolation and quantification in leaf samples were performed using an Acquity UPLC system (Waters, Milford, MA, USA) following the saponification method. Carotenoid extracts were obtained from ground plant material (0.4 g), which was triturated with a mixture of acetone and petroleum ether (1:1). Then, after separation of the tested material, the acetone and the hydrophilic fraction were removed from the extract by washing with water. As a result, the ether extract was obtained with a mixture of carotenoid pigments. The extract thus prepared was concentrated in a vacuum evaporator at 35 °C until an oily residue was obtained, then digested in 2 mL of methanol (Merck, Darmstadt, Germany) and subjected to chromatographic analysis. Lutein, zeaxanthin and β-carotene were determined using Acquity UPLC (Waters, Milford, MA, USA) with a Waters Acquity PDA detector (Waters, Milford, MA, USA). Chromatographic separation was performed on an Acquity UPLC® BEH C18 column (100 mm × 2.1 mm, particle size 1.7 µm) (Waters, Dublin, Ireland). Elution was carried out using solvent A—methanol, B—water and tert-butyl methyl ether (TBME). A gradient was applied at a flow of 0.4 $mL \cdot min^{-1}$. The column and samples were thermostated, the column temperature was 30 °C, the test temperature was 10 °C. During the analysis, the solutions were degassed in a Waters device. The injection volume was

10 µL. The registration was carried out at a wavelength λ = 445 nm. Compounds were identified based on spectra in the range from 200 to 600 nm and retention times compared to standards. The following metabolites were analyzed: zeaxanthin, lutein, β-carotene, α-carotene [65].

Determination of Antioxidant Enzyme Activities

The activities of catalase (CAT; EC 1.11.1.6), guaiacol peroxidase (POD; EC 1.11.1.7), and polyphenol oxidase (PPO) were determined spectrophotometrically in plant extracts. Fresh plant material samples (0.5 g) were homogenized on ice in 5 mL of 50 mM potassium phosphate buffer (pH 7.0) containing 1% (*w*/*v*) polyvinylpolypyrrolidone (PVPP). The homogenates were centrifuged at 12,000 × *g* for 20 min at 4 °C, and the resulting supernatants were collected for enzyme activity assays. CAT activity was determined according to the spectrophotometric principle described by Aebi [66], by monitoring the decrease in absorbance at 240 nm resulting from hydrogen peroxide decomposition. POD activity was determined using guaiacol as a substrate, following the method of Chance and Maehly [67], with modifications. The reaction mixture contained 50 mM phosphate buffer (pH 6.5), 4 mM guaiacol, and 1 mM H_2O_2. Enzyme activity was monitored as the increase in absorbance at 470 nm. PPO activity was determined using catechol as a substrate in phosphate buffer (pH 7.0). The oxidation reaction was monitored spectrophotometrically by recording the increase in absorbance at 420 nm. Enzyme activities were calculated from the initial linear rates of absorbance change, corrected for the corresponding blanks, and expressed as $U \cdot g^{-1}$.

4.2.9. Statistical Analysis

The consistency of the empirical distribution with the normal distribution was tested using the Shapiro–Wilk test [68]. The assessment of homogeneity of variances across genotype-determined groups was performed using Bartlett's test. The two-way analysis of variance was performed to evaluate the effect of genotypes, years and genotype × year interaction on the grain yield. Linear regression analysis was used to assess the impact of the expression of individual genes on the levels of the metabolites under consideration. The significance of associations was evaluated using *p*-values adjusted for multiple testing with the Benjamini–Hochberg procedure. The Benjamini–Hochberg method is a statistical technique that controls the false discovery rate (FDR) during multiple hypothesis testing. It is less stringent than the Bonferroni correction, offering greater power to detect significant differences while reducing the false positive rate. The functional-transcriptional groups were established based on differences in gene expression among maize genotypes, as determined by one-way ANOVA followed by Tukey's honestly significant difference (HSD) post hoc test ($p < 0.05$). All statistical analyses were performed using the Genstat v. 24.2 statistical software package [69].

5. Conclusions

The present study revealed associations between variation in maize grain yield, candidate gene expression profiles, and selected leaf biochemical traits, including phenolic acids, flavonoids, carotenoids, and the activities of selected oxidoreductive enzymes. The integration of phenotypic, transcriptomic, and metabolomic data provided insights into molecular and biochemical patterns that co-occur with differences in yield potential. However, the observed relationships do not establish direct causal mechanisms, and the contribution of individual genes and metabolites to grain yield remains to be experimentally verified. The identified associations therefore provide a basis for further functional investigations into the complex genetic and physiological regulation of maize grain yield. The three-year field assessment confirmed substantial variation in grain yield among the genotypes exam-

ined. The highest mean yield was recorded for line SMH 58024 (14.01 t·ha^{-1}), while the significant genotype × year interaction indicated that genotypes responded differently to the conditions prevailing during individual growing seasons. Analysis of the expression of 33 candidate genes revealed that their transcriptional patterns did not form a uniform profile characteristic of all high-yielding genotypes. The genes G5 (*LOC103635969*), G10 (*LOC100542095*), and G13 (*LOC103643269*) emerged as particularly interesting candidates for further functional validation, as their elevated expression co-occurred with high grain yield in line SMH 58024. At the same time, the high expression of selected genes observed in the low-yielding cultivar Szyldecka confirms that the activity of an individual gene is insufficient to explain the complex grain yield phenotype.

Metabolomic analysis revealed that the SMH lines generally exhibited higher contents of the phenolic acids, flavonoids, and carotenoids analyzed, as well as higher PPO, CAT, and POD activities, than the traditional maize cultivars. Line SMH 58324 was particularly distinctive, exhibiting the highest values for most of the metabolites analyzed, although its mean grain yield was 10.94 t·ha^{-1}, the lowest among the ten SMH lines. In contrast, SMH 58024 achieved the highest yield across the entire genotype panel, despite not exhibiting the most pronounced metabolomic profile in terms of the accumulation of the compounds examined. These findings indicate that elevated contents of secondary metabolites and increased oxidoreductive enzyme activities constitute components of the biochemical characteristics of maize genotypes but cannot be considered independent, direct indicators of grain yield.

The most pronounced relationships in the integrated transcriptomic–metabolomic analysis were identified for genes G1, G15, G26, G29, and G32, whose expression was negatively associated with the metabolic and enzymatic parameters examined. Particular attention was drawn to G32 (*LOC103636139*), whose expression accounted for 42.8–48.7% of the variation in selected biochemical parameters. These results indicate its potential value as a candidate for investigating the relationship between transcriptional regulation and leaf metabolism. However, they do not demonstrate that G32 directly suppresses the metabolic pathways examined or that its expression causally determines grain yield. The unknown molecular function of the encoded protein and the possibility of indirect relationships warrant further functional studies. PCA additionally confirmed substantial variation in gene expression and metabolomic profiles but did not reveal an unambiguous separation of all genotypes according to yield potential.

In conclusion, the results indicate that differences in maize grain yield should be considered a consequence of interactions among multiple genetic, regulatory, and physiological factors rather than the activity of a single gene or the accumulation of a particular group of metabolites. The identified genes G5, G10, G13, G15, G26, G29, and G32 represent candidates for further functional characterization. While some of these genes have putative functional annotations, the molecular functions of G15, G29, and G32 remain unknown and require additional bioinformatic and experimental investigation. Their potential application in marker-assisted selection or precision genome editing should therefore be considered only after their functions and relationships with yield-related traits have been independently validated. Once these relationships have been confirmed, selected genes and markers may be applicable to improving the selection of breeding materials and, in the longer term, to precision genome editing. Before practical implementation, however, the identified relationships must be validated in larger populations and across diverse environments, and it must be demonstrated that modifying gene activity produces reproducible, beneficial effects on grain yield without undesirable pleiotropic effects.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms272210459/s1.

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

Funding: The research presented in this publication was financed as part of the research project "Analysis of genetic determinants of heterosis effect and Fusarium resistance in maize (*Zea mays* L.)". PL: "Analiza genetycznych uwarunkowań związanych z efektem heterozji oraz odpornością na Fuzarium u kukurydzy (*Zea mays* L.)". The project is implemented under the grant from the Ministry of Agriculture and Rural Development, "Biological progress in plant production (recruitment 2020)". Duration of the project: 2021–2026.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Duvick, D.N. The Contribution of Breeding to Yield Advances in Maize (*Zea mays* L.). *Adv. Agron.* **2005**, *86*, 83–145. [CrossRef]
2. Messina, C.D.; Rotundo, J.; Hammer, G.L.; Gho, C.; Reyes, A.; Fang, Y.; van Oosterom, E.; Borras, L.; Cooper, M. Radiation use efficiency increased over a century of maize (*Zea mays* L.) breeding in the US corn belt. *J. Exp. Bot.* **2022**, *73*, 5503–5513. [CrossRef] [PubMed]
3. Cagnola, J.I.; Rotili, D.H.; Otegui, M.E.; Casal, J.J. 50 Years of Breeding to Improve Yield: How Maize Stands up to Climate Change. *Philos. Trans. R. Soc. B Biol. Sci.* **2025**, *380*, 20240250. [CrossRef] [PubMed]
4. Liu, L.; Fang, X.; Ren, S.; Jia, R.; Liu, Q.; Liu, H.; Xiu, L.; Yaqoob, S.; Cai, D.; Liu, J. Targeted Metabolic Reveals Different Part of Maize in Polyphenolic Metabolites during Germination and Hypoglycemic Activity Analysis. *Food Chem. X* **2023**, *19*, 100848. [CrossRef] [PubMed]
5. Li, X.; Zhao, X.; Sun, S.; He, M.; Wang, J.; Xiang, X.; Niu, Y. Genetic Architecture and Meta-QTL Identification of Yield Traits in Maize (*Zea mays* L.). *Plants* **2025**, *14*, 3067. [CrossRef] [PubMed]
6. Sahito, J.H.; Zhang, H.; Gishkori, Z.G.N.; Ma, C.; Wang, Z.; Ding, D.; Zhang, X.; Tang, J. Advancements and Prospects of Genome-Wide Association Studies (GWAS) in Maize. *Int. J. Mol. Sci.* **2024**, *25*, 1918. [CrossRef] [PubMed]
7. Ning, P.; Yang, L.; Li, C.; Fritschi, F.B. Post-Silking Carbon Partitioning under Nitrogen Deficiency Revealed Sink Limitation of Grain Yield in Maize. *J. Exp. Bot.* **2018**, *69*, 1707–1719. [CrossRef] [PubMed]
8. Yue, K.; Li, L.; Xie, J.; Liu, Y.; Xie, J.; Anwar, S.; Fudjoe, S.K. Nitrogen Supply Affects Yield and Grain Filling of Maize by Regulating Starch Metabolizing Enzyme Activities and Endogenous Hormone Contents. *Front. Plant Sci.* **2022**, *12*, 798119, Erratum in *Front Plant Sci.* **2022**, *13*, 882425. [CrossRef] [PubMed]
9. Jin, M.; Yan, S.; Wu, Y.; Zhai, Z.; Alseekh, S.; Chen, Y.; Huang, W.; Ma, N.; Tao, K.; Xiao, K.; et al. Multiomics Analysis of Primary Metabolism Reveals the Genetic Basis of Nitrogen Partitioning Modulated by ZmAVT1A-1 in Maize. *Nat. Genet.* **2026**, *58*, 1722–1733. [CrossRef] [PubMed]
10. Zavala-López, M.; Flint-García, S.; García-Lara, S. Compositional Variation in Trans-Ferulic, p-Coumaric, and Diferulic Acids Levels Among Kernels of Modern and Traditional Maize (*Zea mays* L.) Hybrids. *Front. Nutr.* **2020**, *7*, 600747. [CrossRef] [PubMed]
11. Kolo, Z.; Majola, A.; Phillips, K.; Ali, A.E.E.; Sharp, R.E.; Ludidi, N. Water Deficit-Induced Changes in Phenolic Acid Content in Maize Leaves Is Associated with Altered Expression of Cinnamate 4-Hydroxylase and p-Coumaric Acid 3-Hydroxylase. *Plants* **2023**, *12*, 101. [CrossRef] [PubMed]
12. Chen, W.; Zhang, X.; Lu, C.; Chang, H.; Chachar, Z.; Fan, L.; An, Y.; Li, X.; Qi, Y. Genome-Wide Association Study of Carotenoids in Maize Kernel. *Plant Genome* **2024**, *17*, e20495. [CrossRef] [PubMed]
13. Liu, L.; Gallagher, J.; Arevalo, E.D.; Chen, R.; Skopelitis, T.; Wu, Q.; Bartlett, M.; Jackson, D. Enhancing Grain-Yield-Related Traits by CRISPR–Cas9 Promoter Editing of Maize CLE Genes. *Nat. Plants* **2021**, *7*, 287–294. [CrossRef] [PubMed]

14. Swinnen, G.; Goossens, A.; Pauwels, L. Lessons from Domestication: Targeting Cis-Regulatory Elements for Crop Improvement. *Trends Plant Sci.* **2016**, *21*, 506–515. [CrossRef] [PubMed]
15. Marand, A.P.; Jiang, L.; Gomez-Cano, F.; Minow, M.A.A.; Zhang, X.; Mendieta, J.P.; Luo, Z.; Bang, S.; Yan, H.; Meyer, C.; et al. The Genetic Architecture of Cell Type–Specific Cis Regulation in Maize. *Science* **2025**, *388*, eads6601. [CrossRef] [PubMed]
16. Erenstein, O.; Chamberlin, J.; Sonder, K. Estimating the Global Number and Distribution of Maize and Wheat Farms. *Glob. Food Secur.* **2021**, *30*, 100558. [CrossRef]
17. Erenstein, O.; Jaleta, M.; Sonder, K.; Mottaleb, K.; Prasanna, B.M. Global Maize Production, Consumption and Trade: Trends and R&D Implications. *Food Secur.* **2022**, *14*, 1295–1319. [CrossRef]
18. Martinez-Feria, R.A.; Basso, B. Unstable Crop Yields Reveal Opportunities for Site-Specific Adaptations to Climate Variability. *Sci. Rep.* **2020**, *10*, 2885. [CrossRef] [PubMed]
19. Long, Y.; Wang, C.; Liu, C.; Li, H.; Pu, A.; Dong, Z.; Wei, X.; Wan, X. Molecular Mechanisms Controlling Grain Size and Weight and Their Biotechnological Breeding Applications in Maize and Other Cereal Crops. *J. Adv. Res.* **2024**, *62*, 27–46. [CrossRef] [PubMed]
20. Zhang, H.; Lu, Y.; Ma, Y.; Fu, J.; Wang, G. Genetic and Molecular Control of Grain Yield in Maize. *Mol. Breed.* **2021**, *41*, 18. [CrossRef] [PubMed]
21. Millet, E.J.; Kruijer, W.; Coupel-Ledru, A.; Alvarez Prado, S.; Cabrera-Bosquet, L.; Lacube, S.; Charcosset, A.; Welcker, C.; van Eeuwijk, F.; Tardieu, F. Genomic Prediction of Maize Yield across European Environmental Conditions. *Nat. Genet.* **2019**, *51*, 952–956. [CrossRef] [PubMed]
22. Xu, Y.; Li, P.; Yang, Z.; Xu, C. Genetic Mapping of Quantitative Trait Loci in Crops. *Crop J.* **2017**, *5*, 175–184. [CrossRef]
23. Cooper, M.; Technow, F.; Messina, C.; Gho, C.; Totir, L.R. Use of Crop Growth Models with Whole-Genome Prediction: Application to a Maize Multienvironment Trial. *Crop Sci.* **2016**, *56*, 2141–2156. [CrossRef]
24. Malosetti, M.; Ribaut, J.-M.; van Eeuwijk, F.A. The Statistical Analysis of Multi-Environment Data: Modeling Genotype-by-Environment Interaction and Its Genetic Basis. *Front. Physiol.* **2013**, *4*, 44. [CrossRef] [PubMed]
25. van Eeuwijk, F.A.; Bink, M.C.; Chenu, K.; Chapman, S.C. Detection and Use of QTL for Complex Traits in Multiple Environments. *Curr. Opin. Plant Biol.* **2010**, *13*, 193–205. [CrossRef] [PubMed]
26. Kremling, K.A.G.; Diepenbrock, C.H.; Gore, M.A.; Buckler, E.S.; Bandillo, N.B. Transcriptome-Wide Association Supplements Genome-Wide Association in *Zea mays*. *G3 GenesGenomesGenetics* **2019**, *9*, 3023–3033. [CrossRef] [PubMed]
27. Guo, Z.; Magwire, M.M.; Basten, C.J.; Xu, Z.; Wang, D. Evaluation of the Utility of Gene Expression and Metabolic Information for Genomic Prediction in Maize. *Theor. Appl. Genet.* **2016**, *129*, 2413–2427. [CrossRef] [PubMed]
28. Ajayo, B.S.; Huang, Y.; Hu, Y. From Development to Yield: Genetic and Molecular Regulation of Agronomic Traits in Maize Seeds. *Front. Plant Sci.* **2026**, *17*, 1754331. [CrossRef] [PubMed]
29. Tomkowiak, A.; Nowak, B.; Sobiech, A.; Bocianowski, J.; Wolko, Ł.; Spychała, J. The Use of DArTseq Technology to Identify New SNP and SilicoDArT Markers Related to the Yield-Related Traits Components in Maize. *Genes* **2022**, *13*, 848. [CrossRef] [PubMed]
30. Nowak, B.; Tomkowiak, A.; Sobiech, A.; Bocianowski, J.; Kowalczewski, P.Ł.; Spychała, J.; Jamruszka, T. Identification and Analysis of Candidate Genes Associated with Yield Structure Traits and Maize Yield Using Next-Generation Sequencing Technology. *Genes* **2024**, *15*, 56. [CrossRef] [PubMed]
31. Tomkowiak, A.; Jamruszka, T.; Bocianowski, J.; Sobiech, A.; Jarzyniak, K.; Lenort, M.; Mikołajczyk, S.; Żurek, M. Transcriptomic Characterization of Genes Harboring Markers Linked to Maize Yield. *Genes* **2024**, *15*, 1558. [CrossRef] [PubMed]
32. Lenort, M.; Tomkowiak, A.; Bocianowski, J.; Bobrowska, R.; Kurasiak-Popowska, D.; Mikołajczyk, S.; Kosiada, T.; Weigt, D.; Gawrysiak, P. Using Genome-Wide Association Studies to Reveal DArTseq and SNP Loci Associated with Agronomic Traits and Yield in Maize. *Curr. Issues Mol. Biol.* **2025**, *47*, 1008. [CrossRef] [PubMed]
33. Tamasloukht, B.; Wong Quai Lam, M.S.-J.; Martinez, Y.; Tozo, K.; Barbier, O.; Jourda, C.; Jauneau, A.; Borderies, G.; Balzergue, S.; Renou, J.-P.; et al. Characterization of a Cinnamoyl-CoA Reductase 1 (CCR1) Mutant in Maize: Effects on Lignification, Fibre Development, and Global Gene Expression. *J. Exp. Bot.* **2011**, *62*, 3837–3848. [CrossRef] [PubMed]
34. Ana, L.-M.; Rogelio, S.; Xose Carlos, S.; Rosa Ana, M. Cell Wall Composition Impacts Structural Characteristics of the Stems and Thereby the Biomass Yield. *J. Agric. Food Chem.* **2022**, *70*, 3136–3141. [CrossRef] [PubMed]
35. Schlüter, U.; Bräutigam, A.; Droz, J.-M.; Schwender, J.; Weber, A.P.M. The Role of Alanine and Aspartate Aminotransferases in C4 Photosynthesis. *Plant Biol.* **2019**, *21*, 64–76. [CrossRef] [PubMed]
36. Zhou, Y.; Cai, H.; Xiao, J.; Li, X.; Zhang, Q.; Lian, X. Over-Expression of Aspartate Aminotransferase Genes in Rice Resulted in Altered Nitrogen Metabolism and Increased Amino Acid Content in Seeds. *Theor. Appl. Genet.* **2009**, *118*, 1381–1390. [CrossRef] [PubMed]
37. Baker, R.F.; Leach, K.A.; Boyer, N.R.; Swyers, M.J.; Benitez-Alfonso, Y.; Skopelitis, T.; Luo, A.; Sylvester, A.; Jackson, D.; Braun, D.M. Sucrose Transporter ZmSut1 Expression and Localization Uncover New Insights into Sucrose Phloem Loading. *Plant Physiol.* **2016**, *172*, 1876–1898. [CrossRef] [PubMed]

38. Slewinski, T.L.; Meeley, R.; Braun, D.M. Sucrose Transporter1 Functions in Phloem Loading in Maize Leaves. *J. Exp. Bot.* **2009**, *60*, 881–892. [CrossRef] [PubMed]
39. Carpaneto, A.; Geiger, D.; Bamberg, E.; Sauer, N.; Fromm, J.; Hedrich, R. Phloem-Localized, Proton-Coupled Sucrose Carrier ZmSUT1 Mediates Sucrose Efflux under the Control of the Sucrose Gradient and the Proton Motive Force. *J. Biol. Chem.* **2005**, *280*, 21437–21443. [CrossRef] [PubMed]
40. Trujillo, M. News from the PUB: Plant U-Box Type E3 Ubiquitin Ligases. *J. Exp. Bot.* **2018**, *69*, 371–384. [CrossRef] [PubMed]
41. Trenner, J.; Monaghan, J.; Saeed, B.; Quint, M.; Shabek, N.; Trujillo, M. Evolution and Functions of Plant U-Box Proteins: From Protein Quality Control to Signaling. *Annu. Rev. Plant Biol.* **2022**, *73*, 93–121. [CrossRef] [PubMed]
42. Li, J.; Zhang, Y.; Gao, Z.; Xu, X.; Wang, Y.; Lin, Y.; Ye, P.; Huang, T. Plant U-Box E3 Ligases PUB25 and PUB26 Control Organ Growth in Arabidopsis. *New Phytol.* **2021**, *229*, 403–413. [CrossRef] [PubMed]
43. Shimizu, R.; Ji, J.; Kelsey, E.; Ohtsu, K.; Schnable, P.S.; Scanlon, M.J. Tissue Specificity and Evolution of Meristematic WOX3 Function. *Plant Physiol.* **2009**, *149*, 841–850. [CrossRef] [PubMed]
44. Zhao, Y.; Hu, Y.; Dai, M.; Huang, L.; Zhou, D.-X. The WUSCHEL-Related Homeobox Gene WOX11 Is Required to Activate Shoot-Borne Crown Root Development in Rice. *Plant Cell* **2009**, *21*, 736–748. [CrossRef] [PubMed]
45. Li, D.; Li, Y.; Xing, A.; Zhang, S.; Zhu, Y.; Wei, F.; Zeng, H.; Zheng, T.; Hu, Z.; Wang, C. ZmDLR9 is required for lateral root development via pre-rRNA processing in maize. *Plant J.* **2026**, *128*, e71156. [CrossRef] [PubMed]
46. Walley, J.W.; Sartor, R.C.; Shen, Z.; Schmitz, R.J.; Wu, K.J.; Urich, M.A.; Nery, J.R.; Smith, L.G.; Schnable, J.C.; Ecker, J.R.; et al. Integration of Omic Networks in a Developmental Atlas of Maize. *Science* **2016**, *353*, 814–818. [CrossRef] [PubMed]
47. Cook, J.P.; McMullen, M.D.; Holland, J.B.; Tian, F.; Bradbury, P.; Ross-Ibarra, J.; Buckler, E.S.; Flint-Garcia, S.A. Genetic Architecture of Maize Kernel Composition in the Nested Association Mapping and Inbred Association Panels. *Plant Physiol.* **2012**, *158*, 824–834. [CrossRef] [PubMed]
48. Li, T.; Zhang, Y.; Wang, D.; Liu, Y.; Dirk, L.M.A.; Goodman, J.; Downie, A.B.; Wang, J.; Wang, G.; Zhao, T. Regulation of Seed Vigor by Manipulation of Raffinose Family Oligosaccharides in Maize and *Arabidopsis thaliana*. *Mol. Plant* **2017**, *10*, 1540–1555. [CrossRef] [PubMed]
49. Chander, S.; Guo, Y.Q.; Yang, X.H.; Zhang, J.; Lu, X.Q.; Yan, J.B.; Song, T.M.; Rocheford, T.R.; Li, J.S. Using Molecular Markers to Identify Two Major Loci Controlling Carotenoid Contents in Maize Grain. *Theor. Appl. Genet.* **2008**, *116*, 223–233. [CrossRef] [PubMed]
50. Bocianowski, J. Beyond Protein Abundance: Proteoform Diversity and Phosphoproteome Remodeling in Maize (*Zea mays* L.) Responding to Drought and Heat Stresses. *Proteomes* **2026**, *14*, 44. [CrossRef] [PubMed]
51. Yu, P.; He, X.; Baer, M.; Beirinckx, S.; Tian, T.; Moya, Y.A.T.; Zhang, X.; Deichmann, M.; Frey, F.P.; Bresgen, V.; et al. Plant Flavones Enrich Rhizosphere Oxalobacteraceae to Improve Maize Performance under Nitrogen Deprivation. *Nat. Plants* **2021**, *7*, 481–499. [CrossRef] [PubMed]
52. Da Silva Messias, R.; Galli, V.; Dos Anjos E Silva, S.D.; Rombaldi, C.V. Carotenoid Biosynthetic and Catabolic Pathways: Gene Expression and Carotenoid Content in Grains of Maize Landraces. *Nutrients* **2014**, *6*, 546–563. [CrossRef] [PubMed]
53. Zurak, D.; Grbeša, D.; Duvnjak, M.; Kiš, G.; Međimurec, T.; Kljak, K. Carotenoid Content and Bioaccessibility in Commercial Maize Hybrids. *Agriculture* **2021**, *11*, 586. [CrossRef]
54. Šukalović, V.H.-T.; Veljović-Jovanović, S.; Maksimović, J.D.; Maksimović, V.; Pajić, Z. Characterisation of Phenol Oxidase and Peroxidase from Maize Silk. *Plant Biol.* **2010**, *12*, 406–413. [CrossRef] [PubMed]
55. Eathington, S.R.; Crosbie, T.M.; Edwards, M.D.; Reiter, R.S.; Bull, J.K. Molecular Markers in a Commercial Breeding Program. *Crop Sci.* **2007**, *47*, S-154–S-163. [CrossRef]
56. Prasanna, B.M.; Pixley, K.; Warburton, M.L.; Xie, C.-X. Molecular Marker-Assisted Breeding Options for Maize Improvement in Asia. *Mol. Breed.* **2010**, *26*, 339–356. [CrossRef]
57. Cerrudo, D.; Cao, S.; Yuan, Y.; Martinez, C.; Suarez, E.A.; Babu, R.; Zhang, X.; Trachsel, S. Genomic Selection Outperforms Marker Assisted Selection for Grain Yield and Physiological Traits in a Maize Doubled Haploid Population Across Water Treatments. *Front. Plant Sci.* **2018**, *9*, 366. [CrossRef] [PubMed]
58. Yao, W.; Ku, L.; Wang, B.; Ren, Z.; Su, H.; Li, C.; Liu, H.; Zeng, H.; Liu, Z.; Zhang, D.; et al. Breeding Ideotype Maize with Enhanced Yield through Genomics-Guided Pyramiding of Favorable Alleles. *Nat. Genet.* **2026**, *58*, 916–927. [CrossRef] [PubMed]
59. Mahfouz, M.M.; Cardi, T.; Neal Stewart, C. Next-Generation Precision Genome Engineering and Plant Biotechnology. *Plant Cell Rep.* **2016**, *35*, 1397–1399. [CrossRef] [PubMed]
60. Wang, Y.; Tang, Q.; Pu, L.; Zhang, H.; Li, X. CRISPR-Cas Technology Opens a New Era for the Creation of Novel Maize Germplasms. *Front. Plant Sci.* **2022**, *13*, 1049803. [CrossRef] [PubMed]
61. Chen, F.; Chen, L.; Yan, Z.; Xu, J.; Feng, L.; He, N.; Guo, M.; Zhao, J.; Chen, Z.; Chen, H.; et al. Recent Advances of CRISPR-Based Genome Editing for Enhancing Staple Crops. *Front. Plant Sci.* **2024**, *15*, 1478398. [CrossRef] [PubMed]

62. Oliveira, D.M.; Saleme, M.d.L.S.; Smith, R.A.; Vangeel, T.; Lima, M.d.F.; Chanoca, A.A.; Mota, T.R.; Vanhevel, Y.; Coussens, G.; Van Aelst, K.; et al. CRISPR/Cas9 Editing of p-COUMAROYL-CoA:MONOLIGNOL TRANSFERASE 1 in Maize Alters Phenolic Metabolism, Lignin Structure, and Lignin-First Biomass Processing. *Trends Biotechnol.* **2025**, *43*, 1166–1195. [CrossRef] [PubMed]
63. Przybylska-Balcerek, A.; Szablewski, T.; Szwajkowska-Michałek, L.; Świerk, D.; Cegielska-Radziejewska, R.; Krejpcio, Z.; Suchowilska, E.; Tomczyk, Ł.; Stuper-Szablewska, K. *Sambucus nigra* Extracts–Natural Antioxidants and Antimicrobial Compounds. *Molecules* **2021**, *26*, 2910. [CrossRef] [PubMed]
64. Stuper-Szablewska, K.; Kurasiak-Popowska, D.; Nawracała, J.; Perkowski, J. Response of Non-Enzymatic Antioxidative Mechanisms to Stress Caused by Infection with Fusarium Fungi and Chemical Protection in Different Wheat Genotypes. *Chem. Ecol.* **2017**, *33*, 949–962. [CrossRef]
65. Kurasiak-Popowska, D.; Ryńska, B.; Stuper-Szablewska, K. Analysis of Distribution of Selected Bioactive Compounds in *Camelina sativa* from Seeds to Pomace and Oil. *Agronomy* **2019**, *9*, 168. [CrossRef]
66. Aebi, H. Catalase in Vitro. *Methods Enzymol.* **1984**, *105*, 121–126. [CrossRef] [PubMed]
67. Maehly, A.C.; Chance, B. The Assay of Catalases and Peroxidases. *Methods Biochem. Anal.* **1954**, *1*, 357–424. [CrossRef] [PubMed]
68. Shapiro, S.S.; Wilk, M.B. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* **1965**, *52*, 591–611. [CrossRef]
69. VSN International. *VSN International Genstat for Windows*, 24.2 ed.; VSN International: Hemel Hempstead, UK, 2025.

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.