Genome-wide systematic characterization of HAK/KUP/KT gene family and their expression profiles during plant growth and in response to low K+ stress in Saccharum

CURRENT STATUS: UNDER REVISION

BMC Plant Biology  ▶ BMC Series

Xiaomin Feng
Guangzhou Sugarcane Industry Research Institute

Yongjun Wang
Fujian Agriculture and Forestry University

Nannan Zhang
Guangzhou Sugarcane Industry Research Institute

Zilin Wu
Guangzhou Sugarcane Industry Research Institute

Qiaoying Zeng
Guangzhou Sugarcane Industry Research Institute

Jiayun Wu
Guangzhou Sugarcane Industry Research Institute

Xiaobin Wu
Guangzhou Sugarcane Industry Research Institute

Lei Wang
Guangzhou Sugarcane Industry Research Institute

Jisen Zhang
Fujian Agriculture and Forestry University

Yongwen Qi  yongwen2001@126.com
Guangzhou Sugarcane Industry Research Institute

Corresponding Author

DOI:
10.21203/rs.2.13308/v1

SUBJECT AREAS
Plant Physiology and Morphology  Plant Molecular Biology and Genetics
KEYWORDS
Saccharum, HAK/KUP/KT, evolution, gene expression, low K+ stress
Abstract

Background Plant genomes contain large number of HAK/KUP/KT transporters, and they play important roles in potassium uptake and translocation, osmotic potential regulation, salt tolerance as well as root morphogenesis and plant development. Potassium deficiency in soil of main sugarcane planting area is serious. However, the HAK/KUP/KT gene family remains to be characterized in sugarcane (Saccharum). Results In this study, 30 HAK/KUP/KT genes were identified from Saccharum spontaneum. Phylogenetics, duplication events, gene structure and expression pattern were analyzed. Phylogenetic analysis of HAK/KUP/KT genes from 15 representative plants showed that this gene family were divided into four groups (clade I-IV). Both ancient whole-genome duplication (WGD) and recent gene duplication contributed to the expansion of HAK/KUP/KT gene family. Nonsynonymous to synonymous substitution ratio (Ka/Ks) analysis showed that purifying selection was the main force to drive the evolution of HAK/KUP/KT genes. The divergence time of HAK/KUP/KT gene family was estimated to range from 134.8 to 233.7 Mya based on Ks analysis, suggesting that it is an ancient gene family in plants. Gene structure analysis showed that HAK/KUP/KT genes was accompanied by intron gain/loss in the process of evolution. RNA-seq data analysis demonstrated that HAK/KUP/KT genes from clade II and III mainly displayed constitutive expression in various tissues, while most genes from clade I and IV had no or very low expression in the tested tissues at different developmental stages. SsHAK1 and SsHAK21 displayed upregulated expression in response to low K+ stress. Conclusions This study provided insight into the gene evolutionary history of HAK/KUP/KT genes. HAK7/9/18 were mainly expressed in the high photosynthetic zone and mature zone of stem. Moreover HAK7/9/18/25 were regulated by sunlight. SsHAK1 and SsHAK21 played important role in mediating potassium acquisition under limited K+ supply. Our results provide valuable information and key candidate genes for
further study on the function of HAK/KUP/KT genes in Saccharum.

**Background**

Potassium is one of the essential mineral nutrients for plant growth and development and also the most abundant monovalent cation in plant, accounting for about 2%~10% of plant dry weight [1]. Potassium is involved in many important physiological and biochemical processes such as cell turgor regulation, cell charge balance regulation, enzyme activity regulation and protein synthesis [1]. Symptoms of plant potassium deficiency are usually manifested as: weak stems, easy lodging; decreased tolerance to drought and cold; Proteins and chlorophyll are broken down and the leaves turn yellow, leading to tissue necrosis [2]. So potassium is of great importance to improve crop yield as well as quality.

Sugarcane is an important sugar and energy crop with long growth period, large biomass and large amount of potassium fertilizer absorption. On one hand, it is estimated that sugarcane needs to absorb about 2~2.5 kg of potassium to produce one ton of sugar [3, 4]. On the other hand, sugarcane is mainly cultivated in subtropical and tropical regions, where soil acidification and potassium leaching are common. The contents of total potassium and available potassium in the cultivated layer of these sugarcane area are low.

Plant cell maintain a relatively high and stable $K^+$ concentration (about 100~150 mM) in the cytosol, while $K^+$ concentration is highly variable at the range of 0.01~1 mM [5]. It is generally believed that there are two mechanisms for potassium uptake by plants, namely, high-affinity transport system (HATS) through potassium transporters at low external potassium concentrations ($< 0.2$ mM) and low-affinity transport system (LATS) through potassium channels at high potassium concentrations ($> 0.5$ mM) [6, 7]. According to the structure and function, potassium transporters in plants can be divided
into five families: (1) Shaker channels; (2) TPK (tandam-pore K\(^+\)) channel; (3) HAK (high affinity K\(^+\) transporter) /KUP (K\(^+\) uptake permease) /KT (K\(^+\) transporter); (4) HKT transporters, (5) CPAs (cation-proton antiporters) [2, 8]. Among them, the HAK/KUP/KT family is the largest one, which is widely distributed in bacteria, fungi and plants, but has not been found in animal cells [9].

According to the homology with bacterial KUP and fungal HAK transporters [10], plant HAK/KUP/KT transporters member, AtKUP1 and HvHAK1 were first cloned from Arabidopsis and barley [11, 12]. Both genes could complement the K\(^+\) uptake deficient strains of yeast, indicating that they had potassium transporter activity. Subsequently, several HAK/KUP/KT members were cloned and identified, such as AtKUP3 and AtHAK5 in Arabidopsis, OsHAK1 in rice and CaHAK1 in pepper, which were also proved to be highly compatible potassium transporters [13–16]. Based on comparative genomic methods, 13, 27 and 27 HAK/KUP/KT genes were identified in Arabidopsis, rice and maize [17–19]. And these predicted HAK/KUP/KT transporters were sorted into four clusters. The roles of HAK/KUP/KT K\(^+\) transporters in plants involve K\(^+\) acquisition, K\(^+\) translocation, salt tolerance, osmotic regulation, altering root morphology and shoot phenotype [7]. Expression of OsHAK1 was greatly induced in the root of K\(^+\)-starved rice while OsHAK5 was less expressed in root but abundantly in shoots [20, 21]. Some ions, particularly, Na\(^+\) and NH\(_4^+\) can impose additional effects on the expression of HAK/KUP/KT genes [22, 23].

Transcriptional regulation of HAK/KUP/KT K\(^+\) transporters is an universal mechanism for different plant species responding to K\(^+\)-starvation stress [8]. HAK/KUP/KT genes in cluster I, such as AtHAK5, OsHAK1, CaHAK1 and ThHAK5 display low expression levels both in roots and shoots under control conditions and are highly upregulated in roots upon K\(^+\)-
deficient stress [12–14, 16]. The other HAK/KUP/KT \( K^+ \) transporters especially in clusters II, III and IV exhibit diverse expression pattern [24], most of them do not show transcriptional regulation in response to \( K^+ \)-starvation [25]. In Arabidopsis, several transcription factors including DDF2 (dwarf and delayed flowering 2), JLO (jagged lateral organs), TFII_A (transcription initiation factor II_A gamma chain) and bHLH121 (basic helix-loop-helix 121) have been identified to bind the promoters of \( HAK5 \). The expression of these transcription factors increased and activated \( HAK5 \) under low \( K^+ \) and salt stress [26]. The activity of HAK/KUP/KT \( K^+ \) transporters are also regulated post-transcriptionally and/or post-translationally. \( AtHAK5 \) and its homologs from pepper and tomato can be activated by the CIPK23 (CBL-interacting protein kinase 23)/CBL (calcineurin B-like protein) complex [27].

In summary, tremendous works have been made on the functional research of plant HAK/KUP/KT potassium transporter, and a lot of important progresses have been made. However, the known functional \( HAK/KUP/KT \) genes were mainly identified in a few plants such as Arabidopsis, rice and maize, but their physiological functions and regulatory mechanisms in sugarcane still remains to be known. In this study, based on the newly released \( S. \) spontaneum genome [28], we identified the \( HAK/KUP/KT \) gene family in \( S. \) spontaneum. Phylogenetic relationships among different species, exon/intron organization and gene expression were analyzed. Altogether, these results provides valuable information and robust candidate genes for future functional analysis for genetic improvement of potassium utilization efficiency in sugarcane.

**Results**

**Identification of \( HAK \) genes in**
sugarcane

Based on comparative genomics, 29 members of *SbHAK* genes were identified from *sorghum bicolor* (sugarcane’s nearest relative). Using the protein sequences of sorghum *HAK* genes as reference, 30 distinct *S. spontaneum* *HAK* genes (Table 1), excluding alleles, were identified from the genome of tetraploidy *S. spontaneum* AP85-441 [28]. Each of these genes contain one to four alleles with an average of 3 (Additional file 1). The 30 *SsHAK* genes are distributed on seven *S. spontaneum* chromosome: chromosome 1 contains six genes; chromosome 2 contains seven genes; chromosome 3 contains four genes; chromosome 4 contains two genes; chromosome 5 contains five genes; chromosome 6 and 8 each contains three genes; No *SsHAK* genes were identified only on chromosome 7 (Additional file 1).

All the 30 predicted *SsHAK* proteins have a typical “K_trans” domain (PF02705), which is specific to HAK/KUP/KT potassium transporter family members. For consistency, these *SsHAK* genes were named based on the previously reported *O. sativa* *HAK* nomenclature and phylogenetic relationships [17]. If two *SsHAK* genes were equally close to a single *OsHAK* gene, then the same name was used followed by the letters “a” and “b” (Table 1).

Two paralogous *SsHAK* genes (*SsHAK19a* and *SsHAK19b*) were identified corresponding to the same sorghum gene, Sobic.006G062100, which may be caused by gene loss in sorghum or gene duplication in sugarcane. Amino acid number of the identified 30 *SsHAKs* ranged from 487 to 967, with an average of 758. The predicted isoelectric points (pI) of the *SsHAKs* varied from 5.88 to 9.26, the average pI was 8.15. The molecular weight ranged from 55.84 kDa to 106.49 kDa, with an average of 84.47 kDa (Table 1). Prediction of transmembrane domains in *SsHAK* proteins indicated that most of them contained 11 or 12 transmembrane helices, which was similar to the situation in sorghum. Subcellular
location of the SsHAK proteins predicted by WoLF PSORT showed that the SsHAK proteins were mainly on plasma membrane, which was best suited for their roles as transporters to maintain K⁺ homeostasis in sugarcane. In addition, the SsHAK proteins also showed localization on some organelles, including endoplasmic reticulum, vacuole, cytoplasm, golgi body and chloroplast. Protein sequence alignment of SsHAKs with their orthologs in sorghum showed that *S. spontaneum* and *sorghum bicolor* shared identities ranging from 81% to 98% with an average of 92.5% (Table 1). Four hundred thirty-five pair-wise protein sequence comparisons among these SsHAKs showed that SsHAK19a and SsHAK19b shared the highest identity (96%), other gene pairs had protein sequence similarities ranging from 28% to 82% with an average of 46%, indicating the SsHAKs are an ancient gene family with high sequence divergence (Additional file 2).

To investigate the possible evolutionary functional constraints after the split of sorghum and sugarcane, nonsynonymous to synonymous substitution ratio (Ka/Ks) between SsHAK and its orthologous gene in sorghum was calculated. The results showed that Ka/Ks ratios were less than 0.5, except for SsHAK13, suggesting that purifying selection was the main force to drive the evolution of HAK genes (Fig.1).

**Phylogenic analysis of HAK genes in *S. spontaneum* and some representative angiosperms**

To analyze the evolution of HAK gene family in *S. spontaneum* and different plants, a total of 278 HAK genes from 14 representative angiosperms and a HAK member from *Chlamydomonas reinhardtii* as the outgroup were used to construct phylogenetic tree using the Neighbor-Joining method (Fig. 2, Additional file 3). The 278 HAK genes included
6 in *Amborella trichopoda*, 8 in *Solanum lycopersicum*, 13 in *Vitis vinifera*, 8 in *Carica papaya*, 13 in *Arabidopsis thaliana*, 12 in *Ananas comosus*, 25 in *Brachypodium distachyon*, 27 in *Oryza sativa*, 28 in *Setaria italic*a, 28 in *Setaria viridis*, 27 in *Zea mays*, 29 in *Sorghum bicolor*, 30 in *Saccharum spontaneum* and 24 in *Saccharum* hybrid R570 [29].

Amino acid sequence of the 279 HAK/KUP/KT transporters from 15 representative plant species was uploaded as supplementary data (Additional file 4).

These *HAK* genes could be divided into four clades (I, II, III, IV) based on previously reported *OsHAKs* [17]. In *A. trichopoda*, the earliest diverging angiosperm, there were only 6 *HAK* genes, while in dicots and monocots, the number of *HAKs* ranged from 8 to 30 (Fig. 2, Fig. 3), indicating that the ancient whole-genome duplication (WGD) contributed to the expansion of the *HAK* gene family in both dicots and monocots. Clade II and clade III included *HAK* genes from all 14 angiosperm genomes, indicating that the progenitors of these genes may have already existed prior to the split of angiosperm (Fig. 2, Fig. 3). Clade I and clade IV mainly contained *HAK* genes from monocotyledons. Eighty-three *HAK* genes were identified in clade I, in which only one *HAK* gene was from *A. comosus* (Aco006685, homologous with SsHAK5) and *Arabidopsis* (AtHAK5) respectively, and the other 81 *HAK* genes were from all eight examined *Poaceae* species (Fig. 2, Fig. 3). Twenty-nine *HAKs* were grouped into clade IV, and only 2 out of them were from dicotyledon, these results indicated that the *HAKs* was unevenly distributed.

Based on the pairwise synonymous substitution rates (Ks) in *Sorghum bicolor* and *S. spontaneum* (Additional file 5), the divergence time among four clades of *HAK* family was estimated. The median values of pairwise Ks varied from 1.644 to 2.851, corresponding to a divergence time ranging from 134.8 to 233.7 Mya, suggesting that *HAKs* in the four clades were ancient and divergent. Moreover, the divergence time between two pairs of duplicated SsHAKs (*SsHAK5a/5b* and *SsHAK16a/16b*) ranged from 18.94 to 58.14 Mya.
Exon/intron organization of HAK family in S. spontaneum and other angiosperms

To investigate the structural characteristics and evolution of the HAK gene family, the exon/intron of the HAKs was mapped to the phylogenic tree, and the gene feature and pattern was analyzed (Fig. 2). The exon number in the HAK family of the 15 examined plant species ranged from 2 to 16 with an average of 8.4, and 217 out of 27977.8% HAK genes possessed 8 to 10 exons (Additional file 7 and 8). This result suggested that the last common ancestor (LCA) of angiosperm HAK genes had 8 to 10 exons.

The exon number of SsHAKs varied from 2 to 12, and half of the SsHAKs possessed 8 or 9 exons. The pattern of SsHAKs gene structure was similar to that of HAK genes from sorghum and maize in the same clade, which suggesting that the HAK gene structure in the Panicoideae was relatively conserved. In clade I, exon number of HAK genes varied from 2 to 12, which was also varied the most among these 4 clades. Noteworthy, HAK genes in the subfamily where SsHAK22 located had only 2 to 4 exons, however, the protein size remained consistent, which were likely due to the loss of intron. Clade II had the largest number of HAK genes, with 60 out of 98 HAKs possessed 9 exons, while 5 out of 9 SsHAKs harbored 8 exons. SsHAK3/8/10 had one less exon than their orthologous genes in sorghum; the first exon in SsHAK13 and seventh exon in SsHAK24 were smaller than the corresponding exons in sorghum, both cases caused shorter amino acid sequence in S. spontaneum (Table 1, Fig. 2). In clade III, exon number was relatively conserved, with 61 out of 68 HAK genes possessing 8 to 10 exons, while the gene size varied greatly, which
was mainly due to the different size of introns. In clade IV, exon number ranged from 2 to 8 with an average of 7, which was smaller than in other clades. Noteworthy, HAK genes in the subfamily where SsHAK4 located had only 2 to 5 exons, which was likely caused by intron loss during the process of evolution. The results indicated that HAKs underwent gene structure reconstruction under different evolutionary dynamics in S. spontaneum and other angiosperms in this study.

**Expression analysis of HAK genes in Saccharum species**

To study the expression profiles and potential functions of HAKs in Saccharum, we compared the gene expression patterns according to 4 sets of RNA-seq data: 1) Different developmental stages and tissues; 2) Leaf gradient; 3) Circadian rhythm; 4) Treatment under low potassium stress. FPKM values of HAK1, HAK7 and HAK20b in YT55 at 0 h, 6 h, 12 h, 24 h, 48 h and 72h under K⁺-starvation were verified by RT-qPCR. The relative expression level was positively correlated with FPKM value ($R^2 = 0.8419$, Additional file 9), suggesting the reliability of gene expression based on RNA-seq analysis.

**Expression pattern of HAKs in different tissues at different stages**

To study gene functional divergence among the Saccharum species, transcriptome profiles of HAKs between two Saccharum species, S. officinarum and S. spontaneum were analyzed based on RNA-seq at three developmental stage (seedling, premature and mature stage) in five different tissues including 2 leaf (leaf and leaf roll) and 3 stalks (immature, maturing and mature) (Fig. 4). Among the 30 HAK genes analyzed, 18 genes
showed very low or undetectable expression level in all examined tissues of the two *Saccharum* species. *HAK1* and *HAK2* had different expression pattern in the two *Saccharum* species, *HAK1* had higher expression levels in *S. spontaneum* than in *S. officinarum* and the expression level in leaf were higher than that in stems at three different stages, while *HAK2* had higher expression levels in *S. officinarum* than in *S. spontaneum*, and the expression level in stems were higher than that in leaf. *HAK8* mainly expressed in the upper stems, while the expression level in middle and lower stems were very low. *HAK9* and *HAK10* were observed to have higher expression level in stem than in leaf. *HAK18* was expressed in all examined tissues, with higher expression level especially in leaf at seedling stage and mature stem. Noteworthy, *HAK27* was highly expressed in leaf at all examined three stage, but the expression level in stem was very low or undetectable.

**Expression pattern of HAKs across leaf segment gradient**

To further explore functional divergence of *HAK* genes for photosynthesis in the source tissues, we studied the expression pattern of *HAKs* in continuously developing leaf segment gradient from *S. officinarum* and *S. spontaneum* (Fig. 5). *Saccharum* leaf was divided into four zone: basal zone (sink tissue), transitional zone (undergoing sink-source transition), maturing zone and mature zone (fully differentiated zone with active photosynthesis) following the method described in maize [30]. Consistent with the expression pattern at different developmental stages, 18 *HAK* genes (HAK3/4/5a/5b/12/13/14/15/16a/16b/17/19a/19b/20a/20b/21/22/26) showed very low or undetected expression level in all examined leaf segments, suggesting their limited roles in sugar transport (Fig. 5). *HAK1* and *HAK2* showed higher expression level in basal zone
than in other 3 zones. The expression level of HAK7 increased gradually from the base to
the tip of the leaf of S. spontaneum, while in S. officinarum, HAK7 displayed higher
expression level in the maturing zone than in other 3 zones. The expression level of HAK8
decreased gradually from the base to the tip of leaf both in S. officinarum and S.
spontaneum. HAK9 showed different expression pattern in S. spontaneum and S.
officinarum. In S. spontaneum, the expression level of HAK9 increased gradually from base
zone to maturing zone, then decreased in mature zone; while in S. spontaneum, the
expression level of HAK9 decreased from transition zone to maturing zone then increased
in the mature zone, and the expression level was much higher in S. officinarum,
suggesting the gene functional divergence after the split of this two Saccharum species.
HAK10 showed higher expression level in the transition zone in S. spontaneum, and higher
expression level in the mature zone in S. officinarum. HAK18 displayed higher expression
level in maturing zone both in S. spontaneum and S. officinarum, while HAK23 showed
higher expression level in basal zone both in the two Saccharum species. HAK25 displayed
higher expression level in maturing zone in S. officinarum, but showed higher expression
level in basal zone in S. spontaneum.

Expression pattern of HAKs during the circadian rhythms

Acting as an enzyme activator, potassium ions participate in a series of photosynthesis
process [31]. To analyze the expression pattern of HAKs during the diurnal cycles, we
investigated the transcriptome profiles of the mature leaves in the two Saccharum species
at 2 h intervals over 24 h period and at 4 h intervals over an additional 24 h. Consistent
with the transcriptome profiles at different developmental stages and in leaf segment
gradient, 18 genes (HAK3/4/5a/5b/12/13/14/15/16a/16b/17/19a/19b/20a/20b/21/22/26)
displayed very low or undetectable expression level in the two *Saccharum* species, further supporting their limited roles in growth and development (Fig. 6). Besides, *HAK8* and *HAK24* also showed low expression level over two 24 h period. *HAK1, HAK2, HAK7, HAK18* and *HAK27* showed higher expression level in *S. officinarum* than in *S. spontaneum*, while *HAK9* and *HAK10* displayed higher expression level in *S. spontaneum* than in *S. officinarum*. *HAK1* and *HAK2* were observed to have no diurnal expression pattern in the two *saccharum* species. *HAK7* displayed higher expression level at night than in the daytime and showed the lowest expression level at noon in *S. officinarum*, but showed no diurnal expression pattern in *S. spontaneum*; While *HAK10* displayed higher expression level at night than in the daytime in *S. spontaneum*, but showed no diurnal expression pattern in *S. officinarum*. *HAK9* displayed higher expression level at night than in the daytime in both *saccharum* species. *HAK18* and *HAK27* displayed higher expression in the morning in the two *Saccharum* species. These findings suggested the functional divergence of the *HAK* genes in diurnal rhythms.

**Expression pattern of *HAKs* under *K*⁺-deficient stress**

To investigate the functional divergence of *HAK* genes in response to low potassium stress in sugarcane, we studied the expression profiles of *HAKs* in root from the *Saccharum* hybrid variety YT55 at 0 h, 6 h, 12 h, 24 h, 48 h and 72 h under low *K*⁺ stress (Fig. 7). Among the 30 *HAK* genes analyzed, 14 genes (*HAK3/4/5a/5b/11/13/16a/16b/19a/19b/20a/22/26/27*) showed very low or undetectable expression level before and after exposure to low *K*⁺ stress. Noteworthy, *HAK1* showed strong induction in root under low *K*⁺ condition and reached the highest level at 24 h, then decreased subsequently at 48 h and 72 h. *HAK21* was strongly induced after exposure to
low K$^+$ stress within 12 hours, but was down-regulated to a low expression level subsequently. *HAK20b* was down-regulated within 12 hours, then showed up-regulated and to the highest level at 72 h. *HAK7, HAK10, HAK18* and *HAK24* were down-regulated after exposure to low K$^+$ stress. Some other *HAKs*, such as *HAK12/14/15/25* showed constitutive expression.

**Discussion**

The *HAK/KUP/KT* family of potassium transporters have been widely reported to be associated with K$^+$ transport across membranes in plants. Plant genome contains large number of HAK/KUP/KT transporters whose function involve the K$^+$ absorption and transport, salt tolerance, osmotic potential regulation and in controlling root morphology and shoot phenotyping [7]. However, genome-wide analysis of the *HAK/KUP/KT* gene family have not been conducted in *Saccharum* due to its complex genetic background. Recently released *S. spontaneum* genome allowed us to indentified 30 *HAK* genes from *S. spontaneum*. Besides, 248 *HAK* genes from other 13 representative plant species and an outgroup were used to construct phylogenetic tree and study the evolution of *HAK* genes in *Saccharum*. Furthermore, expression analysis based on RNA-seq revealed spatio-temporal expression and functional divergence in *HAK* family, which provides valuable information and robust candidate genes for future functional analysis.

**Evolution of HAK gene family in Saccharum and representative angiosperms**

WGD or polyploidization, gene loss and diploidization are considered to be important evolutionary forces in plants [32, 33]. Angiosperms, pan-core eudicots and monocots were
originated from the ε, γ and σ WGD event, which have been revealed by rigorous phylogenomic approach [33]. Recent study showed that pineapple had one fewer ancient ρ WGD event than other gramineous plants [34]. A. trichopoda is the earliest known angiosperms to have evolved separately from other angiosperms and has attracted much attention of botanists. In this study, 279 HAKs from 15 plant species representing major WGD events in angiosperms, together with the WGD information allow us to study HAKs gene evolution. HAKs from different plant species could be divided into four clades in duplicated descending order: clade IV, clade I, clade III and clade II. Based on the estimated divergence time among 4 clades of SsHAK gene family (134.8 to 233.7 Mya, Additional file 5), the SsHAK family in angiosperms probably occurred before σ GWD event in angiosperms (about 130 Mya) and after ε GWD event (about 220 Mya) [33]. The HAKs number in four clade varied greatly (from 29 to 98, Fig. 3), which is consistent with previous study that the HAKs were unevenly distributed in different clades among angiosperms [35]. In clade I, only one HAK gene member was from A. comosus and Arabidopsis respectively, while in Poaceae species the HAKs number ranged from 6 to 13. This result indicated that WGD or recent gene duplication contributed greatly to the expansion of HAKs. SsHAK5a/5b, SsHAK16a/16b, SsHAK19a/19b were from tandem duplications while SsHAK20a/20b maybe originated from transposed duplicate. The LCAs of SsHAK5 and SsHAK18 (in clade III) may occurred before the split of monocotyledonous and dicotyledonous plants. HAK5 was speculated to be lost in other dicotyledons except for Arabidopsis, which may be due to the gene functional redundancy of the HAK family. HAK18 was retained in all monocotyledonous and dicotyledonous plants, showing its functional constraint for the HAK family and the expression profiles analysis of HAK18 also confirmed this.

In clade II and clade III, SsHAK2 and SsHAK7 were retained from the ε GWD event, and in
dicotyledons these two orthologous genes were lost. *SsHAK3* and *SsHAK13* originated after *A. trichopoda* had evolved separately from other angiosperms. *SsHAK8, SsHAK9* and *SsHAK10* were assumed to be retained from the ε GWD event; *SsHAK11, SsHAK12, SsHAK15, SsHAK24* and *SsHAK25* were retained from σ GWD event since only monocotyledons contained these genes. *SsHAK14* and *SsHAK23* were assumed to be retained from the ε GWD event, but *HAK14* was probably lost in dicotyledons. Clade IV contained the least number of HAKs, *SsHAK4* and *SsHAK17* were originated before the split of monocotyledons and dicotyledons and after the split of *A. trichopoda* from angiosperms. The LCA of *SsHAK26* originated after the split of the Gramineae and pineapple.

*HAK* gene family in plant showed a less conserved exon/intron structure. The exon number in *Saccharum* ranged from 2 to 12 (Fig. 1, Additional file 7), and the variation range in *Saccharum* was larger than that in rice [17], maize [19] and wheat [36]. Three types of mechanism: exon/intron gain/loss, exonization/pseudo-exonization and insertion/deletion mainly lead to the exon-intron structure difference in paralogous or orthologous genes [37]. Although the gene structure of *SsHAKs* changed greatly, the protein size was relatively conserved, suggesting that exon-intron structure difference in *SsHAKs* was mainly caused by intron gain/loss. Clade I and clade IV belong to the older *HAK* family in *Saccharum*, so the *HAKs* in these two clades were speculated to have more intron gain/loss events based on the ‘introns-early’ theory during the lengthy evolutionary process [38, 39]. The results in this study also support this view since the variation of exon number in clade I and clade IV was much greater than that in clade II and clade III.

**Gene expression and functional divergence of HAKs in Saccharum**

Transcriptional regulation of K⁺ transporters is a common mechanism for plant species
responding to low K\(^+\) stress [8], and expression pattern analysis can provide insight into the potential functions of HAK gene family. In this study, we found that most HAK genes in clade I and clade IV showed low or undetectable expression levels across all examined samples. While most HAK genes in clade II and clade III were strongly expressed in all tested tissues. These results were consistent with the results of previous studies on HAK genes in rice [17], Arabidopsis [25] and wheat [36]. Five OsHAK genes (OsHAK2/10/15/23/25) from clade II and III were expressed in all examined tissues of three different genotype [17]. In Arabidopsis, 12 out of 13 HAK/KUP/KT genes were from clade II and III, most of them were expressed in roots, leaves, siliques and flowers [25]. Similarly, most TaHAKs in wheat belonging to clade II and III were constitutively expressed in all tissues [36].

Low K\(^+\) stress tends to induce the upregulated expression of K\(^+\) transporter genes [40]. previous studies demonstrated that the expression of OsHAK1 in rice [20], TaHAK1 in wheat [36] and PbrHAK1 in pear [41] were induced by K\(^+\) starvation. In this study, the expression level of SsHAK1 increased rapidly under low K\(^+\) stress, and this result is in good agreement with previous studies. Noteworthy, SsHAK21 was upregulated greatly after a short period of K\(^+\) starvation treatment and then rapidly downregulated (transient activation), indicating that SsHAK21 was involved in the low K\(^+\) stress response in sugarcane. And similar results were found in rice that OsHAK21 functions in the maintenance of ion homeostasis and tolerance to salt stress [42]. SsHAK1, SsHAK17 and SsHAK21 displayed upregulated expression, suggesting that they may play important role in maintaining normal growth and mediating potassium acquisition under K\(^+\) deficiency. In addition, nearly half the SsHAK genes were not expressed or had very low levels of expression in all tested tissues, at all stages or even under low K\(^+\) stress, this may be
caused by the gene functional redundancy due to WGD events in sugarcane.

Root system acquire K\(^+\) from soil solution then K\(^+\) is transported among compartment within cells and from root to shoot. A schematic model was proposed based on the expression profiles of the 30 SsHAK genes to illustrate the spatial and temporal gene expression in plant tissues and root hair cell of sugarcane (Fig. 8). HAK7/9/18 were mainly expressed in the tissues of maturing and mature stem and leave, indicating their important roles in K\(^+\) transport in these tissues. HAK7/9/18/25 also showed circadian rhythm expression pattern, suggesting these genes were regulated by sunlight. Low K\(^+\) stress induced up-regulated transcriptional expression of HAK genes. In Arabidopsis, transcription factor such as DDF2, JLO, ARF2, RAP2.11, TFII_A, bHLH121 directly bind the promoter of AtHAK5 to induce its expression and increase tolerance to low K\(^+\) and salt stress [26]. In this study, the expression of HAK1 and HAK21 were greatly up-regulated, which may also be positively regulated by the transcription factors (TFs) such as DDF2 and JLO and further experiment like yeast one-hybrid can be used to screen the TFs. AtHAK5 and its homologs from pepper and tomato can be activated by the CIPK23 (CBL-interacting protein kinase 23)/CBL1 (calcineurin B-like protein) complex [27]. In rice, OsHAK1/19/20 can be phosphorylated by a receptor like protein kinase, RUPO (ruptured pollen tube) [43]. In this study, CBL-CIPK complex and receptor-like kinase RUPO may also act as a regulator of high-affinity potassium transporters, such as HAK1 via phosphorylation-dependent interaction.

Conclusions

In this study, 30 HAK (high affinity K\(^+\) transporter) genes were identified through comparative genomics from sugarcane. Evolution analysis revealed that both ancient whole-genome duplication (WGD) and recent gene duplication contributed to the
expansion of the gene family and purifying selection was the main force to drive the evolution. *HAK/KUP/KT* genes was accompanied by intron gain/loss in the process of evolution. Expression analysis based on RNA-seq under low K⁺ stress and at different developmental stage revealed spatio-temporal expression and functional divergence in *HAK/KUP/KT* gene family. Altogether, these results provides valuable information and robust candidate genes for future functional analysis for genetic improvement of potassium utilization efficiency in sugarcane.

Methods

Plant materials

Two *Saccharum* species, LA-Purple (*S. officinarum*, 2n = 8x = 80, originated in USA and introduced into China, the plants were from Zhang’s lab in Fujian Agriculture and Forestry University) and SES–208 (*S. spontaneum*, 2n = 8x = 64, originated in USA and introduced into China, the plants were from Zhang’s lab in Fujian Agriculture and Forestry University), cultivated in the campus of Fujian Agricultural and Forestry University (Fuzhou, 119°16′48″E, 26°4′48″N, Fujian, China) were sampled to analyze gene expression pattern. For expression pattern analysis at different developmental stages: tissue samples were collected from 9-month-old plants (pre-mature plants) and 12-month-old plants (mature plants), including leaf roll, leaf (fully expanded leaf), top immature internode (i.e. Stem3), premature internode (i.e. stem 9 for LA-Purple and stem 6 for SES–208) and mature internode (i.e. stem 15 for LA-Purple and stem 9 for SES–208). The sugarcane internodes were numbered from top to bottom. Leaf and stem tissue at seedling stage were collected from 35-day-old plants as previously described [44, 45].

For expression pattern analysis of leaf gradient: two *Saccharum* species were grown in the greenhouse with light intensity of 350 μmol/m²/sec, 14:10 L/D, 30°C L/22°C D and 60%
relative humidity. The second leave of 15-day-old LA-Purple and 11-day-old SES208 after planting 3 h into the light period were sampled and cut into 15 1-cm segments. Samples were pooled from an average of 4 plants per biological replicate and three biological replicates in total were prepared, following previous approach described by Li et al [30]. For expression pattern analysis of diurnal cycle: leaves of the mature plants of LA-Purple and SES208 were sampled from a field at Fujian Agriculture and Forestry University for RNA extraction. One plant as one replicate, and three biological replicates were collected at 2 h interval from 6:00 a.m. on 2nd March, 2017 to 4:00 a.m. on 3rd March, 2017, 12 time points (6 a.m., 8 a.m., 10 a.m., noon, 2 p.m., 4 p.m., 6 p.m., 8 p.m., 10 p.m., midnight, 2 a.m., 4 a.m.) were chosen to sample for RNA-seq library construction. Then, mature leaves were sampled at 4 h interval from 6:00 a.m. on 3rd March, 2017 to 6:00 a.m. on 4th March, 2017, 7 time points (6 a.m., 10 a.m., 4 p.m., 8 p.m., midnight, 4 a.m., 6 a.m.) were chosen for RNA-seq library construction. The sunrise and sunset time on 2nd March, 2017 in Fuzhou were 6:25 a.m. and 6:05 p.m. respectively. The tissues collection followed the approach as previously described [34].

For expression pattern analysis at low potassium stress: *Saccharum* hybrid variety YT55 (This variety was bred by Guangzhou Sugarcane Industry Research Institute and was planted in breeding base for sugarcane in Wengyuan, Guangdong Province) was cultured at normal potassium level (3.0 mmol /L) for 20 days in greenhouse and then transferred to the K\(^+\) deficient nutrient solution (0.1 mmol /L) for starvation treatment. Mixed samples of root from 6 plants in a pot (a biological replicate, and three biological replicate in total were collected) were collected at 0 h, 6 h, 12 h, 24 h, 48 h and 72 h after starvation respectively, and stored in liquid nitrogen for total RNA isolation.

**Homology search analysis**
According to previous reports, the protein sequences of 13, 27 and 27 HAK/KUP/KT gene family identified in *Arabidopsis thaliana*, *Oryza sativa* and *Zea mays* [17–19] were obtained from Phytozome V12.1 (https://phytozome.jgi.doe.gov/pz/portal.html). With these protein sequences as queries, putative members of HAK/KUP/KT gene family were searched using BLASTP program in 14 representative angiosperm genomes, including 9 monocotyledons (*Saccharum* hybrid R570 [29], *Saccharum spontaneum*, *Sorghum bicolor*, *Zea mays*, *Setaria viridis*, *Setaria italic*, *Oryza sativa*, *Brachypodium distachyon* and *Ananas comosus*) and 5 dicotyledons (*Arabidopsis thaliana*, *Carica papaya*, *Vitis vinifera*, *Solanum lycopersicum* and *Amborella trichopoda*). Sequences with e-value$\leq 1e^{-10}$ were selected as HAK/KUP/KT candidates. Then, the identified HAK/KUP/KT proteins were subjected to conserved domains validation with the PFAM (https://pfam.xfam.org) and CDD (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi) databases. In addition, a HAK gene from *Chlamydomonas reinhardtii* was selected as the outgroup.

**Sequence and phylogenetic analysis**

Isoelectric points (pI) and relative molecular weight of the HAK/KUP/KT proteins were predicted by ExPASy (https://web.expasy.org/compute_pi/). The exon-intron structures were assessed with the TBtools [46]. TMHHM Server v.2.0 (http://www.cbs.dtu.dk/services/TMHMM/) was used to predict the transmembrane domain of the HAK/KUP/KT proteins. Subcellular location of the HAK/KUP/KT proteins were predicted by WoLF PSORT (https://www.genscript.com/wolf-psort.html).

The evolutionary history of 14 representative angiosperms was inferred by the neighbor-joining (NJ) method [47]. Based on the protein sequence alignment, the phylogenetic tree of the HAK/KUP/KT gene family was constructed using NJ methods. The construction of NJ
The tree was performed using MEGA7 [48] with the “pair deletion” and the “Poisson correction” model. Reliability of internal branches of the tree was valued by bootstrap test (1000 replicates) and the percentage are shown next to the branches. The non-synonymous substitution ratios (Ka), synonymous substitution ratios (Ks) and Ka/Ks ratios of the 30 pairs HAK/KUP/KT orthologous genes from sorghum and sugarcane were calculated by the Easy_KaKs calculation program (https://github.com/tangerzhang/FAFU-cgb/tree/master/easy_KaKs). Meanwhile, Fisher’s exact test for small samples was applied to verify the validity of Ka and Ks calculated by this method [49]. The divergence time was calculated by $T = \frac{Ks}{(2 \times 6.1 \times 10^{-9}) \times 10^{-6}}$ Mya [50].

**Analysis of the expression profiling of HAKs in *Saccharum* based on RNA-seq**

Five μg total RNA of each sample was used to construct cDNA libraries. The cDNA libraries were prepared according to the manufacturer’s protocol (TruSeq™ RNA Sample Preparation Kit, Illumina). The RNA-seq libraries were pooled and sequenced with 100 nt paired-end on Illumina HiSeq2500 platform at the Center for Genomics and Biotechnology, Fujian Agriculture and Forestry University. Raw data was aligned to released *S. spontaneum* AP85-441 reference gene models using TRINITY (https://github.com/trinityrnaseq/trinityrnaseq/wiki). Three independent software modules: Inchworm, Chrysalis and Butterfly combined in Trinity were applied sequentially to process large amount of RNA-seq reads. Fragments per kilobase per million mapped fragments (FPKM) was calculated to represent gene expression levels as previously described [51].

**Validation of HAK gene expression**
levels by RT-qPCR

The expression level of three HAK genes (HAK1, HAK7 and HAK20b) in the root of saccharum hybrid variety YT55 at 6 time points (0 h, 6 h, 12 h, 24 h, 48 h and 72 h) under K⁺-starvation were validated by RT-qPCR (Additional file 10). The reverse transcription, real time PCR program and method to calculate the relative expression levels was performed as previously described [52].

Abbreviations

bHLH121: basic helix-loop-helix 121; CBL: Calcineurin B-like protein; CIPK: CBL-interacting protein kinase; DDF2: Dwarf and delayed flowering 2; HAK/KUP/KT: High affinity K⁺ transporter/ K⁺ uptake permease / K⁺ transporter; JLO: Jagged lateral organs; Ka: Non-synonymous substitution ratio; Ks: Synonymous substitution ratio; LCA: Last common ancestor; FPKM: Fragments per kilobase per million mapped fragments; RT-qPCR: Reverse transcription-quantitative PCR; TF: Transcription factor; TFII_A: Transcription initiation factor II_A gamma chain; WGD: Whole genome duplication.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and material

The datasets supporting the conclusions of this article are included with in the article and its Additional files.
Competing interests
The authors declare that they have no competing interests.

Funding
This work was supported by the fund of GDAS’ Project of Science and Technology Development (2019GDASYL-0103028, used to purchase reagents and consumables for RT-qPCR); the special funds of the Guangdong Academy of Science (2019GDASYL-0104013 and 2017GDASCX-0105, used to pay for RNA-seq); Guangdong Provincial Team of Technical System Innovation for Sugarcane Sisal Hemp Industry (2019KJ104-04, used to pay for technicians to plant sugarcane); the Science and Technology Planting Project of Guangdong Province, China (2014B070705002, used for travel expenses between Fuzhou and Guangzhou); and the China Agricultural Research System (CARS201707, used for article-processing charge).

Authors’ contributions
XF, JZ and YQ conceived the study and designed the experiments. XF, YW, NZ, ZW, QZ, JW, XW, LW and JZ carried out the experiments and analyzed the data. XF wrote the manuscript. All authors read and approved the final paper.

Acknowledgments
We kindly thank Center for Genomics and Biotechnology, Haixia Institute of Science and Technology, Fujian Agriculture and Forestry University for providing access to Saccharum data.

Author details
1. Guangdong Key Lab of Sugarcane Improvement & Biorefinery, Guangdong Bioengineering Institute (Guangzhou Sugarcane Industry Research Institute),
Guangzhou 510316, China

2. Center for Genomics and Biotechnology, Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, College of Crop Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China

3. Guangzhou Guansheng Breeding Research Institute Guangzhou 511453, China

References

1. Leigh RA, Wyn Jones RG. A hypothesis relating critical potassium concentrations for growth to the distribution and function of this ion in the plant cell. New Phytol. 1984;97:1–13.

2. Ashley MK, Grant M, Grabov A. Plant responses to potassium deficiencies: a role for potassium transport proteins. J Exp Bot. 2006; 57(2):425–436.

3. Coale FJ, Izuno FT, Bottcher AB. Nutrient accumulation and removal by sugarcane grown on Everglades Histosols. Agron J. 1993;85:310–315.

4. Wood RA. The roles of nitrogen, phosphorus and potassium in the production of sugarcane in South Africa. Fertilizer Res. 1990;26:89–98.

5. White PJ. Improving potassium acquisition and utilisation by crop plants. J Plant Nutr Soil Sci. 2013;176(3):305–316.

6. Epstein E, Rains DW, Elzam OE. Resolution of dual mechanisms of potassium absorption by barley roots. Proc Natl Acad Sci USA. 1963;49(5):684–692.

7. Li W, Xu G, Alli A, Yu L. Plant HAK/KUP/KT K(+) transporters: Function and regulation. Semin Cell Dev Biol. 2018;74:133–141.

8. Wang Y, Wu WH. Potassium transport and signaling in higher plants. Annu Rev Plant Biol. 2013;64:451–476.

9. Corratge-Faillie C, Jabnoune M, Zimmermann S, Very AA, Fizames C, Sentenac H.
Potassium and sodium transport in non-animal cells: the Trk/Ktr/HKT transporter family. 

Cell Mol Life Sci. 2010;67(15):2511–2532.

10. Schleyer M, Bakker EP. Nucleotide sequence and 3’-end deletion studies indicate that the K(+) uptake protein kup from Escherichia coli is composed of a hydrophobic core linked to a large and partially essential hydrophilic C terminus. J Bacteriol. 1993;175(21):6925–6931.

11. Fu HH, Luan S. AtKuP1: a dual-affinity K+ transporter from Arabidopsis. Plant Cell. 1998;10(1):63–73.

12. Santa-Maria GE, Rubio F, Dubcovsky J, Rodriguez-Navarro A. The HAK1 gene of barley is a member of a large gene family and encodes a high-affinity potassium transporter. Plant Cell. 1997;9(12):2281–2289.

13. Banuelos MA, Garciadeblas B, Cubero B, Rodriguez-Navarro A. Inventory and functional characterization of the HAK potassium transporters of rice. Plant Physiol. 2002;130(2):784–795.

14. Gierth M, Maser P, Schroeder JI. The potassium transporter AtHAK5 functions in K(+) deprivation-induced high-affinity K(+) uptake and AKT1 K(+) channel contribution to K(+) uptake kinetics in Arabidopsis roots. Plant Physiol. 2005;137(3):1105–1114.

15. Kim EJ, Kwak JM, Uozumi N, Schroeder JI. AtKUP1: an Arabidopsis gene encoding high-affinity potassium transport activity. Plant Cell. 1998; 10(1):51–62.

16. Martinez-Cordero MA, Martinez V, Rubio F. Cloning and functional characterization of the high-affinity K+ transporter HAK1 of pepper. Plant Mol Biol. 2004;56(3):413-421.

17. Gupta M, Qiu X, Wang L, Xie W, Zhang C, Xiong L, Lian X, Zhang Q. KT/HAK/KUP potassium transporters gene family and their whole-life cycle expression profile in rice (Oryza sativa). Mol Genet Genomics. 2008; 280(5):437–452.
18. Mäser P, Thomine S, Schroeder JL, Ward JM, Hirschi K, Sze H, et al. Phylogenetic relationships within cation transporter families of Arabidopsis. Plant Physiol. 2001;126:1646–1667.

19. Zhang Z, Zhang J, Chen Y, Li R, Wang H, Wei J. Genome-wide analysis and identification of HAK potassium transporter gene family in maize (Zea mays L.). Mol Biol Rep. 2012;39(8):8465–8473.

20. Chen G, Hu Q, Luo L, Yang T, Zhang S, Hu Y, Yu L, Xu G. Rice potassium transporter OsHAK1 is essential for maintaining potassium-mediated growth and functions in salt tolerance over low and high potassium concentration ranges. Plant Cell Environ. 2015;38(12):2747–2765.

21. Yang T, Zhang S, Hu Y, Wu F, Hu Q, Chen G, Cai J, Wu T, Moran N, Yu L, et al. The role of a potassium transporter OsHAK5 in potassium acquisition and transport from roots to shoots in rice at low potassium supply levels. Plant Physiol. 2014;166(2):945–959.

22. Nieves-Cordones M, Miller AJ, Aleman F, Martinez V, Rubio F. A putative role for the plasma membrane potential in the control of the expression of the gene encoding the tomato high-affinity potassium transporter HAK5. Plant Mol Biol. 2008;68(6):521–532.

23. Rubio F, Nieves-Cordones M, Aleman F, Martinez V. Relative contribution of AtHAK5 and AtAKT1 to K⁺ uptake in the high-affinity range of concentrations. Physiol Plant. 2008;134(4):598–608.

24. Su H, Golldack D, Zhao C, Bohnert HJ. The expression of HAK-type K⁺ transporters is regulated in response to salinity stress in common ice plant. Plant Physiol. 2002;129(4):1482–1493.

25. Ahn SJ, Shin R, Schachtman DP. Expression of KT/KUP genes in Arabidopsis and the role of root hairs in K⁺ uptake. Plant Physiol. 2004; 134(3):1135–1145.
26. Hong JP, Takeshi Y, Kondou Y, Schachtman DP, Matsui M, Shin R. Identification and characterization of transcription factors regulating Arabidopsis HAK5. Plant Cell Physiol. 2013;54(9):1478-1490.

27. Ragel P, Rodenas R, Garcia-Martin E, Andres Z, Villalta I, Nieves-Cordones M, Rivero RM, Martinez V, Pardo JM, Quintero FJ, et al. The CBL-interacting protein kinase CIPK23 regulates HAK5-mediated high-affinity K\textsuperscript+ uptake in Arabidopsis roots. Plant Physiol. 2015;169(4):2863-2873.

28. Zhang J, Zhang X, Tang H, Zhang Q, Hua X, Ma X, Zhu F, Jones T, Zhu X, Bowers J, et al. Allele-defined genome of the autopolyploid sugarcane Saccharum spontaneum L. Nat Genet. 2018;50(11):1565-1573.

29. Garsmeur O, Droc G, Antonise R, Grimwood J, Potier B, Aitken K, Jenkins J, Martin G, Charron C, Hervouet C, et al. A mosaic monoploid reference sequence for the highly complex genome of sugarcane. Nat Commun. 2018;9(1):2638.

30. Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, Kebrom TH, Provat N, Patel R, Myers CR, et al. The developmental dynamics of the maize leaf transcriptome. Nat Genet. 2010;42(12):1060-1067.

31. Lu Z, Xie K, Pan Y, Ren T, Lu J, Wang M, Shen Q, Guo S. Potassium mediates coordination of leaf photosynthesis and hydraulic conductance by modifications of leaf anatomy. Plant Cell Environ. 2019; doi: 10.1111/pce.13553.

32. Edger PP, Pires JC. Gene and genome duplications: the impact of dosage-sensitivity on the fate of nuclear genes. Chromosome Res. 2009;17(5):699-717.

33. Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, et al. Ancestral polyploidy in seed plants and angiosperms. Nature. 2011;473(7345):97-100.

34. Ming R, VanBuren R, Wai CM, Tang H, Schatz MC, Bowers JE, Lyons E, Wang ML, Chen J,
Biggers E et al. The pineapple genome and the evolution of CAM photosynthesis. Nat Genet. 2015;47(12):1435–1442.

35. Nieves-Cordones M, Rodenas R, Chavanieu A, Rivero RM, Martinez V, Gaillard I, Rubio F. Uneven HAK/KUP/KT protein diversity among angiosperms: species distribution and perspectives. Front Plant Sci. 2016; 7:127.

36. Cheng X, Liu X, Mao W, Zhang X, Chen S, Zhan K, Bi H, Xu H. Genome-wide identification and analysis of HAK/KUP/KT potassium transporters gene family in wheat (Triticum aestivum L.). Int J Mol Sci. 2018;19(12).

37. Xu G, Guo C, Shan H, Kong H: Divergence of duplicate genes in exon-intron structure. Proc Natl Acad Sci U S A. 2012;109(4):1187-1192.

38. Jeffares DC, Mourier T, Penny D. The biology of intron gain and loss. Trends Genet. 2006;22(1):16-22.

39. Rogozin IB, Sverdlov AV, Babenko VN, Koonin EV. Analysis of evolution of exon-intron structure of eukaryotic genes. Brief Bioinform. 2005;6(2):118–134.

40. Wang Y, Wu WH. Regulation of potassium transport and signaling in plants. Curr Opin Plant Biol. 2017;39:123–128.

41. Wang Y, Lu J, Chen D, Zhang J, Qi K, Cheng R, Zhang H, Zhang S. Genome-wide identification, evolution, and expression analysis of the KT/HAK/KUP family in pear. Genome. 2018;61(10):755–765.

42. Shen Y, Shen L, Shen Z, Jing W, Ge H, Zhao J, Zhang W. The potassium transporter OsHAK21 functions in the maintenance of ion homeostasis and tolerance to salt stress in rice. Plant Cell Environ. 2015;38(12):2766-2779.

43. Liu L, Zheng C, Kuang B, Wei L, Yan L, Wang T. Receptor-like kinase RUPO interacts with potassium transporters to regulate pollen tube growth and integrity in rice. PLoS Genet. 2016;12(7):e1006085.
44. Chen Y, Zhang Q, Hu W, Zhang X, Wang L, Hua X, Yu Q, Ming R, Zhang J. Evolution and expression of the fructokinase gene family in *Saccharum*. BMC Genomics. 2017;18(1):197.

45. Zhang Q, Hu W, Zhu F, Wang L, Yu Q, Ming R, Zhang J. Structure, phylogeny, allelic haplotypes and expression of sucrose transporter gene families in *Saccharum*. BMC Genomics. 2016;17:88.

46. Chen CJ, Xia R, Chen H, He YH. TBtools, a Toolkit for biologists integrating various HTS-data handling tools with a user-friendly interface. BioRxiv Preprint Mar. 27, 2018.

47. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987; 4(4):406-425.

48. Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33(7):1870-1874.

49. Graham JGU. Fisher’s Exact Test. Journal of the Royal Statistical Society Series A (Statistics in Society). 1992;155(3):395–402.

50. Lynch M, Conery JS. The evolutionary fate and consequences of duplicate genes. Science. 2000;290(5494):1151-1155.

51. Hu W, Hua X, Zhang Q, Wang J, Shen Q, Zhang X, Wang K, Yu Q, Lin YR, Ming R, et al. New insights into the evolution and functional divergence of the SWEET family in *Saccharum* based on comparative genomics. BMC Plant Biol. 2018;18(1):270.

52. Wang Y, Hua X, Xu J, Chen Z, Fan T, Zeng Z, Wang H, Hour AL, Yu Q, Ming R, et al. Comparative genomics revealed the gene evolution and functional divergence of magnesium transporter families in *Saccharum*. BMC Genomics. 2019;20(1):83.

**Table 1**

| Gene          | AA^a  | pI^b  | Mw^c (kDa) | TMS^d | P.L.^e | Gene | AA^a  |
|---------------|-------|-------|------------|-------|--------|------|-------|
| Sobic.00      | 788   | 8.75  | 87.13      | 12    | PM     | SsHAK1 | 780   |
| Gene     | Accession | Value  | Standard Value | Time   | Location  | Protein | Value  |
|----------|-----------|--------|----------------|--------|-----------|---------|--------|
| Sobic.00 | 3G06130   | 783    | 8.91           | 12     | PM        | SsHAK2  | 788    |
| Sobic.00 | 3G1810    | 811    | 8.4            | 10     | PM/ER     | SsHAK3  | 785    |
| Sobic.00 | 3G16440   | 706    | 8.37           | 9      | PM/ER     | SsHAK4  | 702    |
| Sobic.00 | 3G06130   | 775    | 8.78           | 11     | PM        | SsHAK5a | 705    |
| Sobic.00 | 3G1360    | 775    | 8.54           | 11     | PM        | SsHAK5b | 750    |
| Sobic.00 | 3G1370    | 788    | 8.8            | 13     | PM        | SsHAK7  | 818    |
| Sobic.00 | 3G06210   | 805    | 7.36           | 12     | PM/Cyto   | SsHAK8  | 770    |
| Sobic.00 | 3G06130   | 792    | 6.96           | 12     | PM/Cyto   | SsHAK9  | 743    |
| Sobic.00 | 3G1810    | 820    | 8.37           | 10     | PM/ER     | SsHAK10 | 755    |
| Sobic.00 | 3G1370    | 805    | 8.33           | 13     | PM/ER     | SsHAK11 | 719    |
| Sobic.00 | 3G06210   | 790    | 8.21           | 14     | PM        | SsHAK12 | 509    |
| Sobic.00 | 3G06130   | 779    | 8.97           | 12     | PM/Cyto   | SsHAK13 | 757    |
| Sobic.00 | 3G1810    | 843    | 5.71           | 12     | PM/ER     | SsHAK14 | 811    |
| Sobic.00 | 3G1370    | 743    | 8.85           | 12     | PM/ER     | SsHAK15 | 852    |
| Sobic.00 | 3G06210   | 817    | 8.91           | 12     | PM        | SsHAK16a | 487   |
| Sobic.00 | 3G06130   | 810    | 8.61           | 11     | PM/ER     | SsHAK16b | 802   |
| Sobic.00 | 3G1810    | 708    | 8.77           | 12     | PM        | SsHAK17 | 701    |
| Sobic.00 | 3G1370    | 787    | 8.69           | 14     | PM/ER     | SsHAK18 | 788    |
| Sobic.00 | 3G06210   | 746    | 7.29           | 12     | PM/Golgi  | SsHAK19a | 767   |
| Sobic.00 | 3G06210   | 746    | 7.29           | 12     | PM/Golgi  | SsHAK19b | 730   |
| Sobic.00 | 3G16000   | 735    | 8.46           | 12     | PM/ER     | SsHAK20a | 730   |
| Sobic.00 | 3G06170   | 788    | 8.66           | 11     | PM/Cyto   | SsHAK20b | 794   |
| Sobic.00 | 3G06130   | 828    | 8.51           | 11     | PM/Cyto   | SsHAK21  | 818    |
| Sobic.00 | 3G06130   | 931    | 8.61           | 12     | PM/Chlo   | SsHAK22  | 967    |
| Sobic.00 | 3G06130   | 852    | 6.78           | 12     | PM/ER     | SsHAK23  | 846    |
| Sobic.00 | 3G06130   | 773    | 8.39           | 12     | PM/Chlo   | SsHAK24  | 698    |
| Accession | Amino Acids | pI  | Molecular Weight | Membrane Location | Protein Name | Molecular Weight |
|-----------|-------------|-----|------------------|-------------------|--------------|------------------|
| Sobic.00 4G25070 | 774 | 7.34 | 86.29 | 13 | PM/ER | SsHAK25 | 800 |
| Sobic.00 7G20990 | 774 | 9.08 | 82.47 | 10 | PM/Chlo | SsHAK26 | 744 |
| Sobic.00 1G18430 | 814 | 8.32 | 91.82 | 11 | PM/ER | SsHAK27 | 812 |

PM = plasma membrane, ER = endoplasmic reticulum, Vacu = vacuole, Cyto = cytoplasm, Golgi = golgi body, Chlo = chloroplast

a Amino acids number of HAK protein sequence

b Isoelectric point (pI) predicted by ExPASy (https://web.expasy.org/compute_pi/)

c Molecular weight (Mw) predicted by ExPASy (https://web.expasy.org/compute_pi/)

d Number of transmembrane domains possessed by HAK predicted by TMHMM Server v.2.0 (http://www.cbs.dtu.dk/services/TMHMM/)

e Subcellular location of the HAK proteins predicted by WoLF PSORT (https://www.genscript.com/wolf-psort.html)

f Protein sequence similarity between sorghum and sugarcane calculated by BLASTP

Figures
Figure 1: 
Nonsynonymous (Ka) and synonymous (Ks) substitution ratios of SsHAKs and ortholog in sorghum. Ka/Ks ratio was calculated by the Easy_KaKs calculation program (https://github.com/tangerzhang/FAFU-cgb/tree/master/easy_KaKs).
Figure 2

Phylogeny and schematic diagram for intron/exon organization of HAK/KUP/KT genes from 15 plant species. (A) Clade II and Clade III. (B) Clade I and Clade IV.
Figure 3
Phylogenetic relationships of HAK/KUP/KT families based on the current data for angiosperms.

| Species                             | Outgroup | Clade II | Clade III | Clade I | Clade IV | Total |
|-------------------------------------|----------|----------|-----------|---------|----------|-------|
| Saccharum hybrid R570               | 0        | 4        | 3         | 13      | 4        | 24    |
| Saccharum spontaneum                | 0        | 9        | 6         | 12      | 3        | 30    |
| Sorghum bicolor                     | 0        | 9        | 6         | 11      | 3        | 29    |
| Zea mays                            | 0        | 9        | 6         | 9       | 3        | 27    |
| Setaria viridis                     | 0        | 9        | 6         | 11      | 2        | 28    |
| Setaria italica                     | 0        | 9        | 6         | 11      | 2        | 28    |
| Oryza sativa                        | 0        | 9        | 6         | 8       | 4        | 27    |
| Brachypodium distachyon             | 0        | 9        | 6         | 6       | 4        | 25    |
| Ananas comosus                      | 0        | 5        | 4         | 1       | 2        | 12    |
| Arabidopsis thaliana                | 0        | 6        | 6         | 1       | 0        | 13    |
| Carica papaya                       | 0        | 5        | 3         | 0       | 0        | 8     |
| Vitis vinifera                      | 0        | 6        | 6         | 0       | 1        | 13    |
| Solanum lycopersicum                | 0        | 5        | 2         | 0       | 1        | 8     |
| Amborella trichopoda                | 0        | 4        | 2         | 0       | 0        | 6     |
| Chlamydomonas reinhardtii           | 1        | 0        | 0         | 0       | 0        | 1     |
| Total                               | 1        | 98       | 68        | 83      | 29       | 279   |

*α, β, γ, ε, ζ, η, σ, τ: Whole Genome Duplication Events

| Species | Seedling | Pre-mature | Mature |
|---------|----------|------------|--------|
|         | Leaf     | Stem       | Leaf   | Stem   | Leaf   | Stem | Leaf | Stem | Leaf | Stem | Leaf | Stem | Leaf | Stem | Leaf | Stem | Leaf | Stem | Leaf | Stem | Leaf | Stem | Leaf | Stem | Leaf | Stem | Leaf | Stem | Leaf | Stem | Leaf | Stem |
| HAK1    | So 15.10 | 5.95       | 30.59  | 4.82   | 26.84  | 31.26 | 31.92 | 16.49 | 8.37 |
|         | St 56.05 | 9.52       | 54.23  | 5.08   | 45.33  | 44.97 | 45.18 | 36.39 | 32.10 | 16.77 |
| HAK2    | So 28.22 | 54.41      | 12.70  | 11.58  | 59.68  | 23.21 | 24.35 | 68.78 | 53.15 | 65.37 |
|         | St 18.24 | 34.72      | 6.58   | 5.28   | 51.41  | 13.75 | 11.27 | 51.92 | 28.22 | 31.69 |
| HAK3    | So 0.03  | 0.00       | 0.00   | 0.00   | 0.00   | 0.00  | 0.00  | 0.00  | 0.00  |
|         | St 0.02  | 0.00       | 0.00   | 0.00   | 0.00   | 0.00  | 0.00  | 0.00  | 0.00  |
| HAK4    | So 0.00  | 0.05       | 0.00   | 0.00   | 0.00   | 0.00  | 0.00  | 0.00  | 0.00  |
|         | St 0.00  | 0.00       | 0.00   | 0.00   | 0.00   | 0.00  | 0.00  | 0.00  | 0.00  |
| HAK5a   | So 0.15  | 0.07       | 0.00   | 0.00   | 0.00   | 0.00  | 0.00  | 0.00  | 0.00  |
|         | St 1.01  | 0.14       | 0.35   | 0.38   | 0.00   | 0.51  | 0.47  | 0.07  | 0.46  |
| HAK5b   | So 0.13  | 0.23       | 0.16   | 0.48   | 0.00   | 1.29  | 2.25  | 0.57  | 0.72  |
|         | St 0.12  | 0.15       | 0.32   | 0.00   | 0.00   | 1.73  | 2.27  | 0.41  | 0.12  |
| HAK7    | So 6.12  | 6.10       | 12.95  | 4.24   | 25.55  | 24.10 | 27.52 | 11.46 | 43.13 |
|         | St 12.13 | 2.82       | 23.10  | 32.97  | 1.24   | 42.94 | 40.23 | 2.37  | 11.32 |
| HAK8    | So 6.22  | 8.54       | 0.94   | 0.59   | 11.11  | 2.77  | 1.59  | 12.66 | 1.53  |
|         | St 5.80  | 15.88      | 1.39   | 0.59   | 25.16  | 0.94  | 0.71  | 18.60 | 4.80  |
| HAK9    | So 21.00 | 11.54      | 13.12  | 5.45   | 28.02  | 15.39 | 8.60  | 39.24 | 18.75 |
|         | St 36.10 | 30.08      | 33.22  | 17.44  | 43.06  | 12.37 | 12.48 | 42.62 | 29.79 |
| HAK10   | So 9.27  | 16.30      | 2.29   | 1.79   | 12.81  | 6.17  | 2.53  | 13.98 | 14.55 |
|         | St 20.40 | 17.35      | 12.65  | 4.95   | 33.15  | 11.87 | 9.49  | 19.68 | 46.27 |
| HAK11   | So 19.68 | 34.14      | 6.12   | 4.39   | 6.32   | 15.13 | 19.48 | 7.35  | 16.73 |
|         | St 14.02 | 18.61      | 1.99   | 1.77   | 4.61   | 10.92 | 10.70 | 4.98  | 12.93 |
| HAK12   | So 2.06  | 3.87       | 0.59   | 0.15   | 0.94   | 1.50  | 0.59  | 2.77  | 2.74  |
|         | St 1.09  | 1.81       | 1.56   | 1.48   | 5.68   | 6.13  | 5.20  | 5.33  | 7.82  |
| HAK13   | So 0.11  | 0.12       | 0.00   | 0.00   | 0.09   | 0.24  | 0.23  | 0.03  | 0.11  |
|         | St 0.04  | 0.09       | 0.05   | 0.00   | 0.11   | 0.86  | 1.35  | 0.00  | 0.02  |

36
Figure 4

The expression pattern of HAK/KUP/KT genes based on FPKM in different tissues of different stages in S. officinarum and S. spontaneum.

| Gene ID | Basal zone | Transition zone | Maturing zone | Mature zone |
|---------|------------|-----------------|---------------|-------------|
| HAK1    | 35.79      | 19.76           | 6.00          | 4.42        |
| HAK2    | 34.47      | 17.22           | 16.40         | 11.48       |
| HAK3    | 38.73      | 17.83           | 10.31         | 11.00       |
| HAK4    | 0.00       | 0.00            | 0.00          | 0.00        |
| HAK5a   | 0.00       | 0.00            | 0.00          | 0.00        |
| HAK5b   | 0.00       | 0.00            | 0.00          | 0.00        |
| HAK6    | 6.43       | 3.97            | 4.54          | 6.90        |
| HAK7    | 77.52      | 54.47           | 24.73         | 13.85       |
| HAK8    | 31.18      | 36.40           | 10.28         | 15.15       |
| HAK9    | 66.86      | 94.96           | 119.61        | 118.64      |
| HAK10   | 3.51       | 25.95           | 16.29         | 1.13        |
| HAK11   | 15.40      | 21.85           | 11.49         | 14.66       |

FPKM

0 50 100
The expression pattern of HAK/KUP/KT genes based on FPKM across leaf gradients in *S. officinarum* and *S. spontaneum*.
The expression pattern of HAK/KUP/KT genes based on FPKM during the diurnal cycles in *S. officinarum* and *S. spontaneum*.
Figure 7

(A) The expression pattern of HAK/KUP/KT genes based on FPKM under low K+ stress in Saccharum hybrid YT55. (B) relative expression level detected by RT-qPCR
Schematic models for roles of HAKs based on gene expression profiles in sugarcane. In maturing and mature zone of leave and stem, HAK7/9/18 were the mainly expressed genes. Moreover, these genes also presented a diurnal expression pattern. HAK25 was mainly expressed in the maturing and mature zone of leaf tissues; while HAK2 was mainly expressed in the stem. Low K+ stress induced upregulated expression of HAK1 and HAK21. Transcription factors such as DDF2 and JLO may directly bind to the promoters of HAK1/21 to induce gene expression and subsequently promote HAK transporter such as HAK1 and HAK 21 to acquire K+ in root. HAK1 may be phosphorylated and activated by the CBL1-CIPK23 complex or receptor like kinase, RUPO (ruptured pollen tube). K+ concentration in the vacuole is highly varied to maintain the cellular K+
homeostasis. Some HAK transporter, such as HAK10 located in the tonoplast of vacuole may play a role in regulating the K+ concentration in vacuole.

Supplementary Files

This is a list of supplementary files associated with the primary manuscript. Click to download.

Additional file 5.docx
Additional file 6.docx
Additional file 1.xlsx
Additional file 7.xlsx
Additional file 9.docx
Additional file 10.docx
Additional file 2.xlsx
Additional file 8.docx
Additional file 4.txt
Additional file 3.pdf