The genetics and genome-wide screening of perennialism loci in *Zea diploperennis*

Anjun Ma[1,*], Yinjie Qiu[1,*,†], Tajbir Raihan[1], Bimal Paudel[1], Subha Dahal[1,‡], Yongbin Zhuang[1,§], Aravid Galla[1,¶], Donald Auger[1] & Yang Yen[1,#]

Author Affiliation:

[1] Department of Biology and Microbiology, South Dakota State University, Brookings, SD 57007, USA

*: These authors made the equal contributions to this research

†: Current address: Department of Horticultural Science, University of Minnesota, St. Paul, MN 55108, USA

‡: Current address: Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada;

§: Current address: College of Agronomy, Shandong Agricultural University, Taian, Shandong, 271018, China

¶: Current address: Department of Entomology, University of Arkansas, Fayetteville, AR 72701, USA

#: **Corresponding Author**: Dr. Yang Yen, SNP 249C, Box 2140D, Brookings, SD 57007, USA; +1-605-688-4567; yang.yen@sdstate.edu

Classification:

Major: Biological Science

Minor: Plant Biology

Short Title: Perennialism loci in *Zea*

Key Words: Perennial maize, teosinte, *Zea mays*, corn, genetics, perenniality, GBS

Abstract

Perennialism is common among the higher plants, yet we know little about its inheritance. To address this, six hybrids were made by reciprocally crossing perennial *Zea diploperennis* Iltis, Doebley & R. Guzman with three varieties/inbred lines of annual maize (*Z. mays* L. spp. *mays*). We specifically focused on the plant's ability to regrow after flowering and senescence. All the F_1 plants demonstrated senescence and regrowth for several cycles, indicating a dominant effect of the Z. *diploperennis* alleles. The regrowth ability was stably transmitted to progeny of the hybrids in segregation ratios that suggested the trait was controlled by two dominant, complementary loci. Genome-wide screening with genotyping-by-sequencing (GBS) identified two major regrowth loci *reg1* and *reg2* on chromosomes 2 and 7, respectively. GBS results were validated using a larger F_2 population and PCR markers derived from the single nucleotide polymorphisms within the locus intervals. These markers will be employed to select near-isogenic lines for the two loci and to identify candidate genes in the loci in *Z. diploperennis*.

Significance Statement: Our study contributes to our general understanding of inheritance of perennialism in the higher plants. Previous genetic studies of the perennialism in *Zea* have yielded contradictory results. We take a reductionist approach by specifically focusing on the plant's ability to regenerate new shoots after senescence without regard to associated traits, such as rhizome formation, tillering or environmental impacts. Using this criterion, inheritance of perennialism in *Zea* appears to be dominantly and qualitatively inherited. Importantly, our data indicate that there is no major barrier to transferring this trait into maize or other grass crops for perennial crop development, which enhances sustainability of grain crop production in an environmentally friendly way.

Introduction

Perennialism is the phenomenon that a plant can live for more than two years; the ability of doing so is termed perenniality. Plants typically have a life cycle of growth, reproduction (sexual and/or vegetative) and senescence. Annuals and biennials have only one such cycle in their life, leaving behind seeds, bulbs, tubers, etc. to initiate another life cycle. Some perennials maintain juvenile meristematic tissues capable of regrowth after senescence. How perennials do so remains as a mystery. Subterranean stems (such as rhizomes), polycarpy and tuberous roots are often cited as the means by which plants achieve perenniality. However, none of these traits is absolutely required by perennials. For instance, bamboos are essentially monocarpic perennial that regrow from rhizomes. Many perennial temperate grasses, such as switchgrass[1], cordgrass[2] and eastern gamagrass[3], regrow from the crowns instead of rhizomes. On the other hand, some annual/biennial plants, such as radish (*Raphanus sativus*), grow tuberous roots.

Although perennialism is common among higher plants, the study of its genetics and molecular biology is sporadic. So far, the only published research in molecular mechanism of plant perennialism was conducted in Arabidopsis. Melzer et al. successfully mutated this annual herb to show some perennial habits, such as increased woody fiber in the stem by down-regulating two flowering genes coding for MADS-box proteins, SUPPRESSOR OF OVEREXPRESSION OF CONSTANT 1 and FRUITFUL[4]. Unfortunately, this woody mutant was sterile, and no follow-up research was reported. Perennial-related genes and quantitative loci (QTL) have been reported in other species. Major QTL controlling rhizome development, regrowth and tiller number have been mapped on sorghum linkage groups C (chromosome 1) and D (chromosome 4)[5, 6], which are homoeologous to regions of maize chromosomes 1, 4, 5 and 9, respectively[7]. Hu at al. mapped two dominant, complementary QTL *Rhz2*

(*Rhizomatousness 2*) and *Rhz3* that control rhizome production on rice chromosomes 3 and 4 at the loci homoeologous to the sorghum QTL[6]. Tuberous roots in a wild perennial mungbean (*Vigna radiate* ssp. *sublobata*) are conditioned by two dominant, complementary genes[8]. However, after years of effort these perennialism genes have yet to be cloned from any of the species despite that mapping data and complete rice and sorghum genomic sequences are readily available. Therefore, no further research has been reported about these perennialism loci/genes.

In the genus *Zea* L., most species, including maize, are annual. However, two closely related species, tetraploid *Z. perennis* [Hitchc.] Reeves and Mangelsdorf and diploid *Z. diploperennis* Iltis, Doebley & R. Guzman, are perennial. Perenniality of these two teosintes is manifested as regrowth after seed production and senescence, which includes developing juvenile basal axillary buds and rhizomes. Evergreen stalks, bulbils (highly-condensed rhizomes), basal shoot development, stiff stalk and robust root system have all been cited as phenotypic features of perennialism in *Z. diploperennis*[9-11]. For example, evergreen stalks, which was proposed as a component of perennialism in Z. *diploperennis*[9], appears to be linked to *sugary 1* on the short arm of chromosome 4[12].

Conflicting conclusions have been reached in various studies on how perennialism is inherited in *Zea*. Shaver[13] proposed that a triple homozygous recessive genotype is needed for the perenniality in *Zea*. In this model, *pe* (*perennialism*), interacting with *gt* (*grassy tillers*) and *id* (*indeterminate*), plays a key role in conferring totipotency to the basal axillary buds and rhizomes in the perennial teosintes[13, 14]. The nature of *pe* remains unknown and the *Z. perennis*-derived genotype from which *pe* was identified by Shaver[13] was lost and never recovered despite decades of intensive efforts (Shaver, personal communication). Mangelsdorf and Dunn mapped *Pe*-d,* the maize allele of the *pe* homologue in *Z. diploperennis,* to the long arm of

maize chromosome 4[15]. The *gt* gene (aka *gt1*), located on the short arm of maize chromosome 1, encodes a class I homeodomain leucine zipper that promotes lateral bud dormancy and suppresses elongation of lateral ear branches[12]. It appears that *gt1* depends on the activity of a major maize domestication gene, *teosinte branched 1* (*tb1*), and is inducible by shading[16]. The *id* gene (aka *id1*) alters maize's ability to flower[17]. Both *tb1* and *id1* are located on the long arm of maize chromosome 1 and both encode transcription factors with zinc finger motifs[16, 18]. Singleton believed that *id1* inhibits plantlet generation at the upper nodes of a maize stalk[17]. Mangelsdorf et al. proposed that one or two dominant genes control annual growth habit in their *Z. diploperennis*-popcorn hybrid[19]. Murray and Jessup also believed that non-senescence and rhizomatousness are the must-have characteristics of perennial maize[20].

In contrast to the recessive inheritance model, Galinat proposed that perennialism in *Z. diploperennis* is at least partially controlled by two dominant complementary genes[12]. Also, Ting and Yu obtained three perennial F_1 hybrids by pollinating three Chinese field corn varieties with *Z. diploperennis*[21], which indicate that perennial factors are dominant. Unfortunately, there is no further report about these hybrids or their derivatives. Westerbergh and Doebley regarded perennialism in *Z. diploperennis* as a quantitative trait and identified a total of 38 QTL for eight perennial-habit traits from a *Z. diploperennis* x *Z. mays* ssp. *parviglumis* (annual) mapping population[11]. Intriguingly, they did not identify any QTL that shows a singularly large effect.

The various criteria used by previous researchers for what constitutes perennialism in *Zea* may have contributed to the complex and contradictory observations. Traits such as rhizome formation and evergreen stalks may be important adaptive features that support the viability of perennial plants but are not key. In this study, we take a reductionist view and specifically focus on a plant's ability to regrow after senescence. Using this criterion, we have identified two

dominant, complementary loci that control this trait. Here we report the results of our genetic analysis and genome-wide screening of these regrowth loci with genotyping-by-sequencing (GBS) technology.

Results and Discussion

The production and growth of the hybrids

To study perennialism in *Zea*, we made reciprocal crosses of *Z. diploperennis* (Zd, hereafter in a cross combination) with the following three maize lines: B73, Mo17 and Rhee Flint (RF, hereafter in a cross combination). B73 and Mo17 are inbred lines and Rhee Flint is an heirloom maize variety. The first F_1 was made with Rhee Flint in a greenhouse. Rhee flint is small, fast-growing and usually has a few tillers, which affords serial plantings with an increased opportunity of a plant simultaneously flowering with *Z. diploperennis*. Because Rhee Flint is an open-pollinated variety, later F_1s were made with B73 and Mo17 to facilitate molecular analysis. All the F_1 plants are perennial and fertile (Fig. 1), and have completed multiple cycles of growth, reproduction and senescence (Supplementary Fig. S1). Regrowth (as opposed to accidental replanting from seed) of F_1 plants was insured by inspection that new shoots were attached to the base of the F_1 and confirmed by the heterozygosity of polymorphic PCR markers (examples shown in Supplementary Fig. S2). Regrowth of these F_1s originates mainly from basal axillary buds after stem senescence in all the crosses (Figs. 1D, 1E, 1F), but it also can occur at upper nodes of the F_1s when B13 and Mo17 were used as the parent (Fig. 2C). The plantlets regrowing from the upper nodes, however, can only survive if transplanted into soil. This indicates that the senescent stalks do not function to provide the necessary nutrients to the plantlets. Interestingly, some of the basal regrowth immediately developed into a female (Fig. 2A) or a male (Fig. 2B) inflorescence, or a forest of them (Fig. 2D).

Because the F_1 plants and their perennial derivatives are not winter hardy, the regeneration cycles were alternated between the greenhouse and the field (Supplementary Figs. S1 & S3). Interestingly, the ears and kernels of the F_1s of the six crosses all were more teosinte-like (i.e. two rows of oppositely positioned spikelets with paired kernels encased by wooden rachides and glumes) when grown in greenhouse but were more maize-like (i.e. multiple rows of naked kernels with short soft glumes and rachides around a silica-filled soft core) when grown in the field (Fig. 3). In the F_2 and higher generations, ear morphology segregated even under greenhouse condition (Fig. 3). These observations suggest that environmental factors play an important role in the preferential expression of the teosinte or the maize alleles of the genes influencing ear morphogenesis in the hybrids. These observations also indicate that it is possible to breed perennial maize with maize-like ears and kernels.

Some studies have used rhizome development as an indicator of perennialism in *Zea*[13, 14, 19, 22]. We have not observed rhizomes in any of our F_1s and the derived plants; when regrowth occurs, it is always from an axillary bud. Indeed, it is also our observation that the regrowth of *Z. diploperennis* is mainly from basal axillary buds, and only occasionally from rhizomes. Previous conclusions that perennialism in *Zea* is recessive might have resulted from the hypothesis that traits such as tiller number at tasseling (TNT) or rhizome development are indispensable components of perennialism in *Zea*. It is also possible that the perennial teosinte plants used in those studies were heterozygous for one or more perennialism genes. This opinion is supported by the observations of Shaver[9] and Camara-Hernandez and Mangelsdorf[18] that some of their F_1 plants regrew from basal axillary buds after a period of dormancy.

TNT has been associated with perennialism in several studies[11, 14, 23, 24], so we investigated the relationship of TNT with regrowth in the Zd-RF F_2s. One-way ANOVA of TNT

by regrowth (Supplementary Table S1), however, revealed no significant difference of TNT (F = 0.897, p = 0.353) between the regrowth and the non-regrowth F_2s. Indeed, we observed regrowth from several single-stalked hybrid derivatives (Fig. 4A) and non-regrowth of some multi-stalked plants (Fig. 4B). These results suggest that TNT is not essential to perenniality in *Zea*.

The genetics of the hybrids

All our F_1 plants are perennial and have undergone several growth cycles alternatively in greenhouse and field, demonstrating that regrowth is a dominant trait in *Zea*. Brewbaker suggested cytoplasm may contribute to perennialism[25], but our reciprocal F_1s performed similarly, indicating that it does not. To analyze the genetics of regrowth further, 159 Zd-RF F_2s (derived from an F_1 where Zd was the female) and 134 B73-Zd F_2s (derived from an F_1 where B73 was the female) were tested. We did not grow the Mo17-Zd F_2s due to limited resources. Among the 159 Zd-RF F_2s, 90 regrew after senescence and 69 did not (Supplementary Table S2). Similarly, among the 134 B73-Zd F_2s, 81 regrew and 53 did not (Table 1; Supplementary Table S3). Three Zd-RF F_3 populations (Supplementary Table S2) and one B73-Zd F_3 population (Supplementary Table S3), each of which was derived from a single regrowth F_2 plant, were also evaluated for their regrowth.

A chi square (χ^2) test of goodness-of-fit suggests that both of the F_2 populations and one Zd-RF F_3 population best fit the 9:7 regrowth to non-regrowth ratio (Table 2), and the B73-Zd F_3 population and two Zd-RF F_3 populations best fits a 3:1 ratio (Table 2). The simplest model that explains these results is that regrowth in the F_1s and their derivatives is controlled by two dominant, complementary *regrowth* (*reg*) loci. The two dominant, complementary gene model parallels what has been found in other species, such as rice (*Oryza sativa*)[6], Johsongrass (*Sorghum halepense*)[5, 6, 26], basin wildrye (*Leymus cinereus*)[27] and wild mungbean (*Vigna*

radiate ssp. *sublobata*)[8].

The Zd-RF F_1 was also backcrossed to each parental line. All plants from the Zd backcross regrew, while only one of the 20 plants from the RF backcross showed regrowth. Therefore, alternative models, such as one or three dominant complementary genes, are not eliminated but are less probable (Table 2).

We noticed that the number of regrowth plants observed in any generation might be understated, because some plants initially recorded as non-regrowth eventually regrew after about two months of dormancy. It is possible, therefore, that some plants recorded as non-regrowth and discarded to open up greenhouse space may have possessed the ability to regrow. Furthermore, transplanting from the field to the greenhouse and *vise versa* was very stressful to the plants. It is possible that some regrowth plants were killed this way, resulting in a reduced number of regrowth plants. However, the estimated segregation ratios of regrowth to non-regrowth are reliable since they can be verified. For example, the 9:7 ratio of Zd-RF F_2s were verified by the ratios of the Zd-RF F_3s derived from single regrowth F_2 plants (Table 2).

Rice rhizomatousness gene *Rhz2* has been mapped to rice chromosomes 3[6] and sorghum chromosome 1[5, 6, 26], which are both homoeologous to parts of maize chromosome 1[7]. Also, *gt1* and *id1*, which have been implicated with perenniality in *Zea*[8], and *tb1*, which controls *gt1*[16], are all on chromosome 1 in *Zea*[16, 18]. Therefore, we investigated the allele compositions of these three genes in the B73-Zd F_2s (Table 1), and 26 Zd-RF F_2 plants and the three Zd-RF F_3 populations (Supplementary Table S4), and assayed their association with regrowth. Of the 131 regrowth hybrid derivatives, 5, 33 and 115 were homozygous for the maize *gt1*, *tb1* or *id1* alleles, respectively. One Zd-RF F_3 family is homozygous for the *gt1* allele of *Z. diploperennis* (Supplementary Table S4) but segregates approximately 9:7 for regrowth and non-regrowth

(Table 2). Therefore, our results are inconsistent with the model of Shaver[13], and show that *gt1* and *id1* do not control regrowth in our F_1s and their derivatives. *Z. diploperennis*'s *gt1* allele may be helpful to regrowth because the majority of the plants that regrew had at least one copy, but it is not indispensable because many plants regrew without it.

Interestingly, we observed no heterozygosity for *id1* and much less-than-expected heterozygosity for *tb1* in all the hybrid derivatives that were examined, regardless of regrowth (Tables 1; Supplementary Table S4). Of the 134 B73-Zd F_2 plants investigated, only 16 had the *Z. diploperennis id1* allele (Table 1). Similar phenomena were observed in the derivatives of the Zd-RF cross (Supplementary Table S4). It seems that the maize chromosome fragment that carries *id1* was preferentially transmitted into the hybrid derivatives. Excess homozygosity of the maize *id1* allele indicates some sort of selection. It could be that a deficiency or other rearrangement adjacent to the teosinte *id1* allele causes it not to transmit efficiently, or it could be that the teosinte *id1* allele causes the plant not to grow well or flower in South Dakota.

Identifying regrowth loci with genotyping-by-sequencing assay

To identify chromosomal regions that host the two regrowth loci revealed by our genetic analysis, we conducted genome-wide mining of single nucleotide polymorphisms (SNPs) in a randomly selected sub-population of 94 (55 regrowth and 39 non-regrowth) B73-Zd F_2 plants with GBS technology (Supplementary Fig. S4). A total of 2,204,834 (85.14%) Illumina sequencing tags that passed routine quality control filtrations were aligned with B73 reference genome. A total of 714,158 SNPs were then called from 83 (46 regrowth and 37 non-regrowth, labeled in bold in Table 1) of the 94 F_2 plants using TASSEL pipeline[28, 29] (Supplementary Fig. S4). SNP-calling for the excluded 11 plants failed probably due to the failure of barcode addition before sequencing. These SNPs covered all ten chromosomes with an average of 71,416 SNPs

per chromosome (Table 4, Supplementary Fig. S5). As shown in Table 4, these SNPs were first subjected to a two-step filtration to remove those with low minor allele frequency (≤ 0.01) or high missing data rate (>20%) among the F_2 plants. The SNPs that passed the two filtrations were subjected to a χ^2 test for their fit to the two, dominant complementary locus model with the null hypothesis that the observed and the expected are not significantly different ($p \leq 0.05$). We hypothesize that a SNP that is associated with one of the two regrowth factors should be carried by all the regrowth F_2s but one or both are missing from the non-regrowth F_2s. This step kept 946 SNPs that have $\chi^2_{0.05,\,4} < 9.49$. Finally, to simplify the mapping effort, the 946 SNPs were filtered once more by collapsing immediately neighboring SNPs that share the same haplotypes into one. The first SNP in such a cluster was chosen to represent the SNP cluster. This final filtering resulted in 597 SNPs with an overall average distance of 3.52 cM between them in the B73 reference genome. The distribution of these 597 SNPs in the B73 genome are shown in Supplementary Figure S5.

We then conducted locus analysis of the 597 SNPs together with additional 1,969 simulated SNPs, using R/qtl package (version 1.40-8) with the "lodint()" arguments with LOD drop unit of 0.5 cM and the "expandtomarkers" arguments. The results are shown in Figure 5. Using the $LOD_{95\%}$ threshold of 4.17, two candidate *reg* loci were identified with one on B73 chromosome 2 in the interval from 24,244,192 bp (here and hereafter, the nucleotide position in the B73 reference sequence) to 28,975,747 bp with the peak at 27,934,739 bp and one on B73 chromosome 7 in the interval from 2,862,253 bp to 6,681,861 bp with the peak at 5,060,739 bp (Fig. 5). This result supports the genetic model that two major genes control regrowth. Table 5 shows the two representative SNPs for the two candidate *reg* loci on chromosomes 2 (*reg1*) and 7 (*reg2*), and the adjacent maize genes in the B73 reference genome.

Genes *gt1* and *id1* on chromosome 1 were proposed to control perennialism in *Zea*[13], and our LOD analysis located two weak peaks on chromosome 1, assisting regrowth (Fig. 5). One may wonder if these two loci are related to *gt1* and *id1*, respectively. However, these loci are at 82,273,951 bp and 177,235,112 bp, far away from *id1* (around 243,201,405 bp) and *gt1* (around 23,625,801 bp). This observation further indicates that *id1* and *gt1* are irrelevant to regrowth. Previous studies reported that *Z. diploperennis* carried perennialism-related *Pe*-d* and an evergreen gene on chromosome 4[12, 15]. However, our data could not support these observations since no SNP on chromosome 4 significantly associates with regrowth (Fig. 5).

Validation of the candidate SNPs with genetic mapping

To validate the association of the candidate SNPs with the trait of regrowth, we converted two SNPs at the peaks of the two candidate chromosomal intervals on chromosomes 2 and 7 into PCR markers (Table 6). The markers for the peak SNPs were designated S2-2 and S7-1/S7-2, respectively (see Materials and Methods for an explanation of the names). The 134 B73-Zd F_2 population were screened with these PCR markers (Table 1, Supplementary Figs. S6 to S8). The hypothesis is that the PCR markers are linked with the regrowth trait, so a χ^2 test of independence was used to test the alternative hypothesis that the markers segregate independently with the regrowth (Table 3). The test results ($p \leq 0.0001$) indicated that the SNPs are indeed associated with regrowth.

If the S2-2 and the S7-1/S7-2 markers reliably mark the two dominant complementary loci that are necessary and sufficient for regrowth, then no regrowth plant should be homozygous for a maize allele at either locus and all non-regrowth plants should be homozygous recessive for at least one locus. A review of Table 1 indicates some exceptions: 17 regrowth plants are homozygous of the B73 allele for either S2-2 or S7-1/S7-2 and 16 non-regrowth that have at

276 least one *Z. diploperennis* allele at both loci. That is 26.8% of the 123 plants that can be scored
277 for genotype/phenotype exceptions. These exceptions do not necessarily negate the two loci
278 hypothesis because both genome-wide screening and genetic analyses reached the same
279 conclusion. Three possible uncontrolled variables may have caused these exceptions:
280 recombination between the marker and the *reg* locus it represents, mis-scoring of regrowth/non-
281 regrowth phenotypes and mis-scoring of the PCR markers.

282 Recombination may explain some exceptions, but is unlikely to be a major contributor,
283 considering the narrow ranges of the QTL peaks. We reviewed the SNPs among the 83 B73-Zd
284 F_2 plants that were used for the SNP discovery; the estimated maximum rates of recombination
285 between regrowth and the peak SNP represented by S2-2 and S7-2 for the QTL are 0.01% and
286 0.03%, respectively (Supplementary Table S6). Therefore, recombination should not be an issue
287 here.

288 Although the criterion for regrowth phenotyping was simple and reliable, there was still
289 opportunity for mis-scoring. Some plants capable of regrowth may have been scored as non-
290 regrowth because of the abnormality of their regrowth (Fig. 2) or because their regrowth may
291 have been delayed or failed due to pre-mature mortality. Anecdotally, at least one "non-
292 regrowth" plant that was discarded was observed later to have emerging shoots. Alternatively, a
293 non-regrowth plant might have been scored as regrowth because of late developing tillers. The
294 variability in morphology and timing of regrowth shoots indicate that modifiers influence this
295 trait. Even so, unusual regrowth and delayed regrowth were the exceptions.

296 The major contributor to the exceptions is likely the reliability of the PCR markers. For
297 each SNP, primers pair were designed to amplify only one allele. In order to reduce the
298 possibility of amplifying the alternative allele, additional mismatches were incorporated into the

299 primers[30]. While avoiding false positives, this increases the rate of false negatives. Out of 134
300 plants assayed, nine failed to produce a product for either allele using S2-2 (Table 1). An
301 alternative marker for *reg1* on chromosome 2, S2-1, had six failures. Disregarding those failures,
302 the apparent genotypes of S2-2 and S2-1 were different 43 times out of 119 comparisons (36%).
303 Therefore, most differences appear to be due to failure of the marker of one allele or the other to
304 amplify. The S7 primers were designed in a similar fashion as the S2 primers and are likely
305 subject to the same problems. Thus, we believe that most of the genotype/phenotype exceptions
306 are due to the imperfections of these markers.

307 Even so, these PCR markers will be valuable to produce and identify a pair of near-
308 isogenic lines (NILs), each being homozygous dominant for one regrowth locus but homozygous
309 recessive for the alternative. The expectation is that neither NIL is capable of regrowth. Genetic
310 confirmation of the two *reg* loci will be made by a testcross between the NILs, which is expected
311 to produce progeny that demonstrate regrowth. These NILs will also aid in the cloning the
312 functional genes originating from the *Z. diploperennis* loci.

313 In summary, the results presented here indicate that perennialism in *Zea,* when defined as
314 regrowth of shoots from basal axillary buds after senescence, is inherited dominantly and
315 apparently qualitatively. Using this criterion, the inheritance of perennialism in *Zea* does not
316 appear to be as complex as previously thought[11, 13, 14, 22]. Two regrowth loci, *reg1* and *reg2*,
317 were mapped to chromosome 2 and chromosome 7, respectively. Even though our data point to
318 two controlling factors, the data do not discount that perenniality in *Zea* is affected by modifiers
319 and environment. Identification and the functional study of the candidate genes for *reg1* and *reg2*
320 will initiate an understanding about the molecular mechanism of perenniality in *Zea* L.

321 **Materials and Methods**

322 Plant materials and phenotyping

323 *Zea diploperennis* (PI 462368) and *Z. mays* cv. Rhee Flint (PI 213764) were obtained
324 from the USDA North Central Region Plant Introduction Station, Ames, IA. B73 and Mo17
325 inbreds were from the collection of D. Auger and are traceable back to the Maize Genetics
326 Cooperation Stock Center, Urbana/Champaign, IL. In our designations of F1s and their
327 derivatives, the female parental is shown first. Plants were grown and controlled pollinations
328 were made in the greenhouse during the winter and in the field during the summer in Brookings,
329 SD. In the greenhouse, plants were maintained with a 16 h-light/8 h-dark cycle and 20/16 ℃
330 day/night temperature except to induce the floral transition, when two-month old plants were
331 treated with a 10 h light/14 h dark cycle for four weeks.

332 Plants were scored as regrowth if they produced shoots from the basal axillary buds after
333 the original stalks senesced. Occasionally, the hybrid-derived plants developed shoots that
334 terminate in ears ("ear forest") or tassels prior to senescence, these were not scored as regrowth.
335 Rhizome and tuber development were visually investigated on plants that were dug from the soil
336 after senescence. The number of tillers (TNT) per plant was investigated at tasseling stage. Ear
337 and kernel morphology was visually examined and photographed.

338 PCR assay

339 DNA samples were isolated from young leaves using the CTAB procedure[31] and used
340 for PCR-based marker assay. PCR assays were done using GoTaq Green Master Mix (Catalog#
341 M7505, Promega, Madison, WI) at the following conditions: 95°C, 35 cycles of 95°C for 45 sec,
342 55~62°C (primer dependent, see Table 6 for detail) for 1 min and 72°C for 1 min, and 72°C for
343 10 min. The primer sets used in the assays and their annealing temperatures are list in Table 6.
344 The annealing temperatures were determined using a 1°C-touchdown PCR step starting from

65°C. Several primer sets generate only a dominant marker for either the *Z. diploperennis* or *Z. mays* allele, so two primer sets were used in combination to genotype the corresponding locus. This is especially true for the SNP-derived markers S2-1, S2-2, S7-1 and S7-2. In order to reduce the likelihood of false positives, the S2 and S7 primers are designed not to be perfectly complementary to the target sequence[30]. This increases the likelihood of false negatives. For each *reg* locus, the peak SNP and a SNP immediately adjacent to it were chosen for marker development. The marker for the peak SNP of the QTL on the short arm of chromosome 2 is S2-2. A second marker for an adjacent SNP on chromosome 2 was also developed and named S2-1. We could not develop a single PCR marker for both the peak SNP (S7-2) on chromosome 7 QTL so a second one (S7-1) was designed for an immediately adjacent SNP. These two were used in combination as a single marker.

SNP discovery

A GBS assay was conducted according to Elshire et al[28]. The preparation and sequencing of the library were conducted by the University of Wisconsin Biotechnology Center (UWBRC). Generally, DNA samples were digested with *ApeK*I restriction enzyme (RE), and unique barcodes were annealed to each DNA fragments. A single-end 100 bp (1x100bp) sequencing run was carried out on an Illumina HiSeq 2500 platform. The raw data were pooled as a single fastq file and downloaded from UWBRC along with a quality report (FastQC version 0.11.2).

The TASSEL (Trait Analysis by Association, Evolution and Linkage) 3 pipeline was used under the guidance of TASSEL manual[29] for the discovery of SNPs between *Z. diploperennis* and *Z. mays* B73 (Supplementary Fig. S4). TASSEL 4 and 5 pipelines were used if command line was compatible. The barcoded sequence reads were collapsed into a set of

368 unique sequence tags with counts. The tag count files were filtered for a minimum count
369 threshold and merged into the master tag count file. B73_RefGen_V4 reference genome
370 sequence was downloaded from MaizeGDB and processed with Bowtie2 for alignment[32].
371 Master tags were aligned to the B73 reference genome to generate a "Tags On Physical Map"
372 (TOPM) file, which contains the genomic position of each tag with the best unique alignment.
373 The occupancies of tags for each taxon were observed from barcodes information in the original
374 FASTQ files. Tag counts were stored in a "Tags by Taxa" (TBT) file. The TOPM and TBT files
375 were used to call SNPs at the tag locations on the genome. The SNPs were filtered by minimum
376 taxa coverage, minimum locus coverage and minimum minor allele frequency. Fastq files
377 containing sequences of chromosomes 1 to 10 were merged by FASTX_Toolkit and indexed. All
378 commands for SNP discovery were executed in Ubuntu 16.04 LTS platform.

379 SNPs resulted from TASSEL filters plugin with a minimum minor allele frequency of
380 0.01 were filtered again by removing sites that had missing data in more than 20% of the F_2
381 plants. For those SNPs that have missing data in less than 20% of the F_2 plants, the missing data
382 were imputed by treating them as heterozygote since both two alleles can be embodied and
383 considered to be moderate. SNPs were filtered again with χ^2 ($p < 0.05$). The 4th SNP filter was
384 performed by removing SNPs with positions very close to each other, in the range of 100 bp, and
385 showed the exactly same haplotypes, keeping only the first SNP in the cluster. Thus, such a
386 cluster of SNPs was treated as one locus. By removing the redundant SNPs, locus tests can be
387 more precise because repeated SNP sites would affect the LOD value and influence the interval
388 estimation.

389 The filtered SNPs were used for candidate locus estimation. The locus analysis was
390 executed by a standard quantitative trait loci (QTL) procedure in R using the R/qtl package

(version 1.40-8)[33] to better observe the contribution of each SNP and its neighbors. The R codes are listed in Supplementary Table S5. Position simulation was drawn with a maximum distance of 1.0 cM and an error probability of $1x10^{-4}$. The conditional genotype probability (calc.genoprob), as well as simulated genotypes (sim.geno with n.draw=32), were calculated. The "haldane" function was used to convert genetic distances into recombination fractions. Genome scan with a single locus model (scanone) was performed with a binary model using the expectation-maximization algorithm[33]. A permutation test with 1000 replicates was performed in scanone to visualize the LOD thresholds. We determined a locus interval by selecting the first and last SNP sites with significant LOD value. Genes within the intervals were identified by searching the corresponding region on the Gramene website.

Statistical analyses

For statistical analyses, all genotypes and phenotypes were transformed into numeric values. For phenotypes, the regrowth plants were scored as "1" and the non-regrowth plants were scored as "0". For genotypes, the plants that were homozygous to the *Z. diploperennis* allele were scored as "1"; those that were homozygous to the B73 allele were scored as "2"; and those that were heterozygous were scored as "3". When conducting locus analysis, genotype "1" was transformed to "AA", "2" to "BB" and "3" to "AB".

A chi square goodness-of-fit test was used to find the best-fit model or linkage in the genetic analysis and reveal candidate SNPs. To determine if TNT has any correlation with regrowth, a One-Way ANOVA of TNT by regrowth was performed in JMP (JMP® 11.2.0).

Sequencing Data availability

All raw fastq data from this study are available at NCBI data deposition site (https://www.ncbi.nlm.nih.gov/bioproject/) with accession number PRJNA477673.

Acknowledgement This research was partially supported by funds from USDA-NIFA via South Dakota Experiment Station and Department of Biology and Microbiology, South Dakota State University. We greatly appreciate Dr. Frank M. You of Agriculture and Agri-Food Canada for his help in statistics.

Author Contributions Y.Y. designed and supervised this project and all the experiments, and drafted the manuscript; Y.Q., A.M, T.R., B.P., A.G., Y.Z., Y.Y. & D.A. performed the experiments and collected data; Y.Q., A.M., T.R., D.A. & Y.Y. analyzed the data; all authors discussed the results and communicated on and approved the final manuscript.

Competing financial interests The authors declare no competing financial interests.

References

1. Haferkamp MR, Copeland TD (1984) Shoot growth and development of Alamo switchgrass as influenced by mowing and fertilization. *J Range Management* 37: 406-412.
2. Boe A, Owens V, Gonzalez-Hernandez J, Stein J, Lee DK,Koo BC. (2009) Morphology and biomass production of prairie cordgrass on marginal lands. *GCB Bioenergy* 1:240-250.
3. Jackson LL, Dewald CL (1994) Predicting evolutionary consequences of greater reproductive effort in *Tripsacum dactyloides*, a perennial grass. *Ecology* 75:627-641.
4. Melzer S, Lens F, Gennen J, Vanneste S, Rohde A, Beeckman T (2008) Flowering-time genes modulate meristem determinacy and growth form in *Arabidopsis thaliana*. *Nat Genet* 40:1489-1492.
5. Paterson AH, Schertz KF, Lin Y, Liu S, Chang Y (1995) The weediness of wild plants: Molecular analysis of genes influencing dispersal and persistence of johnsongrass, *Sorghum halepense* (L.) Pers. *PNAS* 92:6127-6131.

436 6. Hu FY, Tao DY, Sacks E, Fu BY, Xu P, Li j, Yang Y, McNally K, Khush GS, Paterson AH,
437 Li ZK (2003) Convergent evolution of perenniality in rice and sorghum. *PNAS* 100:4050-
438 4054.
439 7. Wei F, Coe E, Nelson W, Bharti AK, Engler F, Butler E, et al. (2007) Physical and genetic
440 structure of the maize genome reflects its complex evolutionary history. *PLoS Genet* 3: e123.
441 8. Nguyen TD, Lawn RJ, Bielig LM (2012) Expression and inheritance of perenniality and
442 other qualitative traits in hybrids between mungbean cultivars and Australian wild
443 accessions. *Crop & Pasture Sci* 63:619-634.
444 9. Galinat WC (1981a). Evergreen stalks as an indicator of perennialism. *MNL* 55:107.
445 10. Mangelsdorf PC, Dunn ME (1984a) Various expressions and dominance relations of
446 perennialism. *MNL* 58:53.
447 11. Westerbergh A, Doebley J (2004) Quantitative trait loci controlling phenotypes related to the
448 perennial versus annual habit in wild relatives of maize. *Theor Appl Genet* 109:1544-1553.
449 12. Galinat WC (1981b) The inheritance and linkage of perennialism derived from diploperennis.
450 *MNL* 55:107.
451 13. Shaver DL (1967) Perennial maize. *J. Hered* 58:271-273.
452 14. Shaver DL (1964) Perennialism in *Zea*. *Genetics* 50:393-406.
453 15. Mangelsdorf PC, Dunn ME (1984b) Linkage relations of Pe*-d. *MNL* 58:54.
454 16. Whipple CJ, Kebrom TH, Weber AL, Yang F, Hall D, Meeley R, Schmidt R, Doebley J,
455 Brutnell T, Jackson DP (2011) *grassy tillers1* promotes apical dominance in maize and
456 responds to shade signals in the grasses. *PNAS* 108:E506-E512.
457 17. Singleton WR (1946) Inheritance of indeterminate growth in maize. *J Hered* 37:61-64.
458 18. Colasanti J, Yuan Z, Sundaresan V (1998) The *indeterminate* gene encodes a zinc finger

459 protein and regulates a leaf-generated signal required for the transition to flowering in maize.
460 *Cell* 93:593-603.
461 19. Mangelsdorf PC, Roberts LM, Rogers S (1981) Crosses of *Zea diploperennis* with corn.
462 *MNL* 55:19-21.
463 20. Murry SC, Jessup RW. 2013. Breeding and genetics of perennial maize: Progress,
464 opportunities and challenges. In: *Perennial Crops for Food Security, Proceedings of the FAO*
465 *Expert Workshop,* August 28-30, 2013, Rome, Italy, pp103-111.
466 21. Ting YC, Yu MK (1982) Further studies of F1 hybrids of maize x diploid perennial teosinte.
467 *MNL* 56:35-36.
468 22. Camara-Hernandez J, Mangelsdorf PC (1981) Crosses of *Zea diploperennis* with corn. *MNL*
469 55:15-17.
470 23. Doebley J, Stec A, Hubbard L (1997) The evolution of apical dominance in maize. *Nature*
471 386:485-488.
472 24. Camara-Hernandez J, Mangelsdorf PC (1981) Crosses of *Zea diploperennis* with corn. *MNL*
473 55:15-17.
474 25. Brewbaker JL (2007) Grassy tiller and sweet corn. *MNL* 81:12
475 26. Washburn JD, Murray SC, Burson BL, Klein RR, Jessup R. 2013. Targeted mapping of
476 quantitative trait locus regions for rhizomatousness in chromosome SBI-01 and analysis of
477 overwintering in a *Sorghum bicolor* X *S. propinquum* population. *Mol Breed* 31:153-162.
478 27. Yun L, Larson SR, Mott IW, Jensen KR, Staub JE (2014) Genetic control of rhizomes and
479 genomic localization of a major-effect growth habit QTL in perennial wildrye. *Mol Genet*
480 *Genomics* 289:383-397.
481 28. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A

robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6:e19379.

29. Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, et al. (2014) TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. PLoS ONE 9: e90346.

30. Liu J, Huang S, Sun M, Liu S, Liu Y, Wang W, Zhang X, Wang H, Hua W. (2012) An improved allele-specific PCR primer design method for SNP marker analysis and its application. *Plant Methods* 8:34. Doi:10.1186/1746-4811-8-34.

31. Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13-15.

32. Langmead B, Salzberg S (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9:357-359.

33. Broman KW, Wu H, Sen Ś, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19:889-890.

34. Gupta MR, Chen Y (2010) Theory and use of the EM algorithm. *Foundations and Trends in Signal Processing* 4:223-296.

Table legends:

Table 1. The regrowth (R) and the non-regrowth (NR) phenotypes and the marker genotypes of *Zea diploperennis* (Zd), *Z. mays* B73 and their 134 F2 plants*.

Table 2. Results of the χ^2 goodness-of-fit tests of three genetic models.

Table 3. Results of the χ^2 goodness-of-fitness tests for independent assortment of regrowth and the SNP-derived PCR markers among the *Zea mays* B73 - *Z. diploperennis* F_2 population.

Table 4. The actions taken and the numbers of SNPs revealed in each filtration step of SNP analysis of *Zea mays* B13 - *Z. diploperennis* F_2 plants.

505 Table 5. Ranges (bp), peak SNP positions (bp), their LOD, χ^2 and $P(\chi^2)$ of the two candidate
506 *regrowth* (*reg1* and *reg2*) loci**,** and adjacent maize genes per B73RefGen_V4.
507 Table 6. PCR primers and annealing temperatures.

508 **Figure legends**:

509 Figure 1. Photos of *Zea mays* and *Z. diploperennis* (Zd) F_1 plants. **A**: reciprocal Mo17-Zd (right)
510 and Zd-Mo17 (left) F_1 plants; **B**: reciprocal B73-Zd (right) and Zd-B73 (left) F_1 plants; **C**: RF-Zd
511 F_1 plant; **D**: regrowth of a Mo17-Zd F_1 plant; **E**: regrowth of a B73-Zd F_1 plant; and **F**: regrowth
512 of a RF-Zd F_1 plant. B73, Mo17 and RF represent, respectively, inbred lines B73 and Mo17 and
513 cultivar Rhee Flint of *Z. mays*.

514 Figure 2. Photos of abnormal F_1 plants of crosses of *Zea diploperennis* with *Z. mays* inbred lines
515 B73 (A & B) and Mo17 (C) or cv. Rhee Flint (D).

516 Figure 3. Photos of the ears produced from a *Zea mays* cv Rhee Flint x *Z. diploperennis* F_1 plant
517 in different seasons (the upper panel) and from F_2 in summer 2014 in greenhouse (the lower
518 panel).

519 Figure 4. Photos of *Zea mays* Mo17-*Z. diploperennis* F_2 plants, showing regrowth from the basal
520 node of a single-stalked plant (A) or non-regrowth from a multi-stalked plant (B).

521 Figure 5. LOD scores of 597 SNP markers and 1,969 simulated positions for candidate locus
522 determination shown with 95% and 99% LOD thresholds. The thresholds were calculated with
523 1000 permutation. Two significant QTL are indicated by the location of the peak SNPs.

524 **Supplementary documents**:

525 Supplementary Table S1. One way ANOVA Analysis of tiller number at tasseling by regrowth.

Supplementary Table S2. Segregation of regrowth among the *Zea mays* cv Rhee Flint-*Z. diploperennis* F_2s and F_3s.

Supplementary Table S3. Segregation of regrowth among the *Zea mays* B73- *Z. diploperennis* F_2s and F_3s.

Supplementary Table S4. Phenotypes and the *gt1*, *id1* and *tb1* haplotypes of 26 F_2 plants and three F_3 populations of *Zea mays* cv Rhee Flint x *Z. diploperennis* cross.

Supplementary Table S5. *R* codes for candidate loci analysis.

Supplementary Table S6. The SNPs used for marker development, their positions in B73 reference genome and the marker genotypes in *Zea diploperennis* (Zd), *Z. mays* B73 and 83 B73-Zd F_2 plants.

Supplementary Figure S1. A photo showing the growth of *Zea diploperennis* and its F_1 with Z. mays B73 or Mo17 in the field in the Summer 2017.

Supplementary Figure S2. An agarose gel image showing that two molecular markers confirmed the heterozygosity of a *Z. diploperennis*-*Z. mays* cv. Rhee Flint F_1 plant over three life-cycles.

Supplementary Figure S3. A photo showing the regrowth of the B73 - *Z. diploperennis* F_4s in the summer, 2017.

Supplementary Figure S4. An illustration of the general process of TASSEL pipeline used in this study. The barcoded sequence reads are collapsed into a set of unique sequence tags with counts. The tag count files are filtered for minimum count threshold and merged into the master tag count file. Master tags are aligned to the reference genome to generate a "Tags On Physical

546 Map" (TOPM) file, which contains the genomic position of each tag with the best unique
547 alignment. The occupancies of tags for each taxon are observed from barcodes information in the
548 original FASTQ files. The TOPM and TBT files are used to call SNPs at the tag locations on the
549 genome.

550 Supplementary Figure S5. Genetic map of the SNPs on each chromosome of *Zea mays* B73 after
551 the filtration.

552 Supplementary Figure S6. Gel images show presence or absence of the dominant alleles of *Zea*
553 *diploperennis* (Zd) or *Z. mays* B73 of the marker S2-1 for the *reg1* locus among 134 B73-Zd F_2
554 plants. The codes of the plants are listed on the top of the image as 1-1 refers to BZ2-001-1, etc.
555 * 1: homozygous Zd alleles; 2: homozygous B73 alleles; 3: heterozygous; GT: genotype; PT:
556 phenotype; R: regrowth; NR: non-regrowth.

557 Supplementary Figure S7. Gel images showing presence or absence of the dominant allele of *Zea*
558 *diploperennis* (Zd) or *Z. mays* B73 of the SNP marker S2-5 for the *reg1* locus among 134 B73-
559 Zd F_2 plants. The codes of the plants are listed on the top of the image as 1-1 refers to BZ2-001-
560 1, etc.

561 Supplementary Figure S8. Gel images showing presence or absence of the dominant *Zea*
562 *diploperennis* (Zd) allele of marker S7-1 or *Z. mays* B73 allele of marker S7-2 for the *reg2* locus
563 among 134 B73-Zd F_2 plants. The codes of the plants are listed on the top of the image as 1-1
564 refers to BZ2-001-1, etc.

Table 1. The regrowth (R) and the non-regrowth (NR) phenotypes and the marker genotypes of *Zea diploperennis* (Zd), *Z. mays* B73 and their 134 F2 plants*.

Plant	PT	*gt1*	*tb1*	*id1*	S2-1	S2-2	S7-1/ S7-2	Plant	PT	*gt1*	*tb1*	*id1*	S2-1	S2-2	S7-1/ S7-2	Plant	PT	*gt1*	*tb1*	*id1*	S2-1	S2-2	S7-1/ S7-2
Zd	R	1	1	1	1	1	1	**BZ2-006-9**	R	1	1	2	2	3	1	BZ2-009-6	R	1	1	2	1	3	3
B73	NR	2	2	2	2	2	2	**BZ2-006-10**	R	1	3	2	1	1	1	BZ2-009-7	NR	1	-	2	2	2	3
BZ2-001-1	R	1	3	2	3	1	1	BZ2-006-11	NR	3	2	2	3	3	3	**BZ2-009-8**	NR	3	2	2	1	1	2
BZ2-001-2	R	3	3	1	-	3	2	**BZ2-006-12**	NR	3	3	2	3	2	3	**BZ2-009-9**	R	3	1	2	2	3	2
BZ2-001-3	R	1	3	2	1	3	3	BZ2-006-13	NR	3	2	1	-	2	1	**BZ2-009-10**	NR	1	3	2	3	2	2
BZ2-001-4	R	1	3	2	3	3	2	**BZ2-006-14**	R	3	2	2	3	3	3	BZ2-009-11	R	1	3	2	2	3	3
BZ2-001-5	R	1	1	2	1	1	1	**BZ2-007-1**	NR	3	2	2	3	3	2	**BZ2-009-12**	R	3	3	2	3	1	3
BZ2-002-1	NR	1	3	2	1	1	3	BZ2-007-2	NR	3	2	2	2	3	3	**BZ2-010-1**	R	3	2	2	2	3	3
BZ2-002-2	NR	1	2	2	3	1	2	**BZ2-007-3**	R	1	2	2	3	3	3	**BZ2-010-2**	NR	-	3	2	2	2	3
BZ2-002-3	R	3	3	2	1	3	3	BZ2-007-4	R	3	2	2	2	2	3	**BZ2-010-3**	R	3	1	2	2	3	1
BZ2-002-4	R	3	2	1	1	3	3	**BZ2-007-5**	R	3	3	2	2	2	2	BZ2-010-4	R	3	2	1	2	3	1
BZ2-002-5	NR	3	3	2	2	2	3	**BZ2-007-6**	NR	3	3	2	3	3	3	**BZ2-010-5**	NR	3	1	2	2	2	3
BZ2-002-6	NR	3	2	2	1	2	3	**BZ2-007-7**	R	1	3	2	2	3	1	**BZ2-010-6**	R	3	2	2	3	3	3
BZ2-002-7	NR	3	2	2	3	1	2	BZ2-007-8	R	3	1	2	2	2	1	**BZ2-010-7**	NR	3	1	2	2	2	2
BZ2-002-8	R	1	-	2	-	-	2	BZ2-007-9	R	3	1	2	3	3	3	**BZ2-010-8**	R	3	2	2	3	3	3
BZ2-002-9	R	3	2	2	1	1	1	**BZ2-007-10**	NR	1	3	2	3	3	3	**BZ2-010-9**	R	1	3	2	1	1	3
BZ2-002-10	R	3	3	2	3	-	3	BZ2-007-11	NR	3	1	2	2	3	2	**BZ2-010-10**	R	3	2	2	3	2	3
BZ2-002-11	NR	1	2	1	2	-	2	**BZ2-007-12**	R	1	2	2	3	3	3	**BZ2-010-11**	R	2	3	2	3	3	3
BZ2-002-12	R	1	2	2	3	3	2	**BZ2-007-13**	R	3	1	2	3	3	1	**BZ2-010-12**	NR	2	3	2	3	3	3
BZ2-002-13	R	3	3	2	3	3	-	**BZ2-007-14**	NR	1	3	2	3	3	2	**BZ2-010-13**	R	3	3	2	1	3	2
BZ2-002-14	NR	1	1	2	1	1	3	BZ2-007-15	NR	3	1	2	3	3	2	BZ2-010-14	R	1	2	2	3	1	1
BZ2-002-15	R	1	3	1	3	3	2	**BZ2-007-16**	R	1	3	2	1	3	3	**BZ2-010-15**	NR	3	3	2	2	1	3
BZ2-002-16	R	3	3	1	1	1	3	**BZ2-007-17**	NR	1	-	1	2	-	-	**BZ2-010-16**	NR	1	3	2	3	2	3
BZ2-002-17	NR	3	2	2	3	-	-	**BZ2-007-18**	R	1	1	2	2	3	1	BZ2-010-17	R	1	3	2	3	2	3
BZ2-002-18	R	1	2	2	1	1	1	**BZ2-007-19**	R	3	2	2	3	3	3	**BZ2-010-18**	NR	3	3	2	3	3	3
BZ2-002-19	R	1	2	2	3	3	3	**BZ2-007-20**	R	3	2	2	3	3	1	**BZ2-010-19**	R	2	2	1	1	1	1
BZ2-002-20	R	1	2	2	3	3	2	**BZ2-007-21**	NR	3	2	2	2	1	2	**BZ2-010-20**	NR	1	2	1	2	-	-
BZ2-002-21	R	1	2	1	1	-	-	BZ2-008-1	R	1	1	2	3	3	2	BZ2-010-21	NR	3	3	2	3	3	3
BZ2-002-22	R	3	2	2	3	3	3	**BZ2-008-2**	R	3	1	2	3	3	3	BZ2-011-1	NR	1	-	1	-	1	1
BZ2-002-23	R	3	2	1	3	1	1	BZ2-008-3	R	3	1	2	1	3	3	**BZ2-011-2**	R	3	2	2	1	1	3
BZ2-002-24	R	3	2	2	3	3	3	BZ2-008-4	NR	3	2	2	2	3	1	BZ2-011-3	R	2	3	2	3	3	3
BZ2-002-25	NR	1	2	2	2	2	2	**BZ2-008-5**	R	3	3	2	2	2	3	**BZ2-011-4**	R	3	1	2	3	3	3
BZ2-004-1	R	3	1	2	3	3	-	**BZ2-008-6**	NR	3	2	2	3	2	3	BZ2-011-5	R	1	-	1	-	1	1
BZ2-004-2	R	1	3	2	3	1	1	**BZ2-008-7**	NR	3	3	2	2	2	2	**BZ2-011-6**	NR	3	2	2	3	3	2
BZ2-004-3	NR	3	2	2	3	1	-	**BZ2-008-8**	R	3	2	2	3	1	1	**BZ2-011-7**	NR	1	2	2	2	2	3
BZ2-004-4	R	1	3	2	3	3	3	**BZ2-008-9**	NR	3	2	2	2	3	3	**BZ2-011-8**	NR	2	1	1	2	-	1
BZ2-004-5	R	3	2	2	3	3	1	**BZ2-008-10**	R	1	1	2	3	3	1	**BZ2-011-9**	R	3	2	2	3	3	1
BZ2-004-6	R	3	1	2	-	1	2	**BZ2-008-11**	NR	3	2	2	2	1	2	BZ2-011-10	R	3	1	2	3	1	3
BZ2-006-1	NR	3	1	2	3	3	-	**BZ2-008-12**	R	3	3	2	2	3	3	BZ2-011-11	R	1	2	2	3	3	3
BZ2-006-2	NR	1	1	2	2	2	3	**BZ2-008-13**	R	1	2	2	3	3	3	**BZ2-011-12**	R	1	3	2	3	3	2
BZ2-006-3	R	3	3	2	3	3	3	**BZ2-008-14**	NR	1	3	2	2	2	2	**BZ2-011-13**	R	3	1	2	3	3	1
BZ2-006-4	R	3	2	2	3	1	1	**BZ2-009-1**	NR	3	3	2	3	3	3	**BZ2-011-14**	NR	3	1	2	3	3	3
BZ2-006-5	R	3	3	2	3	3	3	BZ2-009-2	R	1	3	1	3	3	3	BZ2-011-15	NR	3	3	2	3	1	3
BZ2-006-6	NR	3	3	2	3	1	3	**BZ2-009-3**	NR	3	2	2	2	2	2	**BZ2-011-16**	NR	1	2	2	2	-	3
BZ2-006-7	R	3	1	2	1	3	1	**BZ2-009-4**	R	1	3	2	1	3	1								
BZ2-006-8	R	3	1	2	2	2	3	**BZ2-009-5**	NR	3	3	2	3	3	2								

* Bold Plants were used for SNP calling in GBS; 1: homozygous Zd allele; 2: homozygous B73 allele; 3: heterozygous; -: missing data. *gt1*: *grassy tillers 1*; *tb1*: *teosinte branched 1*; *id1*: *indeterminate 1*; S2-1 and S2-2: SNP markers for *regrowth 1*; S7-1 and S7-2: SNP markers for *regrowth 2* (The two markers are dominant, respectively, for the parents and thus used as a combined marker for *regrowth 2*).

Table 2. Results of the χ^2 goodness-of-fit tests of three genetic models

Populations	**Observed**			**No. dominant genes (the expected R to NR ratio) and $P(\chi^2)$***		
	Total	R	NR	1 (3:1)	2 (9:7)	3 (27:37)
B73-Zd F_2	134	81	53	0.0001	**0.2964**	0.0001
B73-Zd F_3	72	52	20	**0.5862**	0.0063	0.0001
Zd-RF F_2	160	92	68	0.0001	**0.7499**	0.0001
Zd-RF F_3-3	15	12	3	**0.6547**	0.0639	0.3000
Zd-RF F_3-5	16	9	7	0.0833	**1.0000**	0.2547
Zd-RF F_3-9	16	13	3	**0.5637**	0.0438	0.0016

*: the best fit models are in bold.

Table 3. Results of the $\chi 2$ goodness-of-fitness tests for independent assortment of regrowth and the SNP-derived PCR markers among the *Zea mays* B73 - *Z. diploperennis* F2 plants.

Phenotypes	Genotypes* (ratio)	Numbers of Plants					
		S2-1		S2-2		S7-1/S7-2	
		Obs	Exp	Obs	Exp	Obs	Exp
Regrowth	*R_AA* (3/16)	18	24	18	23.44	26	23.64
	R_Aa (6/16)	45	48	53	46.88	40	47.25
	R_aa (3/16)	14	24	7	23.44	12	23.64
Non-regrowth	*rrAA* (1/16)	4	8	12	7.813	4	7.86
	rrAa (2/16)	24	16	18	15.63	26	15.75
	rraa (1/16)	23	8	17	7.813	18	7.86
$\chi 2$, *df=5*		39.979		25.358		27.220	
$P(\chi^2)$		<0.0001		<0.0001		<0.0001	

* *R* : regrowth allele, *r* : non-regrowth allele; *A* : *Zea diploperennis* SNP allele; *a* : *Z. mays* SNP allele.

Table 4. TThe actions taken and the numbers of SNPs revealed in each filtration step of SNP analysis of *Zea mays* B13 - *Z. diploperennis* F2 plants.

Steps	Actions	Numbers of SNPs on each chromosome										Total
		1	2	3	4	5	6	7	8	9	10	
Raw	Before the filtration	109,543	85,283	81,625	75,832	77,314	58,195	62,280	57,748	57,231	49,107	**714,158**
The 1st filter	Remove SNPs with MAF<0.01	5,751	4,966	4,708	3,376	4,409	2,938	3,108	3,210	2,982	2,477	**37,925**
The 2nd filter	Remove SNPs with missing data rate > 20%	1,628	1,476	1,200	942	1,197	761	877	877	741	732	**10,431**
The 3rd filter	Remove SNPs with $\chi 2 > 9,49$	82	120	120	112	198	87	144	20	29	34	**946**
The 4th filter	Remove immidiate neighoring SNPs with the same haplotype	51	77	75	82	111	49	98	16	16	22	**597**

Table 5. Ranges (bp), peak SNP positions (bp), LOD, χ^2 and $P(\chi^2)$ of the two candidate *regrowth* (*reg1* and *reg2*) loci**,** and adjacent maize genes per B73RefGen_V4.

Loci	Chr	The range (the peak)	LOD	χ^2	$P(\chi^2)$	Gene	Gene position	Annotation*
reg1	2	24,244,192 to 28,975,747 (27,934,739)	4.944	4.097	0.393	GRMZM2G002642	27,769,950 to 27,776,452	Polypeptide: Ankyrin-like protein
						Zm00001d002943	27,769,845 to 27,779,192	NA
						AC199765.4_FGT008	27,777,537 to 27,779,027	Enzyme: Cellulase
						Zm00001d002944	27,779,321 to 27,782,869	NA
						GRMZM2G138125	27,929,562 to 27,938,867	NA
						Zm00001d002950	27,929,335 to 27,938,856	NA
reg2	7	2,862,253 to 6,681,861 (5,060,739)	4.764	7.029	0.134	GRMZM2G048172	5,058,044 to 5,060,380	NA
						Zm00001d018780	5,058,282 to 5,060,333	NA

*NA: no annotation

Table 6. PCR primers and annealing temperatures.

Primers	Sequences	Annealing T
tb1MF	5′-AGTAGGCCATAGTACGTAC-3′	56℃
tb1MR	5′-CTCTTTACCGAGCCCCTACA-3′	
tb1ZF	5′-ACTCAACGGCAGCAGCTACCTA-3′	62℃
tb1ZR	5′-CGTGTGTGTGATCGAATGGT-3′	
tga1cF:	5′-AATAAAATAGAGGAACGTCA-3′	55℃
tga1cR:	5′-TGCTGCAAAGGATTACTGAT-3′	
id1cF	5'-ACCGGACGGATCGAGAGAAA-3'	55℃
id1cR	5′-CCGTACTCACTCGCAGATCG-3′	
mmc0381F:	5′-GTGGCCCTGTTGATGAG-3′	55℃
mmc0381R:	5′-CGACGAGTACCAGGCAT-3′	
gt1-ZF:	5′-TCGCCTACATGACCGAGTAC-3′	60℃
gt1-ZR:	5′-ATACTCTCAGCTGCTACGCG-3′	
gt1-MF:	5′-GAGACCGAGCTGCTGAAGAT-3′	58℃
gt1-MR:	5′-TGTAGCTGTTGTAGGCGTACT-3′	
S2-1MF	5′-CTCTTCGCCTACTGCTAT-3′	60℃
S2-1ZF	5′-CTCTTCGCCTACTGCTAC-3′	
S2-1R	5′-AATGTCAATGCAGACAAGCCT-3′	
S2-2MF	5′-CGATGGTGAACATGATAAACGGA-3′	60℃
S2-2MR	5′-TATGGCTTGATTCGCTCTCTT-3′	
S2-2ZF	5′-CGATGGTGAACATGATAAATAGG-3′	58℃
S2-2ZR	5′-ACGCAAAAAGTATGGCTTGAT-3′	
S7-1MF	5′-CGTATCATCATACGAGCATG-3′	63℃
S7-1MR	5′-TGAATGAGCTCGATTGTGCC-3′	
S7-2ZF	5′-GTGCCTACGCTCCATCCGAA-3′	60℃
S7-2ZR	5′-GTCGCTACCACTGTATCGCA-3′	

Supplementary Table S1. Oneway ANOVA Analysis of tiller number at tasseling (TNT) by regrowth

Source	*Df*	Sum of Squares	Mean Square	F-Ratio	Prob>F
Regrowth	1	3.115385	3.11538	0.8967	0.3531
Error	24	83.38462	3.47436		
Total	25	86.50000			

Supplementary Table S2. Segregation of regrowth among the *Zea mays* cv Rhee Flint-*Z. diploperennis* F2s and F3s*.

Plant	PT	Plant	PT	Plant	PT	Plant	PT	Plant	PT	Plant	PT	Plant	PT
ZR2-001-1	R	ZR2-001-33	NR	ZR2-001-69	NR	ZR2-001-102	R	ZR2-001-132	NR	ZR2-001-162	NR	ZR3-005-7	R
ZR2-001-2	NR	ZR2-001-34	R	ZR2-001-71	R	ZR2-001-103	R	ZR2-001-133	R	ZR2-001-163	R	ZR3-005-8	R
ZR2-001-3	R	ZR2-001-35	NR	ZR2-001-72	NR	ZR2-001-104	R	ZR2-001-134	R	ZR2-001-164	R	ZR3-005-9	R
ZR2-001-4	NR	ZR2-001-36	NR	ZR2-001-73	R	ZR2-001-105	R	ZR2-001-135	NR	ZR2-001-165	NR	ZR3-005-10	R
ZR2-001-5	R	ZR2-001-37	NR	ZR2-001-74	R	ZR2-001-106	R	ZR2-001-136	NR	ZR2-001-166	R	ZR3-005-11	NR
ZR2-001-6	R	ZR2-001-38	NR	ZR2-001-75	R	ZR2-001-107	NR	ZR2-001-137	R	ZR2-001-167	NR	ZR3-005-12	NR
ZR2-001-7	NR	ZR2-001-39	R	ZR2-001-77	NR	ZR2-001-108	NR	ZR2-001-138	R	ZR2-001-168	NR	ZR3-005-13	R
ZR2-001-9	R	ZR2-001-40	NR	ZR2-001-78	NR	ZR2-001-109	NR	ZR2-001-139	R	ZR2-001-169	NR	ZR3-005-14	NR
ZR2-001-10	NR	ZR2-001-42	NR	ZR2-001-79	R	ZR2-001-110	R	ZR2-001-140	NR	ZR2-001-171	R	ZR3-005-15	NR
ZR2-001-11	R	ZR2-001-43	R	ZR2-001-80	R	ZR2-001-111	NR	ZR2-001-141	R	ZR3-003-1	R	ZR3-005-16	NR
ZR2-001-12	R	ZR2-001-44	R	ZR2-001-81	R	ZR2-001-112	R	ZR2-001-142	R	ZR3-003-2	R	ZR3-009-1	R
ZR2-001-13	NR	ZR2-001-45	R	ZR2-001-82	NR	ZR2-001-113	NR	ZR2-001-143	R	ZR3-003-3	R	ZR3-009-2	R
ZR2-001-14	NR	ZR2-001-47	NR	ZR2-001-83	NR	ZR2-001-114	NR	ZR2-001-144	NR	ZR3-003-4	NR	ZR3-009-3	R
ZR2-001-15	R	ZR2-001-48	NR	ZR2-001-84	R	ZR2-001-115	NR	ZR2-001-145	R	ZR3-003-6	R	ZR3-009-4	R
ZR2-001-16	NR	ZR2-001-49	R	ZR2-001-85	R	ZR2-001-116	R	ZR2-001-146	R	ZR3-003-7	R	ZR3-009-5	R
ZR2-001-17	R	ZR2-001-51	NR	ZR2-001-86	R	ZR2-001-117	R	ZR2-001-147	R	ZR3-003-8	R	ZR3-009-6	NR
ZR2-001-18	NR	ZR2-001-53	NR	ZR2-001-87	NR	ZR2-001-118	NR	ZR2-001-148	R	ZR3-003-9	R	ZR3-009-7	R
ZR2-001-19	NR	ZR2-001-54	R	ZR2-001-88	NR	ZR2-001-119	R	ZR2-001-149	R	ZR3-003-10	1	ZR3-009-8	R
ZR2-001-20	R	ZR2-001-55	R	ZR2-001-89	NR	ZR2-001-120	R	ZR2-001-150	R	ZR3-003-11	NR	ZR3-009-9	R
ZR2-001-21	NR	ZR2-001-56	R	ZR2-001-90	R	ZR2-001-121	NR	ZR2-001-151	R	ZR3-003-12	R	ZR3-009-10	R
ZR2-001-22	NR	ZR2-001-57	R	ZR2-001-91	R	ZR2-001-122	R	ZR2-001-152	NR	ZR3-003-13	R	ZR3-009-11	R
ZR2-001-23	R	ZR2-001-58	NR	ZR2-001-92	R	ZR2-001-123	NR	ZR2-001-153	R	ZR3-003-14	R	ZR3-009-12	R
ZR2-001-24	R	ZR2-001-59	R	ZR2-001-93	R	ZR2-001-124	R	ZR2-001-154	R	ZR3-003-15	R	ZR3-009-13	R
ZR2-001-25	NR	ZR2-001-60	R	ZR2-001-94	NR	ZR2-001-125	R	ZR2-001-155	NR	ZR3-003-16	NR	ZR3-009-14	NR
ZR2-001-26	R	ZR2-001-62	R	ZR2-001-95	NR	ZR2-001-126	NR	ZR2-001-156	R	ZR3-005-1	R	ZR3-009-15	NR
ZR2-001-27	NR	ZR2-001-63	R	ZR2-001-97	R	ZR2-001-127	NR	ZR2-001-157	NR	ZR3-005-2	R	ZR3-009-16	R
ZR2-001-28	NR	ZR2-001-64	R	ZR2-001-98	R	ZR2-001-128	NR	ZR2-001-158	NR	ZR3-005-3	NR		
ZR2-001-30	R	ZR2-001-65	R	ZR2-001-99	R	ZR2-001-129	R	ZR2-001-159	NR	ZR3-005-4	NR	PT: phenotype	
ZR2-001-31	R	ZR2-001-67	NR	ZR2-001-100	R	ZR2-001-130	NR	ZR2-001-160	R	ZR3-005-5	R	R: regrowth	
ZR2-001-32	R	ZR2-001-68	NR	ZR2-001-101	R	ZR2-001-131	NR	ZR2-001-161	R	ZR3-005-6	R	NR: non-regrowth	

*F2 are indicated by ZR2 and F3 are indicated by ZR3.

Supplementary Table S3. Segregation of regrowth among the *Zea mays* cv B73 - *Z. diploperennis* F2s and F3s*.

Plant	PT	Plant	PT	Plant	PT	Plant	PT	Plant	PT	Plant	PT	Plant	PT
BZ2-001-1	R	BZ2-004-1	R	BZ2-007-11	NR	BZ2-009-6	R	BZ2-011-3	R	BZ3-010-1-18	R	BZ3-010-1-68	NR
BZ2-001-2	R	BZ2-004-2	R	**BZ2-007-12**	R	BZ2-009-7	NR	**BZ2-011-4**	R	BZ3-010-1-19	R	BZ3-010-1-59	R
BZ2-001-3	R	**BZ2-004-3**	NR	**BZ2-007-13**	R	**BZ2-009-8**	NR	BZ2-011-5	R	BZ3-010-1-20	NR	BZ3-010-1-61	NR
BZ2-001-4	R	BZ2-004-4	R	**BZ2-007-14**	NR	**BZ2-009-9**	R	**BZ2-011-6**	NR	BZ3-010-1-22	R	BZ3-010-1-62	R
BZ2-001-5	R	**BZ2-004-5**	R	BZ2-007-15	NR	**BZ2-009-10**	NR	**BZ2-011-7**	NR	BZ3-010-1-23	R	BZ3-010-1-63	R
BZ2-002-1	NR	**BZ2-004-6**	R	**BZ2-007-16**	R	BZ2-009-11	NR	**BZ2-011-8**	NR	BZ3-010-1-24	R	BZ3-010-1-64	NR
BZ2-002-2	NR	**BZ2-006-1**	NR	**BZ2-007-17**	NR	**BZ2-009-12**	R	**BZ2-011-9**	R	BZ3-010-1-25	R	BZ3-010-1-67	R
BZ2-002-3	R	BZ2-006-2	NR	**BZ2-007-18**	R	**BZ2-010-1**	R	BZ2-011-10	R	BZ3-010-1-26	R	BZ3-010-1-68	R
BZ2-002-4	R	BZ2-006-3	R	**BZ2-007-19**	R	**BZ2-010-2**	NR	BZ2-011-11	R	BZ3-010-1-27	R	BZ3-010-1-69	R
BZ2-002-5	NR	BZ2-006-4	R	**BZ2-007-20**	R	**BZ2-010-3**	R	**BZ2-011-12**	R	BZ3-010-1-30	R	BZ3-010-1-70	R
BZ2-002-6	NR	BZ2-006-5	R	**BZ2-007-21**	NR	BZ2-010-4	R	**BZ2-011-13**	R	BZ3-010-1-31	R	BZ3-010-1-71	R
BZ2-002-7	NR	**BZ2-006-6**	NR	BZ2-008-1	R	**BZ2-010-5**	NR	**BZ2-011-14**	NR	BZ3-010-1-32	R	BZ3-010-1-72	R
BZ2-002-8	R	**BZ2-006-7**	R	**BZ2-008-2**	R	**BZ2-010-6**	R	BZ2-011-15	NR	BZ3-010-1-33	NR	BZ3-010-1-73	R
BZ2-002-9	R	BZ2-006-8	R	BZ2-008-3	R	**BZ2-010-7**	NR	**BZ2-011-16**	NR	BZ3-010-1-34	R	BZ3-010-1-74	R
BZ2-002-10	R	**BZ2-006-9**	R	BZ2-008-4	NR	**BZ2-010-8**	R	BZ3-010-1-1	R	BZ3-010-1-35	R	BZ3-010-1-75	R
BZ2-002-11	NR	**BZ2-006-10**	R	**BZ2-008-5**	R	**BZ2-010-9**	R	BZ3-010-1-2	R	BZ3-010-1-38	NR	BZ3-010-1-76	NR
BZ2-002-12	R	BZ2-006-11	NR	**BZ2-008-6**	NR	**BZ2-010-10**	R	BZ3-010-1-3	R	BZ3-010-1-39	R	BZ3-010-1-77	NR
BZ2-002-13	R	**BZ2-006-12**	NR	**BZ2-008-7**	NR	**BZ2-010-11**	R	BZ3-010-1-4	R	BZ3-010-1-40	R	BZ3-010-1-78	R
BZ2-002-14	NR	BZ2-006-13	NR	**BZ2-008-8**	R	**BZ2-010-12**	NR	BZ3-010-1-5	R	BZ3-010-1-41	NR	BZ3-010-1-79	R
BZ2-002-15	R	**BZ2-006-14**	R	**BZ2-008-9**	NR	**BZ2-010-13**	R	BZ3-010-1-6	R	BZ3-010-1-42	NR	BZ3-010-1-80	R
BZ2-002-16	R	**BZ2-007-1**	NR	**BZ2-008-10**	R	BZ2-010-14	R	BZ3-010-1-7	R	BZ3-010-1-46	NR	BZ3-010-1-81	NR
BZ2-002-17	NR	BZ2-007-2	NR	**BZ2-008-11**	NR	**BZ2-010-15**	NR	BZ3-010-1-8	NR	BZ3-010-1-47	R	BZ3-010-1-82	NR
BZ2-002-18	R	**BZ2-007-3**	R	**BZ2-008-12**	R	**BZ2-010-16**	NR	BZ3-010-1-9	NR	BZ3-010-1-48	R	BZ3-010-1-83	NR
BZ2-002-19	R	BZ2-007-4	R	**BZ2-008-13**	R	BZ2-010-17	R	BZ3-010-1-11	NR	BZ3-010-1-50	R	BZ3-010-1-84	R
BZ2-002-20	R	**BZ2-007-5**	R	**BZ2-008-14**	NR	**BZ2-010-18**	NR	BZ3-010-1-12	R	BZ3-010-1-51	NR	BZ3-010-1-85	R
BZ2-002-21	R	**BZ2-007-6**	NR	**BZ2-009-1**	NR	**BZ2-010-19**	R	BZ3-010-1-13	NR	BZ3-010-1-52	R	BZ3-010-1-87	R
BZ2-002-22	R	**BZ2-007-7**	R	BZ2-009-2	R	**BZ2-010-20**	NR	BZ3-010-1-14	NR	BZ3-010-1-53	R	**Bold: used for GBS**	
BZ2-002-23	R	BZ2-007-8	R	**BZ2-009-3**	NR	BZ2-010-21	NR	BZ3-010-1-15	R	BZ3-010-1-54	R	PT: phenotype	
BZ2-002-24	R	BZ2-007-9	R	**BZ2-009-4**	R	BZ2-011-1	NR	BZ3-010-1-16	R	BZ3-010-1-56	R	R: regrowth	
BZ2-002-25	NR	**BZ2-007-10**	NR	**BZ2-009-5**	NR	**BZ2-011-2**	R	BZ3-010-1-17	R	BZ3-010-1-57	R	NR: non-regrowth	

* F2s are indicated by BZ2 and F3s are indicated by BZ3.

Suplementary Table S4. Phenotypes and the *gt1* , *id1* and *tb1* haplotypes of 26 F2 plants and three F3 populations of Zea mays cv Rhee Flint x Z. diploperennis cross.*

Line	PT	*tb1*	*id1*	*gt1*	Line	PT	*tb1*	*id1*	*gt1*	Line	PT	*tb1*	*id1*	*gt1*
RZ2-001-1	R	3	1	3	RZ3-003-2	R	1	2	1	RZ3-005-13	R	1	2	1
RZ2-001-2	NR	2	2	2	RZ3-003-3	R	1	2	1	RZ3-005-14	NR	1	2	1
RZ2-001-3	R	1	2	3	RZ3-003-4	NR	1	2	3	RZ3-005-15	NR	1	2	1
RZ2-001-4	NR	2	2	2	RZ3-003-6	R	1	2	3	RZ3-005-16	NR	1	2	1
RZ2-001-5	R	1	2	1	RZ3-003-7	R	1	2	1	RZ3-009-1	R	1	2	1
RZ2-001-6	R	3	1	1	RZ3-003-8	R	1	2	3	RZ3-009-2	R	1	2	2
RZ2-001-7	NR	2	2	2	RZ3-003-9	R	1	2	3	RZ3-009-3	R	1	2	2
RZ2-001-9	R	1	2	3	RZ3-003-10	R	1	2	3	RZ3-009-4	R	1	2	3
RZ2-001-10	NR	3	1	3	RZ3-003-11	NR	1	2	3	RZ3-009-5	R	1	2	1
RZ2-001-11	R	2	2	3	RZ3-003-12	R	1	2	1	RZ3-009-6	NR	1	2	2
RZ2-001-12	R	3	1	1	RZ3-003-13	R	1	2	3	RZ3-009-7	R	1	2	3
RZ2-001-13	NR	3	1	3	RZ3-003-14	R	1	2	1	RZ3-009-8	R	1	2	1
RZ2-001-14	NR	2	2	3	RZ3-003-15	R	1	2	1	RZ3-009-9	R	1	2	3
RZ2-001-15	R	3	1	1	RZ3-003-16	NR	1	2	2	RZ3-009-10	R	1	2	3
RZ2-001-16	NR	2	2	1	RZ3-005-1	R	1	2	1	RZ3-009-11	R	1	2	1
RZ2-001-17	R	3	1	3	RZ3-005-2	R	1	2	1	RZ3-009-12	R	1	2	3
RZ2-001-18	NR	3	1	2	RZ3-005-3	NR	1	2	1	RZ3-009-13	R	1	2	1
RZ2-001-19	NR	3	1	2	RZ3-005-4	NR	1	2	1	RZ3-009-14	NR	1	2	3
RZ2-001-20	R	3	1	3	RZ3-005-5	R	1	2	1	RZ3-009-15	NR	1	2	3
RZ2-001-21	NR	2	2	3	RZ3-005-6	R	1	2	1	RZ3-009-16	R	1	2	1
RZ2-001-22	NR	2	2	3	RZ3-005-7	R	1	2	1	RZ3-012-1	R	1	2	1
RZ2-001-23	R	2	2	3	RZ3-005-8	R	1	2	1	RZ3-012-2	R	1	2	1
RZ2-001-24	R	1	2	3	RZ3-005-9	R	1	2	1	RZ3-012-3	R	1	2	1
RZ2-001-25	R	2	1	1	RZ3-005-10	R	1	2	1	RZ3-012-4	R	3	1	1
RZ2-001-26	R	3	2	1	RZ3-005-11	NR	1	2	1	RZ3-012-5	R	3	1	1
RZ3-003-1	R	1	2	3	RZ3-005-12	NR	1	2	1					

* R: regrowth; NR: non-regrowth; 1: homozygous for the *Zea diploperennis* allele; 2: homozygous for the *Z. mays*'s allele; 3: heterozygous; F2s are indicated by RZ2 and F3s are indicated RZ3.

Supplementary Table S5. R codes for candidate locus analysis

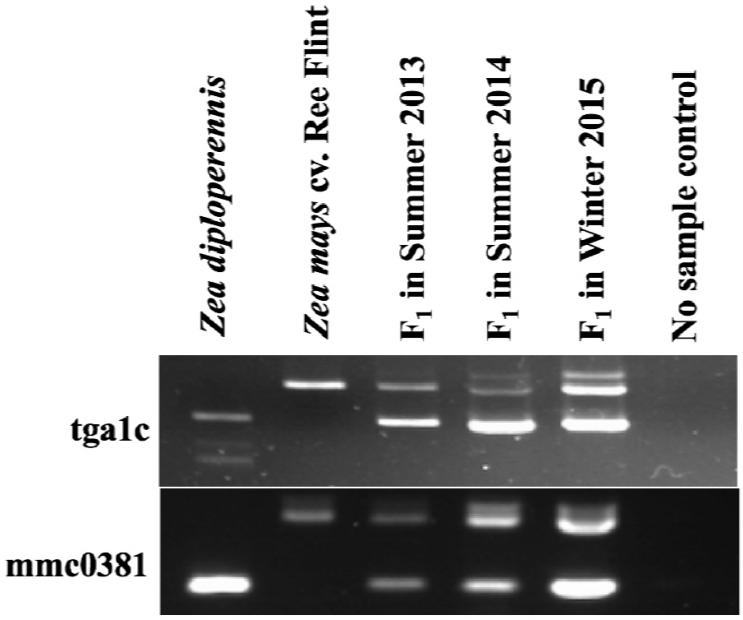
```
library(qtl)

all <- read.cross("csv", file="SNP.csv", genotypes = c("AA","AB", "BB"),
na.strings = "NA", alleles = c("A", "B"))

all <- calc.genoprob(all, step=1.0, off.end = 0.0, error.prob = 1.0e-
4,map.function = "haldane",stepwidth = "fixed")

all <- sim.geno(all, n.draws=32, step=1.0, off.end = 0.0, error.prob = 1.0e-
4,map.function = "haldane",stepwidth = "fixed")

all.scan1 <- scanone(all, pheno.col=2, model="binary", method = "em")

all.scan1.perm <- scanone(all, pheno.col = 2, model = "binary", method="em",
n.perm = 1000)

plot(all.scan1,main="LOD plot of regrowth",ylim = c(0,6))

threshold <- summary(all.scan1.perm, alpha=c(0.1, 0.05, 0.01))

abline(h=threshold[1], lty="dashed", lwd=1, col="blue")

abline(h=threshold[2], lty="dashed", lwd=1, col="yellow")

abline(h=threshold[3], lty="dashed", lwd=1, col="red")

summary(all.scan1, perm=all.scan1.perm, lodcolumn=1, alpha=0.1)

mkname1 <- find.marker(all, chr=2, pos=24.244290)

mkname2 <- find.marker(all, chr=7, pos=5.060739)

effectplot(all,pheno.col=2,mname1= mkname1), ylim=c(0,1))

effectplot(all,pheno.col=2,mname1= mkname2), ylim=c(0,1))

write.csv(all.scan1, "all.scan1.csv",row.names = TRUE)
```

Supplementary Table S6. The SNPs used for marker development, their positions in B73 reference genome and the marker genotypes in *Zea diploperennis* (Zd) and *Z. mays* B73 and 83 B73-Zd F2

Plant	PT	S2-1 (27774017 bp)		S2-2 (7934739 bp)		S7-1 (5835410 bp)		S7-2 (5060739 bp)		S7-1 /S7-2
		SNP	Marker	SNP	Marker	SNP	Marker	SNP	Marker	
Zd	R	C	1	G	1	C/T	1	A	-	1
B73	NR	T	2	A	2	C	-	C	2	2
BZ2-001-4	R	C	3	A/G	3	C	-	C	2	2
BZ2-001-5	R	N	1	G	1	N	1	A	-	1
BZ2-002-7	NR	N	3	A	1	C	-	C	2	2
BZ2-002-9	R	C	1	G	1	C	1	A	-	1
BZ2-002-10	R	C	3	G	-	C	1	N	2	3
BZ2-002-11	NR	N	2	N	-	N	-	C	2	2
BZ2-002-17	NR	N	3	N	-	N	-	C	-	-
BZ2-002-18	R	C	1	G	1	C	1	A	-	1
BZ2-002-21	R	N	1	N	-	N	-	N	-	-
BZ2-002-22	R	C	3	A/G	3	C	1	A/C	2	3
BZ2-002-23	R	N	3	G	1	C	1	A	-	1
BZ2-002-25	NR	T	2	A	2	N	-	N	2	2
BZ2-004-3	NR	N	3	A	1	N	-	C	-	-
BZ2-004-5	R	T	3	N	3	C	1	A	-	1
BZ2-004-6	R	C	-	G	1	N	-	C	2	2
BZ2-006-1	NR	T	3	A/G	3	C	-	C	-	-
BZ2-006-6	NR	T	3	A/G	1	C/T	1	A/C	2	3
BZ2-006-7	R	T	1	G	3	T	1	N	-	1
BZ2-006-9	R	T	2	A/G	3	T	1	A	-	1
BZ2-006-10	R	C	1	G	1	T	1	A	-	1
BZ2-006-12	NR	T	3	A	2	N	1	A/C	2	3
BZ2-006-14	R	C/T	3	A/G	3	C/T	1	A/C	2	3
BZ2-007-1	NR	N	3	A/G	3	C	-	C	2	2
BZ2-007-3	R	N	3	A/G	3	C/T	1	A/C	2	3
BZ2-007-5	R	T	2	A	2	C	-	C	2	2
BZ2-007-6	NR	C	3	G	3	C/T	1	A	2	3
BZ2-007-7	R	C/T	2	A/G	3	T	1	A	-	1
BZ2-007-10	NR	C	3	G	3	T	1	A	2	3
BZ2-007-12	R	N	3	A/G	3	N	1	A/C	2	3
BZ2-007-13	R	C/T	3	G	3	T	1	A	-	1
BZ2-007-14	NR	N	3	A/G	3	C	-	C	2	2
BZ2-007-16	R	C	1	A/G	3	C/T	1	A/C	2	3
BZ2-007-17	NR	T	2	N	-	N	-	C	-	-
BZ2-007-18	R	C	2	N	3	T	1	A	-	1
BZ2-007-19	R	N	3	A	3	T	1	A/C	2	3
BZ2-007-20	R	C/T	3	G	3	T	1	A	-	1
BZ2-007-21	NR	C	2	N	1	C	-	C	2	2
BZ2-008-2	R	C/T	3	A/G	3	C/T	1	A/C	2	3
BZ2-008-5	R	T	2	A	2	C	1	A/C	2	3

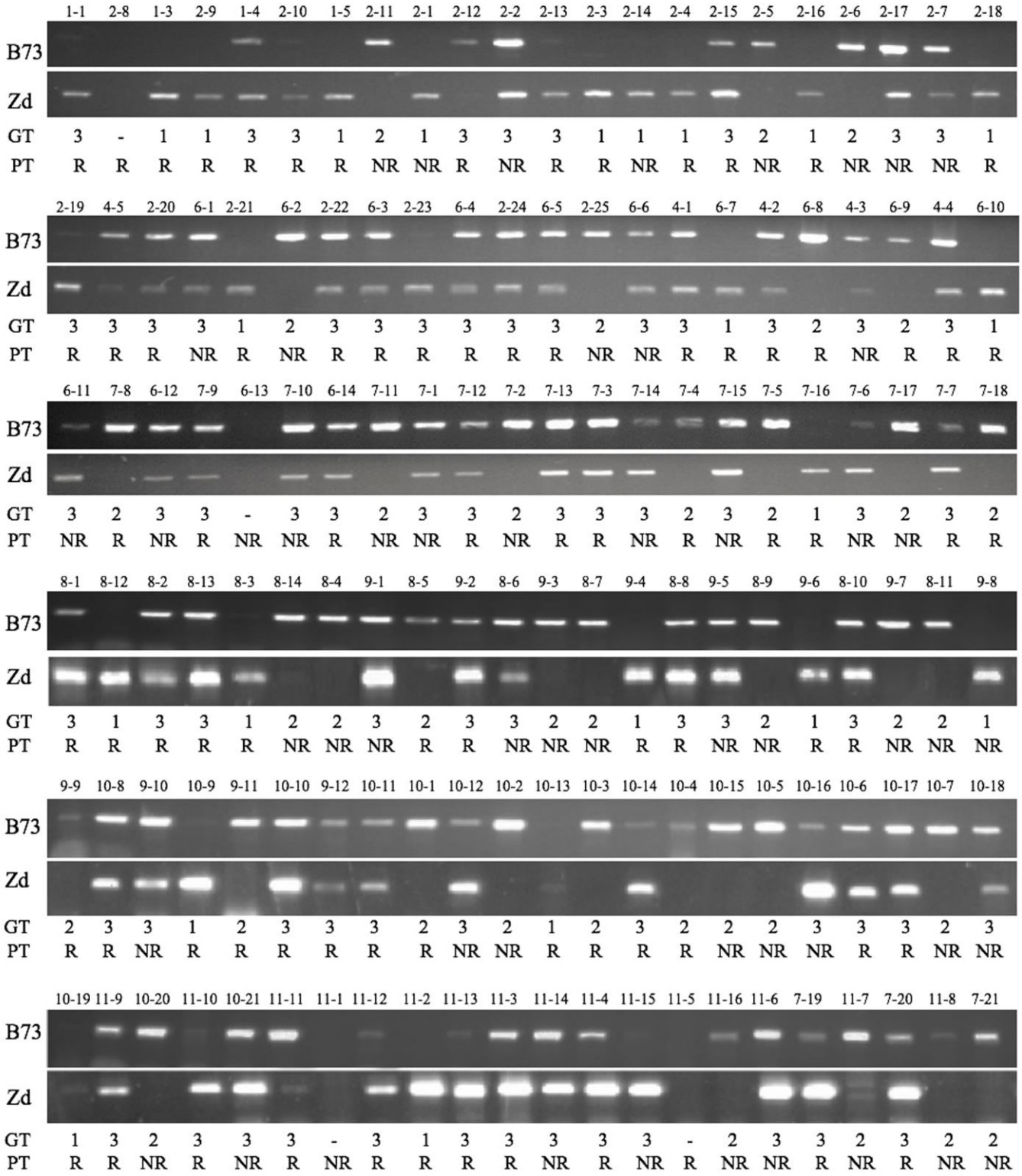
BZ2-008-6	NR	N	3	A/G	2	T	1	A/C	2	3
BZ2-008-7	NR	T	2	A	2	N	-	C	2	2
BZ2-008-8	R	N	3	A/G	1	T	1	A	-	1
BZ2-008-9	NR	T	2	A	3	N	1	N	2	3
BZ2-008-10	R	C	3	A/G	3	T	1	A	-	1
BZ2-008-11	NR	C	2	N	1	T	-	N	2	2
BZ2-008-12	R	C/T	2	A/G	3	C/T	1	A/C	2	3
BZ2-008-13	R	N	3	A	3	C/T	1	A/C	2	3
BZ2-008-14	NR	N	2	A	2	C	-	C	2	2
BZ2-009-1	NR	T	3	A/G	3	C	1	A	2	3
BZ2-009-3	NR	T	2	A	2	N	-	A/C	2	2
BZ2-009-4	R	C	1	G	3	C	1	A	-	1
BZ2-009-5	NR	T	3	A/G	3	C	-	N	2	2
BZ2-009-8	NR	C	1	G	1	C	-	C	2	2
BZ2-009-9	R	N	2	G	3	C	-	C	2	2
BZ2-009-10	NR	N	3	A	2	C	-	C	2	2
BZ2-009-12	R	C/T	3	A/G	1	C	1	A	2	3
BZ2-010-1	R	N	2	A/G	3	T	1	A/C	2	3
BZ2-010-2	NR	T	2	A	2	C/T	1	A/C	2	3
BZ2-010-3	R	N	2	A/G	3	T	1	A	-	1
BZ2-010-5	NR	T	2	A	2	C/T	1	A/C	2	3
BZ2-010-6	R	T	3	A/G	3	T	1	A/C	2	3
BZ2-010-7	NR	T	2	A	2	C	-	C	2	2
BZ2-010-8	R	N	3	N	3	N	1	N	2	3
BZ2-010-9	R	C	1	G	1	C/T	1	A/C	2	3
BZ2-010-10	R	T	3	N	2	C	1	C	2	3
BZ2-010-11	R	C/T	3	A/G	3	C/T	1	A/C	2	3
BZ2-010-12	NR	T	3	A	3	N	1	A	2	3
BZ2-010-13	R	C	1	G	3	C	-	C	2	2
BZ2-010-15	NR	C	2	N	1	C/T	1	A/C	2	3
BZ2-010-16	NR	T	3	A	2	T	1	C	2	3
BZ2-010-18	NR	T	3	N	3	C/T	1	C	2	3
BZ2-010-19	R	N	1	G	1	T	1	A	-	1
BZ2-010-20	NR	N	2	N	-	C	-	N	-	-
BZ2-011-2	R	C	1	G	1	C/T	1	A/C	2	3
BZ2-011-4	R	T	3	A/G	3	C/T	1	A/C	2	3
BZ2-011-6	NR	T	3	N	3	C	-	C	2	2
BZ2-011-7	NR	T	2	A	2	C	1	C	2	3
BZ2-011-8	NR	N	2	N	-	N	1	N	-	1
BZ2-011-9	R	C	3	A/G	3	T	1	A	-	1
BZ2-011-12	R	C	3	G	3	C	-	C	2	2
BZ2-011-13	R	C	3	G	3	T	1	A	-	1
BZ2-011-14	NR	T	3	A/G	3	C	1	A/C	2	3
BZ2-011-16	NR	T	2	N	-	N	1	N	2	3

Supplementary Figure S1. A photo showing the growth of *Zea diploperennis* and its F_1 with Z. mays B73 or Mo17 in the field in the Summer 2017.

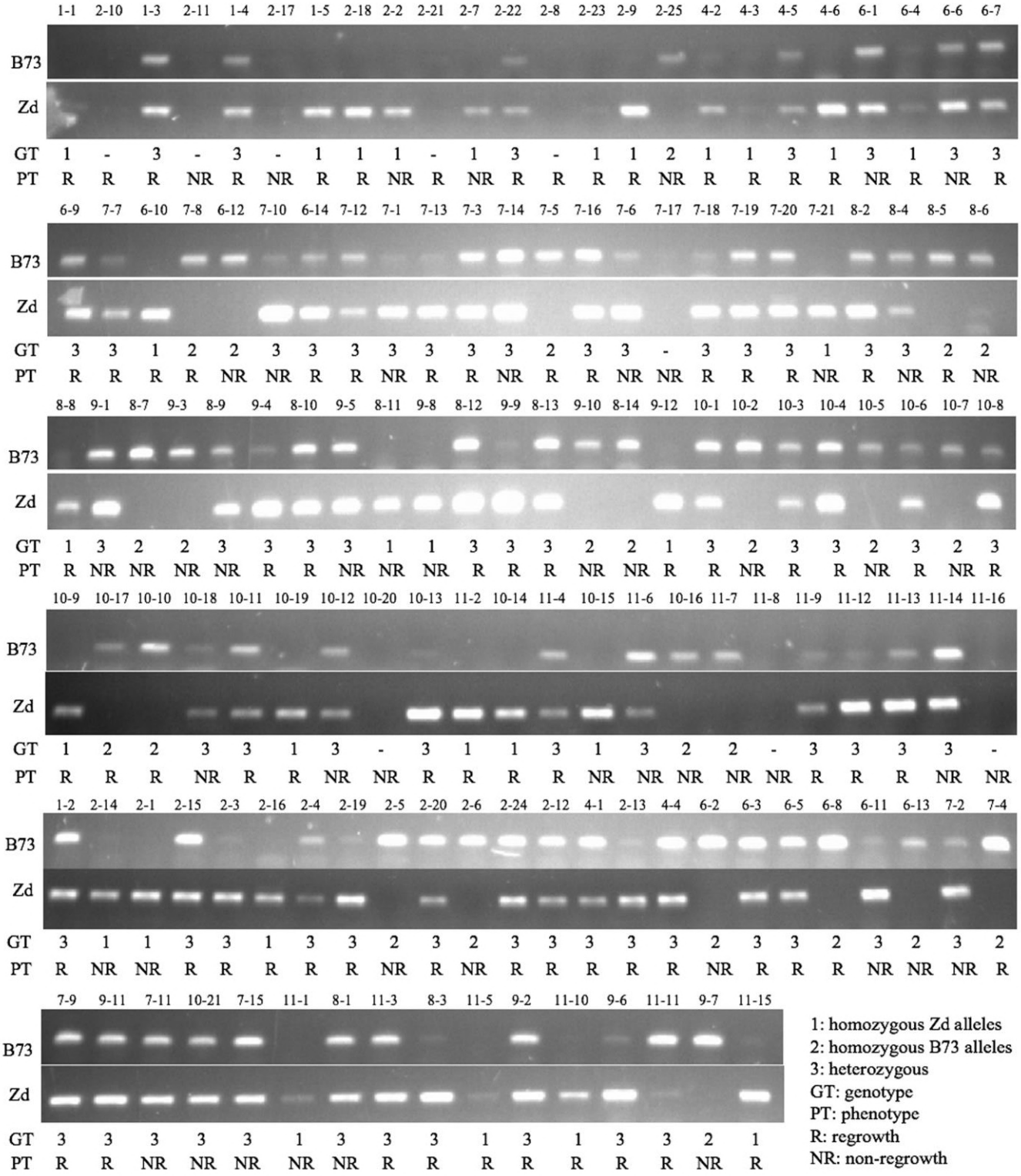


Supplementary Figure S2. An agarose gel image showing that two molecular markers confirmed the heterozygosity of a *Z. diploperennis-Z. mays* cv. Rhee Flint F_1 plant over three life-cycles.

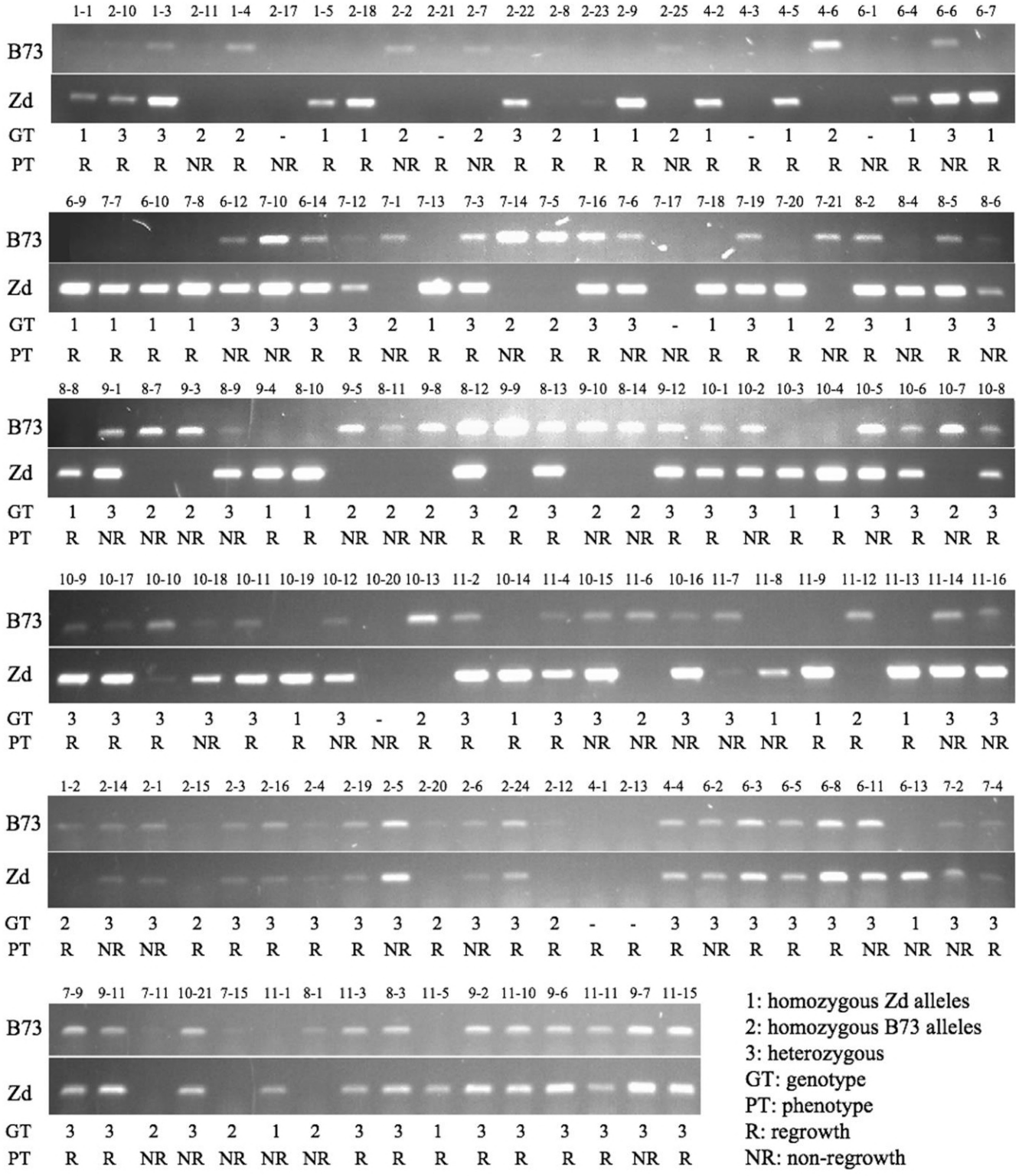
Supplementary Figure S3. A photo showing the growth of the B73 - *Z. diploperennis* regrowth F_4s in the field in summer, 2017.



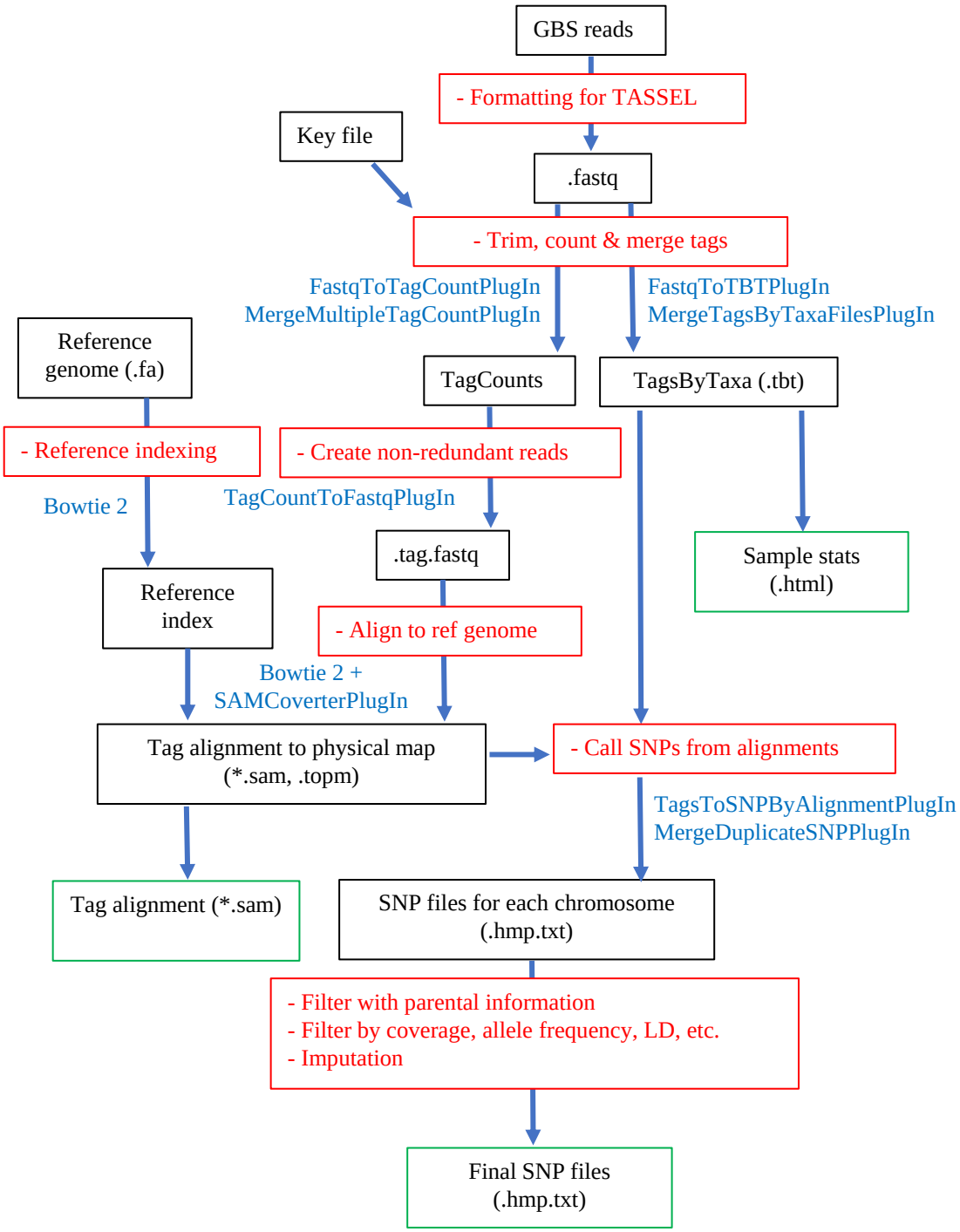
Supplementary Figure S6. Gel images show presence or absence of the dominant alleles of *Zea diploperennis* (Zd) or *Z. mays* B73 of the marker S2-1 for the *reg1* locus among 134 B73-Zd F_2 plants. The codes of the plants are listed on the top of the image as 1-1 refers to BZ2-001-1, etc.. * 1: homozygous Zd alleles; 2: homozygous B73 alleles; 3: heterozygous; GT: genotype; PT: phenotype; R: regrowth; NR: non-regrowth.



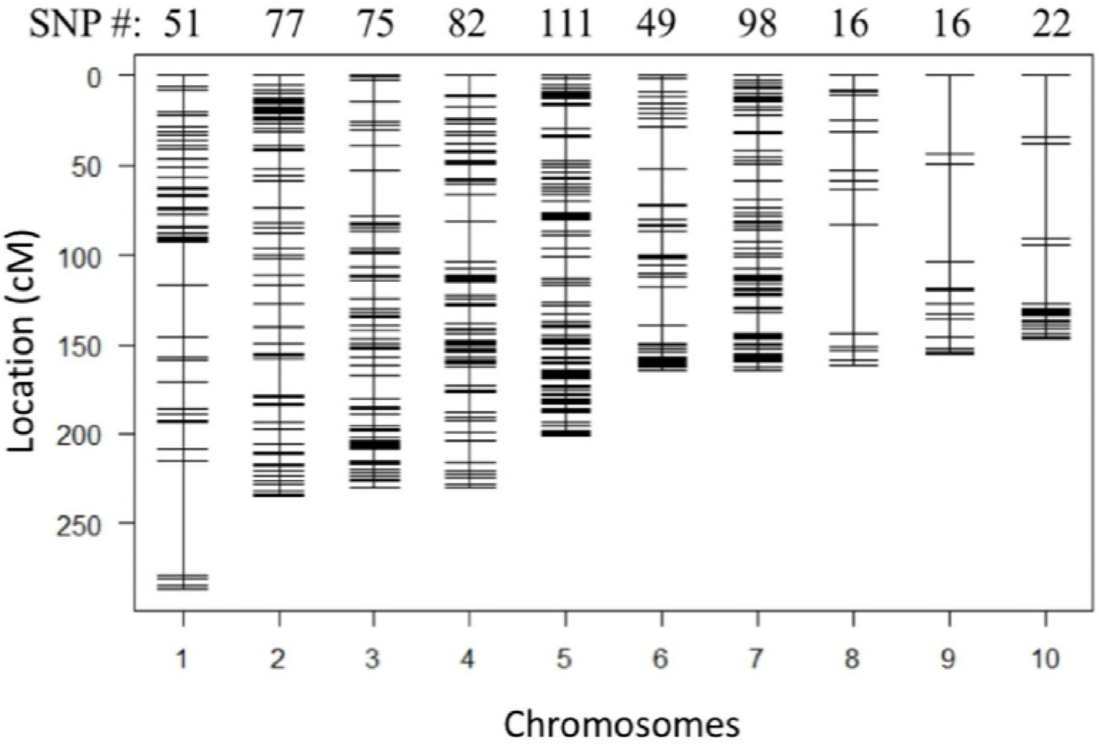
Supplementary Figure S7. Gel images showing presence or absence of the dominant allele of *Zea diploperennis* (Zd) or *Z. mays* B73 of the SNP marker S2-5 for the *reg1* locus among 134 B73-Zd F_2 plants. The codes of the plants are listed on the top of the image as 1-1 refers to BZ2-001-1, etc.



Supplementary Figure S8. Gel images showing presence or absence of the dominant *Zea diploperennis* (Zd) allele of marker S7-1 or *Z. mays* B73 allele of marker S7-2 for the *reg2* locus among 134 B73-Zd F_2 plants. The codes of the plants are listed on the top of the image as 1-1 refers to BZ2-001-1, etc.



Supplementary Figure S4. An illustration of the general process of TASSEL pipeline used in this study. The barcoded sequence reads are collapsed into a set of unique sequence tags with counts. The tag count files are filtered for minimum count threshold and merged into the master tag count file. Master tags are aligned to the reference genome to generate a "Tags On Physical Map" (TOPM) file, which contains the genomic position of each tag with the best unique alignment. The occupancies of tags for each taxon are observed from barcodes information in the original FASTQ files. The TOPM and TBT files are used to call SNPs at the tag locations on the genome.



Supplementary Figure S5. Genetic map of the SNPs on each chromosome of *Zea mays* B73.

Figure 1. Photos of *Zea mays* and *Z. diploperennis* (Zd) F_1 plants. **A**: reciprocal Mo17-Zd (right) and Zd-Mo17 (left) F_1 plants; **B**: reciprocal B73-Zd (right) and Zd-B73 (left) F_1 plants; **C**: RF-Zd F_1 plant; **D**: regrowth of a Mo17-Zd F_1 plant; **E**: regrowth of a B73-Zd F_1 plant; and **F**: regrowth of a RF-Zd F_1 plant. B73, Mo17 and RF represent, respectively, inbred lines B73 and Mo17 and cultivar Rhee Flint of *Z. mays*.

Figure 2. Photos of abnormal F_1 plants of crosses of *Zea diploperennis* with *Z. mays* inbred lines B73 (A & B) and M017 (C) or cv. Rhee Flint (D).

Figure 3. Photos of the ears produced from a *Zea mays* cv Rhee Flint x *Z. diploperennis* F_1 plant in different seasons (the upper panel) and from F_2 in summer 2014 in greenhouse (the lower panel).

Figure 4. Photos of *Zea mays* Mo17-*Z. diploperennis* F_2 plants, showing regrowth from the basal node of a single-stalked plant (A) or non-regrowth from a multi-stalked plant (B).

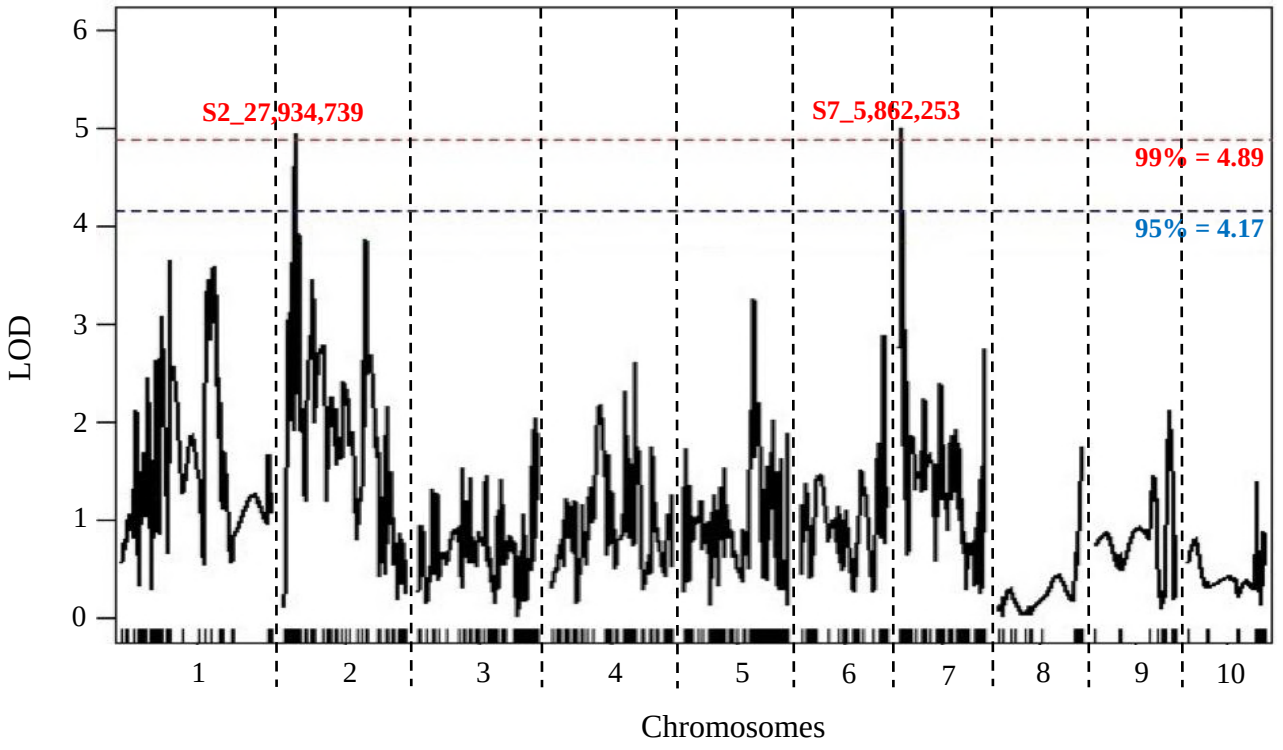


Figure 5. LOD scores of 597 SNP markers and 1,969 simulated positions for candidate locus determination shown with 95% and 99% LOD thresholds. The thresholds were calculated with 1000 permutation. Two significant QTL are indicated by the location of the peak SNPs.