Gene mining and functional analysis related to maize (Zea mays L.) seed size

Syed F.A. GILLANI¹,², Adnan RASHEED³, Yasir MAJEED¹, Shamshad S. BUKHARI⁴, Huma TARIQ⁵, Shafiq-Ur-REHMAN⁵, Zhan-Wu GAO⁶

¹Gansu Agricultural University, College of Agronomy, Lanzhou 730070, China; faheemgillani33@gmail.com; yasirmajeed5453@gmail.com
²Gansu Provincial Key Lab of Arid Land Crop Science, Lanzhou 730070, China
³Jiangxi Agricultural University, Ministry of Education/College of Agronomy, Key Laboratory of Crop Physiology, Ecology and Genetic Breeding, Nanchang 330045, China; adnanbreeder@yahoo.com
⁴University of Poonch Rawalakot Azad Pakistan, Laboratory of Plant Breeding and Molecular Genetics, Azad, Kashmir, Pakistan; huma.tariq006@hotmail.com; shafiqurehman@upr.edu.pk
⁵University of Poonch Rawalakot Azad Pakistan, Laboratory of Plant Breeding and Molecular Genetics, Azad, Kashmir, Pakistan; huma.tariq006@hotmail.com; shafiqurehman@upr.edu.pk
⁶Baicheng Normal University, Tourism and Geographical Science Institute, Baicheng 137000, China; gaozw261@nenu.edu.cn (*corresponding author)

Abstract

Maize has widely been studied as a model of plant-growth promoting rhizobacteria (PGPR). Here, the genome sequences of 9P. The strains, together with 26 other sequenced Maize were comparatively studied. Phylogenetic analysis of the concatenated 244 single-copy core genes suggests that the 9P. The strains and 5 other Paenibacillus spp., isolated from diverse geographic regions and ecological niches, formed a closely related clade (here it is called Poly-clade). Analysis of single nucleotide polymorphisms (SNPs) reveals local diversification of the 14 Poly-clade genomes. SNPs were not evenly distributed throughout the 14 genomes and the regions with high SNP density contain the genes related to secondary metabolism, including genes coding for polyketide. Recombination played an important role in the genetic diversity of this clade, although the rate of recombination was clearly lower than mutation. The distinction among people and different creatures can be gotten by relative examinations. This study reveals that both maize and its closely related species have plant growth promoting traits and they have great potential uses in agriculture and horticulture as PGPR.

Keywords: functional analysis; gene mining; maize; polymorphisms; Rhizobacteria

Introduction

Maize is the most conveyed grain crop all around. Its pack end uses and the effortlessness of advancement over contrasted biological and soil conditions has made it an appealing harvest across the world. Despite human
usage, it is used as feed for trained creatures, rough materials for engineered and food endeavors and as biofuel. To also improve its agronomical characteristics, analysts have interminably endeavored to modify its genome through innate procedures. For the most part, maize characteristics were changed or adjusted through enlighten and substance mutagens. These techniques could introduce changes in the plant genome during transparency and deoxyribonucleic destructive (DNA) fix measures. Change raising is all things considered not accurate. It can provoke both positive and antagonistic outcomes with no control over spaces of the genome to be changed (Gazzafi et al., 2017).

Transposon naming is another occasionally used strategy in maize genetic characteristics, whereby unequivocal transposons are used to cause changes and thus permit quality disclosure. This system is both drawn-out and can be exorbitant. It moreover prompts unpredictable changes and is massive to perform for colossal screens. Notwithstanding that there are many ‘races’ of maize, most of the fiscally grown crossbreeds transmitted two or three huge races. For the inspirations driving discussing the business meaning of maize, maize types can be parcelled into four groupings not related to race. Engraving maize is the fundamental kind of maize filled in the US Corn Belt similarly as in Europe, South Africa, and China. Gouge varieties have been changed through hybridization and assurance to give a wide extent of agronomic and touch characteristics. Exceptional scratch varieties have been conveyed with novel starch characteristics. High-amylose (straight starch) and waxy (fanned starch) maize genotypes have been grown mechanically for quite a while (Rosen et al., 2016).

Maize plants contain both male and female regenerative developments and copy by both cross-treatment and self-preparation. In most modernly appropriate maize genotypes, the female plan projects outward from a central tail, while the male development projects out the most noteworthy mark of the tail. Residue from the tuft is passed on by the breeze to other maize plants, where treatment of the individual pieces on the ear occurs. The ears of maize may go in size from about 2.5 to in excess of 45 cm long (Wanger et al., 2016). The piece size, shape, and concealing similarly shift extensively. Stone maize is innately extraordinary in legacy from gouge maize and is depicted by hard, round bits. Rock maize endosperms involve overwhelmingly of hard or smooth endosperm. The agronomic characteristics of stone maize contrast from gouge maize, fundamentally due to the unique necessities of the chronicled creating regions. Popcorn is a stone kind maize that has been innately decided for its ability to expand or ‘pop’ when warmed. Popping happens when the pieces are immediately warmed to ~240 °C. The thick endosperm limits water rage scattering that makes pressure work inside the part until it explodes (Binder et al., 2015).

**Gene Mining**

Plants are fixed troopers that are attached to one spot, they can’t pursue supplements or escape from herbivores and microorganisms. A portion of these mixtures have been found to battle human dangers also, and cultivators have been scouring these ranges of optional metabolites for their wellbeing advancing properties for quite a long time. Current medication additionally consolidates plant compounds. Around 80% of the total populace as of now depends on ethnobotanical cures and plant drugs, for example, the antineoplastic Taxol, the antimalarial artemisinin, the pain-relieving codeine, the anti diabetic allicin, and the cardiovascular depressant quinidine. The significant expense of new medications, unpalatable results and microbial obstruction are driving a steady and recharged public interest on other option and correlative medication (Qi et al., 2013). Proper mechanism of the process of Gene Mining of Maize given in Figure 1.
Figure 1. Proper mechanism of the process of gene mining of maize (Ayers et al., 1997)

Luckily, sequencing innovation has developed to furnish scientists with the devices to handle virtually these inquiries. Single Molecule, Real-Time (SMRT) Sequencing, which works like a goliath magnifying lens that can in a real sense ‘see’ DNA union progressively, empowers specialists to gather exceptionally coterminous and precise megabase-size extends, or contigs, of plant genomes. These ‘long peruses’ catch undetected primary varieties, completely unblemished qualities and administrative areas inserted in complex constructions that divided draft genomes regularly miss (Ahmadi et al., 2011).

Most genome-wide information is acquired at the degree of quality articulation (i.e., varieties in mRNA amount). It is regularly accepted that every individual quality translates indistinguishable RNA atoms. However, as a general rule, one quality may deliver a few distinctive isoforms by the utilization of elective advertisers, exons and eliminators. During record, elective RNA particles (i.e., isoforms) are frequently created. They can change long and vary uniquely in capacity and articulation design. On the other hand, grafted various record isoforms can significantly expand the protein-coding capability of the genome. Also, grafted isoforms deciphered from a similar quality can have essentially extraordinary and surprisingly hostile impacts (Alexa et al., 2010).

In that capacity, precisely catching isoform action can be critical to understanding quality construction, administrative components and coding areas. Also, covering the whole length of cDNA groupings and records can even empower the revelation of new qualities. Enter the isoform grouping (Iso-Seq) strategy, which uses since quite a while read innovation and requires no get together, making it an inexorably famous apparatus – particularly without reference genomes, which is a reality for some, scientists chipping away at non-model creatures and plants with genomes that are enormous and complex (Bazo et al., 2018).

To address these inconveniences, the Toronto/NY group sequenced the two cannabis assortments utilizing SMRT Sequencing which gave new experiences into the plan of the chromosomes and the cannabinoid biosynthetic qualities, including revelation of considerable adjustment and quality duplications at the intently connected THC and CBD corrosive synthase quality loci. Maybe than resolve the THCAS/CBDAS secret, nonetheless, the hereditary guide brought up more issues. “They are not isoforms at a generally comparable locus, and no likeness THCAS (deactivated or not) is found in hemp”, the creators
composed. Their perceptions recommended that either polymorphisms or differential guideline of sweet-smelling prenyltransferase (AP) adds to cannabinoid creation, apparently by controlling substrate fixation for THCAS and CBDAS (Kang et al., 2015).

Purple Kush has more noteworthy than five-overlay higher record levels of AP than Finola, with no distinction in duplicate number, recommending that AP protein levels might be higher in drug-type plants halfway because of contrasts in record levels. One of home-grown genomics' greatest examples of overcoming adversity additionally represents the limits that accompany deficient genomic and transcriptomic inclusion. Chinese researcher Youyou Tu got a Nobel Prize in Physiology or Medicine in 2015 for her disclosure of the counter jungle fever capacity of artemisinin, an endoperoxide sesquiterpene lactone disengaged from sweet wormwood a yearly spice of the Asteraceae family. Artemisinin-based mix treatments (ACTs), suggested by the World Health Organization for the treatment of simple intestinal sickness brought about by the Plasmodium falciparum parasite, have saved large number of lives. Other helpful impacts have likewise been accounted for artemisinin for infections like malignant growth, tuberculosis and diabetes, so request is high for the compound. Be that as it may, plant-based creation is battling to satisfy the worldwide need because of the low measure of artemisinin delivered in A. annua leaves (0.1%-1.0% of dry weight) (Terole et al., 2015).

Extensive transcriptome sequencing permitted the Choi group to conquer this obstruction. They had the option to distinguish 76 cytochrome P450s and arrange them into remarkable families, which ought to empower their individual capacities to be anticipated with exactness. They were likewise ready to bits together another vital riddle: anti-toxin biosynthesis. This is really where they tracked down the most noteworthy number of isoforms (1,250) and chemicals. Suspected to effect neurodegenerative issues, cardiovascular infections, diabetes and malignant growth, Panax ginseng and Panax notoginseng contain novel saponins called ginsenosides. Study into these glycosylated triterpenes has been hampered, notwithstanding, because of the lethargic development (~4 years/age), long age time, low seed creation and convoluted genome construction of Panax plants (Fu et al., 2012). Basic Mechanism of how the genetic changes occur at different locus position shows in Figure 2.

![Figure 2](https://example.com/figure2.png)

**Figure 2.** Basic Mechanism of how the genetic changes occur at different locus position (Laurent et al., 2000)
The principal all over again get together of a Panax genome – a 2.36 Gbp diploid P. notoginseng with 35,451 protein-encoding qualities – was at long last detailed as a pre-print in July 2018 (5) by a group from the Chinese University of Macau drove by Simon Ming-Yuen Lee. An all over again gathering of a 2.98 Gbp genome (with 59,352 commented on qualities) of the tetraploid P. ginseng cultivar Chunpoong (ChP), created by a group from Seoul National University drove by Tae Jin Yang, followed presently. Sequencing of both DNA and mRNA empowered scientists to bring profound jumps into the ginsenoside biosynthetic hardware, yet in addition its guideline and metabolic use. On account of P. ginseng, Yang et al. (2013) developed genome-scale metabolic organizations covering almost 5,000 quality items, catalyzing 2,194 responses and 2,003 exceptional metabolites (Li et al., 2020).

Ginsenosides gather contrastingly in roots, leaves, stems, blossom buds and berries, in amounts changing with tissue, age, climate and cultivar. Yang's group had the option to decide from whole genome sequencing that the high ginsenoside substance in more established P. ginseng roots are likely the consequence of transportation from shoot tissues instead of dynamic biosynthesis. Co-articulation examination utilizing RNA sequencing information recognized significant chemicals with which ginsenoside creation co-developed. On account of P. notoginseng, two sorts of ginsenosides (PPD and PPT) with contradicting organic exercises (supportive of angiogenesis and against angiogenesis) can be found in a similar plant. Simply by completely portraying the whole genome of the plant, just as isoforms from eight of its constituent parts (Du et al., 2014) was the Lee group ready to verify that the flying parts (e.g., leaf and bloom) contain a higher bounty of PPD contrasted with roots. A long way from being a relic of an old-fashioned past, therapeutic plants and natural cures have educated a lot regarding present day medication and could contribute an extraordinary arrangement to sound, science-based arrangements of things to come (Gao et al., 2013).

We actually have a lot to find out about the hereditary and epigenetic components of these possibly wellbeing advancing plants. Fortunately, current sequencing stages empower us to examine the extraordinary primary association of qualities and the administrative systems fundamental (Cannon et al., 2012) their demeanor designs, permitting the age of inventories of specific digestion in manners inconceivable to the cultivators who initially saddled their recuperating properties hundreds of years prior. As Tessa Moses and Alain Goossens bring up in the Journal of Experimental Botany, all living plant species on the planet together add to a more prominent synthetic variety of bioactive mixtures than any man-made substance library (Garvin et al., 2018). The study was found to be very useful and important for future research.

**Functional genome analysis**

Though human genomes are about 99.9% indistinguishable, the excess 0.1% is the explanation of distinction between individuals brought about by various variations. Since 2003, the total grouping of human genome, its explanation and expanded headway of sequencing innovations (i.e., Sanger and Next-age sequencing; NGS) have given every one of the important conditions to the recognizable proof of all variations in human coding and non-coding grouping. Albeit the strategy for variation recognition is presently turning into an everyday practice, the critical inquiry all through numerous years concerns the capacity of distinguished variations. The asset of significant data about utilitarian genomics are a few enormous scope projects, (Murrey et al., 2012) for example, the ENCODE project, the principle objective of which was to distinguish every one of the useful components, remembering administrative components for both coding and non-coding regions. As per another, the 1000 Genomes Project, there are around 20,000–23,000 variations in interchangeable and no synonymous areas of the human genome. Despite the fact that not every one of them are practically significant, 530–610 of the variations have utilitarian effect by causing in frame cancellations and inclusions, untimely stop codons, frame shifts, or by disturbing join destinations (Gupta et al., 2020).

In spite of various investigations, researchers are as yet confronting a tremendous test in disentangling what the succession implies and in choosing whether or not a discovered variation is pathogenic. A pathogenic
variation can prompt infection or cause various issues. Be that as it may, comprehension of pathogenic components sets out a freedom to forestall extreme outcomes by creating novel demonstrative apparatuses and by planning profoundly successful medicines for the infection. To accomplish this point it is important to perform huge scope utilitarian genome examination that includes various fields of study: genomics, epigenomics, transcriptomics, proteomics, and interactomes. To depict the elements of qualities and proteins just as to investigate the connection between the genotype and the aggregate, an enormous number of different strategies, including model frameworks (e.g., CRISPR-Cas9), can be utilized (Hencock et al., 2020). Different tests performed on number of variants of plants showing NR sequences and the redundancy mentioned in Table 1.

**Table 1.** Different tests performed on number of variants of plants showing NR sequences and the redundancy

<table>
<thead>
<tr>
<th>Sources of ESTs</th>
<th>Number of SSR-EST’s</th>
<th>Consensus sequences</th>
<th>Singletons</th>
<th>Total number of NR sequences</th>
<th>% Reduction in redundancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barley</td>
<td>3462</td>
<td>269</td>
<td>600</td>
<td>869</td>
<td>75</td>
</tr>
<tr>
<td>Wheat</td>
<td>3624</td>
<td>335</td>
<td>302</td>
<td>637</td>
<td>83</td>
</tr>
<tr>
<td>ITEC</td>
<td>2277</td>
<td>230</td>
<td>457</td>
<td>687</td>
<td>70</td>
</tr>
<tr>
<td>Maize</td>
<td>4107</td>
<td>408</td>
<td>385</td>
<td>793</td>
<td>81</td>
</tr>
<tr>
<td>Sorghum</td>
<td>3976</td>
<td>430</td>
<td>498</td>
<td>928</td>
<td>77</td>
</tr>
<tr>
<td>Rice</td>
<td>7160</td>
<td>752</td>
<td>1001</td>
<td>1753</td>
<td>76</td>
</tr>
<tr>
<td>Total</td>
<td>24606</td>
<td>2424</td>
<td>3243</td>
<td>5667</td>
<td>77</td>
</tr>
</tbody>
</table>

The critical way to effective arrangement examination is to adjust the succession important to another grouping whose capacity is referred to (normally name as the reference genome). It very well may be valuable when the quality capacity is obscure however is developmental identified with another quality whose capacity is characterized. In such a case, it very well may be thought that the obscure quality has something very similar or comparable capacity. Moreover, the arrangements may be checked to track down the huge matches (Martin et al., 2011) between the segments of a grouping that have been recently portrayed as enormously affecting the genomics work. To analyze the information, it is important to look for data in changed biomedical data sets. Perhaps the greatest wellspring of biomedical and genomic data is the NCBI (National Center for Biotechnology Information), which gives admittance to different data sets like PubMed, Entrez Gene, OMIM, Variation Viewer, dbSNP, and others (Ince et al., 2014). Pictorial representation of Genomic Functional analysis shows in Figure 3.

There are some of the types that are as follows;
- Epigenomics
- Transcriptomics
- Proteomics and Interactomics
- Functional Genomics Integrating Model Systems
For utilitarian examination, it is critical to consider epigenetic adjustments like DNA methylation and histone alterations, since they influence quality articulation with no progressions in the basic DNA succession. A segment of these combinations has been found to fight human perils additionally, and cultivators have been scouring these scopes of discretionary metabolites (Reddy et al., 2012) for their prosperity propelling properties for a serious long time. Current prescription moreover combines plant compounds. Around 80% of the all-out people as of now relies upon ethnobotanical fixes and plant drugs, for instance, the antineoplastic Taxol, the antimalarial artemisinin, the torment diminishing codeine, the antidiabetic allicin, and the cardiovascular depressant quinidine. The extended sensibility and intricacy of genetic sequencing development is making the total of this possible. To fathom the full metabolic capacity of plants, broad genomic information ought to be gotten together with transcriptomic, proteomic and metabolomics data (Onus et al., 2010).

We ought to have the alternative to react to questions, for instance, the quality duplication, similar to whole genome duplications (WGDs) and neighborhood (couple) duplication (LDs), can in like manner accept a critical part specifically processing, including the outpouring of flavonoid related characteristics. Fortunately, sequencing advancement has created to outfit researchers with the gadgets to deal with practically these requests. Single Molecule, Real-Time (SMRT) Sequencing, which works like a goliath amplifying focal point that can from a genuine perspective ‘see’ DNA association logically (Reddy et al., 2012), enables experts to assemble particularly coterminous and exact mega base-size expands, or contigs, of plant genomes. These ‘long scrutinizes’ catch undetected essential assortments, totally unsullied characteristics and regulatory regions embedded in complex developments that separated draft genomes routinely miss (Burger et al., 2019). Most genome-wide data is gained at the level of value verbalization (i.e., assortments in mRNA sum). It is consistently acknowledged that each individual quality deciphers vague RNA iotas. Be that as it may, when in doubt, one quality may convey a couple unmistakable isoforms by the usage of elective sponsors, exons and eliminators. During record, elective RNA particles (ie, isoforms) are every now and again made. They can change long and fluctuate exceptionally in limit and explanation plan. Then again, joined different record isoforms can

**Figure 3.** Pictorial representation of genomic functional analysis (Bowen et al., 2011)
altogether grow the protein-coding capacity of the genome. Additionally, united isoforms translated from a comparable quality can have basically exceptional and shockingly unfriendly effects (Komada et al., 2003). Epigenomics in maize and detailed mechanism of its working showed in Figure 4.

**Figure 4.** Epigenomics in maize and detailed mechanism of its working (Russell et al., 2013)

**Transcriptomics**

Basically, by totally depicting the entire genome of the plant, similarly as isoforms from eight of its constituent parts, was the Lee gathering prepared to check that the flying parts (e.g., leaf and blossom) contain a higher abundance of PPD appeared differently in relation to roots. They perceived a couple of key characteristics, including a couple seen strangely, similarly as their record factor limiting areas and other related parts drew in with the ginsenosides association pathway. As Yang raises, such information will be fundamental to engaging in silico metabolic planning to anticipate contender characteristics related with overproduction of liked metabolites (Zhang et al., 2010) and in as such accelerate in everyday metabolic planning cycles. Far from being a relic of an older style past, restorative plants and regular fixes have taught a great deal with respect to introduce day medicine (Jia et al., 2017) and could contribute an exceptional course of action to sound, science-based plans of what might be on the horizon (Dutta et al., 2011).

We really have a great deal to get some answers concerning the innate and epigenetic parts of these conceivably prosperity propelling plants. Luckily, current sequencing stages engage us to inspect the
uncommon essential relationship of characteristics and the regulatory frameworks major their disposition plans, allowing the period of inventories of explicit assimilation in habits unfathomable to the cultivators (Grener et al., 2011) who at first outfitted their recovering properties many years earlier. By returning to our basic establishments through roots, and getting old discernments together with present day nuclear mining, we can announce another time of recovering and medicine divulgence (Jones et al., 2002).

Gigantic extension genome projects have fundamentally changed the essence of science. Genomics has habitually been suggested as another field that has provoked an adjustment of viewpoint in the way in which science is performed. Meanwhile the post-genomic period has emerged by misusing the tremendous (Stein et al., 2012) proportion of genome gathering data. It has gotten possible to look at science in a substitute way as in nowadays researchers don’t need to advance toward natural requests in a theory driven way anyway rather can accumulate and examine data in a more non-uneven and more broad style. New sorts of consistent requests can be presented and new kinds of examinations (Bowen et al., 2011) can be performed at a remarkable speed. Continuous inventive advances and the speedy improvement of novel gadgets as of now award the questioning of an all-out genome simultaneously and in a singular preliminary (Levin et al., 2001). Transcriptomics analysis of maize at different genomic level shows in Figure 5.

![Image](image.png)

**Figure 5.** Pictorial diagram of functional gene study in maize (Schmitz et al., 2012)

Knowing the particular progression besides, space of the large number of characteristics of a given animal is figuratively speaking the underlying move towards perceiving how all of the bits of a natural structure collaborate. In such manner reasonable genomics is the imperative (Zhang et al., 2019) method to manage changing sum into quality. Valuable genomics is a general technique toward perceiving how the characteristics of a natural element cooperate by assigning new abilities to cloud characteristics. Information about the speculated limit of a dark quality may be closed (Herr et al., 2014) from its gathering structure using
unquestionably alluded to components of similar characteristics as the justification assessment (Province et al., 2001).

**Functional Genomics Integrating Model Systems**

From the practical perspective, examination of proteomics and interactomes is pretty much as crucially significant as recently portrayed investigation (Noda 2014) of genomics, epigenomics, and transcriptomics, in light of the fact that a few examinations show that quality articulation at DNA or mRNA levels is generously unaltered, in spite of the fact that it influences the protein capacity and the other way around. Proteins play out a huge range of capacities (Feng et al., 2014) inside life forms, however unusual protein articulation that happens because of post-transcriptional changes or protein communication with another protein or nucleic acids upsets cell work (Aliferis et al., 2003).

Contingent upon the plan of the test, there are two notable procedures for protein measurement: immunoassays or neutralizer free discovery techniques. Immunoassay, for example, the catalyst connected immunosorbent examine (ELISA), is a generally utilized technique (Muller et al., 2016) because of its high affectability and solid particularity. In any case, it is only the new variety of these techniques to the simultaneous assessment of tremendous amounts of characteristics, proteins (Keller et al., 2016), substance constituents and oddity plants that has allowed the creation of the current down to earth genomic development stage. A couple of methodologies have viably become standard devices for inspecting quality limit while others are at this point in their beginning phases (Martin et al., 2014). The rapidly emerging field of close to genomics has yielded enthusiastic results. Close to genome examination has gotten conceivable with the openness of different completely sequenced genomes (Wolf et al., 2013). Genomic analysis performed on different variants showed results of different combinations of nitrogenous bases in Table 2. Assessment of complete genomes (Mock et al., 2014) between natural elements consider overall points of view on genome progression and the availability of various completely sequenced genomes fabricates the judicious power in deciphering the privileged information in genome setup, limit and improvement. In like manner, assessment of human characteristics with characteristics (Freeling 2020) from various genomes in a genomic scene could help apportion novel capacities with respect to un-clarified characteristics. Here, we talk about the actually used strategies for relative genomics and their derived allowances in genome biology. As on Jan 25, 2007, 472 genomes are completely sequenced (Bird et al., 2011) yet then another 498 are in progress. The quick headway in genome sequencing demands more relative examination to gain new encounters into extraordinary, biochemical, inherited, metabolic, and physiological pathways (Ogata et al., 2014).

In each genome, and checked districts kept up in both prokaryotic and eukaryotic social occasions of living creatures. Close to genomics not solely can follow out the formative association between life frames yet furthermore differentiations and likenesses inside (Sanbor et al., 2014) and between species. The differentiation among individuals and various animals can be gotten by relative assessments. To report the obvious features of individuals, the most enlightening investigation incorporates standing out individuals from our closest relatives, the chimpanzees (Kim et al., 2012) and apes. Genome correspondence the procedure for choosing the correct correspondence of chromosomal segments and valuable parts across the species contemplated is the underlying stage in relative genomics (Paz et al., 2017).

This incorporates choosing orthologous (characteristics veered after a speciation event) parts of DNA that dive from comparable region in the fundamental begetter (Zhu et al., 2013) of the species took a gander at, and paralogous (characteristics isolated after a duplication event) regions that arose by duplication events going before the uniqueness of the species ponder. The arranging of areas across two genomes can be facilitated without duplication events; one-to many if a region has gone through duplication or incident in one of the creature classes (Jin et al., 2015) or many-to-various if duplication/hardship has occurred in the two heredities. Fitch et al., cultivated a method called BBH (Best Bidirectional Hits), which recognizes quality consolidates
that are best matches of each other as orthologous (Twell et al., 2015). Pictorial diagram of functional gene study in maize presented in Figure 6.

**Figure 6.** Different levels of plant functional genomics. Adapted from Zhang et al. (2012).

### Table 2. Genomic analysis performed on different variants showed results of different combinations of nitrogenous bases

<table>
<thead>
<tr>
<th>SSR Motif</th>
<th>Barely</th>
<th>Wheat</th>
<th>ITEC</th>
<th>Maize</th>
<th>Sorghum</th>
<th>Rice</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA/CT</td>
<td>244</td>
<td>81</td>
<td>112</td>
<td>198</td>
<td>147</td>
<td>373</td>
<td>1155</td>
</tr>
<tr>
<td>CA/GT</td>
<td>37</td>
<td>31</td>
<td>24</td>
<td>40</td>
<td>70</td>
<td>24</td>
<td>226</td>
</tr>
<tr>
<td>AT/TA</td>
<td>14</td>
<td>13</td>
<td>8</td>
<td>51</td>
<td>52</td>
<td>5</td>
<td>2036</td>
</tr>
<tr>
<td>GC/CG</td>
<td>9</td>
<td>0</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>AGC/TGC</td>
<td>106</td>
<td>125</td>
<td>84</td>
<td>109</td>
<td>199</td>
<td>195</td>
<td>818</td>
</tr>
<tr>
<td>AAC/TTC</td>
<td>7</td>
<td>7</td>
<td>3</td>
<td>21</td>
<td>16</td>
<td>25</td>
<td>97</td>
</tr>
<tr>
<td>AAC/TTG</td>
<td>14</td>
<td>13</td>
<td>11</td>
<td>13</td>
<td>24</td>
<td>22</td>
<td>426</td>
</tr>
<tr>
<td>AAT/TTA</td>
<td>44</td>
<td>65</td>
<td>27</td>
<td>24</td>
<td>40</td>
<td>226</td>
<td>662</td>
</tr>
<tr>
<td>GGC/CGG</td>
<td>33</td>
<td>402</td>
<td>171</td>
<td>21</td>
<td>14</td>
<td>21</td>
<td>78</td>
</tr>
<tr>
<td>CCT/GGA</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>42</td>
<td>7</td>
<td>20</td>
<td>2377</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>845</td>
<td>1066</td>
<td>651</td>
<td>721</td>
<td>1196</td>
<td>2502</td>
<td>6981</td>
</tr>
</tbody>
</table>

Understanding the family line of the reasonable segments stood out is central from our game plan and employments of genome assessment (Sano et al., 2014). Most close to methods have focused in on adjusted orthologous areas, anyway it is correspondingly basic to see what segments have gone through duplication events, and which sections were lost since the divergence (Tran et al., 2013) of the species. Taking a gander at divides that arose before the uniqueness of the species may achieve some unsatisfactory understandings of collection protection and contrast. Further, inside seeing quality duplication, a bit of the groundbreaking necessities (Pilu et al., 2013) that a region is under are soothe, and uniform models of headway now don’t get the essential decision for these regions. Thusly, our techniques for choosing quality correspondence ought to
address duplication and adversity events, and assurance that the parts we consider are orthologous (Kun et al., 2016).

**Plant functional genomics**

Enormous scope genome projects have significantly changed the face of science. Genomics has frequently been alluded to as a new field that has prompted (Kranz et al., 2015) a change in outlook in the manner in which science is performed. In the interim the post-genomic period has arisen by exploiting the huge measure of genome grouping information (Hansen et al., 2017). It has gotten conceivable to take a gander at science in an alternate manner as in these days scientists don’t have to move toward organic inquiries in a speculation driven path however rather can gather and investigate information (Smolen et al., 2015) in a more non-one-sided and more extensive style. New kinds of logical inquiries can be posed and new sorts of analyses can be performed at an exceptional speed. At present the mass of genome information is being changed over into gene function information, implying that worth is added to the nucleotide grouping assortments (Cai et al., 2012).

Nonetheless, to nail down the specific capacity of obscure qualities it is important to see every quality’s job in the unpredictable arrangement of all quality exercises in the plant cell (William et al., 2017). Quality capacity examination in this manner requires the investigation of transient and spatial quality articulation designs. The most decisive data about changes in quality articulation levels can be acquired from investigation of the shifting subjective and quantitative changes of courier RNAs, proteins and metabolites (Van et al., 2018). New advances have been created to permit quick and profoundly equal estimations of these constituents of the cell that make up quality action. The essential standards basic the distinctive insightful advances have been known for a very little time (Tohge et al., 2015a).

**Conclusions**

After being domesticated from teosinte, cultivated maize (*Zea mays* ssp. *mays*) spread worldwide and now is one of the most important staple crops. Here, we provide an overview of the history of maize domestication and key genes controlling major domestication-related traits, review the currently available resources for functional genomics studies in maize, and discuss the functions of most of the maize genes that have been positionally cloned and can be used for crop improvement. Finally, we provide some perspectives on future directions regarding functional genomics research and the breeding of maize and other crops. We have defined a family of 33 maize Pap genes which we predict, on the basis of transcript accumulation and similarity to proteins characterized in other plants, to be functionally diverse and to play a role in both the P deprivation response and more generally in maize stress responses and development. From what is known of the post-transcriptional regulation of PAPs in other plants, it is probable that the capacity for functional divergence is even greater than revealed by this first characterization. Furthermore, groups of closely related sequences present dramatically different patterns of transcript accumulation, illustrating a capacity for rapid adoption of new biological roles during the radiation of the maize Pap family. Ultimately, a more complete understanding of the roles of individual maize PAPs will require functional and biochemical analysis. Given the availability of a reference genome and the increasing availability of resources for reverse genetics, it is now feasible to conduct a functional genomics analysis of a maize gene family. On the basis of the characterization presented here, we have selected a number of candidate genes and initiated a program of reverse screening. We anticipate that functional characterization of maize PAPs will facilitate their use as direct targets of selection and manipulation, as well as valuable reporters of plant nutrient and stress status.
**Authors’ Contributions**

SFAG prepared the draft, AR reviewed and improved the paper. YM, SSB, HT, SR provided technical assistance. ZWG supervised the study. All authors read and approved the final manuscript.

**Ethical approval** (for researches involving animals or humans)

Not applicable.

**Acknowledgements**

National Natural Science Foundation of China (No.3177); Key Project of Science and Technology Research in the 13th Five-Year Plan of Education Department of Jilin Province (No.41 of Jijiao Kehe, 2016); Talents of Jilin Province Supported Project (2020047).

**Conflict of Interests**

The authors declare that there are no conflicts of interest related to this article.

**References**


Williams CE, Grotewold E (2017). Differences between plant and animal Myb domains are fundamental for DNA binding activity, and chimeric Myb domains have novel DNA binding specificities. Journal of Biological Chemistry 272:563-571. https://doi.org/10.1074/jbc.272.1.563


