**Supplementary**

**Methods S1 | Methods to get wood coarseness, wood stiffness, and number of cells**

Wood coarseness was predicted with predictor traits of radial tracheid, tangential tracheid, and WD according to a formula wood coarseness = radial tracheid width × tangential tracheid width × WD (Ma *et al.*, 2018). Wood stiffness, expressed as modulus of elasticity (MOE), was predicted using the predictor traits MFA and WD according to a formula (Evans & Elic, 2001). Number of cells was predicted using the predictor traits ring width and radial tracheid width according to a formula (number of cells = ring width ÷ radial tracheid width).

**Methods S2 | SNP calling**

Raw reads based on 2×100 bp sequencing mode were mapped to the *P. abies* reference genome v1.0 using BWA-mem (Langmead & Salzberg, 2012). SAMTools (Li *et al.*, 2009) and Picard (McKenna *et al.*, 2010) were used for sorting and removal of PCR duplicates, and the resulting BAM files were subsequently reduced to include only probe bearing scaffolds (24919) before variant calling. Variant calling was performed using Genome Analysis Toolkit (GATK) HaplotypeCaller (McKenna *et al.*, 2010) in Genome Variant Call Format (gVCF) output format.

**Methods S3 | SNP quality control**

The following seven steps were used to perform SNP quality control using VCFtools.

(1) removing indels, (2) keeping only bi-allelic sites, (3) marking sites with a genotype quality (GQ) < 10 as missing, (4) removing sites with call rate < 70%, (5) removing sites with minor allele frequency (MAF) < 0.005, (6) removing sites with an excess of heterozygotes and deviation from Hardy-Weinberg equilibrium test (HWE-test) for unrelated population (*p*-value < 1.4×10e-7), and (7) removing individuals with call rate < 50%.

**Methods S4 | Adjustment of phenotypic data**

The phenotypic SilviScan data from 1303 trees was adjusted using the following equation:

**[1]**

Where *y* is the vector with tree height, DBH, and one of the 34 SilviScan traits in each of three trait types (Table 1), *b* is the vector of fixed effects including an intercept for three traits, *u* is the vector of random effects including post-block, incomplete block, and additive effects. ***X*** and ***Z*** are the corresponding design matrices of *b* and *u*, is the vector of residuals for the three traits. An unstructured covariance was used in the model fitting for each of the random terms. The phenotypic values entering the GWAS were the original phenotypic data adjusted for the post-block and randomized incomplete block effects.

**Methods S5 | Estimate of narrow-sense heritability**

Pedigree-based narrow-sense heritability () and SNP-based narrow-sense heritability ( were calculated as:

**[2]**

**[3]**

respectively, where and are additive variance based on pedigree and genomic marker models, respectively. and are phenotypic variances for pedigree-based and marker-based models, respectively.

**Methods S6 | BLINK method**

In this study, we performed three Genome-wide association (GWA) methods. BLINK algorithm is developed by Huang et al. (2019) (Huang et al., 2019). Here we slightly modified the definition for the algorithm as below:

The BLINK method conducts two fixed effect models and one filtering process, which selects a set of pseudo quantitative trait nucleotides (QTNs) that are not in LD with each other as covariates. The entire sequence runs repeatedly until all genetic markers are tested and the selection of pseudo-QTNs is optimized. The first fixed-effect model tests *M* genetic markers, one at a time. Pseudo QTNs are included as covariates to simultaneously control false positives and reduce false negatives. Specifically, the first fixed-effect model can be written as follows:

…………………………………………………………[5]

where yi is the observation on the ith individual; , , . . . , are the genotypes of k pseudo QTNs, initiated as an empty set; b1, b2, . . . , bk are the corresponding effects of the pseudo QTNs; is the genotype of the ith individual and jth genetic marker; dj is the corresponding effect of the jth genetic marker; and is the residual having a distribution with a mean of zero and a variance of .

The primary goal of the first fixed effect model is to calculate the *P* values for all *M* testing markers. The second FEM is employed to optimize the selection of pseudo-QTNs. Specifically, the second FEM can be written as follows:

…………………………………………………………… [6]

Equations [5] and [6] differ in two ways. First, the testing marker term in the first FEM is removed from the second FEM; therefore, no testing marker *P* values are output in equation [6]. Second, the number of covariate pseudo-QTNs is varied in the second FEM to select the optimum set of the first k out of t pseudo QTNs. The optimization is performed using BIC, which is twice the negative log-likelihood plus the penalty on the number of parameters, as follows:

………………………………………………………………………………[7]

where LL is the log-likelihood, k is the number of pseudo QTNs, Ln is the natural log, and n is the number of individuals. The available pseudo QTNs, t, are sorted with the most significant at the beginning and the least significant at the end. The first k pseudo QTNs are selected for examination, with k varied from 1 to t.

All markers in equation [5] are candidates for pseudo QTNs in equation [6]. These markers are filtered with two criteria: *P*-value and correlation. All markers are sorted first and then filtered out if their *P* values are larger than a threshold (Bonferroni correction, *α* = 0.01). Of the m SNPs remaining, if their correlation, *r* (Pearson correlation), with the first SNP () is larger than a threshold (0.7), they are also removed. This process is repeated to select , , . . . , until the last SNP, ), is selected.

Because the t remaining markers are sorted and not highly correlated with each other, the first set of k markers is more critical than the second set of k markers. We fit the first k markers in equation [6] and vary k until all possibilities are examined. The set of k markers with the best BIC is used as the set of pseudo QTNs in equation [5]. This process is iterated until the pseudo-QTNs remain the same.

This alternative solution is called the Bayesian-information and linkage-disequilibrium iteratively nested keyway (BLINK) method.

**Methods S7 |** Multi-variate GEMMA (MV-GEMMA) model:

A multivariate linear mixed model was used in GEMMA software (Zhou & Stephens, 2014) as in the following form:

where Y is an *n* by *d* matrix of *d* phenotypes for n individuals; W = (, · · · , wc) is an *n*×*c* matrix of covariates (fixed effects) including a column of 1s; A is a *c* by *d* matrix of the corresponding coefficients including the intercept; x is an *n*-vector of marker genotypes; *β* is a *d* vector of marker effect sizes for the *d* phenotypes; U is an *n* by *d* matrix of random effects; E is an *n* by d matrix of errors; K is a known *n* by *n* relatedness matrix, I*n*×*n* is a *n* by *n* identity matrix, **V**g is a *d* by *d* symmetric matrix of genetic variance component, Ve is a *d* by *d* symmetric matrix of environmental variance component and MNn×d(0, V1, V2) denotes the *n*×*d* matrix normal distribution with mean 0, row covariance matrix V1 (*n* by *n*), and column covariance matrix V2 (*d* by *d*).

**Methods S8 | Estimate of the cumulative proportion of phenotypic variance explained (PVE).**

In the estimate of cumulative PVE by all candidate SNPs for each trait, variance components of random effects were calculated from the following two equations:

…………………………………………………………………………[8]

and

……………………………………………………………………[9]

where y is a vector of adjusted phenotypic values, and are two vectors of corresponding fixed effects including the intercept in equations [8] and [9], respectively. and are two vectors of random additive genetic effects in equations [8] and [9], respectively, following and . *G* is a genomic-based relationship matrix (GRM) based on all SNPs (~134K). and are the additive variances in equations [8] and [9], respectively. ***s*** is a vector of all candidate SNPs as fixed effects for each trait. and are two vectors of residual effects ~ N(0,I) and ~ N(0,) in equations [8] and [9], respectively. is a null vector. *Z*, *W*, and *X* are the related design matrices in both equations:

…………………………………………………[10]

Note: if in some of the cases, the numerator of equation [7] is less than zero, we report it as zero as a percentage of variance explained by an SNP or SNPs could not be less than zero. This also represents that if a genomic prediction model includes the SNP or SNPs as a fixed effect or effects, the prediction model will not be improved in a real case.

**Methods S9 | Validation test for candidate SNP detected by GWAS**

Using ASReml R v4.0, a Wald test was used to test the significance of candidate SNP for each trait as below:

………………………………………………………………………[12]

where y is the vector of adjusted phenotypic values, is a vector of corresponding fixed effect including the intercept, is a vector of random additive genetic effects, following . *G* is a genomic-based relationship matrix (GRM) based on all SNPs (~134K). is the additive variance. s is a single candidate SNP effect. is a vector of residuals. is a null vector. *Z*, *W*, and *X* are the related design matrices.

**References**:

**Evans R, Elic J. 2001.** Rapid prediction of wood stiffness from microfibril angle and density. *Forest Products Journal* **51**(3): 53–57.

**Huang M, Liu X, Zhou Y, Summers RM, Zhang Z. 2019.** BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience* **8**(2): 1–12.

**Langmead B, Salzberg SL. 2012.** Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.

**Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009.** The sequence alignment/map format and SAMtools. *Bioinformatics* **25**(16): 2078−2079.

**Ma T, Inagaki T, Tsuchikawa S. 2018.** Non-destructive evaluation of wood stiffness and fiber coarseness, derived from SilviScan data, via near infrared hyperspectral imaging. *Journal of Near Infrared Spectroscopy* **26**(6): 398-405.

**McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010.** The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**(9): 1297−1303.

**Zhou X, Stephens M. 2014.** Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods* **11**(4): 407–409.