Comparative Systems Biology Reveals Allelic Variation Modulating Tocochromanol Profiles in Barley (Hordeum vulgare L.)

Rebekah E. Oliver1, Emir Islamovic2, Donald E. Obert3, Mitchell L. Wise4, Lauri L. Herrin5, An Hang2, Stephen A. Harrison5, Amir Ibrahim6, Juliet M. Marshall7, Kelci J. Miclaus8, Gerard R. Lazo9, Gongshe Hu2, Eric W. Jackson10

1 Department of Plant Sciences, North Dakota State University, Fargo, North Dakota, United States of America, formerly USDA-ARS, Small Grains and Potato Germplasm Research Unit, Aberdeen, Idaho, United States of America, 2 USDA-ARS, Small Grains and Potato Germplasm Research Unit, Aberdeen, Idaho, United States of America, 3 Limagrain Cereal Seeds, Battle Ground, Indiana, United States of America, formerly USDA-ARS, Small Grains and Potato Germplasm Research Unit, Aberdeen, Idaho, United States of America, 4 USDA-ARS, Cereal Crops Research Unit, Madison, Wisconsin, United States of America, 5 School of Plant, Environmental and Soil Sciences, Louisiana State University, Baton Rouge, Louisiana, United States of America, 6 Department of Soil and Crop Sciences, Texas A&M University, College Station, Texas, United States of America, 7 Department of Plant, Soil, and Entomological Sciences, University of Idaho Research and Extension, Idaho Falls, Idaho, United States of America, 8 SAS Institute Inc., JMP Genomics Development, Cary, North Carolina, United States of America, 9 USDA-ARS, Western Regional Research Center, Genomics and Gene Discovery, Albany, California, United States of America, 10 Crop Biosciences, General Mill, Inc., Kannapolis, North Carolina, United States of America, formerly USDA-ARS, Small Grains and Potato Germplasm Research Unit, Aberdeen, Idaho, United States of America

Abstract

Tocochromanols are recognized for nutritional content, plant stress response, and seed longevity. Here we present a systems biological approach to characterize and develop predictive assays for genes affecting tocochromanol variation in barley. Major QTL, detected in three regions of a SNP linkage map, affected multiple tocochromanol forms. Candidate genes were identified through barley/rice orthology and sequenced in genotypes with disparate tocochromanol profiles. Gene-specific markers, designed based on observed polymorphism, mapped to the originating QTL, increasing R² values at the respective loci. Polymorphism within promoter regions corresponded to motifs known to influence gene expression. Quantitative PCR analysis revealed a trend of increased expression in tissues grown at cold temperatures. These results demonstrate utility of a novel method for rapid gene identification and characterization, and provide a resource for efficient development of barley lines with improved tocochromanol profiles.

Innovation

Tocochromanols, collectively known as vitamin E, are synthesized in chloroplast membranes of plants and certain algae and bacteria [1]. A primary role of these compounds is to protect lipids in photosynthetic membranes and seeds against reactive oxygen species. Additionally, tocochromanols contribute to membrane function and integrity, electron transport, and cell signaling [2–5]. These functions have significance in normal plant growth and development, plant stress tolerance, and seed longevity [6]. Vitamin E is also influential in human health, providing antioxidant protection [7] and contributing to reductions in cardiovascular disease [8], cholesterol [9], and certain forms of cancer [10,11].

The various tocochromanols have significant yet diverse biological activity. To date, a large proportion of vitamin E research (99.2%) has been centered on the tocopherol forms [12], particularly α-tocopherol. This is primarily due to early reports showing retention and distribution of α-tocopherol to be 80% greater than all other tocochromanols in humans and mice [13]. Unfortunately, reports that tocotrienols were more potent oxidative protectors [14,15] did not result in research trends that would give equal emphasis to potency and concentration. Recent studies have shown that tocotrienol bioactivity differs from that of tocopherols. Nanomolar concentrations of α-tocotrienol have been shown to reduce neurodegeneration [16,17], while tocotrienol compounds have been shown to reduce cholesterol [18,19] and oxidative protein damage [20], and suppress human breast cancer [21].

Structurally, the tocochromanols consist of a polar chromanol ring and hydrophobic prenyl side chain (Fig. 1). Classification is based on chemical variations, with tocochromanols divided into
tocopherols and tocotrienols based on the degree of saturation in the prenyl tail. Each tocochromanol type is further broken down into four forms (α, β, γ, and δ) based on the number and position of methyl groups on the chromanol ring [1,22,23].

Biosynthesis of tocochromanols involves two pathways. Homogentisic acid, the aromatic ring common to tocopherols and tocotrienols, is produced by hydroxyphenylpyruvate dioxygenase (HPPD), from products of the shikimate pathway [22,23]. The methylerythritol phosphate pathway produces the prenyl side chains: threefold-unsaturated geranylgeranyldiphosphate for tocotrienols, and fully-saturated phytyldiphosphate for tocopherols [22,23]. Partitioning of tocopherols and tocotrienols occurs during condensation of the chromanol ring and side chain, with preferential addition of the side chains by different enzymes. Homogentisate phytyltransferase (VTE2) primarily produces the tocopherol precursor, while homogentisate geranylgeranyltransferase (HGGT) primarily produces the tocotrienol precursor [22–24].

Remaining biosynthetic reactions are analogous between tocopherols and tocotrienols. Tocopherol cyclase (VTE1) closes the second ring in the head group and functions in synthesis of all four forms [1,22,23]. Form differentiation is caused by two methyltransferase enzymes: VTE3 methylation produces the γ form, which is otherwise identical to δ; and VTE4 methylation converts γ to α, and δ to β. Thus, VTE3 segregates γ and α forms from δ and β, and VTE4 segregates α and β forms from γ and δ (Fig. 1).

Small grains, which have high levels of the powerful and unique tocotrienol forms [24–28], are an excellent model for tocochromanol study. Although loci influencing tocochromanol biosynthesis have been identified on rice, maize, and oat chromosomes [29–32], utility of these QTL is limited by the breadth of the associated genetic region, the requirement for complex statistical models, and

Figure 1. Tocochromanol biosynthesis. Dotted lines represent double bonds in the tocotrienol prenyl side chain. Dual labels correspond to tocopherol and tocotrienol, respectively. doi:10.1371/journal.pone.0096276.g001
difficulty in merging QTL positions with sequence data to identify candidate genes [33–35]. Despite the potential impact of high-vitamin E barley, and the significance of this species as a model for more complex cereal grains, fundamental studies of barley tocochromanol genes have not been published.

Objectives of this research were to quantify tocochromanol accumulation in a population of Falcon x Azhul (FA) barley RILs, characterize candidate genes causing tocochromanol variation using a comparative systems biological approach, and develop predictive assays for novel alleles modulating tocochromanol levels and ratios in barley grain. Here we present genomic sequence data for tocochromanol synthesis genes in wild-type and mutant barley cultivars, correlating differences in tocochromanol profiles with sequence variation in genic and regulatory regions. We also present a genomic-based method for correlating significant genetic loci with the underlying operative gene. Application of this technique could augment the relevancy of the extensive but underutilized QTL data throughout the plant genetics literature. Progress in dissecting the genes identified in this study will facilitate development of barley cultivars with high total tocochromanol levels and desirable compound ratios, with applications for comparative human health studies and subsequent varietal release.

Results
Tocochromanol Analysis
Chemical analysis of the FA barley population detected all eight tocochromanol forms, with relative abundance of the four forms similar in tocopherols (T) and tocotrienols (T3) and total concentrations substantially higher in T3. Concentrations of α forms were highest, followed by γ, β, and δ. The δT form, although detected in 95.2% of samples, was present at low concentrations, with greater variance between replications. Falcon had higher levels of αT than Azhul or either check, although differences were not significant across all environments (Table 1). Tocopherol differences were less pronounced for the remaining three forms, although Azhul tended to have higher concentrations of γ and δ. Azhul accumulated higher levels of all T3 forms, particularly αT3.

Analysis of variance indicated significant influence of location and year (Table S1). With the exception of βT in Tetonia, mean levels of each form were higher in 2009 than in 2008 (Table 1). For both years, concentrations of all forms were higher in Aberdeen than in Tetonia (Table 1). Within the FA RILs, broad sense heritability estimates ranged from 35 to 61 in Aberdeen, and from 41 to 63 in Tetonia, with generally higher heritability for T3.

Comparative Genomics Candidate Gene Identification
Molecular genetic dissection of tocochromanol accumulation in barley grain was achieved through general pathway elucidation using model species, followed by identification of genomic areas associated with acute phenotypic measurements of each bioactive compound. This information was then used to resolve candidate genes based on comparisons with annotated genes from related species, allowing candidate gene validation using wild-type and mutant genotypes.

Genomic regions affecting phenotypes in at least three environments were detected on two barley chromosomes (Table 2 and Fig. 2). Each region encoded enzymes for the production of multiple compounds of either T or T3: chromosome 6H affected γT and δT accumulation, and two loci on 7H affected βT3 and δT3, with the second also significant for γT3. Based on rice and barley orthology, 6H contained a VTE4 sequence while genomic regions of 7H contained sequence for VTE2 and 3-phosphoshikimate 1-carboxyvinyltransferase, which synthesizes a distant precursor of the tocochromanol ring structure (Fig. S1). Additional genomic regions affecting βT phenotypes from two environments were localized on 1H and 3H. Although non-significant, genomic regions were identified on 5H affecting βT and δT3, and on 6H affecting δT3. Barley-rice orthologous relationships showed the genomic region on 5H contained VTE3, and the region on 6H contained VTE1. Since a SNP marker was not mapped at the QTL peak on 5H, analysis of this locus was based on an additional sequence identified through the HarvEST database (Fig. S1). Of the two parental lines, Azhul appeared to have a stronger influence on T3 variation, predominating at both 7H QTL regions (Fig. 2). The Falcon allele appeared to have a greater influence on T accumulation, evidenced by the region on 6H.

Gene sequence and expression analysis
Genes corresponding to the two major QTL regions are substantially different in genetic structure and polymorphism level, but similar in gene expression profile. The full-length VTE4 gene comprised six exons and five introns, ranging in size from 84 to 952 bp (Fig. 3a). Approximately 54% of the total gene length is intronic. Polymorphism was limited to a 2-bp indel in the third intron, and a wobble base in the first exon of Falcon (Table S2). The nucleotide variation changes the amino acid, with the G variant encoding alanine and the T variant encoding serine. Ten clones, based on two different cloning events, were sequenced in an attempt to resolve the nucleotide identity at this position: five clones contained a G and five a T, with sequence identity proportional within each cloning event. Remaining sequence, including the intronic indel, was identical within these clones.

Sequence analysis of 1100 bp upstream of the VTE4 start codon revealed two sequence polymorphisms between Falcon and Azhul, both of which fell within promoter sequences (Fig. 3 and Table S2). Approximately 18 additional types of motifs were annotated at approximately 90 sites within the sequenced promoter region, including motifs for response to low temperature, light, and plant stress (Table 3 and Table S3).

Given the similar function of VTE2 and HOGT, cloning of both genes was attempted. Amplification was not obtained for VTE2; however, HOGT amplified readily. The HOGT full-length gene contained 14 exons and 13 introns, with introns comprising approximately 60% of the total length and ranging in size from 75 to 715 bp in Falcon (Fig. 3b). Polymorphism was pronounced, tending to occur more frequently in introns and in the 5′ half of the gene. Even accounting for relative intron/exon lengths, intronic SNPs occurred at nearly twice the frequency of those in exons. All indels were within introns, including a region of approximately 150 bp in the second intron. HOGT is known for assembling T3; thus, two FA RILs (FA.41 and FA.117) with similar total T levels (13.82 and 14.10 μg/g) and disparate total T3 levels (35.51 and 69.49 μg/g) were also cloned and sequenced. Gene sequences were identical between FA.41 and Falcon, and between FA.117 and Azhul, correlating sequence data with low and high total T3 phenotypes.

Protein structure prediction was performed to understand the effects of allelic variation within HOGT. Of ten FA SNPs within the coding sequence, seven were silent and three caused an amino acid change affecting polarity or pH (Table S4). Although all SNPs sites were predicted to affect secondary protein structure (Fig. 4a), no significant three-dimensional structural or functional differences were predicted based on I-TASSER analysis of Falcon and Azhul protein sequences (Fig. 4b). The best-ranked enzyme
Table 1. Tocochromanol means of Falcon and Azhul parents and two barley checks grown over four location years.

| Environment/Genotype | Tocopherol | Tocotrienol |
|----------------------|------------|-------------|
|                      | α          | β           | γ           | δ        | α           | β           | γ           | δ        |
| AB 2008              |            |             |             |          |             |             |             |          |
| Falcon               | 11.93 A    | 1.11 A      | 1.60 B      | 0.14 C   | 29.00 BC    | 5.49 B      | 5.62 B      | 0.84 B   |
| Azhul                | 10.47 B    | 1.06 A      | 2.29 B      | 0.23 BC  | 35.33 A     | 5.54 B      | 8.37 A      | 1.47 A   |
| Mean                 | 10.91***   | 1.15*       | 2.77***     | 0.38**   | 30.89**     | 6.22***     | 7.36***     | 1.18***  |
| AB 2009              |            |             |             |          |             |             |             |          |
| Falcon               | 13.21 A    | 0.71 B      | 4.07 B      | 0.39 B   | 31.69 B     | 4.78 C      | 9.12 A      | 1.10 B   |
| Azhul                | 11.36 B    | 0.77 AB     | 2.34 C      | 0.23 B   | 37.00 A     | 7.72 B      | 9.89 A      | 1.80 A   |
| Mean                 | 11.95**    | 0.80***     | 3.60**      | 0.45**   | 31.98**     | 8.34**      | 8.66**      | 1.39**   |
| TE 2008              |            |             |             |          |             |             |             |          |
| Falcon               | 12.31 A    | 0.74 B      | 1.59 D      | 0.26 B   | 28.67 C     | 7.13 B      | 6.56 B      | 1.01 B   |
| Azhul                | 11.12 AB   | 0.77 AB     | 2.34 C      | 0.23 B   | 37.00 A     | 7.72 B      | 9.89 A      | 1.80 A   |
| Mean                 | 11.95**    | 0.80***     | 3.60**      | 0.45**   | 31.98**     | 8.34**      | 8.66**      | 1.39**   |

Means and heritability estimates were calculated for Falcon x Azhul (FA) RILs using combined years at each location.

*Concentrations given in μg/g. Values followed by the same letter or number of asterisks are not significantly different within a column (p < 0.05). For means of genotypes within an form, a single asterisk is equivalent to A, two asterisks to B, and three asterisks to C.

+Indicates an intermediate value.

Irrigated field trial in Aberdeen.

Non-irrigated field trial at Tetonia.

Broad sense heritability calculated as genotype variance divided by cumulative variance including error.

doi:10.1371/journal.pone.0096276.t001
| Chrom. | Trait | Genotype | QTL marker | QTL peak (interval) | Original LOD | Add. | R² | With VTE4 LOD | Add. | R² | With HGGT LOD | Add. | R² |
|-------|-------|----------|------------|-------------------|-------------|------|----|--------------|------|----|--------------|------|----|
| 6H    | γT    | AB 2008  | MK_4313-482| 57.2 (56.4, 58.0) | 12.43       | 0.51 | 0.31| 17.42        | 0.65 | 0.50| 12.43        | 0.51 | 0.31|
|       |       | AB 2009  | MK_4313-482| 57.2 (56.9, 56.0) | 15.82       | 0.90 | 0.37| 14.11        | 0.95 | 0.41| 15.57        | 0.90 | 0.37|
|       |       | TE 2008  | MK_4313-482| 57.2 (55.3, 57.6) | 7.76        | 0.37 | 0.21| 13.50        | 0.53 | 0.45| 7.76         | 0.37 | 0.21|
|       |       | TE 2009  | MK_4313-482| 57.2 (56.6, 58.2) | 12.86       | 0.60 | 0.30| 15.12        | 0.71 | 0.43| 12.86        | 0.60 | 0.30|
| δT    |       | AB 2008  | MK_4313-482| 55.2 (55.1, 57.3) | 11.30       | 0.09 | 0.26| 12.64        | 0.11 | 0.37| 11.00        | 0.09 | 0.25|
|       |       | AB 2009  | MK_4313-482| 57.2 (56.4, 58.6) | 8.69        | 0.16 | 0.22| 9.06         | 0.17 | 0.26| 8.69         | 0.16 | 0.22|
|       |       | TE 2008  | MK_1911-55  | 54.7 (53.7, 57.4) | 5.98        | 0.03 | 0.14| 6.58         | 0.04 | 0.23| 5.98         | 0.03 | 0.14|
|       |       | TE 2009  | MK_4313-482| 57.2 (54.2, 57.8) | 10.37       | 0.40 | 0.29| 10.57        | 0.41 | 0.31| 8.89         | 0.09 | 0.20|
| 7H    | βT3   | AB 2008  | BPB-7417_7H | 22.8 (19.9, 23.9) | 11.61       | -0.95 | 0.36| 12.42        | -2.07| 0.36| 12.66        | -1.88| 0.34|
|       |       | AB 2009  | BPB-7417_7H | 21.2 (19.7, 23.2) | 18.19       | -3.27 | 0.42| 18.41        | -3.33| 0.43| 21.23        | -3.51| 0.47|
|       |       | TE 2008  | BPB-7417_7H | 20.9 (19.9, 23.4) | 12.82       | -2.03 | 0.36| 13.12        | -2.03| 0.36| 12.96        | -1.99| 0.35|
|       |       | TE 2009  | BPB-7417_7H | 20.8 (19.0, 23.8) | 10.37       | -0.40 | 0.29| 10.57        | -0.41| 0.31| 9.99         | -0.33| 0.24|
|       |       | AB 2009  | MK_4671-856 | 123.2 (120.1, 126.3) | 2.75        | 0.82  | 0.06| 2.61         | 0.71  | 0.05| 5.12         | 1.04  | 0.10|
|       |       | TE 2008  | MK_5764-430 | 119.7 (117.4, 125.7) | 5.28        | 1.57  | 0.10| 5.36         | 1.57  | 0.10| 7.56         | 1.83  | 0.13|
|       |       | TE 2009  | MK_4671-856 | 123.2 (117.1, 126.0) | 3.47        | 1.00  | 0.06| 2.85         | 0.85  | 0.05| 5.15         | 1.15  | 0.09|
| γT3   |       | AB 2008  | MK_4671-856 | 123.2 (120.1, 126.3) | 4.25        | 0.87  | 0.12| 4.69         | 0.42  | 0.11| 7.99         | 1.19  | 0.19|
|       |       | TE 2008  | MK_4671-856 | 119.7 (117.4, 120.4) | 7.97        | 1.35  | 0.23| 8.04         | 1.35  | 0.23| 11.80        | 1.56  | 0.31|
|       |       | TE 2009  | MK_5764-430 | 117.7 (116.6, 124.4) | 3.43        | 0.80  | 0.09| 3.93         | 0.87  | 0.10| 7.79         | 1.19  | 0.19|
| δT3   |       | AB 2008  | MK_4671-856 | 123.2 (121.6, 124.7) | 5.08        | 0.24  | 0.12| 4.63         | 0.21  | 0.09| 7.80         | 0.27  | 0.17|
|       |       | AB 2009  | MK_4671-856 | 123.2 (120.9, 123.4) | 9.35        | 0.42  | 0.19| 6.77         | 0.34  | 0.11| 11.27        | 0.44  | 0.21|
|       |       | TE 2009  | MK_4671-856 | 123.2 (121.6, 124.9) | 6.03        | 0.26  | 0.12| 5.91         | 0.23  | 0.10| 7.36         | 0.28  | 0.15|

1Indicates the tocochromanol form. T – Tocopherol; T3 – Tocotrienol.
2All indicates the irrigated field trial in Aberdeen; TE indicates the non-irrigated field trial in Tetonia.
3Flanking marker to the left of the QTL peak.
4Interval based on one LOD to each side of the peak. Peak and interval positions are in cM.
5QTL detection was based on a LOD threshold of 2.5 (1000 permutations and type I error of 5%).
6Total phenotypic variation explained by all QTL.
7Percentage of phenotypic variation accounted for by QTL.

Table 2. Tocochromanol quantitative trait locus analysis summary for Falcon x Azhul RILs across four location years.
classification scores were 0.851 for Falcon and 0.856 for Azhul, both matching peptide homolog 1dceA corresponding to the alpha subunit of geranylgeranyltransferase. The active site was predicted at residue 140. Of the remaining best five matches, only 1d8dA and 3dsxA were common to both allele peptides.

The lack of evidence for substantially altered protein structure in HGGT warranted further evaluation of the promoter region. Analysis revealed SNPs in motifs influencing low temperature response, light response, drought tolerance, auxin response, TATA and CAAT boxes, promoter sequences, including a transcription start site, and others (Table 3 and Table S3). The sequenced promoter region included approximately 30 types of promoter motifs at approximately 100 sites (Table S3).
Both VTE4 and HGGT showed a trend of increased expression in response to cold temperatures (Fig. 5). Temperature-based expression differences were pronounced in samples based on Falcon shoot tissue, with an approximately threefold expression difference in VTE4, and more than a twofold difference in HGGT. Samples based on Azhul embryo tissue showed a similar response, with cold temperatures increasing expression twofold in VTE4 and nearly threefold in HGGT. Temperature did not cause a statistically important effect in VTE4 for Falcon embryo tissue or Azhul shoot tissue. However, samples based on Falcon embryos showed more than a twofold increase in HGGT expression in response to cold. Expression of HGGT in Azhul shoots appeared to decrease in response to cold, although the difference was not statistically significant.

Predictive markers and validation of gene identification

Predictive PCR-based markers were designed for VTE4 and HGGT genes (Table S5). The VTE4 marker amplified a 78-bp region containing the 2-bp indel in the third intron; the HGGT marker amplified a 53-bp region surrounding the penultimate SNP, in the eleventh exon (Fig. 3). Both markers generated a 1:1 ratio when genotyped in the FA population (Table S6) and mapped to the respective QTL peaks when appended to the FA genetic linkage map (Fig. 2). Reiteration of marker-trait correlation, including predictive markers, increased the mean LOD score, additive variance, and R² values at the respective loci (Table 2). Inclusion of the VTE4 marker increased the mean R² value from 0.25 to 0.36; addition of the HGGT marker increased the mean R² from 0.13 to 0.19. Mean values for additive variance increased from 0.34 to 0.41 for VTE4, and from 0.34 to 1.00 for HGGT. In addition, inclusion of the HGGT marker allowed detection of additional peaks at the HGGT locus, making the QTL significant for βT3, γT3, and δT3 in all four environments, and for αT3 in all environments except AB 2008.

Table 3. Motifs affected by promoter sequence polymorphism between Falcon and Azhul.

| Adjacent gene | Bp upstream of ATG acposion | strand | Motif function | Motif sequence | Sequence |
|---------------|-----------------------------|--------|----------------|----------------|----------|
| VTE4          | 293                         | +      | Promoter of transcription start | TATA | Falcon: TACA, Azhul: TATA |
|               | 328                         | +      | Cis-acting element, transcription start | CCAAT | Falcon: TCAAT, Azhul: CCAAT |
| HGGT          | 381                         | +      | Cis-acting element, low temperature response | CCGAAA | Falcon: ACGGAAA, Azhul: ACCGAAA |
|               | 452                         | +      | Promoter of transcription start | TATAAA | Falcon: TATATG, Azhul: TATAAA |
|               | 458                         | -      | Promoter of transcription start | TATAAC | Falcon: TATACC, Azhul: TATACA |
|               | 518                         | +      | Light response element | AGAGATG | Falcon: AGAGATG, Azhul: ATG |
|               | 534                         | +      | Cis-acting element, endosperm expression | GTCA | Falcon: GTCA, Azhul: GTCA |
|               | 579                         | +      | Auxin-responsive element | AACGAC | Falcon: AACAC, Azhul: AACGAC |
|               | 581                         | -      | Cis-acting element, MeJA responsiveness | TGAGC | Falcon: TGACC, Azhul: TGAGC |
|               | 620                         | -      | Cis-acting element, low temperature response | CCGAAA | Falcon: CTGAAA, Azhul: CCGAAA |
|               | 824                         | -      | MYB binding site, drought-inducibility | TAACGT | Falcon: TAACGTG, Azhul: TAACGTG |
|               | 1006                        | +      | Cis-acting element, light responsiveness | CACATGG | Falcon: ACACATGG, Azhul: GCACATGG |

doi:10.1371/journal.pone.0096276.t003
Given the complexity of systems biology and the scope of plant genetics research, species-specific elucidation of even the major biochemical pathways is unrealistic; thus, a detailed and finely-curated set of biochemical pathways in Arabidopsis and other models has been critical to assessment of location, structure, and function in corresponding genes of organisms with larger and more complex genomes. Characterization of the tocochromanol pathway in Arabidopsis has accelerated elucidation of vitamin E production in soybean and other plants in which vitamin E is a key trait. However, the comparative genomics approach has not been extended to correlation of genes or sequence data within QTL. Here we demonstrate the application of biochemical and sequence data from model species to directly elucidate corresponding genes in more complex genomes.

Significant variation in field location and year indicate that seed tocochromanol content is impacted by environment. Variation between years appeared to be correlated to temperature and possibly water availability. Mean overall temperatures for both

Discussion

Figure 4. Predicted secondary structure (A) and three-dimensional structure (B) of HGGT proteins in barley cultivars Falcon and Azhul. Sequence polymorphism contributing to folding changes is indicated by the green consensus guide (A). Blue curved arrows represent turns, orange arrows represent beta strands, pink cylinders represent alpha helices, and gray wavy lines represent coils. Pink labels within ribbon diagrams correspond to the predicted active site.

doi:10.1371/journal.pone.0096276.g004
locations were approximately 1°C lower in 2009 than in 2008; thus, cooler conditions within the same location favored increased tocochromanol production. This conclusion concurs with previous reports. For example, environmental impacts on tocochromanol production were reported in a study of wheat, canola, sunflower, and soybean. Although individual field effects were not separated, growth chamber experiments verified a strong correlation between temperature and vitamin E content, with lower temperature causing increased accumulation [36]. Temperatures in Tetonia were colder than in Aberdeen; however, tocochromanol production was lower, possibly reflecting differences in water availability, cultivation practices, or saprophytic fungi or bacteria. Although not fully understood, the trend is consistent with previous data.

Levels of αT and αT3, measured in an oat population grown at the same locations, were significantly lower in Tetonia [32]. Given the sequence variation in promoter motifs related to temperature and drought response, interaction of barley genotypes with environmental impacts is not surprising (Table 1 and Table 3).

As in other barley tocochromanol studies [37,38], αT and αT3 were the predominant forms, with αT3 being the more abundant. Within each tocochromanol type, concentrations of γ forms were approximately one-fourth as abundant, followed by β, and more distantly by δ forms. This indicates that γ and δ precursors are effectively converted to α and β forms in mature barley seed, reflecting activity of VTE4 (Fig. 1). Spatial and quantitative accumulation of the eight forms has been shown to vary during barley seed development, with tissue-specific form distribution, and relative form ratios shifting throughout early kernel formation [39]. Kinetics of tocochromanol forms suggests specific roles throughout seed development and germination, requiring precise control of individual genes.

In this study, major loci affected multiple forms, reflecting activity of genes at key intersections of the tocochromanol biosynthetic pathway. The barley 6H QTL influencing γT and δT corresponded to a rice region containing VTE4, which converts γ to α and δ to β (Fig. 1 and Fig. S1), accounting for the forms mapped to this locus. Minimal FA sequence variation was observed within the VTE4 gene and its upstream regulatory region, although observed polymorphism affected significant regions, modifying two promoter motifs and causing an amino acid change. Based on VTE4 sequences of two FA RILs with divergent levels of γ and δ forms, it does not appear that the intronic indel plays a role in VTE4 activity. However, the indel is heritable, evidenced by the 1:1 ratio produced when genotyping the FA population with the molecular marker based on this polymorphism (Table S6). Despite the sequence similarities, expression patterns of VTE4 differed between Falcon and Azhul, with cold temperatures inducing increased expression primarily in Falcon shoot tissue and Azhul embryo tissue. Since promoter sequence polymorphism affected regions involved in low-temper-
ature response (Table 3 and Table S3), these motifs may affect spatial expression of VTE4. Future work should be performed to explore this possibility, as well as to study additional factors affecting VTE4 gene expression such as methylation patterns and involvement of microRNA.

The QTL accounting for a majority of the phenotypic variance in this study were detected on chromosome 7H. Both regions influence accumulation of multiple tocotrienol forms, the more abundant tocotrienol in barley and some other small grains. The QTL affecting all four T3 forms corresponded to the rice chromosome 6 region containing VTE2 (Fig. 2 and Fig. S1); a gene reported to preferentially add the phytodiphenylside chain to form tocopherols, with analogous formation of tocotrienols by HGGT [24]. Dicots appear to lack a prenyldiphenylase recognizing geranylgeranyl diphenylside, hence they produce little or no tocotrienols. Monocots, however, appear to express two distinct prenyldiphenylases, one favoring phytodiphenylside as substrate, the other favoring geranylgeranyl diphenylside (although there is some overlap in substrate specificity with both enzymes)[24,40]. In this study, a major QTL on chromosome 7H corresponded with increased accumulation of all four tocotrienol forms. Barley SNP sequences within this QTL are syntenous with a rice chromosome 6 region containing VTE2 (OS06g41480). These results suggest that expression may be substantially

Plant material and field experiments

A population of 142 recombinant inbred lines (RILs) was derived from hybridization of six-row barley ‘Falcon’ (F) and ‘Azhul’ (A). Falcon is a hulless feed barley (PI591612) well adapted to Idaho; Azhul is a dwarf, hulless food barley derived from mutation breeding (released by the USDA-ARS and the Arizona Agricultural Experiment Station). The FA population was derived by single-seed descent to the F2 generation, and F8 seed was harvested from individual F2 plants to produce seed for each RIL.

Field trials were performed in a completely randomized block design and included three replications of Falcon, Azhul, and two-row barley checks ‘Baroness’ and ‘CDC Alamo’ (industry leaders in feed and food types, respectively). Trials were conducted over two years (2008 and 2009) at two locations: irrigated at the USDA-ARS in Aberdeen, ID, and non-irrigated at the University of Idaho Teton Research and Extension Center, Teton, ID. Plots consisted of 1.2-meter rows with 0.4 m spacing between rows. Fertilizer and post-emergent herbicide was applied, and seed was combine harvested.

Tocochromanol analysis

Barley seed was ground in a Retsch ZM-1 mill and extracted using a revised hot saponification and extraction method [47]. Briefly, 0.5 g freshly-ground sample was weighed into a 15-mL screw-cap glass test tube and mixed with 0.5 mL potassium hydroxide (600 g/L H2O), 0.5 mL ethanol (95%), 0.5 mL sodium chloride (10 g/L H2O), and 1.25 mL ethanolic pyrogallol (60 g/L EtOH). Solution was mixed with a glass rod. Tubes were then vortexed and saponified in a 70°C water bath for 30 min., with vortexing every 10 min. Tubes were chilled in an ice bath for 15 min. The solution was extracted twice, once with 3.75 mL sodium chloride (10 g/L) and 3.75 mL n-hexane/ethyl acetate (9:1), once with 3.75 mL n-hexane/ethyl acetate, and organic layers were evaporated in a speed vacuum. Dried residue was resuspended in 2 mL hexane. Chromatographic separation was performed with normal phase HPLC, using a Shimadzu (Kyoto, Japan) LC-6A pump, RF10AxL detector, and SCL-10Avp controller. Compounds were separated on a 5-μm 250×4.6 mm Grace Adsorbsil silica column using an isocratic mobile phase (hexane with 2% dioxane and 2% ethyl acetate) at a flow rate of 2 mL/min. Peaks were detected by fluorescence with excitation at 295 nm and emission at 330 nm. Identification of individual forms was based on retention time, and quantification was based on standard curves developed using authentic tocopherols (Metraya LLC, Pleasant Gap, PA). Tocochromanols, having essentially the same fluorescent properties as their corresponding tocopherols [48], were quantified with the same standard curves.

QTL mapping

Genotyping and linkage analysis were performed according to Islamovic et al. [49]. Analysis of variance (ANOVA), means, and broad sense heritability estimates were calculated for each tocochromanol form using the JMP 9.0 software (SAS Institute, Cary, NC). Heritability within the RIL population was calculated by dividing genotypic variance by total variance, calculated from
variance components estimates for genotype, location, and year in a full-factorial random effects model. Based on significant interactions within the model, each genotype, location, and year were treated as separate traits for QTL identification (Table 1).

QTL were detected with WinQTL Cartographer [50], using marker locus positions from an existing Falcon x Azhul SNP map [51]. Experiment-wise thresholds, determined through 1000 permutations of the data, were used to determine significant QTL [52]. Peaks were initially detected using single marker analysis, and QTL, positions, LOD scores, and phenotypic variance were determined with composite interval mapping (CIM)[53], using stepwise regression, a window size of 10 cM, walk speed of 2 cM, and significance threshold of P=0.05. Clustered peaks were counted as separate QTL if individual peaks spanned more than one LOD score.

Orthology of barley QTL to rice sequence

Sequences from SNP markers within one LOD score of significant QTL peaks were used to identify syntenous rice regions, using the BLAST tool of the Rice Genome Annotation Project (http://rice.plantbiology.msu.edu). The Rice Genome Browser was then used to identify candidate rice genes within the region of the QTL peak. For barley regions with widely-spaced markers, additional SNPs within the QTL were identified using the HarvEST:Barley software version 1.77 (http://harvest.ucr.edu). Candidate genes likely to influence trait expression were identified based on map position and function.

Sequencing and annotation of tocochromanol genes and regulatory regions

Primers for gene cloning were designed based on coding sequences of locus AK355075 for VTE4, locus AK366699 for VTE2, and locus AY222860 for HGGT (available in NCBI http://www.ncbi.nlm.nih.gov). Primers for promoter regions (contigs 254212 and 2207823 for VTE4, contig 1025503 for HGGT) were designed from sequences obtained from the morex_rcha database in ViroBLAST, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) http://webblast.ipk-gatersleben.de/barley/viroblast.php). Primer sequences are listed in Table S5. PCR amplification was performed using the Phusion High-Fidelity PCR Kit (Finnzymes, #F-553S, distributed by New England BioLabs, Inc.). Reaction preparation and thermocycling were generally as recommended by the manufacturer, with 40 cycles of 10 s denaturation, 63°C annealing, and 30 s extension. Gel bands corresponding to target regions were purified using a QIAquick PCR Purification Kit (Qiagen, #28104), and purified DNA was cloned using a StrataClone Blunt PCR Cloning Kit (Stratagene, #240207) and plated on LB-Ampicillin agar with 40 μg/ml. White colonies were cultured in 5 ml LB-Ampicillin broth and plasmids were purified with a QuickClean II Plasmid Miniprep Kit (GenScript, #L00420). Sequencing was performed at the Idaho State University Molecular Research Core Facility, using an Applied Biosystems 3130xl instrument and plasmid reactions purified with a BigDye XTerminator Purification Kit (Applied Biosystems, #4376485). Sequences were assembled and analyzed using Geneious Pro v5.5.6 software [54]. Intron regions were identified through comparative alignment of full-length genes with coding sequences. Promoter motif sequences, including transcription start sites, were identified using the Neural Network Promoter Prediction v2.2 software, Berkeley Drosophila Genome Project (http://www.fruitfly.org/seq_tools/promoter.html). Additional promoter motifs were identified using the PlantCARE database of cis-acting regulatory elements (http://bioinformatics.psb.ugent.be/webtools/plantcare/html)[55].

Molecular marker development and application

Gene-specific, PCR-based markers were developed for HGGT and VTE4 based on sequence differences between Falcon and Azhul. Primers were designed to amplify a 50-80-bp region containing the polymorphism, and genotyping of the FA population was performed using high-resolution melt analysis, as described [56]. Markers were appended to the existing map using MultiPoint software (http://www.multiqtl.com) and QTL analysis was repeated separately for addition of each marker, using the same conditions as previously.

Protein structure and gene expression

Secondary protein structure of HGGT was analyzed using coding sequences from Falcon and Azhul and the prediction function in Geneious Pro v5.4. Three-dimensional protein structure and function were analyzed using I-TASSER [57,58].

Tissue for gene expression analysis was collected from etiolated shoots and mature embryos of Falcon and Azhul. Samples of each tissue type were germinated and maintained at two temperature treatments: room temperature and 4°C. RNA was extracted using an UltraClean Plant RNA Isolation Kit (Mo Bio, #13300-50). cDNA was synthesized using a SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen, #18080-051) primed with oligo(dT). Quantitative PCR was performed using 1x SsoFast EvaGreen Supermix (BioRad, #172-5201) with 80 ng cDNA and 0.5 μM forward and reverse primers in a 12.5-μl reaction volume. Thermocycling was performed on a BioRad C1000 thermal cycler with a CFX96 optics module, with initial denaturation at 98°C for 2 min followed by 40 cycles of 98°C for 2 s and 55°C for 5 s, with a fluorescence reading taken at the end of each cycle. Analysis of RT-PCR data was performed using a BioRad algorithm based the method of Vandesompele et al. [59].

Supporting Information

Figure S1 Identification of candidate genes influencing production of tocochromanols in the barley Falcon x Azhul mapping population. Candidate genes were identified in rice chromosome regions syntenous with sequences of barley SNPs within the QTL.

Table S1 Analysis of variance summaries for tocochromanol forms of Falcon, Azhul, Baronesse, and CDC Alamo.

Table S2 FASTA sequences. Sequences for VTE4 and HGGT genes and promoter regions in Falcon, Azhul, and Falcon x Azhul RILs.

Table S3 Annotation of promoter regions. A complete list of promoter motifs identified in the upstream regulatory regions of VTE4 and HGGT.

Table S4 Falcon x Azhul SNPs coding an amino acid change in HGGT.

Table S5 Primer sequences used in cloning of gene and promoter regions, SNP genotyping, and quantitative
PCR analysis in Falcon, Azhul, and the Falcon x Azhul RIL population.

Table S6  Genotype scores of VTE4 and HGGT markers in the Falcon x Azhul RIL population. Genotyping was performed using high-resolution melt analysis evaluation of polymorphic regions within each gene.

(XLS)

Acknowledgments

The authors gratefully acknowledge Irene Shackellford, and the staff at the University of Idaho Research and Extension experiment stations, for expert management of field experiments.

Author Contributions

Conceived and designed the experiments: REO EI DEO GH EWJ. Performed the experiments: REO DEO LLH AH SAH AI JMM. Analyzed the data: REO EI MLW LLH KJM GRL EWJ. Contributed reagents/materials/analysis tools: DEO MLW AH JMM. Wrote the paper: REO MLW EWJ.

References

1. Schultz G, Soll J, Fiedler E, Schulze-Siebert D (1985) Synthesis of prenylquinones in chloroplasts. Physiol Plant 64: 123–129.
2. Kamal-El-Din A, Appelqvist L-A (1996) The chemistry and antioxidant properties of tocopherols and tocotrienols. Lipids 31: 671–701.
3. Azzi A, Stocker A (2000) Vitamin E: Non-antioxidant roles. Prog Lipid Res 39: 231–255.
4. Munut-Bosch S, Alegre L (2002) The function of tocopherols and tocotrienols in plants. Crit Rev Plant Sci 21: 31–57.
5. Munut-Bosch S, Falk J (2004) New insights into the function of tocopherols in plants. Planta 218: 323–326.
6. Satler SE, Gilliland LU, Magdalanes-Lunback M, Pollard M, DellaPenna D (2004) Vitamin E is essential for seed longevity and for preventing lipid peroxidation during germination. Plant Cell 16: 1419–1429.
7. Yoshida Y, Niki E, Noguchi N (2003) Comparative study on the action of tocopherols and tocotrienols as antioxidant: Chemical and physical effects. Chem Phys Lipids 123: 63–75.
8. Raussel AHG, Rahman ARA, Yuen KH, Wong AR (2008) Arterial compliance and vitamin E blood levels with a self emulsifying preparation of tocotrienol rich vitamin E. Arch Pharm Res 31: 1212–1217.
9. Qureshi AA, Sami SA, Salser WA, Khan FA (2001) Synergistic effect of tocotrienol-rich fraction (TRF) of rice bran and lovastatin on lipid parameters in hypercholesterolemic humans. J Nut Bioch 12: 318–329.
10. Chang PN, Yap WN, Lee DTW, Ling MT, Wong YC, et al. (2009) Evidence of HGGT expression in the Falcon x Azhul RIL population. 

PLOS ONE | www.plosone.org 12 May 2014 | Volume 9 | Issue 5 | e96276

11. Sen CK, Khanna S, Roy S, Ryu H, Bahadduri P, Swaan PW, et al. (2003) Molecular basis of vitamin E action. Arch Pharm Res 26: 193–205.
12. Sen CK, Khanna S, Roy S (2006) Tocotrienols: Vitamin E beyond tocopherols. PLOS ONE 1: 1–10.
13. Chang PN, Yap WN, Lee DTW, Ling MT, Wong YC, et al. (2009) Evidence of HGGT expression in the Falcon x Azhul RIL population. 

14. Cooney RW, France AA, Harwood PJ, Hatch-Pigott V, Custer LJ, et al. (1993) Tocopherol and tocotrienol accumulation during development of caryopses from barley. J Amer Oil Chem Soc 70: 2141–2152.
15. Theriault A, Chao JT, Wang Q, Gapor A, Adeli K (1999) Tocotrienol: A review of its therapeutic potential. Clin Biochem 32: 309–319.
16. Khanna S, Roy S, Ryu H, Bahadduri P, Swaan PW, et al. (2003) Molecular basis of vitamin E action. Arch Pharm Res 26: 193–205.
17. Sen CK, Khanna S, Roy S (2006) Tocotrienols: Vitamin E beyond tocopherols. PLOS ONE 1: 1–10.
18. Qureshi AA, Burger WC, Peterson DM, Elson CE (1986) The structure of an oligopeptide modulating barley α-tocopherol biosynthesis. J Amer Oil Chem Soc 73: 2235–2240.
19. Qureshi AA, Sami SA, Salser WA, Khan FA (2001) Synergistic effect of tocotrienol-rich fraction (TRF) of rice bran and lovastatin on lipid parameters in hypercholesterolemic humans. J Nut Bioch 12: 318–329.
20. Chang PN, Yap WN, Lee DTW, Ling MT, Wong YC, et al. (2009) Evidence of HGGT expression in the Falcon x Azhul RIL population. 

1. Schultz G, Soll J, Fiedler E, Schulze-Siebert D (1985) Synthesis of prenylquinones in chloroplasts. Physiol Plant 64: 123–129.
2. Kamal-El-Din A, Appelqvist L-A (1996) The chemistry and antioxidant properties of tocopherols and tocotrienols. Lipids 31: 671–701.
3. Azzi A, Stocker A (2000) Vitamin E: Non-antioxidant roles. Prog Lipid Res 39: 231–255.
4. Munut-Bosch S, Alegre L (2002) The function of tocopherols and tocotrienols in plants. Crit Rev Plant Sci 21: 31–57.
5. Munut-Bosch S, Falk J (2004) New insights into the function of tocopherols in plants. Planta 218: 323–326.
6. Satler SE, Gilliland LU, Magdalanes-Lunback M, Pollard M, DellaPenna D (2004) Vitamin E is essential for seed longevity and for preventing lipid peroxidation during germination. Plant Cell 16: 1419–1429.
7. Yoshida Y, Niki E, Noguchi N (2003) Comparative study on the action of tocopherols and tocotrienols as antioxidant: Chemical and physical effects. Chem Phys Lipids 123: 63–75.
8. Raussel AHG, Rahman ARA, Yuen KH, Wong AR (2008) Arterial compliance and vitamin E blood levels with a self emulsifying preparation of tocotrienol rich vitamin E. Arch Pharm Res 31: 1212–1217.
9. Qureshi AA, Sami SA, Salser WA, Khan FA (2001) Synergistic effect of tocotrienol-rich fraction (TRF) of rice bran and lovastatin on lipid parameters in hypercholesterolemic humans. J Nut Bioch 12: 318–329.
10. Chang PN, Yap WN, Lee DTW, Ling MT, Wong YC, et al. (2009) Evidence of HGGT expression in the Falcon x Azhul RIL population. 

PLOS ONE | www.plosone.org 12 May 2014 | Volume 9 | Issue 5 | e96276

11. Sen CK, Khanna S, Roy S, Ryu H, Bahadduri P, Swaan PW, et al. (2003) Molecular basis of vitamin E action. Arch Pharm Res 26: 193–205.
12. Sen CK, Khanna S, Roy S (2006) Tocotrienols: Vitamin E beyond tocopherols. PLOS ONE 1: 1–10.
13. Chang PN, Yap WN, Lee DTW, Ling MT, Wong YC, et al. (2009) Evidence of HGGT expression in the Falcon x Azhul RIL population. 

14. Cooney RW, France AA, Harwood PJ, Hatch-Pigott V, Custer LJ, et al. (1993) Tocopherol and tocotrienol accumulation during development of caryopses from barley. J Amer Oil Chem Soc 70: 2141–2152.
15. Theriault A, Chao JT, Wang Q, Gapor A, Adeli K (1999) Tocotrienol: A review of its therapeutic potential. Clin Biochem 32: 309–319.
16. Khanna S, Roy S, Ryu H, Bahadduri P, Swaan PW, et al. (2003) Molecular basis of vitamin E action. Arch Pharm Res 26: 193–205.
17. Sen CK, Khanna S, Roy S (2006) Tocotrienols: Vitamin E beyond tocopherols. PLOS ONE 1: 1–10.
18. Qureshi AA, Burger WC, Peterson DM, Elson CE (1986) The structure of an oligopeptide modulating barley α-tocopherol biosynthesis. J Amer Oil Chem Soc 73: 2235–2240.
determination of tocochromanols in cereals. Eur Food Res Technol 215: 353–358.

48. Thompson JN, Hatina G (1979) Determination of tocopherols and tocotrienols in foods and tissues by high performance liquid chromatography. J Liquid Chrom 2: 327–344.

49. Islamovic E, Obert D, Oliver R, Harrison S, Ibrahim A, et al. (2013) Genetic dissection of grain beta-glucan and amylose content in barley (Hordeum vulgare L.) Mol Breed 31: 15–25.

50. Wang S, Baten CJ, Zeng Z-B (2003) Windows QTL Cartographer 2.5. Available at statgen.ncsu.edu/qtlcart/WQTL-Cart.htm. Dept. Statistics, North Carolina State University, Raleigh.

51. Islamovic E, Obert DE, Oliver RE, Marshall JM, Miclaus KJ, et al. (2013) A new genetic linkage map of barley (Hordeum vulgare L.) facilitates genetic dissection of height and spike length and angle. Food Crops Res 154: 91–99.

52. Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138: 963–971.

53. Zeng Z (1994) Precision mapping of quantitative trait loci. Genetics 136: 1457–1468.

54. Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, et al. (2011) Geneious v5.5, available from http://www.geneious.com.

55. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, et al. (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucl Acids Res 30: 325–327.

56. Oliver RE, Lazo GR, Lutz JD, Rubenfield MJ, Tinker NA, et al. (2011) Model SNP development for complex genomes based on hexaploid oat using high-throughput 454 sequencing technology. BMC Genomics 12:77.

57. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols 5:725-738.

58. Zhang Y (2007) Template-based modeling and free modeling by I-TASSER in CASP7. Proteins 69:100-117.

59. Vaesdeompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3: research0034.1 – 0034.11.