



**HAL**  
open science

## **Transcriptional profiling of bud dormancy induction and release in oak by next-generation sequencing**

Saneyoshi Ueno, Christophe C. Klopp, Jean-Charles Leplé, Jérémy Derory, Céline Noirot, Valérie V. Léger, Elodie Prince, Antoine Kremer, Christophe Plomion, Grégoire G. Le Provost

► **To cite this version:**

Saneyoshi Ueno, Christophe C. Klopp, Jean-Charles Leplé, Jérémy Derory, Céline Noirot, et al..  
Transcriptional profiling of bud dormancy induction and release in oak by next-generation sequencing.  
BMC Genomics, 2013, 14, 15 p. 10.1186/1471-2164-14-236 . hal-02650124

**HAL Id: hal-02650124**

**<https://hal.inrae.fr/hal-02650124>**

Submitted on 29 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RESEARCH ARTICLE

Open Access

# Transcriptional profiling of bud dormancy induction and release in oak by next-generation sequencing

Saneyoshi Ueno<sup>1,2,3</sup>, Christophe Klopp<sup>4</sup>, Jean Charles Leplé<sup>5</sup>, Jérémy Derory<sup>2,3</sup>, Céline Noirot<sup>4</sup>, Valérie Léger<sup>2,3</sup>, Elodie Prince<sup>2,3</sup>, Antoine Kremer<sup>2,3</sup>, Christophe Plomion<sup>2,3</sup> and Grégoire Le Provost<sup>2,3\*</sup>

## Abstract

**Background:** In temperate regions, the time lag between vegetative bud burst and bud set determines the duration of the growing season of trees (i.e. the duration of wood biomass production). Dormancy, the period during which the plant is not growing, allows trees to avoid cold injury resulting from exposure to low temperatures. An understanding of the molecular machinery controlling the shift between these two phenological states is of key importance in the context of climatic change. The objective of this study was to identify genes upregulated during endo- and ecodormancy, the two main stages of bud dormancy. Sessile oak is a widely distributed European white oak species. A forcing test on young trees was first carried out to identify the period most likely to correspond to these two stages. Total RNA was then extracted from apical buds displaying endo- and ecodormancy. This RNA was used for the generation of cDNA libraries, and in-depth transcriptome characterization was performed with 454 FLX pyrosequencing technology.

**Results:** Pyrosequencing produced a total of 495,915 reads. The data were cleaned, duplicated reads removed, and sequences were mapped onto the oak UniGene data. Digital gene expression analysis was performed, with both *R* statistics and the R-Bioconductor packages (edgeR and DESeq), on 6,471 contigs with read numbers  $\geq 5$  within any contigs. The number of sequences displaying significant differences in expression level (read abundance) between endo- and ecodormancy conditions ranged from 75 to 161, depending on the algorithm used. 13 genes displaying significant differences between conditions were selected for further analysis, and 11 of these genes, including those for glutathione-S-transferase (GST) and dehydrin xero2 (XERO2) were validated by quantitative PCR.

**Conclusions:** The identification and functional annotation of differentially expressed genes involved in the “response to abscisic acid”, “response to cold stress” and “response to oxidative stress” categories constitutes a major step towards characterization of the molecular network underlying vegetative bud dormancy, an important life history trait of long-lived organisms.

## Background

In temperate zones, woody plants go through seasonal cycles with two stages: a growing period when environmental conditions are favorable and a period of non growth in winter (reviewed in [1]). This second period of non growth is known as dormancy, and is defined as the time between bud set in the fall and bud burst in the spring. Dormancy

allows trees to survive unfavorable conditions in winter [2] and can be divided into two phases of variable duration and intensity [3]: endodormancy and ecodormancy.

- i.) firstly endodormancy: the deepest state of dormancy, in which bud burst is prevented by endogenous factors specific to the meristem. The strength of endodormancy increases from the end of summer, throughout the fall, generally peaking somewhere between the end of October and the end of November in the northern hemisphere depending on the species and the climatic conditions prevailing

\* Correspondence: leprovost@pierron.inra.fr

<sup>2</sup>INRA, UMR 1202 BIOGECO, F-33610 Cestas, France

<sup>3</sup>Univ. Bordeaux, BIOGECO, UMR 1202, F-33400 Talence, France

Full list of author information is available at the end of the article

during the fall [4]. A dormancy breaking test, or forcing test, is the method of choice for assessing the depth of endodormancy in trees [5]. The signals leading to the breaking of endodormancy in perennials are well characterized, temperature playing a major role in both the induction and release of endodormancy [6].

- ii.) secondly ecodormancy: this dormancy state occurs when unfavorable external environmental factors (usually temperature in the case of perennials growing in temperate climates) prevent bud burst in late winter and early spring.

The growing season of trees extends from bud burst in the spring to bud set in the fall. This phenological cycle has been shown to be strongly affected by climatic change, higher temperatures having a negative impact on endodormancy (if the requirement for chilling is not fulfilled, then dormancy break is delayed) and a positive impact on ecodormancy (acceleration of bud cell growth) [7]. On the other hand, Kalcsits *et al.* [8] reported that high temperatures during endodormancy induction in poplar leads to deeper endodormancy and delayed bud burst in the spring [8].

An overall increase in the duration of the growing season has been reported in temperate regions, for both fruit and forest trees (reviewed in [9,10]). In this context, early bud break is a particularly crucial issue in terms not only of the increase in length of the growing season, but also the risk of frost damage [11].

Several quantitative genetic approaches have been used to study the genetic basis of bud phenology in forest trees, focusing essentially on the date of bud burst. Medium to high heritability values have been reported, from 0.43 in *Salix* [12] to 0.98 in *Populus* [13]. Saintagne *et al.* [14] reported repeatability values of 0.15 to 0.52 in pedunculate oak (*Quercus robur* L.), indicating moderate to medium genetic control for this trait, whereas very high heritability values were reported in progeny tests and parent-offspring comparisons for *Quercus petraea*. Derory *et al.* [15] reported the identification of 19 QTLs for bud burst for 13 year  $\times$  site seasonal observations in a cloned full-sib oak pedigree, suggesting polygenic control of this adaptive trait.

However, the molecular mechanism underlying the regulation of bud dormancy in forest trees remains poorly understood. Most investigations of gene expression have been carried out in model plants, such as *Arabidopsis thaliana*, for which functional candidate genes have been identified [16,17]. These studies highlight the role of phytochromes as central players in day-length sensing. Allona *et al.* [18] reported that dormancy induction in forest trees was driven by molecular mechanisms similar to those involved in the photoperiodic

control of flowering time in annuals [19]. However, the mechanisms responsible for the maintenance of dormancy, once it has been established, are also poorly understood. Several molecular studies in perennials, such as kiwi [20], grape [21], potato [22], raspberry [23] and leafy spurge [24], have shown that dormancy signals affect multiple physiological processes, including cell division, secondary metabolism (flavone biosynthesis), hormone signalling (with significant changes in genes responsive to ethylene, auxin, ABA and GA) and oxidative stress. These studies also reported changes in the accumulation of transcripts from many regulatory genes thought to be involved in dormancy transition or regulating specific metabolic pathways involved in dormancy transition. In forest trees, the use of transcriptomic approaches to detect the expression of candidate genes thought to be involved in vegetative bud dormancy remains limited. Derory *et al.* [25] and Rohde *et al.* [17] used targeted approaches (subtractive suppressive hybridization and cDNA-AFLP, respectively) to identify genes displaying differential patterns of expression at different stages of dormancy (induction, maintenance and release). Ruttink *et al.* [26] used high-density microarrays to reveal the existence of gene networks involved in apical bud formation and dormancy induction in poplar. Transgenic studies have also shed light on the molecular mechanisms underlying bud phenology. In comparisons with the peach evergrowing (*evg*) mutant, for example, Jiménez *et al.* [27] identified 23 genes that were upregulated in the wild type, and therefore potentially involved in photoperiod response, growth cessation and the induction of dormancy. In poplar, Ibañez *et al.* [16] reported a potential role of circadian clock-related genes in cold hardiness and delayed bud burst.

In this context, the objective of this study was to characterize the transcriptome of endo- and ecodormant sessile oak (*Quercus petraea*) by whole-transcriptome shotgun sequencing (WTSS or RNA-seq) and to identify genes displaying differential expression between the two phenological stages, with the aim of improving our understanding of dormancy at the molecular level. Firstly, four cDNA libraries were built from total RNA extracted from apical buds harvested during endo- or ecodormancy, from two populations with different bud burst dates (early vs. late flushing populations). The transcriptome was then characterized by pyrosequencing on a Roche 454 FLX platform. In total, 495,915 reads were produced. Reads were mapped onto the oak UniGene sequences generated from 34 cDNA libraries [28]. Digital expression analysis and quantitative PCR identified a set of genes displaying differential regulation between these two developmental phases and possibly inter connected with a gene network. None of these genes was associated with a specific population.

## Results

### Identification of the endodormancy and ecodormancy developmental phases

The three main developmental phases typically involved in dormancy in perennial plants (for review see [1]) were identified in the two oak populations (Figure 1). The endodormancy phase extended roughly from July 30<sup>th</sup> 2007 to October 29<sup>th</sup> 2007, in both populations. Endodormancy levels peaked on the same date in both populations (September 10<sup>th</sup>). A transition phase occurred from October 29<sup>th</sup> to November 26<sup>th</sup>. Finally, the ecodormancy phase began on December 4<sup>th</sup>.

### Dormant bud transcriptome

A total of 495,915 reads were obtained from the four libraries (Table 1). The cleaning process led us to eliminate 143,528 (29.9%) reads, resulting in 352,397 (71.1%) reads which were then assembled into OakContigV1 [28]. Reads from the two endodormancy libraries were distributed into 24,030 contigs and 14,794 singletons, whereas those from the two ecodormancy libraries were distributed into 33,383 contigs and 31,720 singletons (Table 1 and Figure 2). Differential expression was analyzed for 6,471 contigs with corrected read numbers  $\geq 5$  (71, 217 and 6,183 for the endodormancy, ecodormancy and both dormancy libraries, respectively).

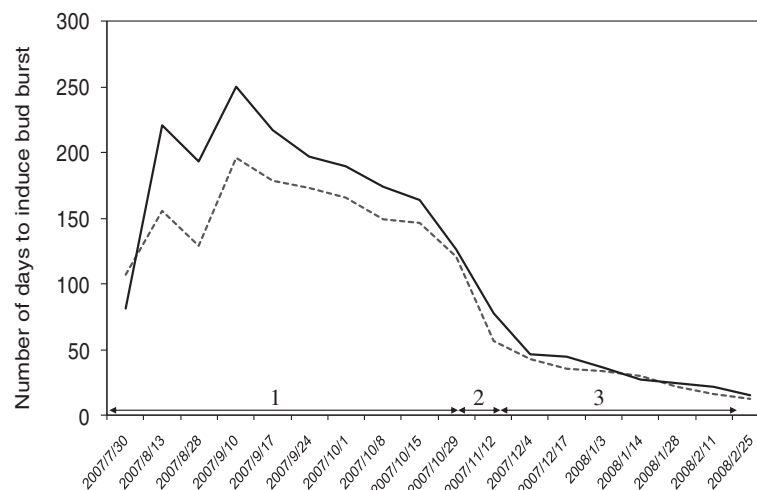
### Detection and analysis of differentially expressed genes

An analysis of differential expression with *R* statistics [29] resulted in 122 (1.89%) contigs (with  $R > 8$ ), suggesting significant differences in expression between the four libraries. The number of contigs for a given value of the *R* statistic was plotted as a function of *R* (see Additional file 1: Figure S1). No genes displaying significant differential

expression between populations were identified, allowing us to treat these two populations as biological replicates. edgeR [30] and DESeq [31] R Bioconductor packages detected 116 (1.8%) and 110 (1.7%) contigs, respectively, displaying significant differential expression between endodormancy and ecodormancy (FDR 20%). These values were higher than those using a 5% FDR (59 (0.91%) and 73 (1.13%) contigs identified as significantly differentially expressed by EdgeR and DESeq, respectively). The differences in relative expression between the endo- and ecodormancy libraries and relative read numbers for each contig are displayed as MA plots (smear-plot) in Figure 3. A Venn diagram displaying the number of differentially expressed genes identified by each of the three programs (*R* statistics, edgeR and DESeq) is shown in Figure 4, where 75 contigs were identified by all three programs (MN set), and 161 contigs (MX set) were detected in total. For 26 contigs, only *R* statistics were consistent with differential expression, whereas only 14 contigs were identified as differentially expressed by edgeR and only nine were identified as differentially expressed by DESeq package alone. Heat maps were drawn for both the MX and MN sets (see Additional file 2: Figure S2).

At least one blast hit against SWISS-PROT was found for 54 (33.5%) contigs and at least one blast hit against TAIR9 was found for 102 (63.4%) contigs (see Additional file 3: Table S1). In total, 96 GO terms were assigned to 45 (28.0%) contigs. These GO terms were mapped against 51 GO-slim terms.

Enrichment analysis for AraCyc Pathways and Gene Ontology Groups (EAPG) was then performed both for the gene up-regulated during endo- (i.e. 34 contigs) and ecodormancy (i.e. 41 contigs) identified by the three programs.



**Figure 1 Variation of apical bud dormancy.** This chart shows the number of days required to induce bud burst in forcing conditions. The dotted line corresponds to the Longchamp population characterized as late flushing in a common garden experiment and the solid line corresponds to the Saint-Jean population (early flushing). Abbreviations correspond to: (1) endodormancy, (2) transition phase and (3) ecodormancy.

**Table 1 454-pyrosequencing and assembly statistics for endo- and ecodormancy libraries**

	Endodormancy		Ecodormancy		Total
	LC1	SJ1	LC2	SJ2	
SRA accession number	SRX019071	SRX019073	SRX019072	SRX019075	
Number of reads	115,050	79,345	137,380	164,140	495,915
Number of reads passed to assembly	70,019	44,732	98,725	138,921	352,397
No. singletons in OakContigV1 from each library (a)	9,937	4,857	15,789	15,931	46,514
No. contigs in OakContigV1 within which reads are assembled (b)		24,030		33,383	41,731*
No. unigene elements in OakContigV1 (a)+(b)		38,824		65,103	88,245*

\* The values are not a simple sum of those for endo- and ecodormancy, because there were common contigs to which reads from both libraries were mapped.

Among ecodormancy regulated genes EAPG enabled us to identify three pathways from AraCyc and 44 groups from the GO commons (see Additional file 3: Table S2). Looking to the AraCyc pathway, EAPG found three pathways enriched: Stachyose biosynthesis, oxidative ethanol degradation and beta alanine biosynthesis. EAPG identified also 44 GO groups enriched for ecodormancy up-regulated genes. The first five groups were linked to response to cold, high light intensity, water deprivation, abscisic acid stimulus and water channel activity.

Sub-network enrichment analysis (FNSE) was performed and identified central hubs and their associated partners. Among ecodormancy up-regulated genes three gene-linked hubs (*DREB2A*, *CBF2* and *DREB1A*), three processes-

linked hubs (drought, cold and heat shock) and two molecule-linked hubs (ABA and  $Ca^{2+}$ ) were identified (see Additional file 3: Table S3 and Figure 5).

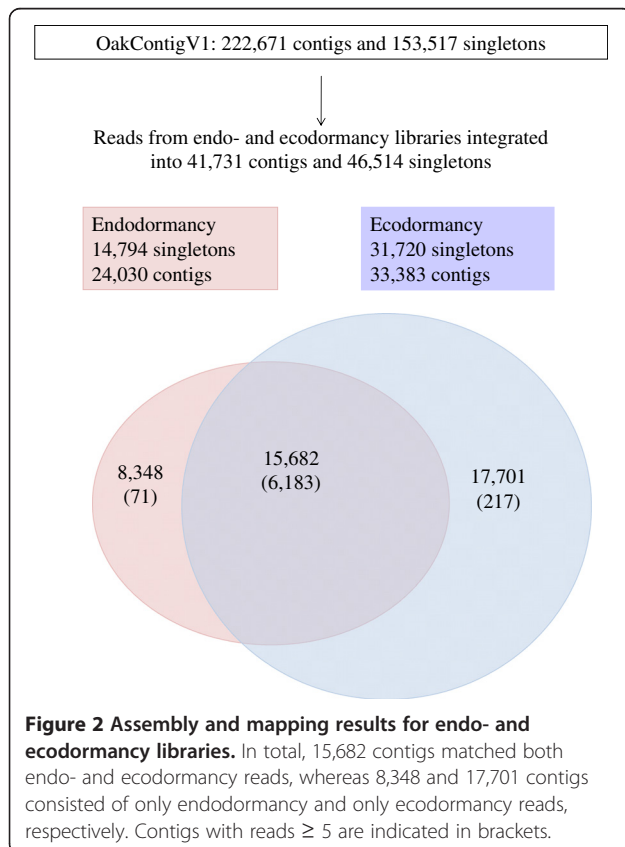
The same analysis was also performed for endodormancy up-regulated genes. EAPG identified one AraCyc enriched pathway (cuticular wax biosynthesis) and 39 enriched GO groups (see Additional file 3: Table S4). The first five group based on their p-value were associated with metal ion binding, cellular transition metal ion homeostasis, fatty acid binding, 2-iron2-sulfur cluster binding and epidermis morphogenesis. No central hubs were identified through the FNSE analysis.

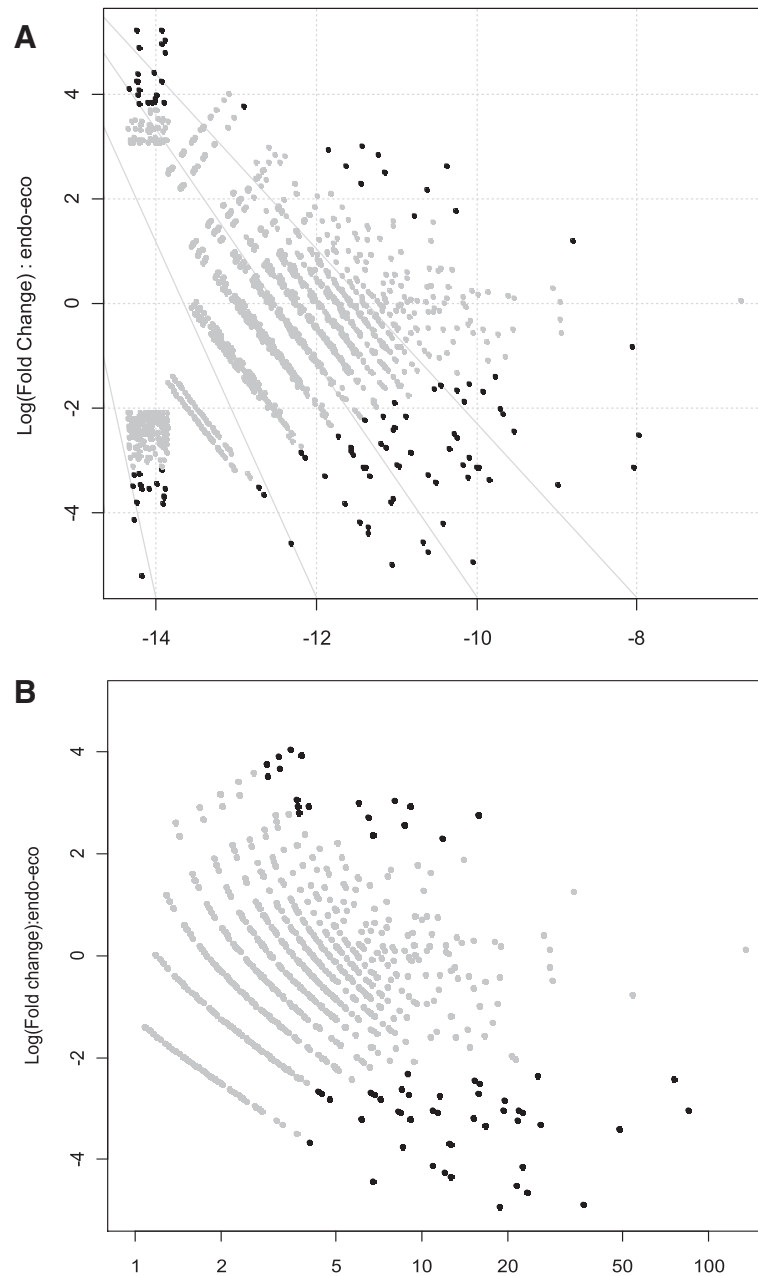
#### Tissue specificity of differentially expressed genes

Significant differential expression of genes in the MN set was detected for 21 genes (see Additional file 4: Table S5) in an *in silico* expression analysis of cDNA libraries for *Quercus* (see Additional file 5: Table S6). The number of reads mapping onto three genes was positive for bud tissue, but zero for other tissue types (root, xylem and leaf tissue). Two of these UniGene elements (F0SUT5C01ARKJ3.lqr.1 and F0SUT5C01A7647.1.qr.1) were identified as up-regulated during endodormancy and ecodormancy, respectively, displayed no similarity to any other protein, whereas the third one (F0SUT5C01BB2R7.lqr.1), was found to be similar to the glutathione S-transferase TAU 19.

#### Validation of candidate genes displaying differential expression, by quantitative PCR

Among the five control genes analysed (actin, tubulin  $\beta$ , ubiquitin, ribosomal protein 18S and elongation factor  $\alpha$ ), only one (ubiquitin) displayed a multi-banding pattern after PCR amplification. Two other genes (actin and tubulin  $\beta$ ) displayed changes in expression over the time course of the experiment, in both populations, and were therefore discarded from the analysis. The two remaining genes (ribosomal protein 18S and elongation factor  $\alpha$ ) were found to display stable expression in both populations, during both dormancy periods, and were therefore used to normalize qPCR data (see Additional file 6: Table S7).

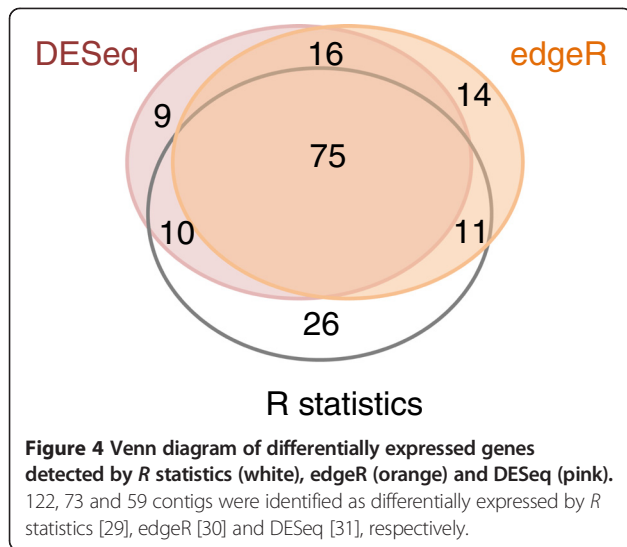




**Figure 3 MA plots for endo- and ecodormancy libraries.** Fold-changes of read counts between endodormancy and ecodormancy are plotted against relative read abundance. Black dots indicate statistically significant contigs displaying differential expression between endo- and ecodormancy libraries as identified with the edgeR (A) and DESeq (B) packages, respectively. If the number of corrected read counts was zero for either endo- or ecodormancy libraries, no dots were generated for (B).

Among the 12 candidate gene analysed, only one encoding a cold-shock protein, was validated in a single population (LC), the 11 remaining genes displayed expression profiles similar to those observed in the 454 sequencing experiment in both populations. Nine of the validated genes were overexpressed during ecodormancy (gibberellin-regulated GASA family protein (*GASA*), glutathione-S-transferase (*GST*), dormancy-associated protein (*DRM1*),

dehydrin xero2 (*XERO2*), plasma membrane-intrinsic protein (*PIP*), AWPM-19-like membrane family protein (*AWPM*), gibberellin receptor *GID1L2* (*GID*), late embryogenesis abundant protein *Lea5* (*LEA*) and early light-induced protein (*ELP*)). The two remaining genes (*basic 7S globulin 2* and *xyloglucan endotransglucosylase/hydrolase9*) were upregulated during the endodormancy period. The expression profiles of the genes analyzed by



qPCR are provided in Additional file 7: Figure S3. For reference and candidate genes, PCR efficiency ranged from 87% (basic 7S globulin 2) to 104% (xyloglucan endotransglycosylase/hydrolase9), which fits with the correct functioning of *Taq* polymerase.

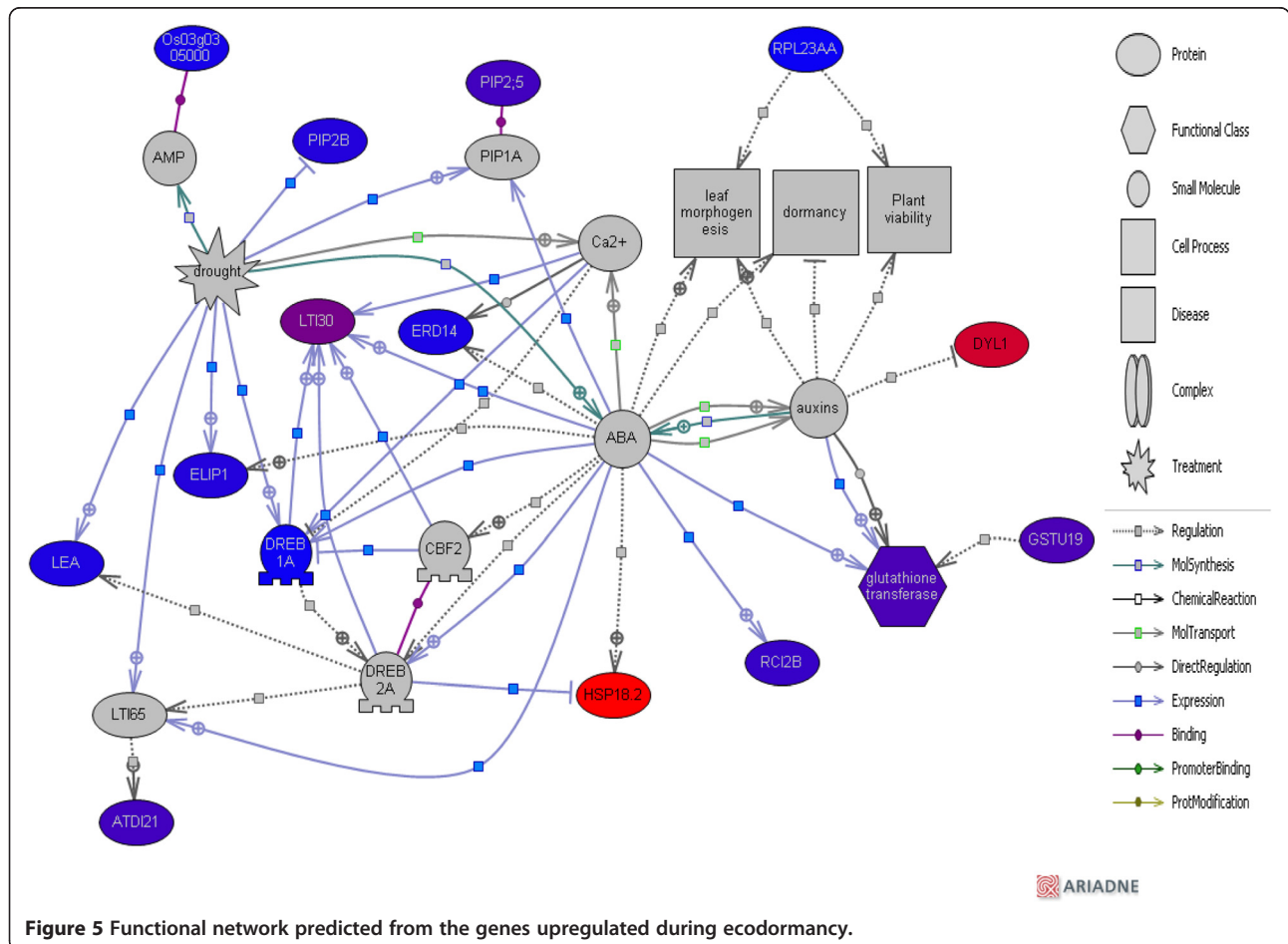
### Discussion

Given the extremely high rate of validation of the *in silico* analysis by qPCR, we will consider for the scope of the discussion the core set of 75 contigs (MN set) showing differential expression with three different algorithms (*R* statistics, edgeR and DESeq). No annotation was available for 22 of these 75 contigs (29.3%). A molecular function was attributed to the 53 remaining contigs (see Additional file 3: Table S1).

Gene ontology analysis showed that this set of genes was enriched in genes responsible for abiotic (i.e. cold) and endogenous stimuli (i.e. ABA) and stress-related genes. Similar trends were observed in analyses of gene expression in other plants. In raspberry, the set of genes displaying differential expression during endodormancy and paradormancy included a high proportion of stress-response/defense/detoxification-related genes [23]. In grape, genes relating to oxidative processes and stress responses within the cell are expressed at the end of dormancy [32].

### Genes upregulated during endodormancy

Plants are sessile and must therefore be able to prepare themselves to survive unfavorable conditions. In temperate



regions, deciduous trees stop growing, lose their leaves in the fall and enter a state of dormancy to allow them to survive the winter. The detailed molecular mechanisms underlying dormancy induction remain unclear, although an analogy with the mechanisms of vernalization and flowering has been suggested [6]. However, some common features of dormancy in buds, seeds and tubers may help to unravel the underlying molecular mechanisms, although some features are specific [1].

Our transcriptomics analysis revealed that 34 genes in the core (MN) set were overexpressed during endodormancy. Among these genes, two were quantified and validated by qPCR. The first gene encodes a *basic 7S globulin 2* gene (see Additional file 7: Figure S3). Globulin is a well known storage protein present in many plant seeds [33,34]. It accumulates during seed maturation and desiccation before dormancy and is consumed during germination. In maize seeds [35], globulin accumulation is triggered by a plant hormone, abscisic acid (ABA), before dormancy. The upregulation of globulin in oak endodormant buds may also be triggered by ABA produced in response to the environmental conditions prevailing in the fall. Short-day conditions or low temperature may initiate ABA biosynthesis, thereby inducing endodormancy and resulting in a cessation of growth, bud set and cold hardening. The accumulation of *globulin* transcripts in endodormant buds in oak may be considered as a response to an abiotic stimulus. Globulin may be used to provide the nutrients required for bud burst, just as seed storage proteins are used for germination [36].

The second gene encodes a *xyloglucan endotransglucosylase/hydrolase 9 (XTH)* (see Additional file 7: Figure S3) belonging to the several ontologies (cell wall GO: 0005618, endomembrane system GO:00012505, hydrolase activity acting on glycosyl bonds GO:00016798 and xyloglucan: xyloglucosyl transferase activity GO:00016762) Overexpression of *XTH* was also reported by Anderson et al. [37] in endo-dormant crown bud of leafy spurge [37]. *XTH* genes constitute a multigene family encoding proteins involved in cell wall construction and disassembly, with diverse levels of expression in *Arabidopsis thaliana* [38]. In tomato, two isozymes of XTH (LeXET and LeXET2) with opposite patterns of expression have been detected [39]. LeXET was abundant in the rapid elongation region, whereas LeXET2 was more abundant in the mature/non elongating region. Enzymatic analysis of xyloglucan endotransglucosylase activity in *Arabidopsis* showed cold tolerance, with optimal activity for this enzyme at temperatures of 12°C to 18°C, consistent with a potential role in cold acclimation or cold-tolerant growth [40]. The *XTH* gene is also strongly expressed in the dormant cambium of *Populus tremula*, suggesting a role for XTH in cell wall modification during dormancy acquisition [41]. Samples of *Prunus percica* buds released from

dormancy are rich in *XTH* mRNA [42], probably due to the remodeling of cell wall structure for resumption of the growth cycle. The abundant expression of the *XTH* gene observed in the oak endodormancy library also suggests a potential role in cold acclimation, through the modification of cell wall structure.

The contigs of known function of the core set upregulated during endodormancy included *MLP423* (GO:0009607, response to abiotic stimulus), *pheophorbide a oxygenase* (PAO, Chlorophyllide *a* oxygenase GO:0010277 and Chloroplast GO:0009507) and *agglutinin*. *MLP423* have been reported to be involved in drought regulation in *Populus* [43], though down-regulated in drought treatment. PAO is an enzyme necessary to break down chlorophyll [44], while *agglutinin* functions as glycan binding in *Castanea crenata* [45].

### Genes upregulated during ecodormancy

Ecodormancy is a short-term phenomenon induced by a prolonged period of cold and long-day conditions, which provide an initial signal triggering a shift from endodormancy [46]. In this section we discuss the genes up-regulated during ecodormancy according to the pathway highlighted through the EAPG and FNSE analysis as well as our qPCR analysis (see Additional file 7: Figure S3). Indeed several pathways known to be regulated during ecodormancy (i.e. response to abscisic acid, response to gibberellin stimulus, response to cold and response to oxidative stress) were differentially regulated in our dataset (see Additional file 3: Table S3 and Figure 5). It should also be borne in mind that because of the low number of contigs, as well as the rather relaxed threshold used to declare a significant differential expressions, the analysis of pathways and networks requires further validation.

### Response to abscisic acid

Abscisic acid (ABA) is an important hormone involved in seed dormancy and germination. Recent studies [47] reported that the interaction between ABA and gibberellin was associated with loss of apical dominance. Several genes belonging to this functional category were illuminated through the EAPG (see Additional file 3: Table S3) and the FNSE (Figure 5). Among these differentially expressed genes, an Early Light Inducible Protein (ELIP) was identified. ELIP protein is a nuclear-encoded chlorophyll a/b binding protein of the thylakoid membranes that is upregulated by illumination [48]. The primary function of this protein is thought to be the prevention of photoinhibition in high light conditions, through the dispersal of absorbed light [49], but ELIP also accumulates in response to various stresses, such as cold [50] and UV-B irradiation, in green leaves [51]. In blueberry floral buds, ELIP mRNA levels are high until the end of



February (1,200 chilling units), with maximal levels of expression in December (400 chilling units) and no expression in October (0 chilling units) in the absence of cold acclimation [52]. The endodormancy library was based on samples from September 17th, 24th and October 1st. We therefore speculate that the higher levels of ELIP expression recorded for ecodormancy libraries (LC2 and SJ2) than for endodormancy libraries (LC1 and SJ1) are consistent with the buds in our endodormancy sample not being cold-acclimated and that ELIP was not induced in the LC1 and SJ1 libraries despite the buds being in a state of endodormancy. Derory *et al.* [25] showed that ELIP was more strongly expressed in *Q. petraea* buds in the vegetative growth phase than during ecodormancy, in macroarray experiments [25]. ELIP is induced not only by cold, but also by high light intensity. The primary stresses responsible for the upregulation of ELIP in this study and in the study by Derory *et al.* [25] may therefore be different.

A low temperature inducible protein LTI30 (Xero 2 or dehydrin xero 2) highly expressed during ecodormancy was also identified. This gene is also regulated by ABA [53]. The LTI30 gene is expressed in *Arabidopsis thaliana* leaves after cold stress [54]. Dehydrins (group 11 family LEA proteins or LEA2) are stress-responsive proteins induced by dehydration in conditions of drought, low temperature, salinity and seed maturation [55]. Dehydrin is upregulated after cold acclimation in blueberry floral buds [52]. In Norway spruce, most dehydrin genes are expressed in ecodormant buds and downregulated towards bud burst [56]. A dehydrin-like protein has also been detected in swelling oak buds [25].

Two Plasma membrane Intrinsic Proteins (*PIP2B*, *PIP1A*) were identified in our transcriptomic approach. These proteins have been shown to be regulated both by ABA and water stress [57]. PIP constitutes a subfamily of aquaporins, membrane proteins that form channels in the membrane, facilitating the movement of water across membranes. The *PIP* gene is also upregulated during vernalization in grape buds [21]. In raspberry, it has also been shown that aquaporin genes were regulated between the different phase of the dormancy [23].

Among ecodormancy up-regulated genes, we identified a Heat Shock Protein (HSP18.2). HSP18.2 (Figure 5) was one of the most expressed genes. HSP 18.2 is also regulated by ABA [58]. HSP have chaperone activity by maintaining protein in their functional conformation and therefore prevent the degradation of proteins damaged during cold stress.

Finally an ATDI21 protein similar to Late Embryogenesis Abundant protein 5 (LEA5) was found to be over-expressed during ecodormancy. LEA5 belongs to a family of LEA originally discovered in cotton seeds during late embryogenesis [59,60]. LEA proteins stabilize cellular components in cells subjected to water deficit, probably because

of their hydrophilic properties [55]. LEA may also protect the cells during environmental stresses by lowering water potential, stabilizing membrane structure and scavenging active oxygen species. In *Q. petraea* buds [25], *LEA5* mRNAs were detected during ecodormancy, and this upregulation during ecodormancy was confirmed by qPCR [25]. Similarly, in this study, *LEA5* transcripts were shown, by qPCR, to be upregulated in ecodormant buds (see Additional file 7: Figure S3). By contrast, an expression analysis on raspberry buds [23], revealed that both dehydrin-like and LEA proteins were produced in larger amounts during endodormancy. Moreover, poplar apical buds also express *LEA14-A* principally during the endodormancy induction phase [17]. In leafy spurge, Horvath *et al.* [61] reported also that dehydration genes were down-regulated following the transition from endo- to ecodormancy [61]. These discrepancies may reflect the diverse functions and expression patterns of LEA proteins in different tissues and during diverse developmental stages [62].

#### Response to cold

In temperate region, perennial species are able to increase their freezing tolerance in response to low temperature observed during ecodormancy. This process is called cold acclimation and the coordinated expression of thousand genes to induce physiological and biochemical changes in the buds [reviewed by [63]] is required. Several cold inducible genes have been isolated from plants and although some are regulated by low temperature, most of them are known to respond both to ABA and water deprivation suggesting a strong interaction between these pathways during ecodormancy. Our functional analyses (EAPG and FNSE) revealed several genes involved in the response to cold (see Additional file 3: Table S2 and S3). First we identified a *RCI2B* gene highly similar to a Rare cold inducible protein 2B. This gene belongs to a small multigene family regulated by low temperature and ABA. It was suggested that this genes is involved in the maintenance of membrane integrity triggered by low temperature that reduces water availability. *RCI2B* genes may interact with other membrane proteins to maintain the water status of the cell [reviewed by [64]].

An *ERD14* gene highly similar to a dehydrin protein was also identified. Dehydrin proteins are a group of proteins with a protective role against stresses. These proteins are synthesized in response to low temperature and ABA. A positive correlation between dehydrin content (i.e. mRNAs and proteins) and freezing tolerance has been reported in *Fragaria* [Reviewed by [65]]. We could hypothesize that the up-regulation of this genes during ecodormancy is associated with a protective role against low temperature observed in late winter.

Finally, a DREB1A genes over-expressed during ecodormancy was identified. Several studies reported that DREB transcription factors are essential for dormancy regulation in plants [47,66]. DREB1A transcription factor contains an AP2/EREB motif and have a binding function to the promoter containing the CRT/DRE (Dehydration responsive element) which exists in many LEA genes including ERD genes. Over-expression of DREB1A transcription factors induces in normal condition the over-expression of these cold-stress related genes and confers an improved tolerance to both water stress and low temperature [67]. Over-expression of DREB1A transcription factor during ecodormancy may be interpreted in light to the over-expression observed for LEA5 and ERD14 genes allowing the bud to face low temperature observed during ecodormancy.

#### Response to gibberellin stimulus

Gibberellins (GAs) are plant hormones that regulate growth and influence various developmental processes, including stem elongation, germination and dormancy. Two genes belong directly to this functional category (T30D6.7/GASA, and ERD14), and one gene indirectly involved was also identified and validated by qPCR (GID1). ERD 14 encodes an early response to dehydration protein (see above). Briefly, ERD proteins belong to a multigene family known to be strongly regulated during dormancy in leafy spurge [68]. This protein accumulates during dehydration and is probably involved in the protection of the membrane.

GASA (gibberellin-regulated gene family from *Arabidopsis thaliana*) [69] was originally identified in a gibberellin-deficit mutant of tomato, as a GA-regulated gene [70]. This gene is expressed in a tissue-specific manner [71,72] and is involved in several developmental processes, such as seed germination, stem elongation and flowering. GA regulates growth after the release of dormancy. In potato tubers, endogenous GA content increases when sprouts begin to grow [73] and GA biosynthesis is controlled by the photoperiod [74].

Finally we identified a GID1 encoding gene (probable GA receptor GID1L2) overexpressed during ecodormancy. GID1 is a soluble protein that interacts directly with GA, triggering the gibberellin response [75]. Gibberellic acids are key hormone involved in dormancy, several authors reported that they could replace chilling of dormant buds (reviewed in [76]). In a microarray analysis of grape buds, GID1 was found to be upregulated during vernalization [21]. On the other hand, a down regulation of the GIB1B gene was observed in endodormant buds of leafy spurge [68]. These authors hypothesized that down regulation of the GIB1B gene in conjunction with DREBs (see above "response to cold") played a significant role in endodormancy release or ecodormancy maintenance.

#### Response to oxidative stress

Three genes (AT3G10020, RPL23AA and GSTU19) belonged to this GO term. GSTU19 is similar to a glutathione-S-transferase (GST). GSTs are found in several animal and plant cytosols, and have diverse functions in detoxification, oxidative stress resistance and light signaling [77]. In some woody plants, GST is upregulated during the release of endodormancy, probably due to the response to oxidative stress [23,78,79] induced by low temperature [80].

#### Other genes identified

The *DYLL1* (or DRM1 dormancy-associated protein involved in response to fructose stimulus) was up-regulated during ecodormancy. DYLL1 is abundantly expressed in axillary buds of pea, with mRNA levels decreasing after decapitation; it is therefore considered a good marker of dormancy [81]. In blueberry buds, *DRM* mRNA levels were highest in a sample collected on December 8th, subsequently decreasing towards bud burst [82]. In grape buds, *DRM* gene expression also decreases progressively towards bud burst, this gene being more strongly expressed during dormancy [32].

Finally, in this study, the expression profile of the *AWPM* (*AWPM*-19-like membrane family protein) gene was also analyzed by qPCR. *AWPM*, a membrane protein that confers cold tolerance, has a sequence similar to that of LEA and is induced by ABA [83]. Quantitative PCR showed also upregulation of the cold-shock protein only in the LC2 library (ecodormant apical bud library from the Longchamp population). This protein has been also detected in plants subjected to cold stress for acclimation [84]. The protein binds to nucleic acids and destabilizes RNA structure, thereby affecting protein translation. These results strengthen once again that ABA and cold acclimation are probably the most important pathways involved in ecodormancy regulation.

#### Conclusions

The timing of bud burst affects the fitness of individual trees. Bud dormancy changes in nature from the fall to the spring. We investigated the molecular basis of dormancy, by analyzing differences in gene expression between two dormancy phases (endodormancy and ecodormancy) in sessile oak (*Quercus petraea* (Matt.) Liebl.) buds by RNA-Seq. We identified 75 UniGene elements displaying differential expression (in terms of read abundance) between the endo- and ecodormancy phases, with a false discovery rate of 20%. Thirty four contigs were over-expressed during endodormancy and 41 were up-regulated during ecodormancy. Enrichment analysis revealed that the molecular mechanisms involved in dormancy regulation were different between the 2 conditions. Indeed, enriched GO groups for response to cold, high light intensity and ABA stimulus were found only during ecodormancy while

other GO groups - metal ion binding, fatty acid binding and epidermis morphogenesis - were specific for endodormancy.

In conclusion, these results demonstrate that RNA-Seq technology and the associated statistical methods are suitable for the detection of differentially expressed genes. We identified promising expression candidate genes for future forward genetics studies. They will serve as a stepping stone to further work on dormancy, particularly as concerns woody plants, for which few data are currently available.

## Methods

### Forcing test and bud sampling

Acorns of sessile oak (*Quercus petraea* (Matt) Lieb.) were collected in 2006 from two populations in northern France: Longchamp (LC: lat. 47° 15' 49"N, long. 05° 18' 37"E) and Saint Jean (SJ: lat. 48° 46' 43"N, long. 06° 43' 42"E), separated by a distance of 150 km. These two populations have repeatedly shown differences in the timing of bud burst in common garden experiments. Bulk harvests of acorns were carried out in the two seed stands in the fall of 2006. The acorns collected were sown, in spring 2007, in 5 l containers filled with a peat and sand mixture (1/1, v/v). Following germination, we obtained 560 seedlings for the SJ population and 468 seedlings for the LC population.

A forcing test was carried out, from July 30<sup>th</sup> 2007 to February 25<sup>th</sup> 2008, in a greenhouse at the INRA forestry research station (south-western France). We selected about 25 seedlings from each population each week. Ten seedlings were used for transcriptome analysis. Apical buds were decapitated from these 10 seedlings and immediately frozen in liquid nitrogen. The 15 remaining seedlings were transferred to the greenhouse in the following conditions: 16-hour photoperiod with a daytime temperature of 25°C and a nighttime temperature of 18°C. For each batch of seedlings, the date of bud burst (defined here as swelling buds, see Additional file 8: Figure S4) was evaluated by observations three times per week, making it possible to determine the number of days required to induce bud burst. The phenological curve obtained in the forcing test (Figure 1) was then used to locate the time points at which endo- and ecodormancy could be differentiated in each population. As endodormancy peaked on the same date in each population and we wished to capture the transcripts expressed at endodormancy in both populations, we calculated the mean number of days to bud burst over the endodormancy period for the two populations together (185 days). This number was then used to determine the sampling date for endodormancy and extended the sampling period to one week before and one week after, for library construction. For ecodormancy, apical buds harvested on January 14<sup>th</sup>, January 28<sup>th</sup> and February 11<sup>th</sup> 2008 were used to generate the library,

because the number of days required to induce bud burst was short in the two populations (16 and 20 days on average for the LC and SJ populations, respectively).

It should be borne in mind that with this forcing test genes identified are related to variations of the dormancy stage or natural environmental variations (mostly light and temperature) or both. However it is not possible to disentangle these effects as light and temperature variations trigger the shift from endo- to ecodormancy.

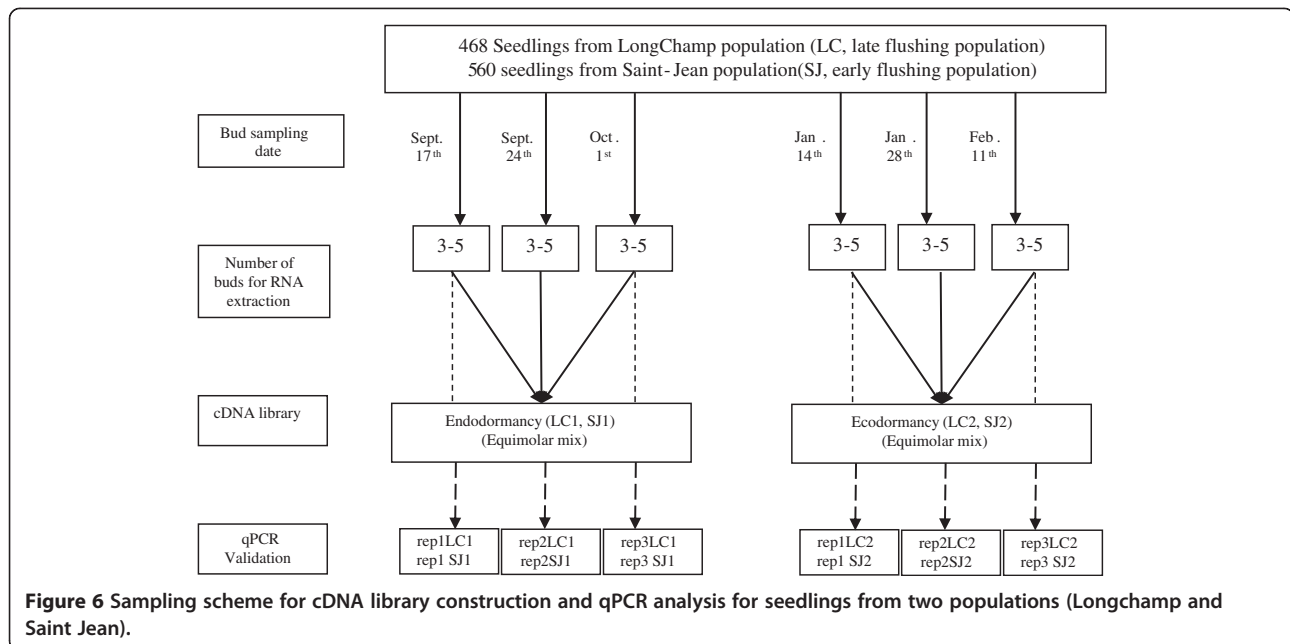
### Total RNA extraction and 454-pyrosequencing

Total RNA were extracted as described by Le Provost *et al.* [85].

RNA quantity and quality were then estimated by both spectrophotometry and visual inspection after electrophoresis in 2% agarose gels. We eliminated any residual DNA, by incubating the RNA with RNase-free DNase RQ1 (Promega, Madison WI, USA). We used two different strategies to quantify transcript accumulation. (1) Before 454-sequencing, we established two cDNA libraries (corresponding to endodormancy and ecodormancy) for each of the two populations (LC and SJ), with the SMART PCR cDNA synthesis kit (Clontech, Palo Alto, CA, USA), according to the manufacturer's instructions. Total RNA was extracted independently from apical buds harvested on September 17<sup>th</sup>, 24<sup>th</sup> and October 1<sup>st</sup> for endodormancy and on January 14<sup>th</sup>, 28<sup>th</sup> and February 11<sup>th</sup> for ecodormancy, and equimolar mixtures were then established, to obtain a representative library for each population and developmental stage (Figure 6). (2) For reverse transcription and quantitative PCR (qPCR), the same RNA samples were used, but without pooling, so as to obtain three biological replicates for each dormancy stage and population. Each library was bar-coded with MID primers and run on a Roche 454 GS-FLX platform. Reads from the four libraries are accessible at **SRA (Short Read Archive) of NCBI** (<http://www.ncbi.nlm.nih.gov/sra>).

### Cleaning, assembly and annotation of the dormant bud transcriptome

Reads were extracted with the sfffile program of Roche 454 Life Science software (Branford, CT, USA). Each 454-read was screened by cross\_match [86] for primers and adaptors and was then masked. The longest non masked region was extracted and further cleaned up with SeqClean [87]. The shorter regions were discarded to eliminate potential chimeras. Cleaned reads were assembled into OakContigV1 [28], which was constructed with ESTs from 20 Sanger- and 14 454-sequencing libraries, including the four libraries studied here (LC1, LC2, SJ1 and SJ2, Figure 6), with the TGICL pipeline (see [28] for detailed procedures).



### Detection and annotation of differentially expressed genes

We assumed that the read counts corresponding to each contig for each library were indicative of the level of gene expression. We corrected for artificial duplication biases during emulsion PCR, by developing an in-house tool called 'pyrocleaner' [88]. We defined duplicated reads as clusters in megablast [89] analyses with a minimum hit score of 100, a percent identity cutoff of 98, with an alignment of reads starting at exactly the same position and ending within 70 bp of the end of the longest sequence. Duplicated reads were not taken into account when counting the number of reads in each contig.

We corrected the number of reads in each contig for the coverage of each read and the length of the corresponding contig (Method II in [90]). We first counted the number of reads for each nucleotide position (position score) in the contig, for each library. A total score was obtained by summing the position scores over the entire contig for each library. This total score was divided by the length (base pairs) of the contig. The resulting values were rounded off for use as relative levels of gene expression (Cval: corrected value) for each library. Contigs with a total Cval over four libraries  $\geq 5$  were used for further analysis. We used three statistical methods to determine the significance of differences in gene expression between endo- and ecodormancy libraries. In the first method, *R* statistics [29] were calculated. This method can be used for the comparison of differential expression between multiple cDNA libraries. Contigs with  $R > 8$  were considered statistically significant. We also used two R Bioconductor packages: edgeR [30] and DESeq [31]. We

applied a false discovery rate (FDR) of 0.2 as a type I error for significant differential expression. The genes identified as differentially expressed with the three methods were combined, to define a minimum (MN) and a maximum (MX) set of differentially expressed genes. The MN set was defined as the intersection of the genes identified as differentially expressed by the three programs, whereas the MX set was defined as all the genes identified as differentially expressed by at least one of these techniques. Heat maps were drawn for both the MN and MX sets, for visualization of the relationships between differentially expressed genes by gplots libraries in R [91]. We assessed the statistical significance of GO slim terms for differentially expressed genes, by using the goseq [92] R Bioconductor package to overcome the length bias inherent to RNA-seq data. The FDR was set at 0.05. We used both the MN and MX sets for goseq analysis.

Homology searches were carried out with BlastX [93] against the SWISS-PROT [94] and TAIR9 protein databases. The e-value cutoff for the Blast searches was set at  $1 \times 10^{-5}$ . Gene ontology (GO) [95] annotations were based on the top Blast hit against SWISS-PROT database. Blast2GO software [96] was used to summarize GO, which were mapped onto plant-specific GO-slim terms.

### Enrichment analyses

Enrichment Analysis for Pathways and groups (EAPG) and Sub-networks (FNSE) for selected genes were performed using Ariadne Pathway Studio 9 Desktop edition Software and the Resnet Plant Version 4 database (Ariadne Genomics Inc., Rockville, MD, USA). As described on Ariadne web site (<http://www.ariadnegenomics.com>) the "Find Pathways/

Groups enriched with selected entities” function (named here as EAPG for simplicity) execute a Fisher’s exact test on each pathway or group and returns informations relative to overlapping entities together with the p-value of the statistical test. EAPG was executed against Gene Ontology as well as AraCyc plant metabolic pathway [97,98]. The “Find sub-networks Enriched with selected entities” function (FSNE) was used to identify the set of entities (sub-networks) organized by specific relationship using Fisher’s Exact Test.

#### Detection of bud-specific genes by *in silico* analysis of differentially expressed genes

Eleven cDNA libraries (see Additional file 5: Table S6) constructed from four tissue types (bud, leaf, root and xylem) were used to analyze the bud-specific expression of genes identified as differentially expressed between the endo- and ecodormancy phases (75 contigs in the MN set, 34 of which were upregulated during endodormancy, the other 41 being upregulated during ecodormancy). For a detailed description of these cDNA libraries, see [28]. We counted the number of reads mapping to the MN set from these libraries, with OakContigV1 [28], for each tissue type and used *R* statistics [29] to analyze differential expression between tissue types regarded as virtual cDNA libraries. Values of *R* > 8 were considered to indicate statistically significant differences in expression between tissue types.

#### Reverse transcription and quantitative real-time PCR (qPCR)

The following criteria was applied to select candidate genes: (1) first, genes displaying the strongest differential expression in previous statistical analyses were selected, and (2) the molecular function of the genes selected in bud development in other species was also investigated using the *Arabidopsis* database available at the following URL (<http://www.arabidopsis.org>). Using this strategy 13 candidate genes based on expression and function data (see Additional file 6: Table S7) were identified.

Among the 13 genes analyzed, only one encoding an 18.1 kDa class I heat-shock protein, displayed a multi-banding pattern after PCR amplification and was discarded from the analysis. Relative quantification was therefore carried out for the remaining 12 genes.

Five control genes were also selected, for which we analyzed expression profiles in our experimental design.

For quantification, 1 µg of total RNA were reverse transcribed using the Improm-II™ reverse transcription system (Promega®), according to the manufacturer’s instructions. The cDNA generated was diluted by a factor of 10 for qPCR analysis. We performed qPCR and quantification on a Chromo4™ Multicolor Real-Time PCR Detection System (Bio-Rad Laboratories, Inc. Hercules, CA, USA), according to a procedure described elsewhere

[99]. PCR primer pairs were designed with Primer3 software [100]. Primers were designed to have an optimal size of 20 bp (18–22 bp), a GC content of 40 to 60%, and a Tm of 60°C. Other criteria, such as primer self-annealing, were also taken into account. Oligonucleotides were synthesized by Eurogentec (Liege, Belgium). Primer pairs are listed in Additional file 6: Table S7.

We performed qPCR and data analysis as previously described [99]. Briefly, data were analyzed with the Excel (Microsoft) macro GENEX v1.10 (Gene Expression Analysis for iCycle iQ® Real-Time PCR Detection System, v1.10, 2004, Bio-Rad Laboratories), using the methods derived from published algorithms [101]. Expression levels were normalized with respect to the stable levels of expression of genes encoding an elongation factor and an 18 S ribosomal protein. GeNorm software was used to assess the stability of the genes used as internal control.

#### Additional files

**Additional file 1: Figure S1** Distribution of *R* statistics. Description of data: *R* was calculated as described by Stekel *et al.* [29]. Differentially expressed contigs are shown in the gray area.

**Additional file 2: Figure S2.** Heat maps for genes displaying differential expression between the endo- and ecodormancy phases. Description of data: A) MN and B) MX sets, as defined in the methods section.

**Additional file 3: Table S1.** Candidate genes displaying differential expression between the endo- and ecodormancy phases. Description of data: Differentially expressed genes were detected by three programs (*R* statistics, edgeR and DESeq). Their annotations with SWISS-PROT and TAIR9 databases and gene ontology were based on the top hit in BLASTX homology searches. 1: Differentially expressed genes (upregulation during ecodormancy) at a false discovery rate of 20%; -1: differentially expressed genes (upregulation during endodormancy) at a false discovery rate of 20%; NA: No annotations available. **Table S2.** EAPG results for ecodormancy up-regulated genes. **Table S3.** FNSE results for ecodormancy up-regulated genes. **Table S4.** EAPG results for endodormancy up-regulated genes.

**Additional file 4: Table S5.** Tissue-specific *in silico* expression analysis of genes differentially expressed in the endo- and ecodormancy phases. Description of data: Genes differentially expressed between tissues were identified by *R* statistics. *R* values > 8 were considered statistically significant (cells in yellow). Cells in green indicate contigs to which the reads from the bud library were mapped but to which no reads from other libraries were mapped.

**Additional file 5: Table S6.** cDNA libraries for European oak used for *in silico* analysis of tissue-specific expression. Description of data: The tissue type, the number of total reads in the OakContigV1, the library code used by Ueno *et al.* [28] and the Evoltree library code are given in the table.

**Additional file 6: Table S7.** Primer pairs used for qPCR. Description of data: List of the genes validated by qPCR in our experiment. Abbreviations are as follows: Tm: annealing temperature, NA: not available.

**Additional file 7: Figure S3.** Relative expression profiles of the genes analyzed by qPCR. Description of data: For each dormancy stage, the expression levels of the genes were estimated over three biological replicates. Error bars represent standard deviations (N=3). Abbreviations are as follows: LCendo: Longchamp endodormancy, LCeco: Longchamp ecodormancy, SJendo: Saint-Jean endodormancy and SJeco: Saint-Jean ecodormancy.

**Additional file 8: Figure S4.** Detailed phenological scale of bud development in *Quercus petraea* (Matt) Lieb. Description of data: Stage 0: quiescent bud, scales visible, stage 1: swelling bud, stage 2: bud opening, stage 3: leaves visible, stage 4: internode expansion.

#### Competing interest

The authors declare that they have no competing interests.

#### Authors' contributions

SU analyzed the data and drafted the manuscript with GLP, JCL AK and CP. GLP and JD conceived the study. GLP prepared the libraries with VL and performed qPCR with EP. CK and CN carried out the bioinformatic analysis. Pathway Studio analysis was performed by JCL. AK coordinated the project. All the authors read and approved the final manuscript.

#### Acknowledgments

The study was carried out with financial support from the European Commission (FP6-2004-GLOBAL-3, Network of Excellence EVOLTREE "Evolution of Trees as Drivers of Terrestrial Biodiversity", N°016322) and INRA (AIP Bioressource) to A. Kremer. SU received postdoctoral support from INRA. VL and EP were recruited to work on the EVOLTREE project. This research was also supported by grants from Agence Nationale de la Recherche, Plateformes Technologiques du Vivant (BOOST-SNP project, 07PFTV002).

#### Author details

<sup>1</sup>Forestry and Forest Products Research Institute, Department of Forest Genetics, Tree Genetics Laboratory, 1 Matsunosato, Tsukuba, Ibaraki 305-8687 Japan. <sup>2</sup>INRA, UMR 1202 BIOGECO, F-33610 Cestas, France. <sup>3</sup>Univ. Bordeaux, BIOGECO, UMR 1202, F-33400 Talence, France. <sup>4</sup>Plateforme bioinformatique Genotoul, UR875 Biométrie et Intelligence Artificielle, INRA, Castanet-Tolosan 31326 France. <sup>5</sup>INRA, UR0588 Amélioration Génétique et Physiologie Forestières, Orléans F-45075 France.

Received: 22 November 2012 Accepted: 4 April 2013

Published: 10 April 2013

#### References

- Rohde A, Bhalerao RP: Plant dormancy in the perennial context. *Trends Plant Sci* 2007, **12**:217–223.
- Faust M, Erez A, Rowland LJ, Wang SY, Norman HA: Bud dormancy in perennial fruit trees: physiological basis for dormancy induction, maintenance, and release. *HortSci* 1997, **32**:623–629.
- Lang GA, Early JD, Martin GC, Darnell RL: Endo-, Para-, and Ecodormancy: physiological terminology and classification for dormancy research. *HortSci* 1987, **22**:371–377.
- Naor A, Flaishman M, Stern R, Moshe A, Erez A: Temperature effects on dormancy completion of vegetative buds in apple. *J Amer Soc Hort Sci* 2003, **128**:636–641.
- Guerriero R, Monteleone P, Viti R: Evaluation of end of dormancy in several apricot cultivars according to different methodological approaches. *Acta Hort* 2006, **701**:99–103.
- Horvath DP, Anderson JV, Chao WS, Foley ME: Knowing when to grow: signals regulating bud dormancy. *Trends Plant Sci* 2003, **8**:534–540.
- Chuine, Cour P: Climatic determinants of budburst seasonality in four temperate-zone tree species. *New Phytol* 1999, **143**:339–349.
- Kalcsits L, Silim S, Tanino K: Warm temperature accelerates short photoperiod-induced growth cessation and dormancy induction in hybrid poplar (*Populus* × spp.). *Trees - Structure and Function* 2009, **23**:971–979.
- Chmielewski FM, Muller A, Bruns E: Climate change and trends in phenology of fruit trees and field crops in Germany, 1961–2000. *Agric For Meteorol* 2004, **121**:69–78.
- Menzel A, Fabian P: Growing season extended in Europe. *Nature* 1999, **397**:659.
- Ducousso A, Guyon JP, Kremer A: Latitudinal and altitudinal variation of bud burst in western populations of sessile oak (*Quercus petraea* (Matt) Liebl). *Annales Des Sciences Forestieres* 1996, **53**:775–782.
- Tsarouhas V, Gullberg U, Lagercrantz U: Mapping of quantitative trait loci controlling timing of bud flush in *Salix*. *Hereditas* 2003, **138**:172–178.
- Bradshaw HD Jr, Stettler RF: Molecular genetics of growth and development in populus. IV. Mapping QTLs with large effects on growth, form, and phenology traits in a forest tree. *Genetics* 1995, **139**:963–973.
- Saintagne C, Bodenes C, Barreneche T, Pot D, Plomion C, Kremer A: Distribution of genomic regions differentiating oak species assessed by QTL detection. *Heredity* 2004, **92**:20–30.
- Derory J, Scotti-Saintagne C, Bertocchi E, Le Dantec L, Graignic N, Jauffres A, Casasoli M, Chancerel E, Bodenes C, Alberto F, Kremer A: Contrasting relationships between the diversity of candidate genes and variation of bud burst in natural and segregating populations of European oaks. *Heredity* 2010, **104**:438–448.
- Ibanez C, Kozarewa I, Johansson M, Ogren E, Rohde A, Eriksson M: Circadian clock components regulate entry and affect exit of seasonal dormancy as well as winter hardiness in *Populus* trees. *Plant Physiol* 2010, **153**:1823–1833.
- Rohde A, Ruttink T, Hostyn V, Sterck L, Van Driessche K, Boerjan W: Gene expression during the induction, maintenance, and release of dormancy in apical buds of poplar. *J Exp Bot* 2007, **58**:4047–4060.
- Allona I, Ramos A, Ibanez C, Contreras A, Casado R, Aragoncillo C: Molecular control of winter dormancy establishment in trees. *Span J Agric Res* 2008, **6**:201–210.
- Bohlenius H, Huang T, Charbonnel-Campaa L, Brunner AM, Jansson S, Strauss SH, Nilsson O: CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science* 2006, **312**:1040–1043.
- Walton E, Wu R, Schaffer R, Thodey K, Janssen B, Richardson A, Hellens R, Rae G, Wood M: Genetic regulation of budbreak in kiwi. *Acta Hort* 2007, **753**:561–566.
- Mathiason K, He D, Grimplet J, Venkateswari J, Galbraith DW, Or E, Fennell A: Transcript profiling in *Vitis riparia* during chilling requirement fulfillment reveals coordination of gene expression patterns with optimized bud break. *Funct Integr Genomics* 2009, **9**:81–96.
- Campbell M, Seegar E, Beers L, Knauber D, Suttle J: Dormancy in potato tuber meristems: chemically induced cessation in dormancy matches the natural process based on transcript profiles. *Funct Integr Genomics* 2008, **8**:317–328.
- Mazzitelli L, Hancock RD, Haupt S, Walker PG, Pont SD, McNicol J, Cardle L, Morris J, Viola R, Brennan R, et al: Co-ordinated gene expression during phases of dormancy release in raspberry (*Rubus idaeus* L.) buds. *J Exp Bot* 2007, **58**:1035–1045.
- Horvath DP, Sung S, Kim D, Chao W, Anderson J: Characterization, expression and function of DORMANCY ASSOCIATED MADS-BOX genes from leafy spurge. *Plant Mol Biol* 2010, **73**:169–179.
- Derory J, Leger P, Garcia V, Schaeffer J, Hauser MT, Salin F, Luschnig C, Plomion C, Glossl J, Kremer A: Transcriptome analysis of bud burst in sessile oak (*Quercus petraea*). *New Phytol* 2006, **170**:723–738.
- Ruttink T, Arend M, Morreel K, Storme V, Rombaux S, Fromm J, Bhalerao RP, Boerjan W, Rohde A: A molecular timetable for apical bud formation and dormancy induction in poplar. *Plant Cell* 2007, **19**:2370–2390.
- Jimenez S, Li Z, Reighard G, Bielenberg D: Identification of genes associated with growth cessation and bud dormancy entrance using a dormancy-incapable tree mutant. *BMC Plant Biol* 2010, **10**:25.
- Ueno S, Le Provost G, Leger V, Klopp C, Noiroit C, Frigerio JM, Salin F, Salse J, Abrouk M, Murat F, et al: Bioinformatic analysis of ESTs collected by Sanger and pyrosequencing methods for a keystone forest tree species: oak. *BMC Genomics* 2010, **11**:650.
- Stekel DJ, Git Y, Falciani F: The comparison of gene expression from multiple cDNA libraries. *Genome Res* 2000, **10**:2055–2061.
- Robinson MD, McCarthy DJ, Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2009, **26**:139–140.
- Anders S, Huber W: Differential expression analysis for sequence count data. *Genome Biol* 2010, **11**:R106.
- Pacey-Miller T, Scott K, Ablett E, Tingey S, Ching A, Henry R: Genes associated with the end of dormancy in grapes. *Funct Integr Genomics* 2003, **3**:144–152.
- Casey R: Distribution and some properties of seed globulins. In *Seed proteins*. Edited by Shewry P, Casey R. Dordrecht: Kluwer Academic Publishers; 1999:159–169.
- Shewry PR, Halford NG: Cereal seed storage proteins: structures, properties and role in grain utilization. *J Exp Bot* 2002, **53**:947–958.

35. Rivin CJ, Grudt T: **Abscisic Acid and the developmental regulation of embryo storage proteins in maize.** *Plant Physiol* 1991, **95**:358–365.
36. Shutov AD, Vaintraub IA: **Degradation of storage proteins in germinating seeds.** *Phytochemistry* 1987, **26**:1557–1566.
37. Anderson JV, Gesch RW, Jia Y, Chao WS, Horvath DP: **Seasonal shifts in dormancy status, carbohydrate metabolism, and related gene expression in crown buds of leafy spurge.** *Plant Cell Environ* 2005, **28**:1567–1578.
38. Yokoyama R, Nishitani K: **A comprehensive expression analysis of all members of a gene family encoding cell-wall enzymes allowed us to predict cis-regulatory regions involved in cell-wall construction in specific organs of *Arabidopsis*.** *Plant Cell Physiol* 2001, **42**:1025–1033.
39. Catala C, Rose JK, York WS, Albersheim P, Darvill AG, Bennett AB: **Characterization of a tomato xyloglucan endotransglycosylase gene that is down-regulated by auxin in etiolated hypocotyls.** *Plant Physiol* 2001, **127**:1180–1192.
40. Purugganan MM, Braam J, Fry SC: **The *Arabidopsis* TCH4 xyloglucan endotransglycosylase. Substrate specificity, pH optimum, and cold tolerance.** *Plant Physiol* 1997, **115**:181–190.
41. Schrader J, Moyle R, Bhalerao R, Hertzberg M, Lundeberg J, Nilsson P, Bhalerao RP: **Cambial meristem dormancy in trees involves extensive remodelling of the transcriptome.** *Plant J* 2004, **40**:173–187.
42. Leida C, Terol J, Marti G, Agusti M, Llacer G, Badenes ML, Rios G: **Identification of genes associated with bud dormancy release in *Prunus persica* by suppression subtractive hybridization.** *Tree Physiol* 2010, **30**:655–666.
43. Cohen D, Bogeat-Triboulot MB, Tisserant E, Balzergue S, Martin-Magniette ML, Lelandais G, Ningre N, Renou JP, Tamby JP, Le Thiec D, Hummel I: **Comparative transcriptomics of drought responses in *Populus*: a meta-analysis of genome-wide expression profiling in mature leaves and root apices across two genotypes.** *BMC Genomics* 2010, **11**:630.
44. Hortensteiner S: **Chlorophyll degradation during senescence.** *Annu Rev Plant Biol* 2006, **57**:55–77.
45. Nomura K, Takahashi N, Hirose M, Nakamura S, Yagi F: **Overall carbohydrate-binding properties of *Castanea crenata* agglutinin (CCA).** *Carbohydr Res* 2005, **340**:2004–2009.
46. Horvath D: **Bud Dormancy and Growth.** In *Plant Developmental Biology - Biotechnological Perspectives*. Edited by Pua EC, Davey MR. Berlin Heidelberg: Springer; 2010:53–70.
47. Anderson J, Doğramaci M, Horvath D, Foley M, Chao W, Suttle J, Thimmapuram J, Hernandez A, Ali S, Mikel M: **Auxin and ABA act as central regulators of developmental networks associated with paradormancy in Canada thistle (*Cirsium arvense*).** *Funct Integr Genomics* 2012, **12**:515–531.
48. Meyer G, Kloppstech K: **A rapidly light-induced chloroplast protein with a high turnover coded for by pea nuclear DNA.** *Eur J Biochem* 1984, **138**:201–207.
49. Montane MH, Kloppstech K: **The family of light-harvesting-related proteins (LHCs, ELIPs, HLIPIs): was the harvesting of light their primary function?** *Gene* 2000, **258**:1–8.
50. Montané M-H, Dreyer S, Triantaphylides C, Kloppstech K: **Early light-inducible proteins during long-term acclimation of barley to photooxidative stress caused by light and cold: high level of accumulation by posttranscriptional regulation.** *Planta* 1997, **202**:293–302.
51. Savenstrand H, Olofsson M, Samuelsson M, Strid A: **Induction of early light-inducible protein gene expression in *Pisum sativum* after exposure to low levels of UV-B irradiation and other environmental stresses.** *Plant Cell Rep* 2004, **22**:532–536.
52. Dhanaraj AL, Slovin JP, Rowland LJ: **Analysis of gene expression associated with cold acclimation in blueberry floral buds using expressed sequence tags.** *Plant Sci* 2004, **166**:863–872.
53. Chung S, Parish RW: **Combinatorial interactions of multiple cis-elements regulating the induction of the *Arabidopsis* XERO2 dehydrin gene by abscisic acid and cold.** *Plant J* 2008, **54**:15–29.
54. Jung SH, Lee JY, Lee DH: **Use of SAGE technology to reveal changes in gene expression in *Arabidopsis* leaves undergoing cold stress.** *Plant Mol Biol* 2003, **52**:553–567.
55. Close TJ: **Dehydrins: Emergence of a biochemical role of a family of plant dehydration proteins.** *Physiol Plant* 1996, **97**:795–803.
56. Yakovlev I, Asante D, Fossdal C, Partanen J, Junttila O, Johnsen Ø: **Dehydrins expression related to timing of bud burst in Norway spruce.** *Planta* 2008, **228**:459–472.
57. Yang L, Zheng B, Mao C, Yi K, Liu F, Wu Y, Tao Q, Wu P: **cDNA-AFLP analysis of inducible gene expression in rice seminal root tips under a water deficit.** *Gene* 2003, **314**:141–148.
58. Kovtun Y, Chiu WL, Tena G, Sheen J: **Functional analysis of oxidative stress-activated mitogen-activated protein kinase cascade in plants.** *Proc Natl Acad Sci USA* 2000, **97**:2940–2945.
59. Dure L 3rd, Greenway SC, Galau GA: **Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by in vitro and in vivo protein synthesis.** *Biochemistry* 1981, **20**:4162–4168.
60. Galau GA, Hughes DW, Dure L: **Abscisic acid induction of cloned cotton late embryogenesis-abundant (Lea) mRNAs.** *Plant Mol Biol* 1986, **7**:155–170.
61. Horvath DP, Chao WS, Suttle JC, Thimmapuram J, Anderson JV: **Transcriptome analysis identifies novel responses and potential regulatory genes involved in seasonal dormancy transitions of leafy spurge (*Euphorbia esula* L.).** *BMC Genomics* 2008, **9**:536.
62. Hundertmark M, Hincha DK: **LEA (late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*.** *BMC Genomics* 2008, **9**:118.
63. Sakai A, Larcher W: **Frost survival of plants: responses and adaptation to freezing stress.** Berlin: Springer-Verlag; 1987.
64. Medina J, Catala R, Salinas J: **Developmental and stress regulation of RC12A and RC12B, two cold-inducible genes of *Arabidopsis* encoding highly conserved hydrophobic proteins.** *Plant Physiol* 2001, **125**:1655–1666.
65. Davik J, Koehler G, From B, Torp T, Rohloff J, Eidem P, Wilson RC, Sonstebj A, Randall SK, Alsheikh M: **Dehydrin, alcohol dehydrogenase, and central metabolite levels are associated with cold tolerance in diploid strawberry (*Fragaria* spp.).** *Planta* 2013, **237**:265–277.
66. Doğramaci M, Horvath D, Christoffers M, Anderson J: **Dehydration and vernalization treatments identify overlapping molecular networks impacting endodormancy maintenance in leafy spurge crown buds.** *Funct Integr Genomics* 2011, **11**:611–626.
67. Cong L, Zheng H-C, Zhang Y-X, Chai T-Y: ***Arabidopsis* DREB1A confers high salinity tolerance and regulates the expression of GA dioxygenases in Tobacco.** *Plant Sci* 2008, **174**:156–164.
68. Doğramaci M, Horvath DP, Chao WS, Foley ME, Christoffers MJ, Anderson JV: **Low temperatures impact dormancy status, flowering competence, and transcript profiles in crown buds of leafy spurge.** *Plant Mol Biol* 2010, **73**:207–226.
69. Herzog M, Dorne AM, Grellet F: **GASA, a gibberellin-regulated gene family from *Arabidopsis thaliana* related to the tomato GAST1 gene.** *Plant Mol Biol* 1995, **27**:743–752.
70. Shi L, Gast RT, Gopalraj M, Olszewski NE: **Characterization of a shoot-specific, GA3- and ABA-regulated gene from tomato.** *Plant J* 1992, **2**:153–159.
71. Aubert D, Chevillard M, Dorne AM, Arlaud G, Herzog M: **Expression patterns of GASA genes in *Arabidopsis thaliana*: the GASA4 gene is up-regulated by gibberellins in meristematic regions.** *Plant Mol Biol* 1998, **36**:871–883.
72. Roxrud I, Lid SE, Fletcher JC, Schmidt ED, Opsahl-Sorteberg HG: **GASA4, one of the 14-member *Arabidopsis* GASA family of small polypeptides, regulates flowering and seed development.** *Plant Cell Physiol* 2007, **48**:471–483.
73. Suttle JC: **Involvement of endogenous gibberellins in potato tuber dormancy and early sprout growth: a critical assessment.** *J Plant Physiol* 2004, **161**:157–164.
74. Garcia-Martinez JL, Gil J: **Light regulation of gibberellin biosynthesis and mode of action.** *J Plant Growth Regul* 2001, **20**:354–368.
75. Ueguchi-Tanaka M, Nakajima M, Motoyuki A, Matsuoka M: **Gibberellin receptor and its role in gibberellin signaling in plants.** *Annu Rev Plant Biol* 2007, **58**:183–198.
76. Rinne PL, Welling A, Vahala J, Ripel L, Ruonala R, Kangasjarvi J, van der Schoot C: **Chilling of dormant buds hyperinduces FLOWERING LOCUS T and recruits GA-inducible 1,3-beta-glucanases to reopen signal conduits and release dormancy in *Populus*.** *Plant Cell* 2011, **23**:130–146.
77. Dixon DP, Laphom A, Edwards R: **Plant glutathione transferases.** *Genome Biol* 2002, **3**:REVIEWS3004.1-3004.10.
78. Halaly T, Pang X, Batikoff T, Crane O, Keren A, Venkateswari J, Ogrodovitch A, Sadka A, Lavee S, Or E: **Similar mechanisms might be triggered by alternative external stimuli that induce dormancy release in grape buds.** *Planta* 2008, **228**:79–88.

79. Hedley PE, Russell JR, Jorgensen L, Gordon S, Morris JA, Hackett CA, Cardle L, Brennan R: **Candidate genes associated with bud dormancy release in blackcurrant (*Ribes nigrum* L.).** *BMC Plant Biol* 2010, **10**:202.
80. Prasad T: **Mechanisms of chilling-induced oxidative stress injury and tolerance in developing maize seedlings: changes in antioxidant system, oxidation of proteins and lipids, and protease activities.** *Plant J* 1996, **10**:1017–1026.
81. Stafstrom JP, Ripley BD, Devitt ML, Drake B: **Dormancy-associated gene expression in pea axillary buds. Cloning and expression of PsDRM1 and PsDRM2.** *Planta* 1998, **205**:547–552.
82. Naik D, Dhanaraj AL, Arora R, Rowland LJ: **Identification of genes associated with cold acclimation in blueberry (*Vaccinium corymbosum* L.) using a subtractive hybridization approach.** *Plant Sci* 2007, **173**:213–222.
83. Koike M, Takezawa D, Arakawa K, Yoshida S: **Accumulation of 19-kDa plasma membrane polypeptide during induction of freezing tolerance in wheat suspension-cultured cells by abscisic acid.** *Plant Cell Physiol* 1997, **38**:707–716.
84. Guy C: **Molecular responses of plants to cold shock and cold acclimation.** *J Mol Microbiol Biotechnol* 1999, **1**:231–242.
85. Le Provost G, Herrera R, Paiva J, Chaumeil P, Salin F, Plomion C: **A micromethod for high throughput RNA extraction in forest trees.** *Biol Res* 2007, **40**:291–297.
86. *Documentation for phrap and cross\_match.* <http://bozeman.mbt.washington.edu/phrap.docs/phrap.html>.
87. *SeqClean.* <http://sourceforge.net/projects/seqclean/>.
88. Jerome M, Noirot C, Klopp C: **Assessment of replicate bias in 454 pyrosequencing and a multi-purpose read-filtering tool.** *BMC research notes* 2011, **4**:149.
89. Zhang Z, Schwartz S, Wagner L, Miller W: **A greedy algorithm for aligning DNA sequences.** *J Comput Biol* 2000, **7**:203–214.
90. Wilhelm BT, Landry JR: **RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing.** *Methods* 2009, **48**:249–257.
91. R Development Core Team: *A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing; 2010.
92. Young MD, Wakefield MJ, Smyth GK, Oshlack A: **Gene ontology analysis for RNA-seq: accounting for selection bias.** *Genome Biol* 2010, **11**:R14.
93. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**:403–410.
94. Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, *et al*: **The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003.** *Nucleic Acids Res* 2003, **31**:365–370.
95. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, *et al*: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet* 2000, **25**:25–29.
96. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M: **Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research.** *Bioinformatics* 2005, **21**:3674–3676.
97. Mootha VK, Lindgren CM, Eriksson K-F, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, *et al*: **PGC-1[alpha]-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes.** *Nat Genet* 2003, **34**:267–273.
98. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP: **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.** *Proc Natl Acad Sci USA* 2005, **102**:15545–15550.
99. Paiva JA, Garces M, Alves A, Garnier-Gere P, Rodrigues JC, Lalanne C, Porcon S, Le Provost G, Perez Dda S, Brach J, *et al*: **Molecular and phenotypic profiling from the base to the crown in maritime pine wood-forming tissue.** *New Phytol* 2008, **178**:283–301.
100. Rozen S, Skaletsky HJ: **Primer3 on the WWW for general users and for biologist programmers.** In *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Edited by Krawetz SA, Misener S. Totowa/NJ: Humana Press; 2000:365–386.
101. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: **Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.** *Genome Biol* 2002, **3**: RESEARCH0034.

doi:10.1186/1471-2164-14-236

**Cite this article as:** Ueno *et al.*: Transcriptional profiling of bud dormancy induction and release in oak by next-generation sequencing. *BMC Genomics* 2013 **14**:236.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

