

RESEARCH ARTICLE

Open Access

# Transcriptome analysis of bitter acid biosynthesis and precursor pathways in hop (*Humulus lupulus*)

Shawn M Clark<sup>1</sup>, Vinidhra Vaitheeswaran<sup>1</sup>, Stephen J Ambrose<sup>1</sup>, Randy W Purves<sup>1</sup> and Jonathan E Page<sup>1,2\*</sup>

## Abstract

**Background:** Bitter acids (e.g. humulone) are prenylated polyketides synthesized in lupulin glands of the hop plant (*Humulus lupulus*) which are important contributors to the bitter flavour and stability of beer. Bitter acids are formed from acyl-CoA precursors derived from branched-chain amino acid (BCAA) degradation and C5 prenyl diphosphates from the methyl-D-erythritol 4-phosphate (MEP) pathway. We used RNA sequencing (RNA-seq) to obtain the transcriptomes of isolated lupulin glands, cones with glands removed and leaves from high  $\alpha$ -acid hop cultivars, and analyzed these datasets for genes involved in bitter acid biosynthesis including the supply of major precursors. We also measured the levels of BCAAs, acyl-CoA intermediates, and bitter acids in glands, cones and leaves.

**Results:** Transcripts encoding all the enzymes of BCAA metabolism were significantly more abundant in lupulin glands, indicating that BCAA biosynthesis and subsequent degradation occurs in these specialized cells. Branched-chain acyl-CoAs and bitter acids were present at higher levels in glands compared with leaves and cones. RNA-seq analysis showed the gland-specific expression of the MEP pathway, enzymes of sucrose degradation and several transcription factors that may regulate bitter acid biosynthesis in glands. Two branched-chain aminotransferase (BCAT) enzymes, *HIBCAT1* and *HIBCAT2*, were abundant, with gene expression quantification by RNA-seq and qRT-PCR indicating that *HIBCAT1* was specific to glands while *HIBCAT2* was present in glands, cones and leaves. Recombinant *HIBCAT1* and *HIBCAT2* catalyzed forward (biosynthetic) and reverse (catabolic) reactions with similar kinetic parameters. *HIBCAT1* is targeted to mitochondria where it likely plays a role in BCAA catabolism. *HIBCAT2* is a plastidial enzyme likely involved in BCAA biosynthesis. Phylogenetic analysis of the hop BCATs and those from other plants showed that they group into distinct biosynthetic (plastidial) and catabolic (mitochondrial) clades.

**Conclusions:** Our analysis of the hop transcriptome significantly expands the genomic resources available for this agriculturally-important crop. This study provides evidence for the lupulin gland-specific biosynthesis of BCAAs and prenyl diphosphates to provide precursors for the production of bitter acids. The biosynthetic pathway leading to BCAAs in lupulin glands involves the plastidial enzyme, *HIBCAT2*. The mitochondrial enzyme *HIBCAT1* degrades BCAAs as the first step in the catabolic pathway leading to branched chain-acyl-CoAs.

## Background

The female inflorescences ('cones') of *Humulus lupulus* L. (hop, Cannabaceae) contain prenylated acylphloroglucinols and prenylchalcones (e.g. xanthohumol) that are important for the brewing industry and have potential medicinal uses [1–3]. The acylphloroglucinols, referred to as bitter acids because they contribute the characteristic

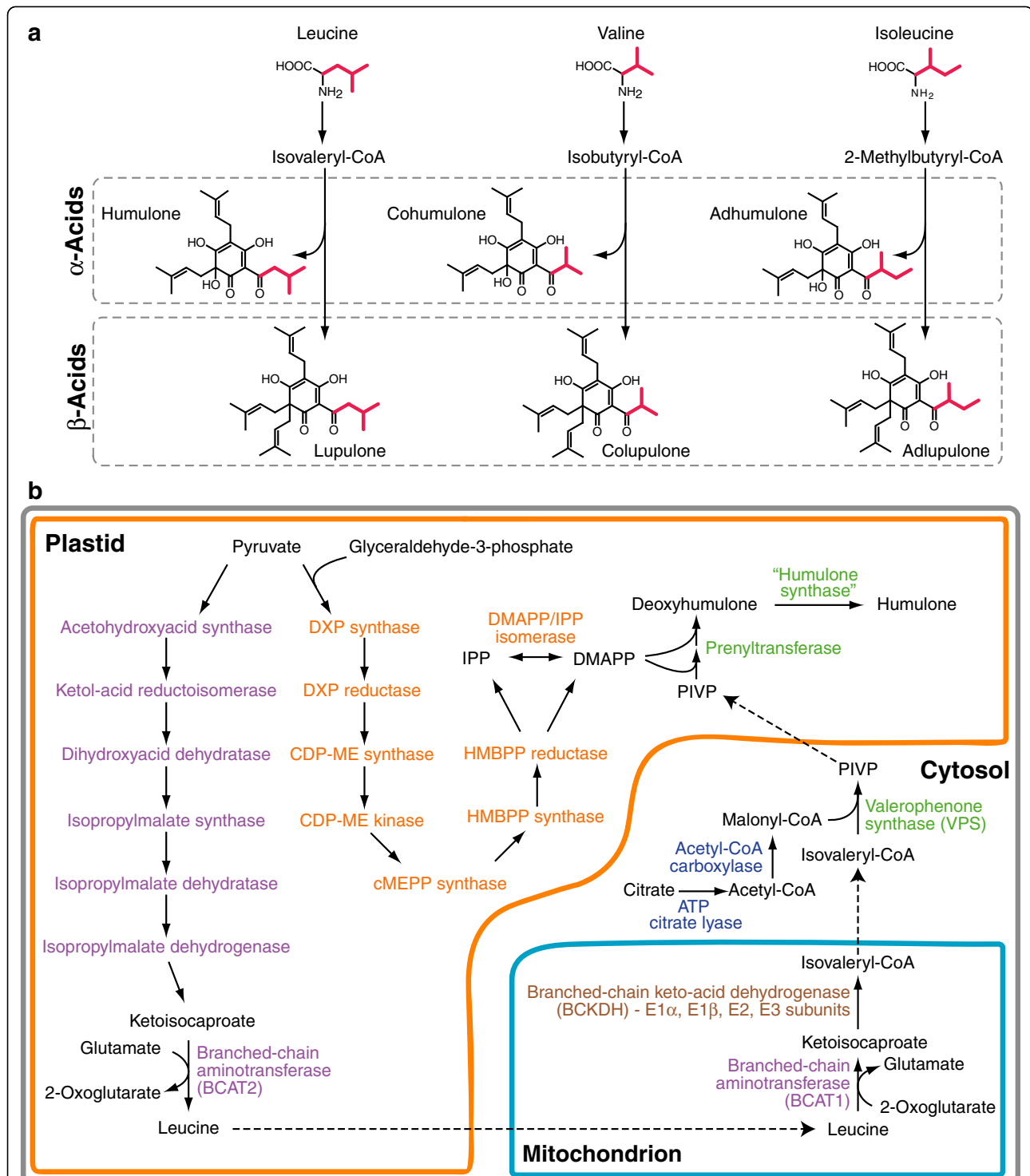
bitter flavour of beer, include the  $\alpha$ -acid humulone and its acyl-side chain variants cohumulone and adhumulone (Figure 1a). Antimicrobial  $\beta$ -acids (e.g. lupulone, colupulone and adlupulone), which differ from  $\alpha$ -acids by the presence of an additional dimethylallyl prenyl group, are also abundant. Hop cultivars vary widely in their content and composition of bitter acids with some "super-alpha" cultivars containing greater than 20% humulone by dry weight in cones [4].

The bitter acids and other prenylated polyketides occur mainly in lupulin glands, which are glandular trichomes present on hop cones and leaves [5].

\* Correspondence: jon.page@nrc-cnrc.gc.ca

<sup>1</sup>National Research Council of Canada, 110 Gymnasium Place, Saskatoon, SK S7N 0W9, Canada

<sup>2</sup>Department of Biology, University of Saskatchewan, 112 Science Place, Saskatoon, SK S7N 5E2, Canada



**Figure 1 Biosynthesis of hop bitter acids from branched-chain amino acid (BCAA) precursors. a** The acyl side-chains of humulone/lupulone are derived from leucine, cohumulone/colupulone from valine and adhumulone/adlupulone from isoleucine via the corresponding acyl-CoA thioesters. The carbon chain of the BCAAs is shown in red to indicate their incorporation into the α- and β-acids. **b** Schematic depiction of the major pathways contributing to the biosynthesis of the major α-acid product, humulone, in lupulin glands. Pathways are coloured to match Figure 2b: BCAA biosynthesis, purple; BCKDH complex, brown; MEP pathway, orange; malonyl-CoA biosynthesis, blue; bitter acid biosynthesis, green. The movement of metabolic intermediates between cellular compartments is indicated by dashed line. Note: the subcellular localization of the putative humulone synthase enzyme is not known.

Glandular trichomes typically have little or no photosynthetic activity and often contain only a few highly active biosynthetic pathways, which has made transcriptome analysis of isolated trichome secretory cells useful for the identification of the genes involved in the production of specialized metabolites [6,7]. Enzymes involved in the biosynthesis of xanthohumol and terpenoids in hop lupulin glands [5,8] and cannabinoids in glandular trichomes of the closely related species *Cannabis sativa* (Cannabaceae) [9,10] have been identified using a combination of trichome-targeted EST analysis and biochemical assays.

The precursors for bitter acid biosynthesis are branched-chain acyl-CoAs thioesters, which function as primers for polyketide synthesis, malonyl-CoA, which is the polyketide extender and dimethylallyl diphosphate (DMAPP), which serves as a prenyl donor. The acyl-CoAs, isovaleryl-CoA, isobutyryl-CoA and 2-methylbutyryl-CoA, are obtained by the degradation of the branched-chain amino acids (BCAA) leucine, valine and isoleucine, respectively (Figure 1a). The study of BCAA metabolism is of particular interest in hop since it is one of the few plants that produces large amounts of BCAA-derived natural products in specialized secretory structures. Some Solanaceae species utilize BCAAs in the trichome-localized production of acyl sugars [11,12].

BCAA biosynthesis in plants consists of eight enzymatic steps to produce leucine and valine from pyruvate, or isoleucine from threonine and pyruvate (Figure 1b) (reviewed in Binder *et al.* [13]). Branched-chain aminotransferase (BCAT) enzymes catalyze both the final step in BCAA biosynthesis and the first step in BCAA degradation. These two phases are physically separated, with biosynthesis occurring in plastids and catabolism in mitochondria. BCATs have been studied in both *Arabidopsis* and tomato with each containing six and seven isoforms, respectively [14–16]. BCAA derived CoA-esters are generated by the branched-chain ketoacid dehydrogenase complex (BCKDH) which is made up of four subunits [17]. Three of these subunits, ketoacid dehydrogenase E1 $\alpha$  and E1 $\beta$ , and dihydrolipoyl acyltransferase (E2), are specific to the BCKDH complex while the dihydrolipoyl dehydrogenase (E3) subunit also functions in other metabolic pathways [18].

Labeling studies have shown that the DMAPP required for the prenylation of bitter acids is provided by the methyl-D-erythritol 4-phosphate (MEP) pathway localized in plastids [19]. MEP-derived isoprenoids are also used in the biosynthesis of mono- and sesquiterpenes, which co-occur with bitter acids in lupulin glands [8]. The MEP pathway is initiated with glyceraldehyde-3-phosphate and pyruvate; the latter is also used in BCAA biosynthesis (Figure 1b). Previous investigations have shown that transcripts for the seven

MEP pathway enzymes and DMAPP/IPP isomerase are abundant in lupulin glands but their expression was not compared with leaves or other organs [5,8].

The final steps in bitter acid biosynthesis are unique to hop. The type III polyketide synthase enzyme valerophenone synthase (VPS) condenses a BCAA-derived acyl-CoA starter molecule with three molecules of malonyl-CoA to form the polyketide core (e.g. phlorisovalerophenone, PIVP) [20] (Figure 1b). One or more aromatic prenyltransferase enzymes perform two C5 prenylations to yield diprenyl intermediates (e.g. deoxyhumulone) [21]. At this point the  $\alpha$ - and  $\beta$ -acid pathways diverge with a third prenylation forming the  $\beta$ -acids while an oxygenation reaction, catalyzed by an as yet unknown humulone synthase enzyme, yields the  $\alpha$ -acids [22].

Here we used RNA-seq to assemble and compare the transcriptomes of lupulin glands, cones (with glands removed) and leaves. We analyzed the expression of genes involved in precursor biosynthesis, including BCAAs, MEP-derived isoprenoids and malonyl-CoA, and the final steps in bitter acid formation. We focused on the BCAT enzymes which play key roles in the formation of branched-chain acyl-CoAs. Two BCAT-encoding genes were present, *HIBCAT1* and *HIBCAT2*, with transcripts for both more abundant in lupulin glands compared with leaves and cones. *HIBCAT1* and *HIBCAT2* were experimentally-localized using GFP fusion proteins to the mitochondrion and plastid, respectively. Assays of recombinant *HIBCAT1* and *HIBCAT2* showed they catalyze biosynthetic and catabolic reactions. These findings provide evidence for lupulin gland-specific BCAA and isoprenoid metabolism to produce precursors for bitter acid biosynthesis.

## Results

### Transcriptome sequencing and assembly

Lupulin glands were isolated from the mature hop cones of four commercial cultivars grown for their high  $\alpha$ -acid content: 'Taurus', 'Nugget', 'Magnum' and 'Apollo'. Leaves and cones were also collected from 'Taurus' and 'Apollo' cultivars. The cone samples consisted of the photosynthetic bracts and bracteoles that remained after the removal of lupulin glands. RNA purified from eleven samples (five gland, three cone, three leaf) was reverse-transcribed and the cDNAs sequenced on a single lane of an Illumina HiSeq 2000 instrument as a multiplexed sample. Each cDNA produced 15.7–24.7 million reads (Additional file 1: Table S1). The reads obtained from all samples were assembled using Trinity [23] and refined using Cdhit-EST [24] to produce a transcriptome of ~170 000 contigs with an average read length of 745 bp. This transcriptome was used as the reference for RNA-seq analysis of the individual tissues using

CLC bio. At least 87% of the reads from each sample were mapped to the reference transcriptome for each tissue (Additional file 1: Table S1).

#### Comparison of metabolic transcripts in lupulin glands, cones and leaves

We compared the expression of mRNAs encoding primary metabolic enzymes (Calvin cycle, tricarboxylic acid (TCA) cycle and glycolysis) between the three tissues (Figure 2a, Additional file 2: Table S2). As expected, Calvin cycle genes were predominantly represented in the photosynthetic tissue (leaves and cones) but not in the lupulin glands. Notable exceptions were the Calvin cycle enzymes involved in glycolysis, which are important for the flow of carbon from sucrose to pyruvate in the glands. An isoform of the Rubisco small subunit was one of the most abundant genes in the leaf transcriptome, while the Rubisco large subunit was not highly expressed. TCA cycle genes showed only minor differences in expression levels between the three tissues. Transcripts encoding glycolytic enzymes were also represented at comparable levels in leaves and lupulin glands, however, some genes within this pathway showed specificity to one tissue over the other. Several genes encoding enzymes required for sucrose utilization (e.g. *fructokinase*, *neutral alkaline/invertase* and *sucrose synthase*) were more highly expressed in the gland versus leaf tissue.

MEP pathway genes were more highly expressed in lupulin glands compared with green tissues. As expected from the need for malonyl-CoA for polyketide synthesis, transcripts encoding acetyl-CoA carboxylase [25] and ATP-citrate lyase [26] were abundant in the glands (Figure 2b).

Most genes in the BCAA biosynthetic pathway showed higher expression in lupulin glands compared to cones and leaves (Figure 2b), with the transcript for ketol-acid reductoisomerase showing very high abundance. Exceptions were the low expression of threonine deaminase, which is only required for isoleucine biosynthesis, and the regulatory subunit of acetohydroxyacid synthase, which had minimal expression in all three tissues (Additional file 2: Table S2). Four putative BCAT-encoding genes were identified in the transcriptome. Two of these genes, denoted *Humulus lupulus BCAT1* (*HIBCAT1*) and *HIBCAT2*, were abundant in the lupulin glands and *HIBCAT1* showed gland-specific expression (Figure 2b). Formation of the BCAA derived CoA-esters is carried out by the BCKDH complex (Figure 1b). Two isoforms of the E1 $\alpha$  subunit were identified in the dataset and the expression of both isoforms along with the E1 $\beta$  and E2 subunits were increased in the lupulin glands (Figure 2b). The expression of the E3 subunit was similar in all tissues.

*VPS* was one of the most abundant mRNAs in the lupulin gland dataset, dwarfing all other genes involved in secondary metabolism examined here (Figure 2b). Two aromatic prenyltransferase genes, the previously characterized *HIPT1* [21] and *HIPT2*, showed similar gland-specific expression.

#### Lupulin gland-specific transcription factors

We also analyzed the RNA-seq data for transcription factors that may be involved in regulating bitter acid biosynthesis or other lupulin gland-specific metabolic activities. Forty-eight transcription factors were significantly more abundant in lupulin glands compared to the leaves (Additional file 3: Table S3). Of these, a subset of 29 was also higher in glands compared to cones. Eleven transcription factors were expressed at more than ten-fold higher levels in the lupulin glands compared to cones and leaves (Table 1).

#### Validation of gene expression by quantitative RT-PCR

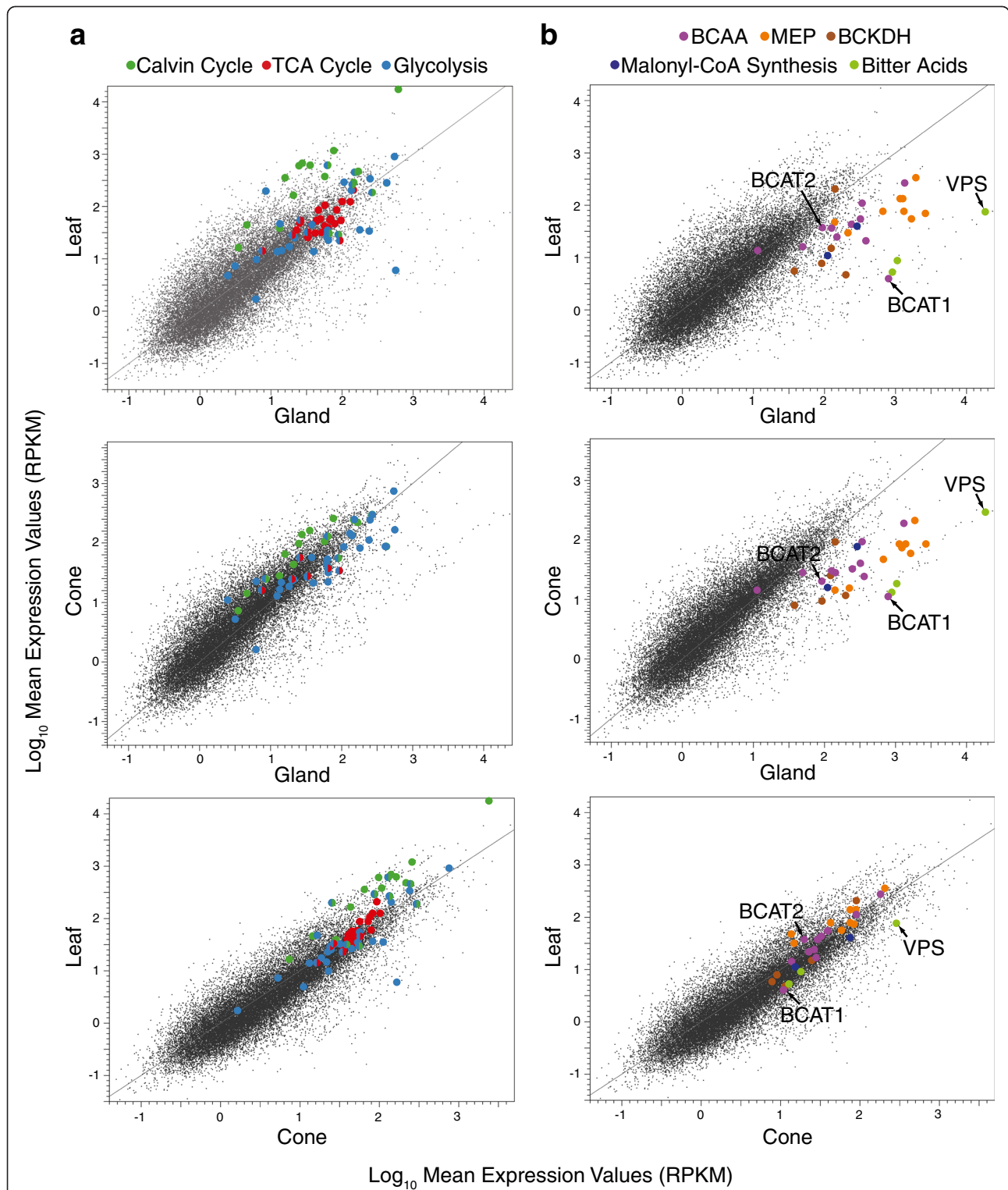
The RNA-seq analysis of the BCKDH complex, lupulin gland-specific BCATs, *VPS* and *HIPT1* was confirmed using qRT-PCR. *HIBCAT1* was lupulin gland specific while *HIBCAT2* was expressed in both lupulin glands and leaves (Additional file 4: Figure S1). *E1 $\alpha$* , *E1 $\beta$*  and *E2* of the BCKDH complex were more highly expressed in lupulin glands while the *E3* subunit was represented equally in all tissues. The genes encoding *VPS* and *HIPT1* were highly expressed in the lupulin glands while essentially non-existent in the leaf tissue. All of these findings were in agreement with the RNA-seq data (Figure 2b).

#### Metabolite profiling of BCAAs, acyl-CoAs and bitter acids

We quantified BCAA levels in lupulin glands, cones and leaves using HPLC with fluorescence detection of their derivatized adducts. Leucine, valine and isoleucine levels were similar in lupulin glands and leaves, with valine the most prominent (Figure 3a). BCAA levels in the cone tissue were highly variable and, in the case of leucine and isoleucine, were also elevated compared to glands and leaves.

Using a sensitive LC-MS method with multiple reaction monitoring, we measured the abundance of the acyl-CoAs that are used as substrates in bitter acid biosynthesis (malonyl-CoA, isovaleryl-CoA, isobutyryl-CoA and 2-methylbutyryl-CoA). Although acetyl-CoA is not used by *VPS* directly, it is converted to malonyl-CoA by acetyl-CoA carboxylase and therefore we included it in the analysis. Both acetyl-CoA and malonyl-CoA exhibit elevated levels in lupulin glands (Figure 3b). The starter substrates for the *VPS* reaction, isovaleryl-CoA, isobutyryl-CoA and 2-methylbutyryl-CoA, accumulated in glands while remaining largely absent in other tissues.





**Figure 2** Scatter plot comparison of gene expression of metabolic pathways between lupulin glands, leaves and cones. Data is presented as reads per kb exon model per million mapped reads (RPKM). Corresponding gene names and expression values can be found in Table S2. **a**) Expression of genes encoding enzymes of primary metabolic pathways (Calvin cycle, TCA cycle and glycolysis). **b**) Expression of genes encoding enzymes of BCAA metabolism (BCAA, BCKDH complex, *HIBCAT1* and *HIBCAT2*), MEP pathway, malonyl-CoA synthesis and bitter acid biosynthesis.

**Table 1 Transcription factors that are ten-fold more abundant in the lupulin glands compared to leaves and cones**

Feature ID	Description	Fold change	
		Gland <sup>a</sup> vs Leaf <sup>b</sup>	Gland <sup>a</sup> vs Cone <sup>b</sup>
comp76122_c0_seq4	MYB	498.2	21.63
comp50231_c0_seq1	WRKY	138.5	30.8
comp68974_c0_seq1	WRKY	127.2	34.3
comp53904_c0_seq1	MYB (CAN883 homolog <sup>c</sup> )	75.6	14.5
comp69082_c0_seq1	MYB	67.3	41.5
comp70463_c0_seq1	AP2-like ethylene response	56.6	36.4
comp58890_c0_seq1	Homeobox-leucine zipper	54.9	23.0
comp27968_c0_seq1	Zinc finger	50.0	10.1
comp72691_c0_seq1	Homeobox-leucine zipper	45.3	16.7
comp73886_c0_seq1	WRKY	39.9	23.4
comp63432_c0_seq1	MYB (CAN738 homolog <sup>c</sup> )	18.8	23.1

Fold change represents the difference between mean RPKM values for gland compared to leaf and cone.

<sup>a</sup> n=5.

<sup>b</sup> n=3.

<sup>c</sup> Marks et al. [38].

Isovaleryl-CoA was five times more abundant than the other branched-chain acyl-CoAs.

HPLC analysis of the major bitter acids and their immediate precursors showed, as expected, that the  $\alpha$ -acids accumulated to higher levels than the  $\beta$ -acids and that humulone was the most abundant end-product (Figure 3c). These results correlated well with the levels of acyl-CoA precursors (Figure 3b). We also quantified the diprenylated precursors (e.g. deoxyhumulone) of the bitter acids. Small amounts of deoxyhumulone and deoxycohumulone were present while deoxyadhumulone could not be detected (Figure 3c, inset). The absence of deoxyadhumulone may reflect the overall lower levels of isoleucine-derived bitter acids. Although we attempted to identify the products from the VPS reaction such as phlorisovalerophenone, as well as the monoprenylated forms, these metabolites were undetectable.

#### HIBCAT1 and HIBCAT2 catalyze aminotransferase reactions with BCAAs

To further investigate BCAA metabolism in lupulin glands, we characterized the biochemical properties of the two gland-specific BCAT enzymes. HIBCAT1 is 393 amino acids long with calculated molecular weight of 43.2 kDa; HIBCAT2 is a 44.5 kDa (408 amino acid) protein. Nucleotide sequences for both enzymes are

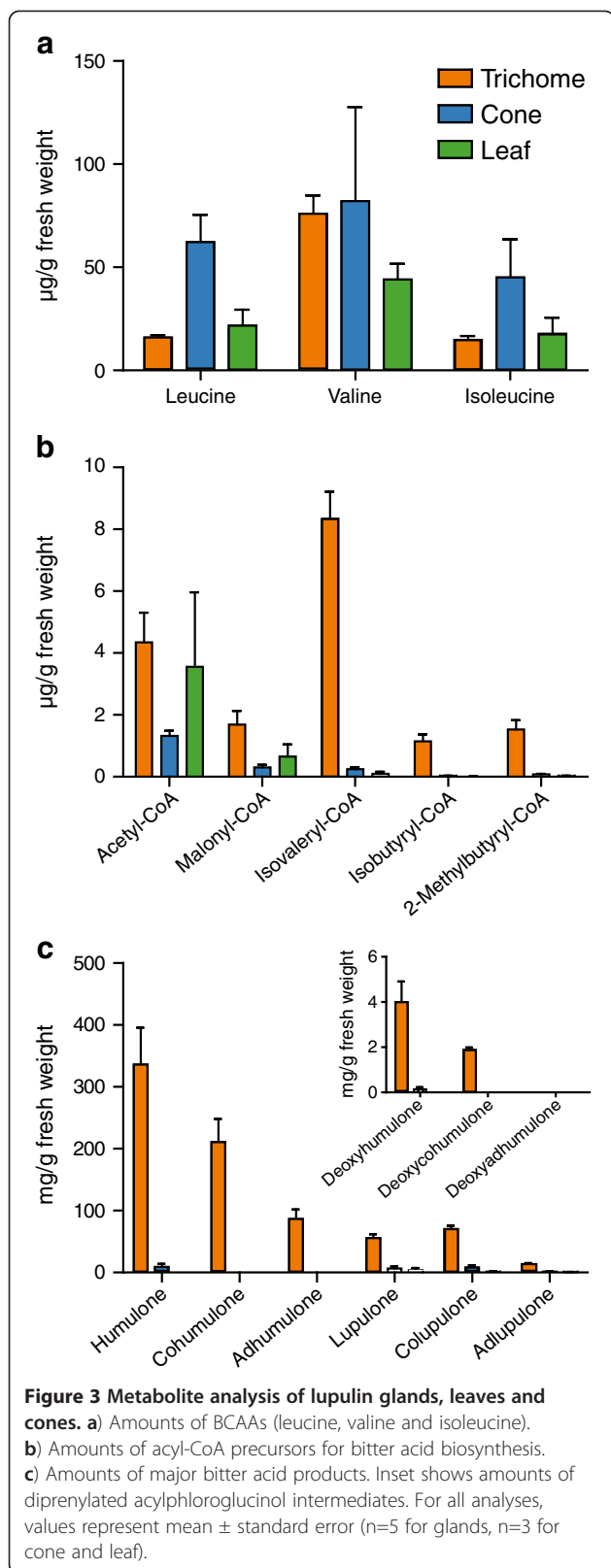
available in GenBank [*HIBCAT1*, JQ063073; *HIBCAT2*, JQ063074]. The mature forms (i.e. lacking predicted transit peptides) of HIBCAT1 and HIBCAT2 were expressed as his-tagged recombinant proteins in *E. coli* and purified. Enzyme assays were performed in both the forward (anabolic, e.g. leucine to 2-ketoisocaproate) and reverse (catabolic, e.g. 2-ketoisocaproate to leucine) directions to test catalytic activities and to determine kinetic parameters (Table 2). Both enzymes functioned as BCATs and showed reversibility, as would be expected based on their ping pong reaction mechanism [27]. Unlike previous work in tomato, where the BCATs exhibited kinetic preferences corresponding to anabolic or catabolic functions [15], the hop BCATs had similar  $K_m$  values in either direction.

#### Subcellular localization of hop BCATs

HIBCAT1, HIBCAT2 and HIBCAT3 contain putative N-terminal signal peptides while HIBCAT4 does not. TargetP [28] analysis predicted that HIBCAT1 and HIBCAT3 are localized to the mitochondria, and HIBCAT2 to the plastids. We tested the subcellular localization of the two BCATs with gland-specific expression, HIBCAT1 and HIBCAT2, using transient expression in *Nicotiana benthamiana* leaves. Hop leaves were not used for this experiment because they proved difficult to infiltrate with *Agrobacterium* solution. HIBCAT1-GFP constructs with signal peptide (Figure 4) or signal peptide with mature protein (Additional file 5: Figure S2) co-localized with the mitochondrial MTRK marker [29]. The full-length HIBCAT2 failed to localize to a recognizable compartment, however its signal peptide fused to GFP colocalized with the plastid PTRK marker (Figure 4, Additional file 5: Figure S2). HIBCAT1 and HIBCAT2 constructs without signal peptide accumulated in the cytosol (Additional file 5: Figure S2).

#### Phylogenetic analysis of hop BCATs

A phylogenetic tree was constructed using BCAT protein sequences from hop, *Arabidopsis*, tomato, *Cannabis sativa* and other plants obtained using the Phytozome database [22]. We observed a separation of dicot BCATs into two clades, which could be classified as biosynthetic/plastidial and catabolic/mitochondrial based on the presence of experimentally-localized proteins in each group (Figure 5). HIBCAT1 and 3 grouped into the mitochondrial BCAT clade and HIBCAT2 grouped into the plastidial clade, which agrees with their predicted localizations by TargetP and our analysis of intracellular localization (Figure 4). HIBCAT4 was grouped with the other BCATs which do not appear to localize to either the plastid or mitochondrion. Although the monocot BCATs clustered into two distinct groups that were



**Table 2 Kinetic parameters for recombinant HIBCAT1 and HIBCAT2 enzymes in anabolic (forward) and catabolic (reverse) directions**

HIBCAT	$K_m$ (µM) <sup>a</sup>	$V_{max}$ (µmol/mg/min) <sup>a</sup>
<b>HIBCAT1</b>		
<b>Anabolic:</b>		
Glutamate	720 ± 55.59	5.59 ± 0.13
Ketoisocaproate	35 ± 3.27	4.50 ± 0.04
Ketoisomethylvalerate	99 ± 10.75	2.95 ± 0.09
Ketoisovalerate	99 ± 6.53	4.23 ± 0.08
<b>Catabolic:</b>		
2-Oxoglutarate	30 ± 5.39	4.50 ± 0.16
Leucine	40 ± 9.20	4.84 ± 0.23
Isoleucine	230 ± 26.94	19.40 ± 0.85
Valine	340 ± 90.94	12.60 ± 0.96
<b>HIBCAT2</b>		
<b>Anabolic:</b>		
Glutamate	690 ± 88.51	135.90 ± 6.03
Ketoisocaproate	120 ± 11.43	119.20 ± 2.97
Ketoisomethylvalerate	160 ± 11.63	84.68 ± 1.70
Ketoisovalerate	200 ± 11.72	96.38 ± 1.45
<b>Catabolic:</b>		
2-Oxoglutarate	210 ± 58.86	36.92 ± 2.68
Leucine	280 ± 69.56	43.43 ± 2.55
Isoleucine	250 ± 41.07	88.08 ± 3.83
Valine	130 ± 27.74	23.29 ± 1.16

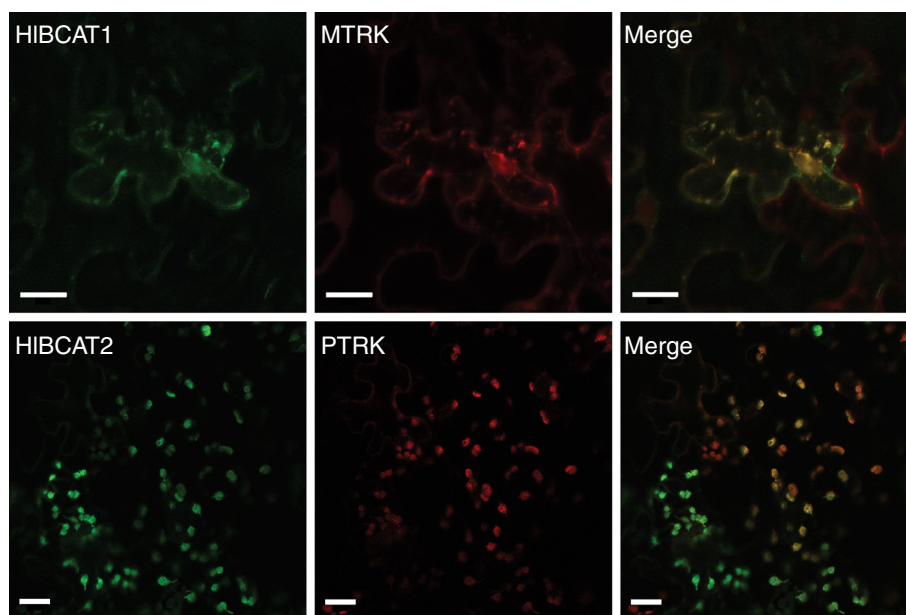
<sup>a</sup>mean ± standard error (n=3).

positioned between the dicot clades, there is no experimental localization data for monocot BCATs at this time.

## Discussion

### Illumina sequencing provides new genomic resources for hop

Although the transcriptome of lupulin glands has been extensively studied using ESTs obtained through Sanger sequencing [5,8,30] and several 454 pyrosequencing datasets are available in NCBI, an in-depth comparison of the lupulin gland transcriptome with green tissues (i.e. leaves, cones) has not been reported. Schillmiller *et al.* [31] previously used 454 pyrosequencing to compare gene expression in tomato trichomes, leaves and stems, and RNA-seq analysis was used to identify terpene synthases in tomato trichomes [32]. Microarrays have also been used to compare the trichome and stem transcriptome of alfalfa [33]. The Illumina sequences reported here significantly expand the genomic resources available for hop, in particular leaves and cones for which limited sequences are available. The high number of contigs in the de novo transcriptome assembly probably relates to the presence of “broken contigs”, as evidenced



**Figure 4** Subcellular co-localization of BCAT signal peptides fused to GFP with organelle specific markers using transient expression in *Nicotiana benthamiana* leaves. The HIBCAT1 signal peptide fused to GFP colocalized with the mitochondrial MTRK marker (upper panel) while the HIBCAT2 signal peptide-GFP fusion colocalized with the plastidial PTRK marker (lower panel). Plasmid constructs in *Agrobacterium* were agro-infiltrated into *N. benthamiana* leaves and fluorescent protein expression visualized using confocal microscopy. Scale bars represent 20  $\mu\text{m}$ .

by the relatively short contig length (745 bp). This may have resulted from our use of combined RNA-seq data from several hop cultivars, the lack of sufficient sequencing depth and the presence of introns from alternative splicing events in some transcripts. In addition to the further development of transcriptome resources for hop, an important next step is to link the transcriptome with a whole genome sequence and genetic map. However, the estimated 2.8 Gbp size of the *Humulus lupulus* genome [34] and the highly heterozygous nature of this dioecious species present obvious challenges for genome assembly.

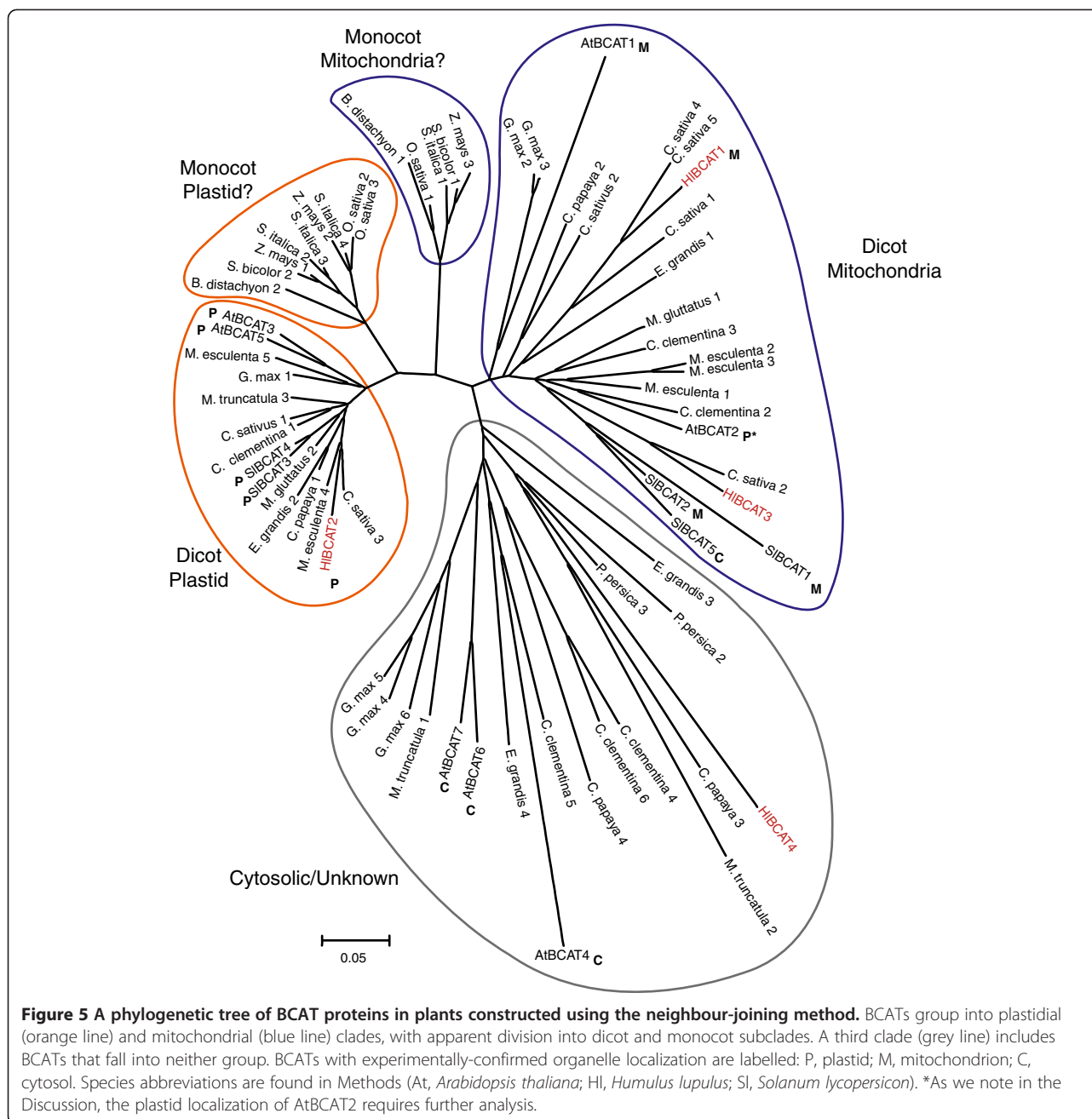
#### Lupulin glands synthesize precursor metabolites for bitter acid production

The expression pattern for genes related to photosynthesis (e.g. Calvin cycle), and the absence of bitter acids from the cone minus gland samples (Figure 3c), indicates that separation of the non-photosynthetic glands from photosynthetic cones was largely complete. Analysis of gene expression for primary metabolic pathways (TCA cycle and glycolysis) showed the reliance of lupulin glands on surrounding green tissues for the supply of sugars. The genes involved in the Calvin cycle were more abundant in the leaves and cones compared to lupulin glands (Figure 2a). The only exceptions to this pattern were genes encoding enzymes that function in both the Calvin cycle and glycolysis. The TCA cycle had similar expression levels across the leaf, cone and

lupulin gland samples which can be attributed to its role in supplying intermediates for many pathways. Among the glycolysis transcripts three genes (*fructokinase*, *neutral/alkaline invertase* and *sucrose synthase*) stood out for their high expression in the lupulin glands compared to leaves. All three are involved in sucrose mobilization, underscoring the exogenous supply of sugars to the lupulin gland. This is in agreement with the previous suggestion that sucrose would be an ideal energy source and precursor for secondary metabolism in glandular trichomes [6].

The high amounts of bitter acids produced by elite hop cultivars represents a large carbon commitment within the lupulin glands. Given that sugars are imported, we analyzed the RNA-seq dataset to determine if other precursors (e.g. BCAAs and prenyl diphosphates) are imported or synthesized in situ. Both the MEP and BCAA pathways were more abundantly represented in the lupulin glands (Figure 2b), a finding that provides evidence for the de novo biosynthesis of BCAAs and DMAPP in these cells. BCAA biosynthesis in lupulin glands is logical given the demand for the acyl side-chains of the BCAAs but not the amino group, and the need to prevent the accumulation of excess nitrogen. This finding is in agreement with Schillmiller *et al.* [6] who noted that while phenylpropanoid-derived metabolites are synthesized in trichomes (e.g. phenylpropenes in basil) the import of phenylalanine into glandular trichomes was minimal.





Transcripts corresponding to the enzymes of the polyketide portion of the bitter acid pathway (e.g. *VPS*) were among the most abundant in the glandular trichome dataset (Figure 2b). The presence of two highly expressed aromatic prenyltransferases is notable since HIPT1 has only been shown to catalyze the first prenylation step in bitter acid biosynthesis [21]. The  $\alpha$ - and  $\beta$ -acids require two and three prenyl groups, respectively, and the second prenyltransferase enzyme (HIPT2, 45% amino acid identity to HIPT1) may catalyze these steps.

**Lupulin gland-specific expression of transcription factors**  
 RNA-seq analysis also revealed a number of lupulin gland-specific transcription factors which may be involved in regulating bitter acid biosynthesis or other metabolic or developmental processes in glandular trichomes (Table 2). Matouček and co-workers have identified several transcription factors that may control flavonoid biosynthesis in hop through the regulation of chalcone synthase (CHS) (e.g. HlMYB1, HlMYB2, HlMYB3, HlMYB7, HlbHHLH2 and HIWDR1 [35–37]). Although all of these transcription factors could be

identified in the hop transcriptome reported here, only transcripts of *HlMYB7* and *HlbHLH2* were more abundant in lupulin glands compared to leaves (Additional file 3: Table S3). This expression pattern makes sense since *CHS*, which plays a role in both the gland-specific production of xanthohumol and in the formation of flavonoids, is expressed throughout the plant. Indeed our RNA-seq analysis shows that *CHS\_H1* is expressed in lupulin glands and leaves (Additional file 2: Table S2), with the three-fold higher expression levels in glands likely reflecting the production of xanthohumol. Two MYB transcription factors, *CAN883* and *CAN778*, have been shown to be highly expressed in cannabis glandular trichomes and implicated in regulating trichome-specific metabolic processes [38]. Both have close homologs to hop transcription factors that are over ten-fold more abundant in the lupulin glands (Table 2). *CAN738*, which has 70% amino acid identity to the hop MYB protein encoded by *comp63432\_c0\_seq1*, was proposed to be involved in regulating cannabinoid formation. Since *Cannabis* and *Humulus* are closely related genera with glandular trichome-localized polyketide pathways, *comp63432\_c0\_seq1* may be a candidate for regulating the bitter acid pathway.

#### **Lupulin glands contain both mitochondrial and plastidial BCATs**

The biosynthesis and subsequent degradation of the BCAAs requires both metabolic and catabolic BCATs. The lupulin gland datasets contained two BCATs, *HIBCAT1* and *HIBCAT2*, which were abundant in the glandular trichomes. Based on RNA-seq and qRT-PCR analysis (Figure 2b, Additional file 4: Figure S1) *HIBCAT1* is exclusively expressed in the lupulin glands. Transient expression of the *HIBCAT1*-GFP fusion construct confirmed its mitochondrial targeting (Figure 4). *HIBCAT2* with N-terminal signal peptide failed to localize to the plastid and appeared to form aggregates in the *N. benthamiana* epidermal cells (Additional file 5: Figure S2), however, the predicted *HIBCAT2* signal peptide fused to GFP was targeted to the plastid (Figure 4). These localizations were further supported through the construction of a phylogenetic tree comparing BCATs from a variety of plant species (Figure 5). Problems with the experimental localization of hop plastid proteins have been previously reported for a GPP synthase small subunit (*GPP-ssu*) identified in lupulin glands [8].

#### **Kinetic analysis suggests lack of specialization for hop BCATs**

Kinetic analysis has provided equivocal evidence for the catalytic specialization of BCATs for BCAA biosynthesis or catabolism. In tomato, biosynthetic BCATs (*SIBCAT3*

and 4) had  $K_m$  values with the keto acids that were an order-of-magnitude less than the catabolic enzymes (*SIBCAT1* and 2), indicating specialization of the former for forward (biosynthetic) reactions [15]. In contrast, Arabidopsis *AtBCAT1*, which has been shown to be a catabolic enzyme, had  $K_m$  values for the keto acids (0.036-0.8 mM) that were much lower than for BCAA substrates (0.4-6 mM), suggesting a preference for the biosynthetic reaction. Indeed, the  $K_m$  values for *AtBCAT1* with the keto acids were lower than those reported with the plastidial, biosynthetic enzyme *AtBCAT3* [39]. The kinetic properties of the hop BCATs appear to be more similar to the Arabidopsis enzymes. The catabolic enzyme *HIBCAT1* exhibited lower  $K_m$  values for keto acids, and therefore an apparent preference for the forward (biosynthetic) reaction, compared with than *HIBCAT2* (Table 2). It is noteworthy that *HIBCAT1* showed a lower  $K_m$  for leucine compared with valine and isoleucine which may reflect this enzyme's preferred role in synthesizing isovaleryl-CoA for humulone biosynthesis. The lower  $V_{max}$  values for the *HIBCAT1* may also explain why it is more highly expressed compared to *HIBCAT2* in the lupulin gland tissue (Figure 2).

#### **Phylogenetic analysis indicates separation of biosynthetic and catabolic BCATs**

To further analyze the function of the hop BCATs, we constructed a phylogenetic tree of diverse BCATs including well characterized enzymes from Arabidopsis and tomato. This analysis was successful in separating biosynthetic (plastidial) and catabolic (mitochondrial) BCATs into separate clades. BCATs are also reported to accumulate in alternative cellular compartments such as the cytosol and vacuole for an as of yet unknown function [15,16] and these genes likely make up the third group that includes several cytosolic enzymes and *HIBCAT4* (Figure 5). An anomalous grouping was observed for *AtBCAT2*, which has been shown to be plastid localized [6] but which was placed in the mitochondrial clade. A previous study found *AtBCAT2* to be the only Arabidopsis gene to be co-expressed with other genes involved in BCAA catabolism [40], a role one would expect to be occupied by a mitochondrial BCAT. In addition, TargetP analysis of *AtBCAT2* predicts mitochondrial rather than plastidial localization.

#### **Conclusions**

This study provides an in-depth analysis of the transcriptional activity of primary and specialized metabolism leading to hop bitter acids. RNA-seq analysis showed that both the MEP and BCAA pathways were highly expressed in lupulin glands together with genes required for remobilization of carbon from sucrose.

Branched-chain acyl-CoAs and bitter acids are present in higher levels in isolated lupulin glands compared with cones (with glands removed) and leaves. A plastidial BCAT enzyme involved in BCAA biosynthesis and a mitochondrial BCAT catalyzing BCAA degradation both show lupulin gland-specific expression. Together these findings indicate that the glandular trichomes are the site of biosynthesis for the precursors for bitter acid biosynthesis. The deep transcriptome sequencing reported here significantly expands the genomic resources available for hop. This new information, and our analysis of BCAA metabolism, will be useful for the further elucidation of enzymes and regulatory proteins involved in bitter acid production and for breeding hop cultivars with increased or modified bitter acid content.

## Methods

### Hop transcriptome assembly and RNA-seq analysis

RNA was isolated from lupulin glands sampled from hop cultivars 'Taurus', 'Nugget' and 'Magnum' (collected in Mainburg, Germany in August 2005), from 'Taurus' (collected in Saskatoon, SK in September 2011) and 'Apollo' (collected in Yakima, WA in August 2009). RNA was isolated from leaves and cones-minus-glands from 'Apollo' (collected Yakima 2009) and 'Taurus' (collected Mainburg 2005 and Saskatoon 2011). Glands were separated from cones by stirring in liquid nitrogen followed by filtration through a 1 mm screen to separate cone material. Total RNA was isolated using a CTAB method and an RNeasy kit with on-column DNaseI treatment (Qiagen) [41]. RNA was prepared for sequencing by the National Research Council Canada DNA Technologies Unit using the Illumina TrueSeq RNA sample preparation platform v.2 with multiplex labeling and subjected to paired-end sequencing using an Illumina HiSeq 2000 with a read length of 101 nt. Trimmomatic 0.15 [42] was used to remove adaptor sequences and trim bases with quality lower than 20 (Phred 33 quality scores). The transcriptome assembly was performed with Trinity [23] and refined using cd-hit-EST [24] to combine identical reads. CLC bio v.5 was used for RNA-seq analysis and statistical analysis via the Baggerly's test [43]. Annotation of the transcriptome was performed using Blast2GO [44].

The Illumina data for all samples have been deposited in NCBI under Bioproject PRJNA175602 with the transcriptome available via accession number GAAW00000000. RPKM values for all transcripts is found in Additional file 6: Table S4.

### qRT-PCR analysis

cDNA was synthesized from 0.5 µg total RNA from lupulin glands of 'Taurus' and 'Nugget' sampled in

Mainburg, Germany and 'Taurus' sampled in Saskatoon, Canada. RNA was isolated from leaves from 'Taurus' sampled in Mainburg, Germany in 2005, Saskatoon, Canada and 'Apollo' from Yakima, USA using the QuantiTect reverse transcription kit (Qiagen). Primers (Additional file 7: Table S5) were designed to amplify products of 101–110 bp. PCR reactions were performed with a StepOnePlus real-time PCR system (Applied Biosystems) using the QuantiFast SYBR Green qPCR kit (Qiagen).  $\beta$ -tubulin was used a reference gene. StepOnePlus software was used to calculate  $\Delta\Delta C_T$  values [45].

### Metabolite analysis

For BCAA analysis, samples were weighed and ground in 5 vol (based on 1 mL per mg fresh weight of plant material) of 0.1 N HCl, centrifuged at 20,000 g for 10 min and derivatized using the AccQ-Tag reagent (Waters). Derivatized amino acids were analyzed with a Waters 2695 separation module using a Waters Symmetry C18 column (5 µm, 2.1 × 100 mm) and a Waters 2475 fluorescence detector as described by Cohen and Michaud [46]. Amino acids were quantified using standard curves.

For CoA analysis, samples were ground in liquid nitrogen and extracted twice with 10 vol of 5% TFA with internal standards ( $^{13}C_3$  malonyl-CoA, butenoyl-CoA and benzoyl-CoA). Samples were evaporated, resuspended in 900 µL of acetonitrile: isopropyl alcohol: 100 mM  $KH_2PO_4$ : aqueous acetic acid (9:3:4:4, v/v/v/v) and solubilized by vortexing and sonication. Oasis WAX weak anion exchange cartridges (Waters) were primed with methanol, equilibrated with acetonitrile: isopropyl alcohol: water: acetic acid (9:3:4:4, v/v/v/v), the samples loaded and washed with the equilibration mixture. Samples were eluted with 250 mM ammonium hydroxide in 80% methanol, dried and stored at  $-20^\circ C$ . The residue was reconstituted in 100 µL of water: acetonitrile (95:5, v/v) containing 5 mM triethylamine (TEA) and 3 mM acetic acid, vortexed, sonicated and centrifuged. The injection volume was 10 µL. CoAs were separated on a Waters Acquity UPLC system with a Waters BEH C18 column (1.8 µm, 2.1 × 100 mm) with a Vanguard guard column. Mass detection used a Quattro Ultima (Waters) in positive ion, multiple reaction monitoring (MRM) mode. Solvents were A=5 mM TEA and 3 mM acetic acid in water, and B=5 mM TEA and 0.3 mM acetic acid in acetonitrile: water (95:5, v/v). A flow rate of 200 µL/min was used with an initial solvent composition of 98:2 A:B changing to 87:13 A:B over 19 min. Detection of the acyl-CoAs using MRM involved neutral loss of 507 Da from each protonated molecular ion [47]. Quantification by standard curves used MassLynx 4.1 (Waters). Since there was no standard for

2-methylbutyryl-CoA, quantification was based on the calibration curve constructed for n-valeryl-CoA.

For polyketide analysis, samples were ground in 10 vol of 80% MeOH, filtered through a 0.45  $\mu$ m Spin-X filter (Corning) and separated on a Waters 2695 separation module using a Waters Symmetry C18 column (5  $\mu$ m, 2.1  $\times$  100 mm). A flow rate of 0.3 mL/min was used with an initial solvent composition of acetonitrile: water (46:54, v/v) with 0.05% formic acid changing to acetonitrile: water (95.5:4.5, v/v) with 0.05% formic acid over 25 min. Metabolites were detected using a Waters photodiode array detector at 237 nm and 3100 Mass Detector in SIR with positive electrospray and a cone voltage of 30 V. Metabolite levels were determined with standard curves of deoxyhumulone, humulone and lupulone.

#### BCAT characterization

Oligonucleotides used in cloning procedures are found in Additional file 7: Table S5. Signal peptides were predicted using Mitoprot [48] and ChloroP [49]. The open-reading frames of HIBCAT1 and HIBCAT2 lacking signal peptides were amplified using *Pfu* Ultra II polymerase (Agilent), the amplification products cloned into pCR8/GW/TOPO (Invitrogen) and recombined into pHIS8/GW using LR Clonase (Invitrogen). HIBCAT1 and HIBCAT2 in pHIS8/GW were transformed into *E. coli* BL21(DE3) Rosetta2 pLysS cells (Novagen). Cells were grown in Luria-Bertani media containing 50  $\mu$ g/L kanamycin and 34  $\mu$ g/L chloramphenicol to an OD<sub>600</sub> of 0.6 at 37°C, cooled to 18°C, induced with 0.3 mM IPTG and incubated for 4 h at 18°C. Bacterial cultures were pelleted and frozen at -80°C. Pellets were thawed, resuspended in 20 mM Tris HCl (pH 8.0), 300 mM NaCl, 100  $\mu$ M pyridoxal-5-phosphate (PLP), 10% glycerol, 5 mM imidazole, 5 mM  $\beta$ -mercaptoethanol, 0.2% triton X-100, 0.1% CHAPS, 0.2% sarkosyl, 0.5 mM PMSE, 1  $\mu$ g/ml pepstatin A and 1  $\mu$ g/ml leupeptin, and lysed by sonication. Recombinant proteins were purified from the cell lysate using Talon resin (Clontech). The resin-bound protein was washed with 20 mM Tris HCl (pH 8.0), 300 mM NaCl, 10% glycerol, 10 mM imidazole, 5 mM  $\beta$ -mercaptoethanol and eluted with the same buffer composition containing 150 mM imidazole. Samples were desalted and concentrated using Amicon Ultra filters (Millipore) and stored at -80°C. Protein concentration was determined using the Bradford assay [50] and protein purity was assessed at >95% via SDS-PAGE gel analysis with Coomassie brilliant blue stain.

Anabolic reactions were linked to glutamate dehydrogenase to allow monitoring with NADH turnover as described by Prohl *et al.* [51]. Anabolic assays contained 66  $\mu$ g/ml HIBCAT1 or 12  $\mu$ g/ml HIBCAT2, 250 mM Tris HCl (pH 8.5), 50  $\mu$ M PLP, 1.5 mM dithiothreitol, 0.2 mM

EDTA, 100 mM ammonium chloride, 0.2 mM NADH, 3 units/mL glutamate dehydrogenase (Sigma), 2 mM glutamate for branched-chain keto-acid kinetics and 0.5 mM ketoisovalerate for glutamate kinetics. Assays were run in triplicate at 22°C and observed in real time at 340 nm with an Ultraspec 3100 pro spectrophotometer (Biochrom). Time point assays were employed for catabolic reactions. Catabolic assays contained 5.4  $\mu$ g HIBCAT1 or 1.2  $\mu$ g HIBCAT2, 20 mM Tris HCl (pH 8.5), 50  $\mu$ M PLP, 1.5 mM dithiothreitol, 0.2 mM EDTA, 2 mM 2-oxoglutarate for BCAA kinetics and 2 mM leucine for 2-oxoglutarate kinetics. Assays were run in triplicate at 22°C for 2 min and halted with the addition of HCl. Samples were centrifuged at 20,000 *g* for 10 min, derivatized using the AccQ-Tag reagent (Waters) and analyzed as described above.

#### Analysis of subcellular localization

Three constructs in pCR8/GW/TOPO were prepared for each gene: full-length including predicted signal peptide, the mature protein lacking the signal peptide, and the signal peptide only. Each was recombined into pEarleyGate103 to generate C-terminal GFP fusion constructs under the control of the 35S promoter [52]. Plasmid constructs of HIBCATs, and mitochondrial (MTRK) or plastid (PTRK) markers [29], were transformed into *Agrobacterium* C58 and infiltrated into *Nicotiana benthamiana* leaves as previously described [53]. Plants were grown for 2–4 days and the leaves directly observed with a Carl Zeiss Axiovert 200M/ConfoCor 2/LSM 510 confocal laser scanning microscope. GFP was observed with an argon laser with excitation at 488 nm and emission at 505–530 nm. mCherry was observed with a helium laser with excitation at 543 nm and emission at 560–615 nm.

#### Phylogenetic tree

BCAT protein sequences were obtained from the Phytozome database [22] for *Manihot esculenta*, *Populus trichocarpa*, *Cucumis sativus*, *Prunus persica*, *Medicago truncatula*, *Glycine max*, *Carica papaya*, *Citrus clementina*, *Eucalyptus grandis*, *Mimulus guttatus*, *Sorghum bicolor*, *Zea mays*, *Setaria italica*, *Oryza sativa*, and *Brachypodium distachyon*. Previously characterized Arabidopsis and tomato proteins were included along with those from the *Cannabis sativa* genome [54]. All sequences were aligned using ClustalW [55] with variable N-terminal regions removed when present. Mega5 [56] was used to construct a neighbour-joining tree with 10,000 bootstrap replicates.

#### Additional files

**Additional file 1: Table S1.** Total Illumina sequencing reads for each RNA sample and the number of reads that mapped to the reference transcriptome.



**Additional file 2: Table S2.** RNA-seq analysis of metabolic genes compared between lupulin gland and leaf, lupulin gland and cone (minus glands), and leaf and cone (minus glands). Data is represented in reads per kb exon model per million mapped reads (RPKM). Note that each tissue comparison is located on a separate Excel sheet.

**Additional file 3: Table S3.** Transcript abundance of transcription factors compared between lupulin gland and leaf, lupulin gland and cone (minus glands) and leaf and cone (minus glands). Data is represented in RPKM.

**Additional file 4: Figure S1.** Relative abundance of select metabolic transcripts in lupulin gland and leaf samples as determined by qRT-PCR. Data is presented using relative quantitation (RQ) based on  $\Delta\Delta C_T$  values. Samples are normalized using  $\beta$ -tubulin. Data represents the mean and standard error,  $n=3$ . Transcript abundance between panels a, b and c is not directly comparable in this figure. **a)** qRT-PCR analysis of *HIBCAT1* and *HIBCAT2*. A single gland sample was used as reference. **b)** qRT-PCR analysis of genes encoding enzymes of the BCKDH complex (ketoacid dehydrogenase E1 $\alpha$  and E1 $\beta$ , dihydrolipoyl acyltransferase (E2), and dihydrolipoyl dehydrogenase (E3)). A single leaf sample was used as reference. **c)** qRT-PCR analysis of *VPS* and *HIPT1*. A single leaf sample was used as reference.

**Additional file 5: Figure S2.** Subcellular co-localization of GFP fusion proteins of *HIBCAT1* and *HIBCAT2*, with and without N-terminal signal peptides, compared with organelle specific marker proteins using transient expression in *Nicotiana benthamiana* leaves. MTRK and PTRK are mitochondrial and plastidial makers. Scale bars represent 20  $\mu$ m.

**Additional file 6: Table S4.** RPKM values for lupulin glands, cones (minus glands) and leaves.

**Additional file 7: Table S5.** Sequences of oligonucleotides used.

#### Abbreviations

BCAA: Branched-chain amino acid; BCAT: Branched-chain aminotransferase; BCKDH: Branched-chain ketoacid dehydrogenase complex; CHAPS: 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate; CoA: Coenzyme A; DMAPP: Dimethylallyl diphosphate; EDTA: Ethylenediaminetetraacetic acid; GFP: Green fluorescent protein; IPTG: Isopropyl  $\beta$ -D-1-thiogalactopyranoside; MEP: Methyl-D-erythritol 4-phosphate; qRT-PCR: Quantitative reverse transcription polymerase chain reaction; PLP: Pyridoxal-5-phosphate; PMSF: Phenylmethylsulfonyl fluoride; RPKM: Reads per kb exon model per million mapped reads; sp: Signal peptide; VPS: Valerophenone synthase.

#### Competing interests

Partial funding for this work was provided by Hopsteiner Inc, a major hop processing company.

#### Authors' contributions

SMC isolated RNA for RNA-seq, assembled the hop transcriptome, analyzed transcriptome data, performed qRT-PCR analysis, characterized the BCATs, measured BCAA and bitter acid levels and wrote the manuscript. VV cloned and expressed BCATs. SA and RP quantified acyl-CoAs. JEP conceived of the study, and participated in its design and coordination and wrote the manuscript. All authors read and approved the final manuscript.

#### Acknowledgements

We thank P. Matthews, N. Pitra and J. Nagel for hop sample collection, D. Klassen and D.L. Fancy for Illumina sequencing, D. Cram and C. Matsalla for bioinformatics assistance, and G. Liu for technical support with confocal laser microscopy. This work was funded by Hopsteiner Inc, the National Research Council of Canada, and Genome Prairie/Province of Saskatchewan via the PhytoMetaSyn project.

Received: 1 October 2012 Accepted: 12 December 2012

Published: 24 January 2013

#### References

1. Cleemput MV, Heyerick A, Libert C, Swerts K, Philippé J, Keukeleire DD, Haegeman G, Bosscher KD: **Hop bitter acids efficiently block inflammation independent of GR $\alpha$ , PPAR $\beta$ , or PPAR $\gamma$ .** *Mol Nutr Food Res* 2009, **53**:1143–1155.

2. Verzele M, Steenbeke G, Verhagen LC, Strating J: **Preparative liquid chromatography of hop and beer bitter acids.** *J Chromatogr A* 1989, **484**:361–368.
3. Stevens JF, Page JE: **Xanthohumol and related prenylflavonoids from hops and beer: to your good health!** *Phytochemistry* 2004, **65**:1317–1330.
4. Kavalier AR, Litt A, Ma C, Pitra NJ, Coles MC, Kennelly EJ, Matthews PD: **Phytochemical and morphological characterization of hop (*Humulus lupulus* L.) cones over five developmental stages using high performance liquid chromatography coupled to time-of-flight mass spectrometry, ultrahigh performance liquid chromatography photodiode.** *J Agric Food Chem* 2011, **59**:4783–4793.
5. Nagel J, Culley L, Lu Y, Liu E, Matthews PD, Stevens JF, Page JE: **EST analysis of hop glandular trichomes identifies an O-methyltransferase that catalyzes the biosynthesis of xanthohumol.** *Plant Cell* 2008, **20**:186–200.
6. Schillmiller AL, Last RL, Pichersky E: **Harnessing plant trichome biochemistry for the production of useful compounds.** *Plant J* 2008, **54**:702–711.
7. Tissier A: **Glandular trichomes: what comes after expressed sequence tags?** *Plant J* 2012, **70**:51–68.
8. Wang G, Tian L, Aziz N, Broun P, Dai X, He J, King A, Zhao PX, Dixon RA: **Terpene biosynthesis in glandular trichomes of hop.** *Plant Physiol* 2008, **148**:1254–1266.
9. Stout J, Boubakir Z, Ambrose S, Purves R, Page J: **The hexanoyl-CoA precursor for cannabinoid biosynthesis is formed by an acyl-activating enzyme in trichomes of *Cannabis sativa*.** *Plant J* 2012, **71**:353–365.
10. Gagne SJ, Stout JM, Liu E, Boubakir Z, Clark SM, Page JE: **A cannabinoid pathway cyclase enzyme reveals a new catalytic route to plant polyketides.** *Proc Natl Acad Sci USA* 2012, **109**:12811–12816.
11. Slocombe SP, Schauvinhold I, McQuinn RP, Besser K, Welsby NA, Harper A, Aziz N, Li Y, Larson TR, Giovannoni J, Dixon RA, Broun P: **Transcriptomic and reverse genetic analyses of branched-chain fatty acid and acyl sugar production in *Solanum pennellii* and *Nicotiana benthamiana*.** *Plant Physiol* 2008, **148**:1830–1846.
12. Walters DS, Steffens JC: **Branched-chain amino acid metabolism in the biosynthesis of *Lycopersicon pennellii* glucose esters.** *Plant Physiol* 1990, **93**:1544–1551.
13. Binder S, Knill T, Schuster J: **Branched-chain amino acid metabolism in higher plants.** *Physiol Plant* 2007, **129**:68–78.
14. Diebold R, Schuster J, Däschner K, Binder S: **The branched-chain amino acid transaminase gene family in Arabidopsis encodes plastid and mitochondrial proteins.** *Plant Physiol* 2002, **129**:540–550.
15. Maloney GS, Kochevenko A, Tieman D, Tohge T, Krieger U, Zamir D, Taylor MG, Fernie AR, Klee HJ: **Characterization of the branched-chain amino acid aminotransferase enzyme family in tomato.** *Plant Physiol* 2010, **153**:925–936.
16. Kochevenko A, Klee HJ, Fernie AR, Araújo WL: **Molecular identification of a further branched-chain aminotransferase 7 (BCAT7) in tomato plants.** *J Plant Physiol* 2012, **169**:437–443.
17. Mooney BP, Miernyk JA, Randall DD: **The complex fate of  $\alpha$ -ketoacids.** *Ann Rev Plant Biol* 2002, **53**:357–375.
18. Lutziger I, Oliver DJ: **Characterization of two cDNAs encoding mitochondrial lipoamide dehydrogenase from Arabidopsis.** *Plant Physiol* 2001, **127**:615–623.
19. Goese M, Kammhuber K, Bacher A, Zenk MH, Eisenreich W: **Biosynthesis of bitter acids in hops. A  $^{13}C$ -NMR and  $^2H$ -NMR study on the building blocks of humulone.** *Eur J Biochem* 1999, **263**:447–454.
20. Paniego NB, Zuurbier KW, Fung SY, van der Heijden R, Scheffer JJ, Verpoorte R: **Phlorisovalerophenone synthase, a novel polyketide synthase from hop (*Humulus lupulus* L.) cones.** *Eur J Biochem* 1999, **262**:612–616.
21. Tsurumaru Y, Sasaki K, Miyawaki T, Uto Y, Momma T, Umemoto N, Momose M, Yazaki K: **HIPT-1, a membrane-bound prenyltransferase responsible for the biosynthesis of bitter acids in hops.** *Biochem Biophys Res Comm* 2012, **417**:393–398.
22. Zuurbier KWM, Paniego NB, Scheffer JJC, Verpoorte R: **Conversion of deoxyhumulone into the hop  $\alpha$ -acid humulone.** *Phytochemistry* 1997, **44**:1047–1053.
23. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Maudeli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A: **Full-length transcriptome assembly from RNA-Seq data without a reference genome.** *Nat Biotechnol* 2011, **29**:644–652.

24. Li W, Godzik A: Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 2006, **22**:1658–1659.
25. Nikolau BJ, Ohlrogge JB, Wurtele ES: Plant biotin-containing carboxylases. *Arch Biochem Biophys* 2003, **414**:211–222.
26. Fatland BL, Ke J, Anderson MD, Mentzen WI, Cui LW, Allred CC, Johnston JL, Nikolau BJ, Wurtele ES: Molecular characterization of a heteromeric ATP-citrate lyase that generates cytosolic acetyl-coenzyme A in Arabidopsis. *Plant Physiol* 2002, **130**:740–756.
27. Martinez-Carrion M, Hubert E, Iriarte A, Mattingly JZS: Mechanism of aminotransferase action. In *Transaminase*. Edited by Christen P. MD New York: Wiley; 1985:307–317.
28. Emanuelsson O, Nielsen H, Brunak S, von Heijne G: Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* 2000, **300**:1005–1016.
29. Nelson BK, Cai X, Nebenführ A: A multicolored set of *in vivo* organelle markers for co-localization studies in Arabidopsis and other plants. *Plant J* 2007, **51**:1126–1136.
30. Wang G, Dixon RA: Heterodimeric geranyl(geranyl)diphosphate synthase from hop (*Humulus lupulus*) and the evolution of monoterpene biosynthesis. *Proc Natl Acad Sci USA* 2009, **106**:9914–9919.
31. Schillmiller AL, Miner DP, Larson M, McDowell E, Gang DR, Wilkerson C, Last RL: Studies of a biochemical factory: tomato trichome deep expressed sequence tag sequencing and proteomics. *Plant Physiol* 2010, **153**:1212–1223.
32. Bleeker P, Spyropoulou E: RNA-seq discovery, functional characterization, and comparison of sesquiterpene synthases from *Solanum lycopersicum* and *Solanum habrochaites* trichomes. *Plant Mol Biol* 2011, **77**:323–336.
33. Aziz N, Paiva NL, May GD, Dixon RA: Transcriptome analysis of alfalfa glandular trichomes. *Planta* 2005, **221**:28–38.
34. Zonneveld BJM, Leitch IJ, Bennett MD: First nuclear DNA amounts in more than 300 angiosperms. *Ann Bot* 2005, **96**:229–244.
35. Matoušek J, Kocábek T, Patzak J, Füssy Z, Procházková J, Heyerick A: Combinatorial analysis of lupulin gland transcription factors from R2R3Myb, bHLH and WDR families indicates a complex regulation of chs\_H1 genes essential for prenylflavonoid biosynthesis in hop (*Humulus lupulus* L.). *BMC Plant Biol* 2012, **12**:27.
36. Matousek J, Vrba L, Novák P, Patzak J, Keukeleire JD, Skopek J, Heyerick A, Roldán-Ruiz I, Keukeleire DD: Cloning and molecular analysis of the regulatory factor HIMyB1 in hop (*Humulus lupulus* L.) and the potential of hop to produce bioactive prenylated flavonoids. *J Agric Food Chem* 2005, **53**:4793–4798.
37. Matousek J, Kocábek T, Patzak J, Skopek J, Maloukh L, Heyerick A, Fussy Z, Roldán-Ruiz I, Keukeleire DD: HIMyB3, a putative regulatory factor in hop (*Humulus lupulus* L.), shows diverse biological effects in heterologous transgenotes. *J Agric Food Chem* 2007, **55**:7767–7776.
38. Marks MD, Tian L, Wenger JP, Omburo SN, Soto-Fuentes W, He J, Gang DR, Weiblen GD, Dixon RA: Identification of candidate genes affecting  $\Delta^9$ -tetrahydrocannabinol biosynthesis in *Cannabis sativa*. *J Exp Bot* 2009, **60**:3715–3726.
39. Knill T, Schuster J, Reichelt M, Gershenzon J, Binder S: Arabidopsis branched-chain aminotransferase 3 functions in both amino acid and glucosinolate biosynthesis. *Plant Physiol* 2008, **146**:1028–1039.
40. Mentzen WI, Peng J, Ransom N, Nikolau BJ, Wurtele ES: Articulation of three core metabolic processes in Arabidopsis: fatty acid biosynthesis, leucine catabolism and starch metabolism. *BMC Plant Biol* 2008, **8**:76.
41. Meisel L, Fonseca B, González S, Baeza-Yates R, Cambiazo V, Campos R, González M, Orellana A, Retamales J, Silva H: A rapid and efficient method for purifying high quality total RNA from peaches (*Prunus persica*) for functional genomics analyses. *Biol Res* 2005, **38**:83–88.
42. Shao Z, Zhao H, Zhao H: DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res* 2009, **37**:e16.
43. Baggerly KA, Deng L, Morris JS, Aldaz CM: Differential expression in SAGE: accounting for normal between-library variation. *Bioinformatics* 2003, **19**:1477–1483.
44. Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A: High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* 2008, **36**:3420–3435.
45. Livak KJ, Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  Method. *Methods* 2001, **25**:402–408.
46. Cohen SA, Michaud DP: Synthesis of a fluorescent derivatizing reagent, 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate, and its application for the analysis of hydrolysate amino acids via high-performance liquid chromatography. *Anal Biochem* 1993, **211**:279–287.
47. Perera MADN, Choi S-Y, Wurtele ES, Nikolau BJ: Quantitative analysis of short-chain acyl-coenzymeAs in plant tissues by LC-MS-MS electrospray ionization method. *J Chromatogr B Analyt Technol Biomed Life Sci* 2009, **877**:482–488.
48. Claros MG, Vincens P: Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur J Biochem* 1996, **241**:779–786.
49. Emanuelsson O, Nielsen H, von Heijne G: ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci* 1999, **8**:978–841.
50. Bradford MM: A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 1976, **72**:248–254.
51. Prohl C, Kispal G, Lill R: Branched-chain-amino-acid transaminases of yeast *Saccharomyces cerevisiae*. *Methods Enzymol* 2000, **324**:365–375.
52. Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song K, Pikaard CS: Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J* 2006, **45**:616–629.
53. Sparkes IA, Runions J, Kearns A, Hawes C: Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat Protoc* 2006, **1**:2019–2025.
54. van Bakel H, Stout JM, Cote AG, Tallon C, Sharpe AG, Hughes TR, Page JE: The draft genome and transcriptome of *Cannabis sativa*. *Genome Biol* 2011, **12**:R102.
55. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG: Clustal W and Clustal X version 2.0. *Bioinformatics* 2007, **23**:2947–2948.
56. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S: MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011, **28**:2731–2739.

doi:10.1186/1471-2229-13-12

Cite this article as: Clark et al.: Transcriptome analysis of bitter acid biosynthesis and precursor pathways in hop (*Humulus lupulus*). *BMC Plant Biology* 2013 **13**:12.

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

