Introduction

Hybrid poplars have been intensively cultivated in North America as a short-rotation woody crop species for bioenergy and pulp and paper industries [1,2,3,4]. The recent focus on lignocellulosic biofuels from plant biomass as a complement to fossil fuels has led to increased interest in the genetic characteristics of *Populus* as a rapidly growing biomass feedstock [5]. Several factors account for *Populus* success as a feedstock for biofuels and pulp and paper industries, foremost of which is the interspecific hybridization resulting in hybrids with marked improvement in performance, hybrid vigor (i.e. heterosis or over-dominance) compared to parental genotypes [6]. Being a genetically diverse genus which displays considerable variation among its species in such adaptive traits as stem growth, crown architecture, and disease resistance, the genetic diversity can be captured through the ease of hybridization among the approximately 30 species of *Populus*. The most desirable clones from these hybrid combinations can then be easily propagated by the well-developed vegetative systems inherent to *Populus*. Inter-american hybrids [7] generated from crosses between *Populus trichocarpa* Torr. & Gray and *Populus deltoides* Bartr. ex Marsh. (i.e., TxD hybrids) are estimated to produce as much as 35 Mg ha$^{-1}$ yr$^{-1}$ of aboveground biomass at age four [8]. Much of the success of the TxD hybrids is thought to result from the complementary combination of desirable traits inherited from each of the parental species in conjunction with the associated hybrid vigor [9]. Although hybrids exhibit superior performance at the overall phenotypic level, out-breeding depression or under-dominance, at the individual locus level resulting from combining alleles that result in poorer performance of the hybrid relative to parental genotypes is also a known genetic phenomenon [10], and may prevent hybrids from achieving maximum possible performance. Therefore, it is of primary interest to identify specific genomic loci contributing toward such hybrid vigor as well as those that contribute toward out-breeding depression for marker-assisted pyramiding of beneficial loci.

Biomass productivity traits are generally quantitative in nature involving numerous genes and genetic pathways whose activity may be modified by the environment leading, sometimes, to environment specific trait expression [11]. Understanding the genotype x environmental interactions of these genetic elements may enable targeted introgression of beneficial loci through ideotype breeding strategies [12]. Numerous studies have identified quantitative trait loci (QTLs) that are involved in biomass...
accumulation using hybrid *Populus* pedigrees. Results of these studies were generally reproducible among different studies and environments [13–17] suggesting that economically beneficial genes can be isolated for *Populus* improvement.

With the first genome sequence of any tree species [18] in addition to a well developed DNA marker resource, *Populus* has mature genomic resources that should allow for in-depth characterization of loci of interest. Although QTL intervals typically include tens or hundreds of genes, candidate gene mapping has provided some insight into potentially valuable gene targets for improving hybrid *Populus* production. In addition, prioritization of marker saturation can be accurately guided by knowing the physical interval compared to using cM distances whose relationship with physical distance is not always linear due to the heterogeneity of recombination rates across the *Populus* genome [19]. To date, numerous strategies have been applied in mapping efforts to identify potentially viable loci down to the gene level. These strategies include differential gene content in syntenic genomic regions where QTL presence or absence may suggest genetic determinants of economically important traits [20].

The overarching objective of this study was to build on previous work that identified and characterized genetically-driven variation in growth phenotypes within an F2, *P. trichocarpa* × *P. deltoides* pedigree. In this work, estimates of QTL numbers [12], QTL positions on a genetic linkage map [14], and aspects related to hybrid vigor, genotype-by-environment interaction (G×E), genetic correlations, and broad-sense heritability at the phenotypic level [15] were characterized. However, due to the molecular anonymity of RFLP, STS, and RAPD markers used in these studies and the unavailability of a reference genome at that time, the genomic location and genetic features associated with these loci remained anonymous. The completion of the *Populus* genome assembly [18] and the incorporation of SSR markers with known physical positions in the genetic map [21] now provide an opportunity for further characterization of loci described in the precedent work.

Therefore, the goals of this study were to reanalyze phenotypic data for the F2 pedigree to map QTLs segregating for stem height and diameter using an updated genetic map that incorporated SSR markers with known physical positions. In addition, we sought to provide estimates of the role of G×E interactions and hybrid vigor at the individual QTL level. Finally, we sought to utilize flanking SSR markers to delimit genomic intervals that encompass these QTLs thereby enabling the characterization of genetic features associated with these loci.

**Results**

**QTL Analysis and Detection Across Contrasting Environments**

Nine putative and three suggestive QTLs were detected for stem height and diameter on eleven linkage groups (LGs) of the family 331 genetic map (Tables 1 and 2). Among the 9 QTLs, 5 were significant in at least one experiment when the more stringent genomewide LOD threshold was used as the cut-off point (Table 1). All QTLs identified in this study were associated with both height and diameter in the same experiment or across different experiments. Of these, six QTLs on LGs I, II, VII, VIII, XIII, and XIV were detected in both Boardman and Clatskanie experimental sites. Three of these on LGs I, II, and XIV were detected in all 5 datasets analyzed. QTLs BM-3 and BM-8 exhibited the highest level of location specificity with QTL × environment interactions showing significance at Prob (>F) <0.1 (Table 3). All QTLs mapped reproducibly in the same map interval and peak positions were typically associated with no more than three markers in close proximity (Table 1). Figure 1 shows a graphical example of QTL BM-2 and associated peak on LG II highlighting the close agreements between three different phenotypic datasets used to identify the QTL. Averaged across experiments, the percent phenotypic variance explained ranged from 3.2 to 8.5% for each QTL.

**Additive and Dominant Effects**

One putative and one suggestive QTL exhibited consistent evidence of over-dominance across experiments and traits, whereas two putative QTLs exhibited consistent under-dominance across traits and environments (Tables 1 and 2). QTL BM-7 on LG XIV and a suggestive QTL on LG I were detected in all five datasets and exhibited over-dominance in each case (Tables 1 and 2). Putative QTLs BM-1 and BM-2 on LGs I and II, respectively, were also detected in all 5 datasets but each exhibited over-dominance in four of the five instances (Table 1). On the other hand, QTLs BM-4 and BM-6 on LGs VII and XIII, respectively, exhibited under-dominance across different environments for both height and diameter (Table 1).

**Genome Anchoring of QTL Intervals**

In this study, seven of the nine putative QTLs BM-1, BM-2, BM-3, BM-4, BM-5, BM-6, and BM-8 were successfully anchored on the *Populus* genome assembly (Figures 2 and 3 and Table 4). QTL BM-9 located on LG XIX could not be anchored due to lack of flanking SSR markers with known physical positions. Figure 2 illustrates the use of two SSR markers flanking a QTL interval peaking at marker P_422 on LG II to anchor and estimate the QTL peak position on the genome assembly. The genome assembly position of QTL BM-7 on LG XIV was reported previously by Ranjan et al. [20]. Interestingly, both markers associated with BM-7 on LG XIV, GTACG-N1 and AGCGA-14 (Table 1), were associated with a QTL previously identified for root lignin percentage on the same map position [21]. The lignin QTL, RL-5, was subsequently anchored on the genome assembly and analyzed for candidate genes by Ranjan et al. [20]. Further, the QTL exhibited the same pattern of over-dominance for the lignin phenotype as it did for height and diameter in the current study. Genomic intervals covered by individual QTLs ranged from 1.3 to 8.8 Mb (Table 4).

Based on the QTL intervals defined here and the results of the SSR marker placement on the genome assembly, we identified 197 previously unmapped SSR markers that occurred within QTL intervals identified in this study (Table S1). The number of additional SSR markers ranged from 9 to 36 within individual QTLs (Tables 4 and S1).

**Candidate Gene Identification and Characterization**

Intervals spanning the genomic regions summarized in Table 4 were used to identify all genes occurring within the 8 genome-anchored QTLs (Table S2). The number of genes in each interval ranged from 37 for QTL BM-7 to 721 for QTL BM-8. All together, there were 3,031 genes within the 8 genome-anchored QTL intervals. Out of these, 1,892 (62%) had annotations based on PFAM domains and these fell into 290 unique annotations (Table S2). Of the 1,892 annotated gene models, 1,313 (72%) had at least one duplicate in a QTL interval mapping on a different scaffold (Table S3). These represented 255 (88%) of the 290 unique annotations (Table S3).
Discussion

The genus *Populus* is an economically important tree crop widely grown as feedstock for lignocellulosic biofuels and pulp and paper products, in part, for its rapid growth and ability to thrive in marginal lands that are not suitable for food crop production. Aside from its key importance as an industrial feedstock, *Populus* is also a biological model system for perennial tree crops because of its relatively compact genome, high level of interspecies diversity, and ease of experimental manipulation compared to other genera [4,18]. Increasing biomass productivity using genetic manipulation is of major interest in contemporary efforts to make biofuels production from *Populus* economically viable. Results presented here provide a basis for the isolation of specific genetic determinants that mediate expression of key productivity traits such as height and diameter. Additionally, this work represents a valuable resource in identification and prioritization of genomic intervals that may be targeted for marker-assisted breeding programs in *Populus*. Below, we discuss specific attributes of this work that should facilitate the application of results presented here in genetic improvement of *Populus* feedstocks.

Given the environmental contrast between the Boardman and Clatskanie experimental sites, QTL detection and expression for the stem-growth traits was remarkably consistent across sites. Only three of the twelve QTLs identified in this study exhibited evidence of location specificity. Therefore, these results indicate that the pattern of phenotypic response from each genotypic class was relatively consistent across the two environments. The generally robust detection of QTL regardless of environment is consistent with findings by Rae et al. [11] who evaluated progeny from the same pedigree in France, United Kingdom, and Italy and identified virtually the same QTL intervals reported in this current study. Given their detection across different geographical environments, QTLs segregating in this pedigree offer valuable targets for further characterization and utilization in improving *Populus* productivity.

Despite the overall hybrid vigor observed for both height and diameter, varying levels of locus-specific QTL mode of action were observed, ranging from under-dominance to over-dominance. The presence of loci exhibiting apparent out-breeding depression (under-dominance) on linkage groups VII and XIII highlight the potential for further improvement of hybrid performance using targeted exclusion of such loci in ideotype breeding approaches. Interestingly, Tschaplinski et al. [10] identified a QTL for osmotic potential exhibiting under-dominance on LG XIII that mapped in the same interval as QTL *BM-6*. Conversely, the QTL peak for *BM-7* on LG XIV, with consistent evidence of over-dominance for both height and diameter, co-located with the same markers which were associated with a root lignin QTL, *RL-5*, which, in a previous study, also exhibited the same over-dominance mode of action [20,21]. Identification of the same QTL interval and the consistency of QTL mode of action between traits suggest that both height and diameter are largely influenced by the same genes. Such pleiotropic effects have been widely described in *Populus* for both related and diverse traits [12,20,21]. Additionally, the association of particular genomic intervals previously associated with phenotypes other than height and diameter suggest that pleiotropy may also exist between apparently non-related traits in *Populus*, although genetic linkage cannot be ruled out based on available data.
<table>
<thead>
<tr>
<th>QTL (LG)</th>
<th>Marker at QTL peak</th>
<th>LOD Score</th>
<th>Trait</th>
<th>% PVE</th>
<th>mu (ac)</th>
<th>mu (ad)</th>
<th>mu (bc)</th>
<th>mu (bd)</th>
<th>a</th>
<th>d</th>
<th>lodc</th>
<th>d/lod</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM-1 (I)</td>
<td>AGCTG-1 (I)</td>
<td>3.26</td>
<td>2.9</td>
<td>4.3</td>
<td>3.28</td>
<td>1.17</td>
<td>1.17</td>
<td>1.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM-2 (II)</td>
<td>T_D03_07 (II)</td>
<td>3.66</td>
<td>2.8</td>
<td>4.3</td>
<td>3.66</td>
<td>1.23</td>
<td>1.23</td>
<td>1.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM-3 (V)</td>
<td>AGCTG-7 (V)</td>
<td>7.02</td>
<td>2.7</td>
<td>4.3</td>
<td>7.02</td>
<td>1.90</td>
<td>1.90</td>
<td>1.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM-4 (VII)</td>
<td>T_AE20_120 (VII)</td>
<td>3.37</td>
<td>2.6</td>
<td>4.3</td>
<td>3.37</td>
<td>0.90</td>
<td>0.90</td>
<td>0.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM-5 (VIII)</td>
<td>TCCGT-16 (VIII)</td>
<td>4.94</td>
<td>2.3</td>
<td>4.3</td>
<td>4.94</td>
<td>1.35</td>
<td>1.35</td>
<td>1.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM-6 (XIII)</td>
<td>CCCCT-N3 (XIII)</td>
<td>2.83</td>
<td>2.7</td>
<td>4.3</td>
<td>2.83</td>
<td>0.75</td>
<td>0.75</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM-7 (XIV)</td>
<td>CTACG-N1 (XIV)</td>
<td>4.34</td>
<td>2.7</td>
<td>4.3</td>
<td>4.34</td>
<td>1.25</td>
<td>1.25</td>
<td>1.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM-8 (XVII)</td>
<td>T_L09_105 (XVII)</td>
<td>2.50</td>
<td>2.4</td>
<td>4.3</td>
<td>2.50</td>
<td>0.75</td>
<td>0.75</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM-9 (XIX)</td>
<td>ACCTG-24 (XIX)</td>
<td>3.72</td>
<td>2.5</td>
<td>4.3</td>
<td>3.72</td>
<td>0.55</td>
<td>0.55</td>
<td>0.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

%PVE = percent phenotypic variance explained; Mean associated with heterozygous genotypes 'ac' and 'bd' where alleles are derived from the same species; Mean associated with heterozygous genotypes 'ad' and 'bc' where alleles are derived from different species; a = additive; d = dominance; d/|a| = QTL mode of action.
Although informative loci have been identified in other studies based on the candidate gene approach in diverse systems such as Populus [20], Loblolly pine [22], and cowpeas [23,24], the large genomic intervals encompassed in some of the QTL intervals can make the narrowing down of candidate genes difficult. This limitation was evident in our analyses where only 1 QTL interval (BM-7) encompassed less than 100 genes. Despite this limitation, the unique genome duplication event which resulted in synteny between different Populus chromosomal segments [18] provides a reasonable starting point at narrowing down the list of candidate genes. The extensive duplication and paralogous relationship between genes found in QTL intervals located on different chromosomes suggests that these QTLs may have been derived from the same ancestral sequences and is consistent with the high levels of inter-chromosomal synteny described previously by Tuskan et al. [18]. This assumption would imply that functional activity for these genes was conserved post-duplication leading to existence of paralogous QTLs on different Populus chromosomes. The opposite scenario, where functional divergence or gene loss occurred after the duplication event resulting in deferential QTL presence on otherwise syntenic chromosome intervals was harnessed to identify unique candidate genes for cell-wall chemistry in a previous study [20]. QTL relationships based on ancestral sequences and candidate gene analyses will benefit from further improvements in the genome assembly and annotation of all gene models. At present, 38% of genes present within the 8 QTL intervals did not have adequate annotation information. Although informative loci have been identified in other studies based on the candidate gene approach in diverse systems such as Populus [20], Loblolly pine [22], and cowpeas [23,24], the large genomic intervals encompassed in some of the QTL intervals can make the narrowing down of candidate genes difficult. This limitation was evident in our analyses where only 1 QTL interval (BM-7) encompassed less than 100 genes. Despite this limitation, the unique genome duplication event which resulted in synteny between different Populus chromosomal segments [18] provides a reasonable starting point at narrowing down the list of candidate genes. The extensive duplication and paralogous relationship between genes found in QTL intervals located on different chromosomes suggests that these QTLs may have been derived from the same ancestral sequences and is consistent with the high levels of inter-chromosomal synteny described previously by Tuskan et al. [18]. This assumption would imply that functional activity for these genes was conserved post-duplication leading to existence of paralogous QTLs on different Populus chromosomes. The opposite scenario, where functional divergence or gene loss occurred after the duplication event resulting in deferential QTL presence on otherwise syntenic chromosome intervals was harnessed to identify unique candidate genes for cell-wall chemistry in a previous study [20]. QTL relationships based on ancestral sequences and candidate gene analyses will benefit from further improvements in the genome assembly and annotation of all gene models. At present, 38% of genes present within the 8 QTL intervals did not have adequate annotation information. Although informative loci have been identified in other studies based on the candidate gene approach in diverse systems such as Populus [20], Loblolly pine [22], and cowpeas [23,24], the large genomic intervals encompassed in some of the QTL intervals can make the narrowing down of candidate genes difficult. This limitation was evident in our analyses where only 1 QTL interval (BM-7) encompassed less than 100 genes. Despite this limitation, the unique genome duplication event which resulted in synteny between different Populus chromosomal segments [18] provides a reasonable starting point at narrowing down the list of candidate genes. The extensive duplication and paralogous relationship between genes found in QTL intervals located on different chromosomes suggests that these QTLs may have been derived from the same ancestral sequences and is consistent with the high levels of inter-chromosomal synteny described previously by Tuskan et al. [18]. This assumption would imply that functional activity for these genes was conserved post-duplication leading to existence of paralogous QTLs on different Populus chromosomes. The opposite scenario, where functional divergence or gene loss occurred after the duplication event resulting in deferential QTL presence on otherwise syntenic chromosome intervals was harnessed to identify unique candidate genes for cell-wall chemistry in a previous study [20]. QTL relationships based on ancestral sequences and candidate gene analyses will benefit from further improvements in the genome assembly and annotation of all gene models. At present, 38% of genes present within the 8 QTL intervals did not have adequate annotation information. Although informative loci have been identified in other studies based on the candidate gene approach in diverse systems such as Populus [20], Loblolly pine [22], and cowpeas [23,24], the large genomic intervals encompassed in some of the QTL intervals can make the narrowing down of candidate genes difficult. This limitation was evident in our analyses where only 1 QTL interval (BM-7) encompassed less than 100 genes. Despite this limitation, the unique genome duplication event which resulted in synteny between different Populus chromosomal segments [18] provides a reasonable starting point at narrowing down the list of candidate genes. The extensive duplication and paralogous relationship between genes found in QTL intervals located on different chromosomes suggests that these QTLs may have been derived from the same ancestral sequences and is consistent with the high levels of inter-chromosomal synteny described previously by Tuskan et al. [18]. This assumption would imply that functional activity for these genes was conserved post-duplication leading to existence of paralogous QTLs on different Populus chromosomes. The opposite scenario, where functional divergence or gene loss occurred after the duplication event resulting in deferential QTL presence on otherwise syntenic chromosome intervals was harnessed to identify unique candidate genes for cell-wall chemistry in a previous study [20]. QTL relationships based on ancestral sequences and candidate gene analyses will benefit from further improvements in the genome assembly and annotation of all gene models. At present, 38% of genes present within the 8 QTL intervals did not have adequate annotation information. Although informative loci have been identified in other studies based on the candidate gene approach in diverse systems such as Populus [20], Loblolly pine [22], and cowpeas [23,24], the large genomic intervals encompassed in some of the QTL intervals can make the narrowing down of candidate genes difficult. This limitation was evident in our analyses where only 1 QTL interval (BM-7) encompassed less than 100 genes. Despite this limitation, the unique genome duplication event which resulted in synteny between different Populus chromosomal segments [18] provides a reasonable starting point at narrowing down the list of candidate genes. The extensive duplication and paralogous relationship between genes found in QTL intervals located on different chromosomes suggests that these QTLs may have been derived from the same ancestral sequences and is consistent with the high levels of inter-chromosomal synteny described previously by Tuskan et al. [18]. This assumption would imply that functional activity for these genes was conserved post-duplication leading to existence of paralogous QTLs on different Populus chromosomes. The opposite scenario, where functional divergence or gene loss occurred after the duplication event resulting in deferential QTL presence on otherwise syntenic chromosome intervals was harnessed to identify unique candidate genes for cell-wall chemistry in a previous study [20]. QTL relationships based on ancestral sequences and candidate gene analyses will benefit from further improvements in the genome assembly and annotation of all gene models. At present, 38% of genes present within the 8 QTL intervals did not have adequate annotation information. Although informative loci have been identified in other studies based on the candidate gene approach in diverse systems such as Populus [20], Loblolly pine [22], and cowpeas [23,24], the large genomic intervals encompassed in some of the QTL intervals can make the narrowing down of candidate genes difficult. This limitation was evident in our analyses where only 1 QTL interval (BM-7) encompassed less than 100 genes. Despite this limitation, the unique genome duplication event which resulted in synteny between different Populus chromosomal segments [18] provides a reasonable starting point at narrowing down the list of candidate genes. The extensive duplication and paralogous relationship between genes found in QTL intervals located on different chromosomes suggests that these QTLs may have been derived from the same ancestral sequences and is consistent with the high levels of inter-chromosomal synteny described previously by Tuskan et al. [18]. This assumption would imply that functional activity for these genes was conserved post-duplication leading to existence of paralogous QTLs on different Populus chromosomes. The opposite scenario, where functional divergence or gene loss occurred after the duplication event resulting in deferential QTL presence on otherwise syntenic chromosome intervals was harnessed to identify unique candidate genes for cell-wall chemistry in a previous study [20]. QTL relationships based on ancestral sequences and candidate gene analyses will benefit from further improvements in the genome assembly and annotation of all gene models. At present, 38% of genes present within the 8 QTL intervals did not have adequate annotation information. The approach described above, however, does not negate the need for additional marker saturation and fine-mapping of these intervals before candidate genes are selected for molecular validation. To that end, we identified 197 unmapped SSR markers that occurred within the 8 QTL intervals. These markers represent a potential source of SSR markers for use in saturating QTL intervals. Since QTLs identified in this study were independently veriﬁed in other studies, they potentially harbor key genes mediating biomass productivity as well as the expression of heterosis in Populus hybrids. The cumulative results presented in this study provide a basis for further genomic characterization of these high-value QTLs for subsequent use in improving biomass productivity in Populus.

Materials and Methods

Experiments were conducted in field sites located in Boardman and Clatskanie, Oregon. These two sites contrast in water...
availability and provide opportunity for characterizing the influence of differential water availability in trait expression [10]. Details related to the mapping population, experimental design, and phenotypic measurements were described previously [12,14,15]. Briefly, an interspecific T×D hybrid population (F1 Family 53) was generated in 1981 from a cross between a female P. trichocarpa (T, black cottonwood clone 93–968) and a male P. deltoides (D, eastern cottonwood clone ILL-129). Two resulting hybrids (F1 clones 53–246 and 53–242, female and male, respectively) were sib-mated in 1988 and again in 1990 to generate an F2 mapping population, Family 331; of approximately 375 individuals [13,15]. Phenotypic data for stem height and diameter for 310 of these individuals were reanalyzed in this study. Specifically, height and diameter data collected after 4 years of growth were reanalyzed for each site and height data collected after 8 years of growth was reanalyzed for the Clatskanie site.

Genomic Resources

We used the family 331 and Populus consensus genetic maps described by Yin et al. [21,25] for QTL identification and genome anchoring. Briefly, the family 331 map was based on 841 AFLP, RAPD, RFLP and SSR markers. Of these, 155 SSR markers were shared with the consensus map and were used to align the family 331 map to the 19 LGs corresponding to the Populus haploid chromosome number. Detailed map characteristics, marker nomenclature, and the resulting syntenic relationship between the genetic maps were summarized in Yin et al. [21].

Additionally, we used 2,524 SSR markers with known positions on the Populus V2.2 genome assembly (http://www.phytozome.net/cgi-bin/gbrowse/poplar/) to establish the genome assembly framework used for anchoring QTL. Procedure for assigning SSR markers to physical positions on the genome assembly was previously described by Ranjan et al. [20]. Briefly, the physical position of SSR markers in Populus genome sequence was obtained by BLAST search of the corresponding forward and reverse primers. Additional checks were done to ensure that the predicted SSR length based on BLAST result was same as the length of the actual sequenced SSR marker. MapChart 2.2 [26] was used to graphically represent synteny and collinearity between the genetic maps and the genome assembly.

QTL Analysis

The Multiple-QTL Model (MQM) package of MapQTL 6.0 [27] with automatic cofactor selection was used to map putative and suggestive QTL intervals on the family 331 genetic map. The non-parametric Kruskal-Wallis (KW) analysis with a significance threshold of 0.005 was used as secondary confirmation of detected QTL. Mean phenotypic values across ramets and replicates were analyzed separately for each trait and experiment. The criteria for declaring QTL was based on the MQM analysis in which results were subjected to permutation tests [28]. 1000 permutations were conducted separately for each trait and experiment to determine linkage groupwise and genomewise LOD significance thresholds at the 0.05 significance level. A putative QTL was declared when it was detected in at least two experiments or in the same experiment for both traits, with at least one of those instances exceeding the linkage groupwise LOD significance threshold. A suggestive QTL was declared when detected in at least two experiments or in one
experiment for both traits, with LOD scores above 2.0 but failing to exceed the LOD significance threshold in any one instance [29]. Results of the KW analysis were used for QTL verification only and were not reported in this manuscript. RAPDS and RFLP markers were excluded in the QTL analysis due to the small population size with available genotypic information for these markers as described by Yin et al. [21]. Nomenclature for naming putative biomass QTL was based on the abbreviation BM.

![Figure 3. Genome anchored QTL positions on the Populus V2.2 assembly.](image)

Blue bars represent SSR marker coverage for each scaffold, red bars indicate scaffold intervals between flanking SSR markers used for genome anchoring, vertical green lines represent QTL intervals and estimated peak position. Scaffold intervals are represented in Mb.

doi:10.1371/journal.pone.0054468.g003

<table>
<thead>
<tr>
<th>QTL</th>
<th>Scaffold</th>
<th>Marker at peak</th>
<th>Genome assembly interval (Mb)</th>
<th>QTL peak Position on assembly (Mb)</th>
<th>Interval size (Mb)</th>
<th>Additional SSRs¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM-1</td>
<td>1</td>
<td>AGCTG-1</td>
<td>9.30–14.50</td>
<td>13.10</td>
<td>5.20</td>
<td>35</td>
</tr>
<tr>
<td>BM-2</td>
<td>2</td>
<td>P_422</td>
<td>3.50–6.30</td>
<td>4.80</td>
<td>2.80</td>
<td>12</td>
</tr>
<tr>
<td>BM-3</td>
<td>5</td>
<td>AGCTG-7</td>
<td>22.00–23.30</td>
<td>22.20</td>
<td>1.30</td>
<td>12</td>
</tr>
<tr>
<td>BM-4</td>
<td>7</td>
<td>T_AE20_120</td>
<td>3.60–8.70</td>
<td>7.20</td>
<td>5.10</td>
<td>32</td>
</tr>
<tr>
<td>BM-5</td>
<td>8</td>
<td>TCCGT-16</td>
<td>3.40–6.60</td>
<td>4.60</td>
<td>3.20</td>
<td>35</td>
</tr>
<tr>
<td>BM-6</td>
<td>13</td>
<td>CCCCT-N3</td>
<td>0.60–4.10</td>
<td>1.90</td>
<td>3.50</td>
<td>36</td>
</tr>
<tr>
<td>BM-7²</td>
<td>14</td>
<td>AGCGA-14/CTACG-N1</td>
<td>0.01–0.50</td>
<td>0.24</td>
<td>0.49</td>
<td>9</td>
</tr>
<tr>
<td>BM-8</td>
<td>17</td>
<td>T_LO9_105</td>
<td>4.30–13.1</td>
<td>7.50</td>
<td>8.80</td>
<td>32</td>
</tr>
</tbody>
</table>

¹See Table S1.
²Physical position of the QTL was reported previously by Ranjan et al. (2010).
doi:10.1371/journal.pone.0054468.t004
Genome Anchoring of QTL Intervals

We used the genome anchoring strategy described by Ranjan et al. [20] with minor variation to establish QTL intervals on the genome assembly. In addition to establishing QTL intervals, we also estimated the physical position of the marker closest to the QTL peak within the QTL interval. In this regard, we used the QTL to physical distance ratio determined from SSR markers flanking the QTL to calculate the approximate position of the QTL peak relative to their cM distance from each flanking SSR marker. Where more than one marker was associated with the QTL peak in different experiments, the cM position of the marker with the closest peak association or the marker with the highest LOD score was used to estimate the QTL peak position on the genome assembly.

Candidate Gene Identification and Characterization

Gene models lying within genome-anchored QTL intervals were identified from the Populus genome assembly V2.2 in the Phytozome database (http://www.phytozome.net/cgi-bin/gbrowse/poplar/). Annotations based on PFAM domains were used to exclude gene models with unknown function from the analysis and to establish duplication relationships between genes occurring in different QTL intervals. Paralogous relationships were verified based on information available for each gene model in the GRAMENE database (http://www.gramene.org).

Supporting Information

Table S1 SSR markers mapping in QTL intervals based on physical positions on the Populus genome assembly. (XLSX)
Table S2 Populus gene models within QTL intervals identified for stem height and diameter in an F2 pedigree. (XLSX)
Table S3 Gene duplication within and between QTL intervals identified from stem height and diameter in an F2 Populus pedigree. (XLSX)

Acknowledgments

The authors wish to thank Dr. H.D. Bradshaw for sharing the phenotypic data with us.

Author Contributions

Conceived and designed the experiments: MMS TJT GAT. Performed the experiments: MMS LEG TJT GAT. Analyzed the data: WM MMS PR TY. Wrote the paper: WM MMS.

References


Additive and Dominant Effects

Locus specific additive and dominant effects were calculated from mean phenotypic values of the alternate homozygous genotypes and heterozygotes at each QTL. Specifically, locus-specific phenotypic means for heterozygous loci carrying alleles derived from the same species (mu(ac), mu(bd)) and for heterozygous loci carrying alleles derived from both species (mu(ad), mu(bc)) were computed from the MQM mapping procedure using MapQTL 6.0 [27]. Additive (a), dominance (d) effects were calculated as:

\[ a = \frac{[\text{mu(ac)} - \text{mu(bd)}]}{2}; \]
\[ d = \frac{[\text{mu(ad)} + \text{mu(bc)}]}{2} - \frac{[\text{mu(ac)} + \text{mu(bd)}]}{2}; \]

QTL mode of action was calculated as the ratio of dominance over additivity, \( d/a \) [14,31]. \( d/a \) ratios less than 1 were regarded as reflecting under-dominance, ratios between 0 and 1 reflected partial dominance, and ratios greater than 1 reflected over-dominance as suggested by Hua et al. [31].

\[ Y_{gsr} = \mu + \alpha g + \beta s + \sum_n \alpha_{gsr}^n \delta_m + \rho_{gs} + \epsilon_{gsr} \]

Where \( Y_{gsr} \) is the observed LOD score of \( g^th \) genotype (marker) in \( e^{th} \) environment for \( r^{th} \) replicate;
\( \mu \) = grand mean; \( \alpha_g \) = deviation of mean of \( g^th \) marker from grand mean; \( \beta_s \) = deviation of mean of the \( e^{th} \) environment from the grand mean; \( \delta_m \) = singular value for the \( n^{th} \) interaction principal component axis (PCA); \( \gamma_m \) = the marker eigenvector for the \( n^{th} \) PCA; \( \delta_n \) = the environment eigenvector values for the \( n^{th} \) PCA; \( \rho_{gs} \) = residual effects; and \( \epsilon_{gsr} \) is the error term.

Genome Anchored Biomass QTLs in Hybrid Populus

PLOS ONE | www.plosone.org 8 January 2013 | Volume 8 | Issue 1 | e54468
interspecific hybrid inbred F$_2$ poplar pedigree growing under contrasting environments. Tree Physiology 26: 395–604.


