



OPEN

Comparative de novo transcriptome analysis identifies salinity stress responsive genes and metabolic pathways in sugarcane and its wild relative *Erianthus arundinaceus* [Retzius] Jeswiet

P. Vignesh^{1,6}, C. Mahadevaiah^{1,6}✉, R. Parimalan^{2,3}, R. Valarmathi¹, S. Dharshini¹, Singh Nisha^{4,5}, G. S. Suresha¹, S. Swathi¹, H. K. Mahadeva Swamy¹, V. Sreenivasa¹, K. Mohanraj¹, G Hemaprabha¹, Ram Bakshi¹ & C. Appunu¹✉

Erianthus arundinaceus [Retzius] Jeswiet, a wild relative of sugarcane has a high biomass production potential and a reservoir of many genes for superior agronomic traits and tolerance to biotic and abiotic stresses. A comparative physiological, anatomical and root transcriptome analysis were carried out to identify the salt-responsive genes and metabolic pathways associated with salt-tolerant *E. arundinaceus* genotype IND99-907 and salinity-sensitive sugarcane genotype Co 97010. IND99-907 recorded growth of young leaves, higher proline content, higher relative water content, intact root anatomical structures and lower Na⁺/K⁺, Ca²⁺/K⁺ and Mg²⁺/K⁺ ratio as compared to the sugarcane genotype Co 97010. We have generated four de novo transcriptome assemblies between stressed and control root samples of IND99-907 and Co 97010. A total of 649 and 501 differentially expressed genes (FDR<0.01) were identified from the stressed and control libraries of IND99-907 and Co 97010 respectively. Genes and pathways related to early stress-responsive signal transduction, hormone signalling, cytoskeleton organization, cellular membrane stabilization, plasma membrane-bound calcium and proton transport, sodium extrusion, secondary metabolite biosynthesis, cellular transporters related to plasma membrane-bound trafficking, nucleobase transporter, clathrin-mediated endocytosis were highly enriched in IND99-907. Whereas in Co 97010, genes related to late stress-responsive signal transduction, electron transport system, senescence, protein degradation and programmed cell death, transport-related genes associated with cellular respiration and mitochondrial respiratory chain, vesicular trafficking, nitrate transporter and fewer secondary metabolite biosynthetic genes were highly enriched. A total of 27 pathways, 24 biological processes, three molecular functions and one cellular component were significantly enriched (FDR≤0.05) in IND99-907 as compared to 20 pathways, two biological processes without any significant molecular function and cellular components in Co 97010, indicates the unique and distinct expression pattern of genes and metabolic pathways in both genotypes. The genomic resources developed from this study is useful for sugarcane crop improvement through development of genic SSR markers and genetic engineering approaches.

¹ICAR-Sugarcane Breeding Institute, Coimbatore, India. ²ICAR-National Bureau of Plant Genetic Resources, New Delhi, India. ³Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Brisbane, Australia. ⁴ICAR-National Institute for Plant Biotechnology, New Delhi, India. ⁵Institute for Genomic Diversity, Cornell University, Ithaca, NY 14853, USA. ⁶These authors contributed equally: P. Vignesh and C. Mahadevaiah. ✉email: C.Mahadevaiah@icar.gov.in; cappunu@gmail.com

Abbreviations

BUSCO	Benchmarking universal single-copy orthologs
DEGs	Differentially expressed genes
FDR	False discovery rate
GO	Gene ontology
SEM-EDX	Scanning electron microscopy—energy dispersive X-ray analysis
ICP-OES	Inductively coupled plasma—optical emission spectrometry
KEGG	Kyoto encyclopaedia of genes and genomes
NCBI-nr	National Center for biotechnology information-non-redundant database
Pfam	Protein family database
qRT-PCR	Quantitative real-time polymerase chain reaction
QUAST	Quality assessment tool for genome assemblies
TF	Transcription factors

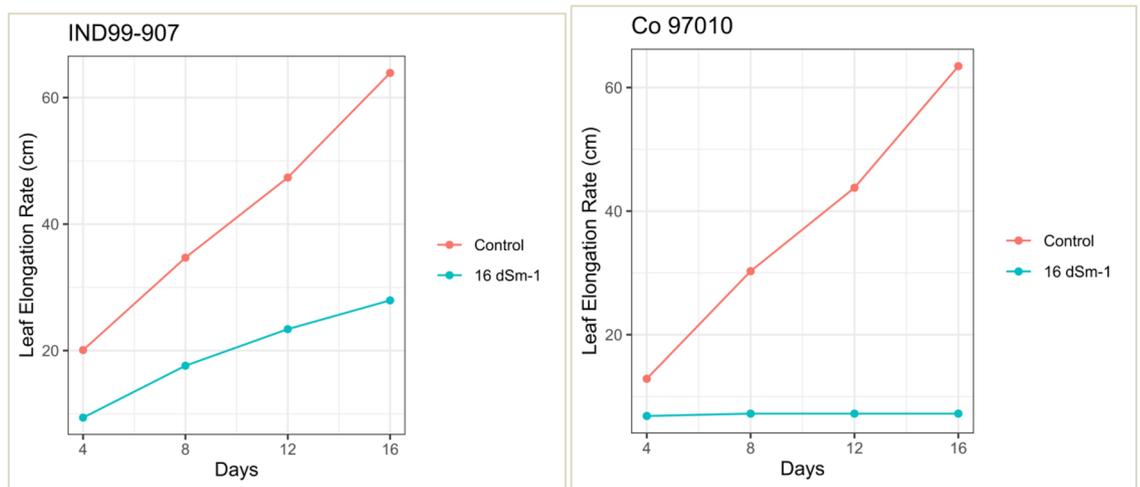
Sugarcane is an industry-oriented crop grown in the tropics and sub-tropics in the world. Sugarcane is a potential source of basic material for the manufacturing of sugar, ethanol, bioenergy, and biodegradable products. Traditional sweeteners such as jaggery, khandsari, and brown sugars having immense medicinal value are also produced from sugarcane^{1,2}. Globally, sugarcane was cultivated in 28.19 million hectares which produced 2059.74 million tonnes of canes with a productivity of 72.80 t/ha during 2019³. In India, sugarcane was cultivated in 5.06 million hectares and 405.42 million tonnes was produced with a productivity of 80.10 tonnes/hectare during 2019³. The growth, productivity, and juice quality of sugarcane are affected by abiotic stresses viz., cold, salinity, and drought. Nearly 10% of arable land or 25–30% of irrigated lands were affected by salinity in the world⁴. After the advent of modern agriculture, soil salinity has become a major environmental issue and underground water used for irrigation also contributes to the soil salinity⁵. Sugarcane is highly sensitive to salinity⁶ and affects both biomass accumulation and juice quality parameters⁷. Therefore, the development of saline tolerant varieties, gene pools, and genomic resources are helpful in sugarcane crop improvement through conventional and biotechnological approaches.

The sessile plants are sensing the salinity stress-induced osmotic and ionic stresses through the activation of calcium signalling and salt-overly sensitive pathways for exclusion of sodium^{8,9}. Glycosyl inositol phosphoryl ceramide sphingolipids are also involved in the sensing of salinity stress¹⁰. The salinity stress firstly reduces the water uptake causing the salinity-induced osmotic stress and secondly, increases the concentration of cytotoxic ions which causes ionic stress^{5,10}. The plant possesses several mechanisms to tolerate the salinity stress such as (i) accumulation of low molecular weight, water-soluble free state compounds (proline, betaine, water-soluble sugars) which help in the maintenance of osmotic adjustment and plant metabolic activities (ii) selective ion-uptake such as the exclusion of sodium and maintaining higher cytosolic K^+/Na^+ (iii) nullifying the effect of ROS by enzymes such as catalase, SOD and APX (iv) salinity-tolerant genes associated with sodium-hydrogen antiporter activities of vacuolar membrane, plasma membrane and, genes related to ROS scavenging enzymes¹¹. Besides, many gene networks related to signal transduction, hormone signalling, biosynthesis of secondary metabolites, amino acids and transporters significantly contribute to the salinity tolerance mechanisms in plants¹². The comparative global gene expression studies are certainly helping to dissect the various genes and metabolic pathways associated with salinity or abiotic stress tolerances in plants¹³.

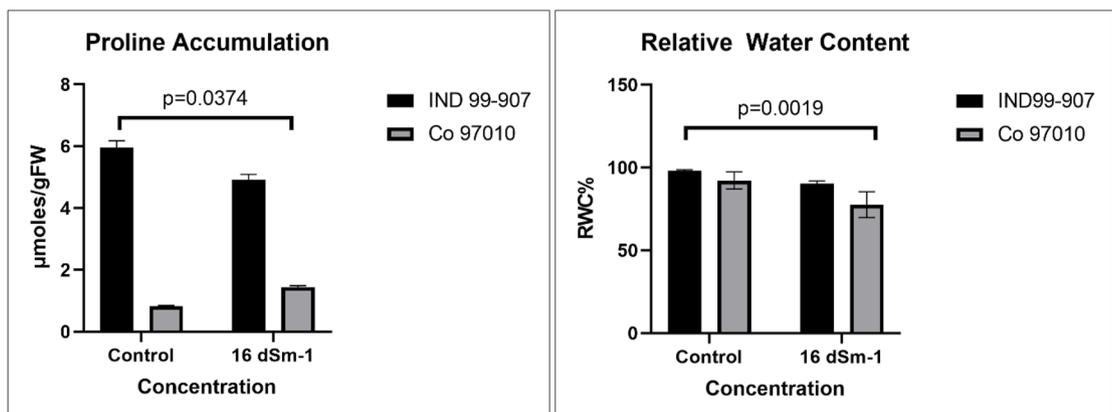
Sugarcane wild relative *E. arundinaceus* is a potential donor for many genes related to biomass and tolerance to biotic and abiotic stresses^{14–16}. Introgression of many genes from *E. arundinaceus* into sugarcane through conventional and biotechnological approaches has improved the agronomic performance of sugarcane varieties^{17,18}. We have isolated, characterized, and overexpressed many genes from *E. arundinaceus* such as heat shock protein 70^{14,19}, DREB2²⁰, glyoxalase gene¹⁵, α -expansin 1²¹ and chilling tolerant divergence 1 (COLD1) gene²², which shows the genetic importance of *E. arundinaceus* in improving the tolerances to biotic and abiotic stresses in sugarcane. There were several RNASeq studies in sugarcane for many traits such as agronomic traits, cold stress, sucrose, lignin, red stripe and smut disease except for salinity tolerance^{23–28}. Hence, we performed the comparative salt transcriptome studies to identify the salinity stress-responsive genes and metabolic pathways in salt-tolerant *E. arundinaceus* accession IND99-907 collected from saline soils Ernakulum district, Kerala, India²⁹ and salt-sensitive genotype Co 97010^{30,31}. From this study, we have identified many differentially expressed genes, enriched metabolic pathways and GO terms associated with salt tolerance in salt-tolerant *E. arundinaceus* accession IND99-907. Our studies showed the enrichment of 27 pathways, 24 biological processes, three molecular functions and one cellular component in IND99-907 as compared to 20 pathways, two biological processes without any significant molecular function and cellular components in Co 97010 (FDR ≤ 0.05), which specifies the unique and distinct expression pattern of genes and metabolic pathways regulating the salinity stress in IND99-907 and Co 97010. The genomic resources developed from this study are useful in sugarcane crop improvement through development of genic markers and advanced biotechnological approaches.

Results

Physiological and root anatomical assay. Sugarcane wild relative *E. arundinaceus* genotype IND99-907 was collected from the salinity affected area of Padiyattu Kadavu, Ernakulum, India under germplasm exploration programme during 1999 and adopted to high salinity stress with normal plant growth²⁹ and a potential source for harnessing the salt stress-responsive genes. On the other hand, the sugarcane genotype Co 97010 is highly sensitive to salinity and has been used as a salt-sensitive standard in many studies^{30,31}. The 90 days old plants of IND99-907 and Co 97010 were imposed with salinity stress by irrigating with 175 mM saline water. The leaf elongation rate (LER) was significantly higher in IND99-907 under salinity stress and growth was ceased in



a) Leaf elongation rate under control and 175 mM salt stress in IND99-907 and Co 97010



b) Proline Accumulation in roots and relative water content of leaves in IND99-907 and Co 97010 under salinity stress

Figure 1. Phenotypic and physiological responses of IND99-907 and Co 97010 under salinity stress (A) Leaf elongation rate, (B) Proline content of roots and Relative water content in leaves. Error bars indicates the mean \pm SD.

Co 97010 (p value = 0.0251) (Fig. 1a). The physiological assay for proline content in roots showed that IND99-907 recorded significantly higher proline content of 8.908 $\mu\text{moles/gram}$ of root samples as compared to 1.438 $\mu\text{moles/gram}$ in Co 97010 under salinity stress (p value = 0.0374) (Fig. 1b). The relative water content revealed that IND99-907 has retained the higher water content under the stressed condition as compared to Co 97010 (p value = 0.0019).

The root anatomical studies under light microscope revealed the clear differences in thickening of cells around the metaxylem in IND99-907 and Co 97010. The vacuolization of the cortex was observed in Co 97010 as compared to the normal cortex in IND99-907. The thickening of endodermis and deformed protoxylem vessels were observed in Co 97010 as compared to no significant change in IND99-907 (Fig. 2a). The same results were confirmed in Scanning Electron Microscope which showed cell-wall thickening in protoxylem and thickening with circular xylem vessels in Co 97010. IND99-907 revealed xylem vessels without a significant change in protoxylem and metaxylem (Fig. 2b). Similar changes in the root cellular structure under drought and salinity stress in the tolerant genotypes are attributed to be an adaptive mechanism of resistant genotypes^{32,33}.

The elemental composition of stressed and control samples were analyzed by SEM- EDX and ICP-OES methods. The differential accumulation pattern of sodium, calcium, magnesium, and potassium under both control and stressed root samples were observed in IND99-907 and Co 97010. The lowest Na^+/K^+ ratio of 0.64 was observed in IND99-907 as compared to 1.50 in Co 97010 by ICP-OES analysis. A similar trend of lowest Na^+/K^+ was observed in IND99-907 (2.18) as compared to 7.64 by Co 97010 (Table 1 and Fig. 2b) in SEM-EDX analysis. Similarly, IND99-907 also recorded the lowest $\text{Ca}^{2+}/\text{K}^+$ (1.05) and $\text{Mg}^{2+}/\text{K}^+$ (0.23) ratio as compared to 4.52 and 0.73 in Co 97010 respectively under salinity stressed condition.

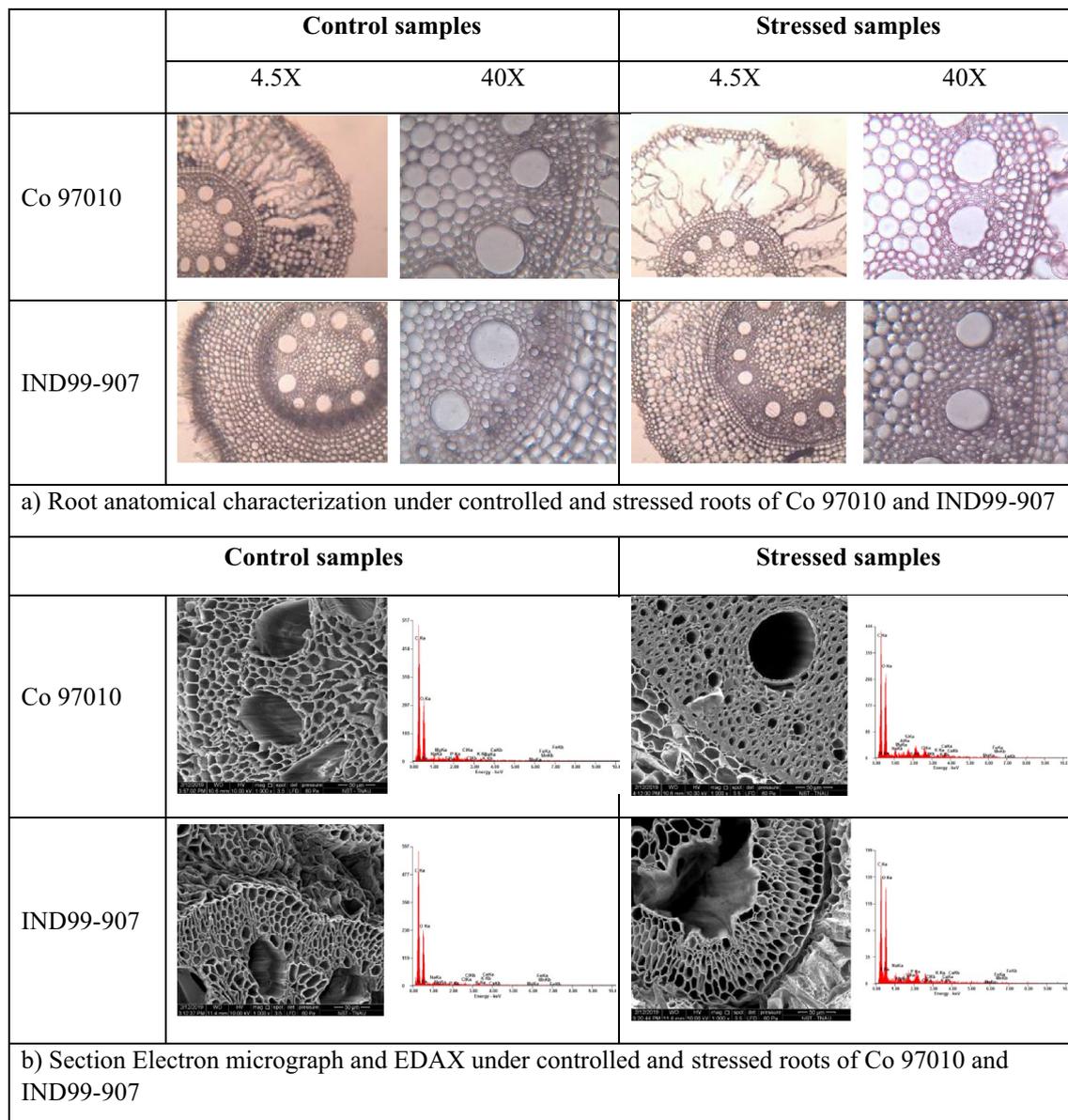


Figure 2. Root anatomical changes in responses of IND99-907 and Co 97010 under salinity stress (A) Root anatomical images showing the vacuolization of cortex and thickening of metaxylem vessels in Co 97010 as compared to no significant changes in IND99-907. (B) Scanning Electron Micrograph showing thickening of protoxylem with circular xylem vessels in Co 97010 as compared to irregular shaped xylem vessels in IND99-907 and, EDAX analysis showing elemental composition of roots under salinity stress. (a) Root anatomical characterization under controlled and stressed roots of Co 97010 and IND99-907. (b) Section Electron micrograph and EDAX under controlled and stressed roots of Co 97010 and IND99-907.

Genotypes	Na ⁺ /K ⁺ ratio		Ca ⁺ /K ⁺ ratio		Mg ⁺ /K ⁺ ratio	
	S	C	S	C	S	C
(i) ICP-OES analysis						
Co 97010	7.64	1.88	4.52	2.30	0.79	0.51
IND99-907	2.18	0.75	1.05	0.45	0.23	0.18
(ii) SEM-EDX analysis						
Co 97010	1.50	0.80	3.24	1.06	0.33	0.71
IND99-907	0.64	1.03	0.00	0.87	0.18	1.17

Table 1. Estimated Sodium–Potassium ratio by ICP-OES and SEM–EDX method at 175 mM for 15 days in pot culture experiment.

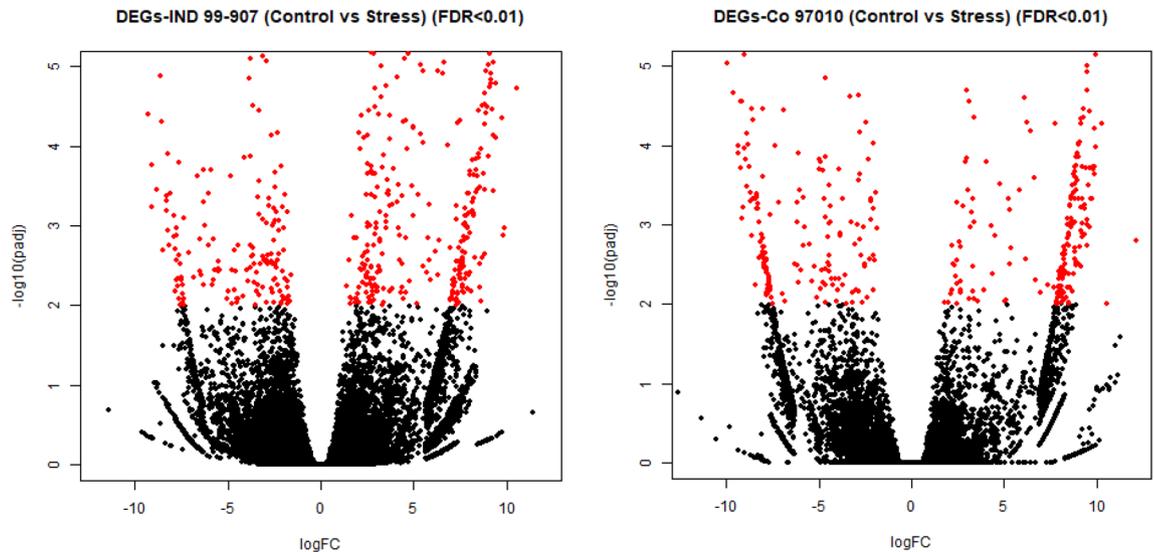


Figure 3. Volcano plots displaying the distribution of DEGs between control vs stress conditions of (a) IND99-907 and (b) Co 97010 based on FDR < 0.01.

Transcriptome data overview. The comparative transcriptome analysis was performed between stressed and control samples of saline-tolerant *E. arundinaceus* accession IND99-907 and saline-sensitive sugarcane genotype Co 97010 with three biological replications. A total of 305, 244, 264, and 263 million raw reads were obtained from the stressed and control libraries of IND99-907 (marked as 907S and 907C) and Co 97010 (marked as 97010S and 97010C) respectively (Supplementary Table S1). The pre-processing of raw reads by removal of adapter sequences, rRNA sequences, and ambiguous bases with high-quality sequences with a Phred score of ≥ 30 (Q30) resulted in 226, 182, 196, and 198 million clean reads from stressed and control libraries of IND99-907 and Co 97010 respectively. We obtained more than 74% of the high-quality clean reads after pre-processing of the raw data resulting in 802 million clean reads which were good enough to perform the RNASeq analysis.

De novo assembly of transcripts and unigenes clustering. The high-quality reads of stressed and control libraries of IND99-907 and Co 97010 were assembled into four de novo assemblies for 907C, 907S, 97010C and 97010S using Trinity v2.8.5. The quality assessment and comparison of assemblies were performed using Quast and BUSCO tools^{34,35}. (i) The Quast analysis revealed the total number of contigs of four assemblies ranging from 205,717 to 345,745, with the highest contig N_{50} of 1727 bp observed in 907S assembly and the largest contig of 29,724 bp was recorded in 97010S assembly. (ii) BUSCO analysis showed the completeness percentage ranging from 69.8% (907C) to 80% (907S) with a maximum duplication of 42.1% was observed in 97010C assembly. After unigenes clustering, the number of contigs ranged from 179,304 to 242,390 with the highest contig N_{50} of 2102 bp in 907S unigenes assembly. The merging of control and stressed unigene assemblies for each genotype resulted in the identification of 315,611 unigenes with a contig N_{50} of 2071 bp and 426,820 unigenes with N_{50} of 1800 bp in clustered unigene assembly of IND99-907 and Co 97010 respectively (Supplementary Table S1). The high quality clustered assemblies of IND99-907 and Co 97010 were used for transcript quantification and annotation processes.

Transcript abundance estimation and differential gene expression. The coding DNA sequences were identified from the clustered assemblies of IND99-907 and Co 97010 by using the TransDecoder tool and the transcript quantification was performed by using RSEM v1.2.25. The transcript counts generated by RSEM were used to perform the differential gene expression (DGE) analysis by using DESeq2 for both IND99-907 and Co 97010 separately. The DEGs were identified based on absolute $\log_2FC > 2$ for upregulated genes and $\log_2FC < -2$ for downregulated genes. A total of 7731 DEGs from IND99-907 (Control vs Stress) DEG library and 6159 DEGs from Co 97010 (Control vs Stress) library were obtained at significance level $p < 0.05$. The FDR corrections were performed through the Benjamini–Hochberg method³⁶ to reduce the false positives and, a stringent FDR < 0.01 was used to further reduce the number of DEGs which yielded a total of 649 DEGs in IND99-907 and among them, 425 genes were upregulated and 197 were downregulated. A total of 501 DEGs (FDR < 0.01) were identified in Co 97010 and out of them, 283 genes were upregulated and 213 genes were downregulated (Fig. 3). These DEGs (FDR < 0.01) identified in IND 99-907 and Co 97010 were used for functional annotation.

Functional mapping of DEGs. The functional annotation of DEGs (FDR < 0.01) of IND99-907 and Co 97010 were carried out by using tools such as Blast2GO, HMMer & Command-line Blast against NCBI-nr, Swissprot, Pfam, and Uniprot Plant databases. The results against the NCBI-nr database for total DEGs (FDR < 0.01) showed 88.4% and 90.6% annotations for IND99-907 and Co 97010 respectively. Gene Ontology (GO) annotations were obtained for all DEGs. Upregulated and downregulated genes were also annotated separately using Blast2GO for both IND99-907 and Co 97010. In IND99-907, 88.2% of the upregulated and 93.4% of the down-

regulated genes were blasted with hits respectively and for Co 97010, 91.2% of upregulated and 93.4% of down-regulated genes had blast annotations. The top hit species of DEGs were *Sorghum bicolor*, *Zea mays*, *Setaria italica*, and *Setaria viridis*. A total of 81 DEGs in IND99-907 had fold change values (\log_2FC) greater than 20, which includes ATP synthase, abscisic stress ripening protein 2, E3 ubiquitin-protein ligase AIRP2, serine/threonine-protein kinase WNK1 and ABC Transporter F (Supplementary Table S2). In total, five stress-related proteins, nine heat-shock proteins (HSPs), 25 transporters, 46 kinases, and 37 transcription factors (TFs) were significantly differentially regulated ($FDR < 0.01$) under salt stress. Similarly, for Co 97010, only eight DEGs had fold change values > 20 which included three hypothetical proteins, two ribosomal proteins, and two dehydrogenases. Three stress-related DEGs were identified in Co 97010, namely salt stress-induced hydrophobic peptide ESI3 (ESI3), heat stress transcription factor C-2a (HSFC2A) and abscisic stress ripening protein 2 (ASR2), out of which ESI3 and HSFC2A were upregulated and ASR2 was downregulated under stress conditions. Further, a total of six heat shock proteins, 18 transporters, 18 kinases, and 19 TFs were identified in Co 97010. Gene families were identified for both DEGs of IND99-907 and Co 97010, which showed 413 gene families present in 649 DEGs of IND 99-907 and 327 gene families present in 501 DEGs of Co 97010.

Gene ontology (GO) enrichment. Gene Ontology (GO) annotations for DEGs were categorized into three GO functional classifications as cellular component (CC), molecular function (MF) and biological process (BP)³⁷. A total of 78.7% hits in IND99-907 and 78.2% in Co 97010 were mapped using GO analysis. GO enrichment analysis showed significant enrichment ($FDR < 0.05$) of 17 different processes for 649 DEGs in IND99-907. The significantly enriched stress-related processes ($FDR < 0.05$) were given in Supplementary Table S3, S4 & Supplementary Figure S1A. Similarly, the GO enrichment of MF and CC were also studied to identify the significant enrichments at $FDR \leq 0.05$ correspondingly which showed two molecular functions namely transporter (GO:0005215) and transferase (GO:0016740) activity (Supplementary Table S3 & Supplementary Figure S1B) and one cellular component namely vacuole (GO:0005773) (Supplementary Figure S1C) under salinity stress in IND99-907. Whereas in Co 97010, two biological processes namely metabolite and energy precursor generation (GO:0006091) and response to stress (GO:0006950) were significantly enriched among 501 DEGs at $FDR \leq 0.05$ (Supplementary Table S4) and no significant enrichment at $FDR \leq 0.05$ was observed for cellular components and molecular functions.

KEGG annotation and enrichment. KEGG annotations and Enzyme Code (EC) mapping were performed using Blast2GO. A total of 79 different KEGG pathway hits were obtained for 649 DEGs in IND99-907 and 67 KEGG pathway hits were obtained for 501 DEGs in Co 97010 with both of them having major hits in Glycolysis/Gluconeogenesis (ko00010), Pyruvate metabolism (ko00620), Galactose metabolism (ko00052) and Starch and Sucrose metabolism (ko00500). Further, KEGG enrichment showed a total of 27 statistically significant ($FDR < 0.05$) enriched KEGG pathways in IND99-907 (Supplementary Table S5) including Endocytosis (ko04144), Lysine degradation (ko00310), Plant hormone signal transduction (ko04075), Oxidative phosphorylation (ko00190) and MAPK signalling pathway (ko04016). In contrast, DEGs of Co 97010 showed 20 significantly enriched KEGG pathways ($FDR < 0.05$) (Supplementary Table S6) such as Oxidative phosphorylation (ko00190), Metabolic pathways (ko01100), and Photosynthesis (ko00195).

Endocytosis. Endocytosis, an important pathway for the intracellular vesicle trafficking system was significantly enriched ($FDR = 3.18E-09$) in total DEGs of IND99-907 under salinity stress. A total of nineteen transcripts were differentially regulated and among them, seven were upregulated and twelve were downregulated. Among the upregulated gene hits in the pathway, clathrin interactor epsin-2 showed the highest upregulation ($\log_2FC = 20.31$). Other genes and their fold change values were ras-related RabE1d protein ($\log_2FC = 8.77$), Arf-GAP ($\log_2FC = 7.72$), phospholipase D ($\log_2FC = 7.990$ and HSP70 ($\log_2FC = 3.49$). It was observed that the same pathway was significantly enriched ($FDR = 0.001157$) in the DEGs Co 97010 with two of the proteins ADP-ribosylation factor 1 ($\log_2FC = -9.21$) and vacuolar protein sorting associated protein ($\log_2FC = -3.49$) being highly downregulated under salinity stress (Supplementary Figure S2).

Oxidative phosphorylation. Oxidative phosphorylation is a pathway associated with the formation of ATP through phosphorylation of NADH or FADH₂ by electron transport chain. This pathway was significantly enriched at $FDR = 0.005099$ in DEGs of IND 99-907 with a total of nine transcripts involving in the pathway. The highest upregulation was observed in the protein ATP synthase ($\log_2FC = 22.66$) which catalyses the formation of ATP. Several ATPase hits were observed in the pathway like calcium-transporting, plasma membrane, V-type proton ATPases and the highest upregulated was observed in ATPase subunit 6 ($\log_2FC = 21.09$). The same pathway was also significantly enriched ($FDR = 7.45E-16$) in the DEGs of Co 97010 with a total of twenty transcripts involving in the pathway including dehydrogenase, cytochrome c oxidase, and ATP synthase. The highest down regulation was observed in cytochrome c oxidase (COX) ($\log_2FC = -26.86$) followed by ATP synthase CF1 beta subunit ($\log_2FC = -25.00$) (Supplementary Figure S3).

qRT-PCR validation of transcriptome data. A total of 43 DEGs representing the various metabolic pathways in both IND 99-907 and Co 97010 were chosen for qRT-PCR validation. The details of the genes selected for validation and their primers are given in Table 2. Dehydrin DHNs are a group of LEA proteins involved in membrane stabilization during abiotic stresses³⁸ was upregulated in both IND99-907 and Co 97010 and the same was reconfirmed in qRT-PCR. NADH dependent malic enzyme upregulated in response to abscisic acid signalling and drought³⁹ was upregulated in IND99-907 and downregulated in Co 97010. Dirigent-like protein associated with lignin biosynthesis and cellular stabilization⁴⁰ was upregulated in IND99-907 and nil

Pathway	Putative functions	Gene name	Protein name	Unigene ID		IND99-907		Co 97010	
				IND99-907	Co 97010	qRT-PCR	RNA Seq	qRT-PCR	RNA Seq
Membrane Stabilization	LEA proteins involved in membrane stabilization ³⁸	DHN1	Dehydrin DHN1	ERI-C_DN10750_c0_g1_i1.p1	SUG-S_DN859_c0_g1_i3.p1	5.8	5.81	8.12	6.28
Membrane Stabilization	lignin biosynthesis and cellular stabilization ⁴⁰	DIR1	Dirigent-like protein	ERI-S_DN657_c0_g1_i9.p1		2.51	3.14	-0.76	-
Membrane Stabilization	Microtubules organization ⁴¹	KTN80.4	Katanin p80 WD40 repeat-containing subunit B1	ERI-S_DN8743_c0_g1_i22.p1		10.25	3.72	0.26	-
Membrane-bound trafficking	clathrin-mediated endocytosis ⁹²	EPSIN2	Clathrin interactor EPSIN 2	ERI-S_DN1711_c0_g1_i13.p1		2.55	20.3	-0.59	-
Membrane-bound trafficking	cellular cargo located on golgi bodies and plasma membrane ⁹⁵	RABE1D	Ras-related protein RABE1d-like	ERI-S_DN1918_c0_g1_i19.p2		3.14	8.77	-0.47	-
Metabolic Process	Carbohydrate metabolism ¹¹⁸	PFK6-like	ATP-dependent 6-phosphofructokinase 6-like	ERI-S_DN1551_c0_g1_i5.p1	SUG-C_DN1191_c0_g1_i35.p1	-1.05	-2.22	-0.46	2.46
Metabolic Process	Carbohydrate metabolism ¹¹⁹	SUS2	Sucrose synthase 2	ERI-C_DN1377_c0_g1_i34.p3		-1.34	-11.04	-0.2	-
Metabolic Process	Carbohydrate metabolism ⁷¹	DLAT2	Dihydrolipoyllysine-residue acetyltransferase component 2 of pyruvate dehydrogenase complex mitochondrial	ERI-C_DN16417_c0_g1_i16.p1		1.58	20.43	0.03	-
Metabolic Process	secondary metabolites ⁷²	GAD1	glutamate decarboxylase	ERI-S_DN184_c0_g1_i30.p1	SUG-S_DN212_c0_g2_i1.p1	11.4	10.87	1.49	1.45
Metabolic Process	secondary metabolites ⁷⁵	LKR	Lysine-ketoglutarate reductase/saccharo-pine dehydrogenase1	ERI-S_DN817_c0_g1_i119.p1		5.89	6.19	0.94	-
Metabolic Process	Sterol synthesis ¹²⁰	SMO2-1	methylsterol monooxygenase 1-1 isoform X2	ERI-S_DN4522_c0_g1_i1.p1		5.96	5.77	0.02	-
Metabolic Process	Sterol synthesis	STE1	Delta(7)-sterol-C5(6)-desaturase 1	-	SUG-S_DN756_c0_g1_i20.p2	-0.58	-	4.32	8.99
Metabolic Process	Monoterpenoid/secondary metabolite biosynthesis	7DGT	7-deoxyloganetin glucosyltransferase	ERI-S_DN2625_c1_g2_i2.p2	SUG-C_DN47939_c0_g1_i1.p3	1.25	2.74	-1.81	-3.70
Metabolic Process	arabinogalactan-protein biosynthesis—cell wall expansion	GLCAT14A	beta-glucuronosyltransferase GlcAT14A-like	ERI-S_DN7775_c0_g1_i3.p2		1.69	5.24	-0.94	-
Metabolic Process	Hypusine biosynthesis	DHS	deoxyhypusine synthase	ERI-S_DN2053_c0_g1_i2.p2		4.7	20.44	-0.48	-
Metabolic Process	Cell wall organization	GDPDL3	Glycerophosphodiester phosphodiesterase GDPDL3	ERI-C_DN1124_c0_g1_i5.p1		2.31	7.42	0.91	-
Metabolic Process	Osmoprotectant	HSP23.6	23.6 kDa heat shock protein, mitochondrial	ERI-S_DN974_c0_g1_i3.p2	SUG-S_DN78847_c0_g1_i1.p1	1.2	3.31	5.03	6.08
Metabolic Process	Secondary metabolites ⁷⁴	AKR1B1	aldose reductase	ERI-S_DN5048_c1_g1_i18.p1	SUG-S_DN8120_c0_g1_i6.p1	4.59	4.31	3.52	2.44
Metabolic Process	Secondary metabolites- proanthocyanidin biosynthesis	BAN	Anthocyanidin reductase	-	SUG-S_DN242_c0_g1_i1.p1	-0.26	-	1.5	3.28
Metabolic Process	Osmo-protectant	LEA3	late embryogenesis abundant protein 3	ERI-S_DN1898_c0_g1_i22.p1	SUG-C_DN6306_c0_g1_i16.p2	7	9.44	2.86	3.5
Signal Transduction	Phospholipid signalling	PLDZETA1	phospholipase D zeta 1 isoform X1	ERI-S_DN2802_c0_g1_i47.p1	SUG-C_DN8702_c0_g1_i47.p2	2.23	8	4.16	8.02
Signal Transduction	Signalling ⁹³	CPK15	Calcium-dependent protein kinase 15	ERI-S_DN497_c0_g1_i36.p1		23.71	20.22	-0.12	-
Signal Transduction	Abscisic acid mediated signalling ⁹⁴	SIP1	stress inducible protein coi6.1	ERI-S_DN921_c0_g1_i3.p2		8.75	8.84	-0.75	-
Signal Transduction	Abscisic acid biosynthesis	NCED3	9-cis-epoxycarotenoid dioxygenase 1, chloroplastic	ERI-S_DN1264_c0_g2_i5.p1	SUG-S_DN43107_c0_g1_i6.p1	3.68	3.06	1.61	3.35
Signal Transduction	Abscisic acid mediated signalling	PLDalpha1	phospholipase D alpha 1	ERI-C_DN805_c0_g1_i2.p1		1.69	7.74	-0.06	0
Signal Transduction	transcriptional activator	ARF7	Auxin response factor 7	ERI-S_DN1671_c0_g1_i6.p1	SUG-C_DN1108_c0_g1_i50.p1	9.62	20.53	2.84	7.34

Continued

Pathway	Putative functions	Gene name	Protein name	Unigene ID		IND99-907		Co 97010	
				IND99-907	Co 97010	qRT-PCR	RNA Seq	qRT-PCR	RNA Seq
Signal Transduction	abscisic acid mediated signalling ³⁹	NADP-ME1	NADP-dependent malic enzyme	ERI-S_DN965_c0_g1_i69.p1	SUG-C_DN4380_c0_g1_i1.p1	5.75	20.47	-1.13	-4.73
Signal Transduction	Abscisic acid mediated signalling	AIRP2	E3 ubiquitin-protein ligase AIRP2-like	ERI-S_DN2132_c0_g1_i8.p1		2.98	21.09	-0.3	-
Signal Transduction	Signal Transduction	Osmo-protectant	Embryogenesis transmembrane protein-like	ERI-S_DN415_c0_g1_i6.p3	SUG-C_DN1841_c0_g1_i2.p1	12.61	7.14	-3.3	-1.98
Signal Transduction	benzothiadiazole (BTH)- mediated cell death-Defence mechanism	SNL6	Cinnamoyl-CoA reductase-like SNL6	ERI-C_DN1176_c1_g1_i2.p1	SUG-C_DN1270_c0_g1_i10.p1	-1.08	-2.43	11.89	8.83
Solute Transport	calcium fluxes ⁸²	ACA10	calcium-transporting ATPase 10, plasma membrane-type	ERI-S_DN107_c0_g1_i2.p1		1.39	20.56	0.52	-
Solute Transport	active proton transport and sodium extrusion ⁸⁴	PMA1	plasma membrane ATPase 1	ERI-S_DN2505_c0_g1_i2.p1	SUG-S_DN3426_c0_g1_i35.p2	14.48	9.15	1.92	3.99
Transcriptional Regulation	splicing of pri-mRNA	matK	maturase K	-	SUG-C_DN8_c1_g1_i7.p1	0.04	0	4.22	13.99
Transcriptional Regulation	Transcriptional factor-Flowering pathway ⁵³	MADS26	MADS-box transcription factor 26	ERI-S_DN1719_c0_g1_i2.p1		6.98	10.3	0.34	0
Transcriptional Regulation	Transcriptional factor-drought and salt tolerance	NAC74	NAC domain-containing protein 74	ERI-S_DN392_c0_g1_i7.p1		3.44	20.26	0.12	0
Transcriptional Regulation	nuclear import receptor for serine-arginine rich proteins	MOS14	Transportin MOS14	ERI-S_DN7275_c0_g1_i16.p1		5.96	7.98	0.42	0
Transferase Activity	Metalic homeostasis	naat-A	Nicotianamine aminotransferase A	ERI-S_DN110_c0_g1_i3.p1		4.2	3.18	0.02	0
Transporter Activity	Transport ⁹¹	CAT2	Cationic amino acid transporter 2, vacuolar	ERI-C_DN2868_c0_g1_i3.p3		23.99	20.48	-0.31	0
Transporter Activity	secondary metabolites ⁷³	KCS11	3-ketoacyl-CoA synthase 11-like	ERI-S_DN506_c0_g1_i8.p1		3.02	10.85	-0.39	0
Transporter Activity	Sugar transporter	CSTLP3	CMP-sialic acid transporter 3	ERI-S_DN14571_c0_g2_i1.p1		5.77	7.19	-0.18	0
Transporter Activity	iron transporter	IRT2	fe(2+) transport protein 2	ERI-C_DN1057_c0_g1_i4.p1	SUG-S_DN5534_c0_g1_i2.p1	-1.1	-4.74	-6.94	-2.05
Transporter Activity	Efflux of glucose	GLT	probable plastidic glucose transporter 2	ERI-S_DN3048_c0_g1_i50.p1		1.13	3.92	0.15	0

Table 2. Validation of candidate genes through qRT-PCR.

expression was observed in Co 97010. Katanin is a microtubule⁴¹ and its regulating gene Katanin p80 WD40 repeat was upregulated in IND99-907 and nil expression was observed in Co 97010. Total of 43 genes were validated using qRT-PCR and correlation of 0.53 was observed between RNAseq and quantitative gene expression by qRT-PCR.

Discussion

Sugarcane is a glycophyte and grows poorly in saline soils⁶. Salinity stress is a main restrictive feature and significantly affect the growth and biomass accumulation^{42,43}. Sugarcane wild relative *Erianthus arundinaceus* is a potential crop for the production of high biomass and reservoir of many genes related to agronomic, biotic and abiotic stresses. We have isolated and characterized many genes from *E. arundinaceus*^{14,15,19-22} and many of genes are still unexplored. Besides, transcriptome studies were reported for many traits such as cold transcriptome²³, sucrose²⁴, agronomic traits²⁵, lignin²⁸, red stripe disease²⁶, smut disease²⁷ except for salinity stress. Hence, we performed the comparative transcriptome analysis to identify the total salt-responsive genes and metabolic pathways associated with salt tolerance in *E. arundinaceus* genotype IND99-907 and salt-sensitive sugarcane genotype Co 97010.

Physiological analyses like leaf elongation rate, relative water content, proline accumulation, root anatomical structure and Na⁺/K⁺ ratio revealed the salinity tolerance level of IND99-907. The growth of young leaves was ceased in Co 97010 under salinity stress as compared to growing young leaves in IND99-907. The growth of young leaves in tolerant genotypes was also reported in other crops such as maize⁴⁴ and sorghum⁴⁵. The accumulation of proline was very high in IND99-907 under salinity stress as compared to Co 97010. Proline acts as an osmolyte and protects the sub-cellular structures by scavenging of free-radicals and optimizing the osmotic potential during abiotic stresses⁴⁶. IND99-907 maintained the higher level of relative water content under salinity

stress as compared to Co 97010 and this is in confirmation with previous reports maintaining the higher relative water content in tolerant genotypes¹⁹.

Soil-salinity exhibits higher osmotic potential in soil and creates physiological drought for plants⁵. The root protoxylem and metaxylem thickness also change during exposure to abiotic stresses⁴⁷. In our study, Co 97010 showed a clear lignification of cell wall around the metaxylem constituting to the cell-wall thickening, vacuolization of the cortex and damaged protoxylem in Co 97010 under salinity stress as compared to normal root structures in IND99-907. The higher concentration of sodium in tissues interfere with the metabolic processes and causes cellular death⁵. The tolerant genotypes minimize the cytotoxic effect of sodium and other cations through ionic exclusion, compartmentalization of ions and maintenance of turgor potential through increased uptake of K⁺ ions⁵. In our study, both EDAX and ICP-OES analysis showed the lower Na⁺/K⁺ ratio in IND99-907 and the lower Na⁺/K⁺ ratio is essential for the maintenance of optimal physiological activities such as stomatal movement, photosynthesis and transpiration under salinity stress⁴⁸.

The genome-wide expression of genes for agronomic traits, biotic and abiotic stresses are studied through NGS based transcriptome sequencing in many crops including non-model crops¹³. The comparative transcriptome analysis was performed to identify the salt-responsive genes and metabolic pathways associated with the salt-tolerant *E. arundinaceus* genotype IND99-907 and sugarcane salt-sensitive genotype Co 97010. We have generated 107.71 GB of raw data and 80.28 GB of clean data from IND99-907 and Co 97010. We assembled the high-quality clean reads (Q30) into four different assemblies with the N₅₀ values of 1727 bp, 1506 bp, 1389 bp, and 1359 bp respectively for 907S, 907C, 97010S and 97010C, which is found to be superior or on par with the previous transcriptome studies in sugarcane such as cold transcriptome^{23–28}. In our study, we identified 315,611 and 426,820 unigenes for both IND99-907 and Co 97010 respectively. The stringent computational protocol such as p-value adjustment or false discovery rate (FDR ≤ 0.01) based on the Benjamini–Hochberg method was adopted to control the false positives during differential gene expression analysis³⁶. A total of 649 and 501 DEGs (FDR < 0.01) were identified from IND99-907 and Co 97010 respectively reveals the salt-responsive genes and metabolic pathways associated with salt tolerance in *E. arundinaceus* and sugarcane.

Transcription factors are the major driving factors in the regulation of gene expression and downstream activation of the cascade of metabolic pathways in biotic and abiotic stresses⁴⁹. In our study, DREB1C was upregulated (log₂FC = 8.70) in IND99-907 and, its overexpression in tobacco and *Arabidopsis* has enhanced the tolerance levels to salt and freezing stresses⁵⁰. The DREB1H was upregulated (log₂FC = 8.52) in IND99-907 and it is in confirmation with other crops for cold and salt stresses⁵¹. The overexpression of G-box-binding factor 3-like conferred the tolerance to salt stress in *Arabidopsis*⁵². MADS26, associated with meristem response and growth⁵³ and it was also upregulated (log₂FC = 10.30) in IND99-907. MYB related transcriptional factors regulating abiotic stress tolerance such as MYB-308-like, MYB4-like, MYBAS2^{54–57} were all upregulated in IND99-907.

The GO enrichment analysis is a statistical approach to identify the enriched or over-represented and depleted group of genes and to compare the functional profile of DEGs and metabolic pathways⁷⁹. In our study, 'GO term: Response to stress' is significantly enriched in both IND99-907 (FDR = 1.60E-07) and Co 97010 (FDR = 0.031). It is in confirmation with the previous report of GO term: response to stress under salinity stress⁸⁰. In salt-tolerant IND99-907, most of the enriched genes are associated with regulation of cell wall and cytoskeleton reorganization such as (i) Actin (log₂FC = 21.48) which dynamically regulates the membrane permeability and cytoskeleton organization⁸¹; (ii) Plasma membrane-type calcium-transporting ATPase 10 (log₂FC = 20.56) regulating the calcium fluxes, plant immunity and developmental processes⁸²; (iii) cellulose synthase A7 regulating the secondary cell wall synthesis⁸³ (iv) plasma membrane ATPase 1 (log₂FC = 9.15) regulating the active proton transport and sodium extrusion under salt-stress and its upregulation has been observed to boost the salt tolerance in transgenic events⁸⁴. We have validated two DEGs namely Plasma membrane-type calcium-transporting ATPase 10 and plasma membrane ATPase 1 through qRT-PCR in IND99-907. Whereas, the GO enriched genes in salt-sensitive Co 97010 are majorly representing the metabolic pathways viz., (i) cytochrome c maturase subunit C (log₂FC = 12.14) associated with electron transport system⁸⁵, (ii) 2,3-bisphosphoglycerate-independent phosphoglycerate mutase-like (log₂FC = 11.00) associated with glycolysis⁸⁶; (iii) cysteine proteinase 1 (log₂FC = 9.52) associated with senescence, protein degradation and apoptosis⁸⁷ and (iv) Histidine kinase 4 associated with cytokinin signalling⁸⁸. This has shown that genes and metabolic pathways were uniquely and differentially enriched in salt-tolerant IND99-907 and salt-sensitive Co 97010 for GO term: Response to stress.

The 'GO term: Transport' is significantly enriched in IND99-907 (FDR = 0.01) as compared to non-significant enrichment (FDR = 0.06) in Co 97010. Membrane transport plays a major role in plant systems and endomembrane trafficking for adjusting to the changing environmental stresses⁸⁹. Most of the enriched genes in IND99-907 were related to membrane-bound trafficking and stress-related genes. Genes such as (i) nucleobase-ascorbate transporter 3 (log₂FC = 20.48), also known as nucleobase:cation-symporter 2 associated with nucleobase transportation⁹⁰ (ii) vacuolar cationic amino acid transporter 2 (log₂FC = 20.48), which is located on tonoplasts and regulate the soluble amino acid concentration in the cytoplasm⁹¹ (iii) clathrin interactor EPSIN-2 (log₂FC = 20.30) associated with clathrin-mediated endocytosis⁹²; (iv) Calcium-dependent protein kinase 15 (log₂FC = 20.22) interacts with WRKY transcriptional factors and regulation of defense signaling⁹³; (v) stress-inducible protein coi6.1 (log₂FC = 8.84) associated with cold stress tolerance⁹⁴ (vi) ras-related protein RABE1d-like (log₂FC = 8.77) is a cellular cargo located on golgi bodies and plasma membrane⁹⁵. In our study, DEGs such as nucleobase-ascorbate transporter 3, clathrin interactor EPSIN-2, stress-inducible protein coi6.1 and ras-related protein RABE1d-like were validated through qRT-PCR in IND99-907. Whereas, the enriched genes in Co 97010 were more related to metabolic pathways such as (i) NADH dehydrogenase subunit 6 (log₂FC = 13.99) associated with cellular respiration and mitochondrial respiratory chain⁹⁶; (ii) nitrate regulatory gene2 protein (log₂FC = 9.51) regulating the nitrate transporter⁹⁷ (iii) exocyst complex component EXO70B1 (log₂FC = 8.65) regulating the vesicular trafficking of plasma membrane-bound kinase FLS2⁹⁸ and (iv) Enhancer of AG-4 protein 2 (log₂FC = 8.56) regulating the anthocyanin production pathways⁹⁹. GO term: transport enriched genes showed

that endomembrane and stress-related transport genes were enriched in IND99-907 and basic metabolic pathway related transport genes were enriched in Co 97010.

KEGG enrichment analysis identifies the statistically significant enriched pathways from a set of differentially expressed genes from RNAseq data⁵⁸. In this study, the KEGG enrichment analysis showed the significant enrichment of MAPK signalling pathway in IND99-907 (FDR = 0.0023) as compared to non-significant expression in Co 97010 (FDR = 0.2042). The sessile plants perceive the stresses by the cascades of mitogen-activated kinase-related pathways^{59,60} and activate the downstream regulatory genes. A total of 15 mitogen-activated kinases (MAPKs), five mitogen-activated kinase kinase (MAPKKs) and 16 mitogen-activated kinase kinase kinase (MAP3Ks) were regulating the biotic and abiotic stresses in sugarcane commercial varieties (*Saccharum spp. hybrids*)⁶¹. In our study, ABA signalling genes viz., MAP3K1 ($\log_2FC = 8.04$), and protein phosphatase-2C (PP2C) ($\log_2FC = 7.54$) were highly enriched in IND99-907. PP2C negatively regulates the ABA signalling through inhibition and phosphorylation of SNRKs⁶². MAP3K regulates the intra- and inter-cellular genes associated with the homeostasis of reactive oxygen species⁶³ and activation of ABA signalling pathways through phosphorylation of SNRK2 kinases⁶⁴. The RAN1 copper-transporting P-type ATPase was downregulated in IND99-907, which is linked to the ethylene-response pathway⁶⁵ and negative regulation of ethylene response enhances the plant growth⁶⁶. Whereas in Co 97010, the late responsive genes such as MAPK ($\log_2FC = 3.36$), serine/threonine-protein kinase SAPK1 ($\log_2FC = 8.2$), and PR1 ($\log_2FC = -2.06$) are significantly enriched. These genes were previously reported for their expression under abiotic stresses in many cultivated crops^{67,68}.

The KEGG enrichment pathway: Biosynthesis of secondary metabolites was significantly enriched in both salt-tolerant IND99-907 (FDR = 3.56E-10) and salt-sensitive Co 97010 (FDR = 8.40E-05). Secondary metabolites are essential for physiological adaptive responses in various stimuli⁶⁹. Secondary metabolites such as anthocyanins, flavones, and phenolics inactivate the cytotoxic ions by binding to them, thus protecting the cells from ionic oxidative damage⁷⁰. In our study, highly expressed genes involved in the pathway biosynthesis of secondary metabolites in IND99-907 were (i) DLAT2 of pyruvate dehydrogenase complex ($\log_2FC = 20.43$) regulating glycolytic and citric acid pathways, and its upregulation conferring tolerance to abiotic stresses is in confirmation with previous reports⁷¹; (ii) glutamate decarboxylase ($\log_2FC = 10.87$) confer tolerance to multiple stresses by regulating the γ -amino butyrate content in plants⁷² (iii) 3-ketoacyl-CoA synthase 11-like ($\log_2FC = 10.85$) gene regulates the elongation of fatty acid and confer tolerance to salt stress⁷³ (iv) aldose reductase ($\log_2FC = 4.31$) involved in biosynthesis of osmo-compatible compounds such as sorbitol⁷⁴ (v) lysine-ketoglutarate reductase/saccharopine dehydrogenase 1 ($\log_2FC = 4.41$) regulating the signalling of abscisic acid, jasmonate, sugar and starch starvation related pathways⁷⁵ was enriched in IND99-907 and these genes are validated through qRT-PCR. The enriched genes in Co 97010 were also stress-related genes such as (i) acetyl-CoA acetyltransferase, cytosolic 1-like isoform X2 ($\log_2FC = 8.60$) associated with drought tolerance⁷⁶; (ii) chloroplastic solanesyl-diphosphate synthase 3 ($\log_2FC = 8.52$) associated with prenylquinones⁷⁷; (iii) UDP-glycosyltransferase ($\log_2FC = 8.31$) involved in glycosylation or inactivation of aglycones such as hormones, secondary metabolites, and its overexpression confer tolerance to abiotic stress tolerance⁷⁸. Though there are common genes expressed in two genotypes, many genes were associated with the biosynthesis of secondary metabolites in IND99-907 as compared to Co 97010.

Many DEGs related to various metabolic pathways such as cytoskeleton organization (DHN1, DIR1 and KTN80.41), cellular membrane stabilization (EPSIN-2 and RAB1d), metabolic pathways related to biosynthesis of secondary metabolites, late embryogenesis proteins, osmolytes and sterols (PFK6-Like, SUS2, DLAT2, GAD1, LKR, SMO2-1, STE1, 7DGT, GLCAT14A, DHS, GDPDL3, HSP23.6, AKR1B1, BAN and LEA3), signal transduction (PLDZETA1, CPK15, SIP1, NCED3, PLDalpha1, ARF7, NADP-ME1, AIRP2, SNL6, ETP and SNL6), solute transport (ACA10 and PMA1), transcriptional regulation (matK, MADS26, NAC74 and MOS14), transferase or metallic homeostasis (NAAT-A) and transporters (CAT2, KCS11, CSTLP3, IRT2 and GLT) were validated through qRT-PCR in IND99-907 and Co 97010.

Conclusion

In summary, a comparative analysis of transcriptome was carried out between two genotypes *Erianthus arundinaceus* genotype IND99-907 and salt-sensitive sugarcane genotype Co 97010 to elucidate the salt-responsive genes and metabolic pathways associated with salt-tolerance. Our findings showed clear and distinct metabolic pathways and genes involved in pathways such as MAPK signal transduction, plant hormone signalling and biosynthesis, secondary metabolite biosynthesis and hormones, biosynthesis of amino acid, transporters, and many other pathways were significantly enriched. The gene ontology enrichment analysis also revealed enrichment of many stress-related biological processes such as response to abiotic stimulus, stress, stimulus and endogenous stimulus, generation of precursor metabolites and energy, and many other biological processes. Our study has shown the metabolic pathways and genes linked with salt tolerance in IND99-907 at the transcriptome level. Further studies are required to overexpress the differentially expressed candidate genes from IND99-907 in the background of sugarcane cultivars through biotechnological approaches and development of functional markers associated with salinity tolerance.

Materials and methods

Plant materials and treatment. Salinity-tolerant *Erianthus arundinaceus* genotype IND99-907 collected from the saline lands from Ernakulum, India²⁹ and salinity-sensitive sugarcane genotype Co 97010^{30,31} were used in this study. Single buds from each genotype with five replications were grown in pots containing river sand: red soil: farmyard manure in 1:1:1 proportion. The experiment was conducted in the polyhouse facility with a 16-h photoperiod, 30/25 °C day/night, 380–400 $\mu\text{molm}^{-2}\text{s}^{-1}$. Salt stress was imposed during the formative stage on the 90th day after planting by irrigating with salinized water containing NaCl, CaCl₂ and Na₂SO₄ in the ratio of

2:2:1³⁰. Salinity stress was imposed by adding 25 mM of saline water on the first day followed by incrementally adding 25 mM on every day till reaches 175 mM¹⁰⁰.

Physiological assay of IND99-907 and Co 97010 under salinity stress. The young leaf elongation was measured at an interval of 3, 6, 9, 12 and 15th days after imposing the salinity stress¹⁹. Samples were collected, cut into 10 cm segments and leaf fresh weight was weighed. The leaves were soaked overnight in distilled water. After soaking, excess water on the leaf surface was removed and turgid weight was quantified. The leaves were oven-dried at 37 °C and dry weight was recorded and RWC was calculated¹⁰¹. Proline was estimated by the colourimetric method from the root samples¹⁰².

To assess the cellular changes in the root cortical and vascular tissues, root samples of 5 cm from the root tip were sampled from the 16th day of stressed and control treatments. The hand sections were made, mounted on the slide and the observed image was captured from the light microscope under 4.5X and 40X resolutions. The ultrastructure of cellular changes in selected genotypes was analysed using a Scanning Electron Microscope (FEIQuanta250, Icon analytical, FEI, USA). Root sections of 4–5 mm thick were gold coated with a sputter coater and mounted on round aluminum stubs with the aid of double side adhesive tape. The samples were scanned and selected regions were photographed. The elemental distribution of Na⁺, K⁺, Ca²⁺ and Mg²⁺ in the root sections were analysed using a line scan of EDX.

For ionic analysis in roots, the oven-dried and finely grounded roots (100 mg each) were digested with 10 ml of HNO₃:HClO₄ (3:1) di-acid mixture for overnight. The flasks were gently heated on a hot plate till the development of intense white fumes and the process continued till the digested mixture remains transparent. Leftover digested material made the final volume of 50 ml by adding distilled water. This solution was filtered with Whatman filter paper no. 42 and the extract were used for estimation of Na⁺, K⁺, Ca²⁺ and Mg²⁺ content by ICP-OES (ICP 9000, Shimadzu, Japan).

Transcriptome sequencing. The gene expression patterns are differing with respect to the duration and concentration of exposure to salt stress. The isolation of short-term salt responsive gene requires few hours of exposure to salt stress and long-term salt stress gene isolation requires exposure longer duration for about a week or more. The gene expression patterns are also differing for salt shock and gradual incremental increase in salt stress. The maximum number of genes are expressed within a few hours under salt shock and after a week under gradual exposure to salt stress¹⁰⁰. Hence, our study is aimed at isolation of long-term salt responsive genes under gradual exposure to salinity stress and hence, we opted incremental increase of salt stress of 25 mM till the final concentration reaches 175 mM and the total RNA from stressed and control samples of IND99-907 and Co 97010 were sampled after a week.

Total RNA was isolated from control and salt-stressed roots sampled on 10th day of salinity stress from IND99-907 and Co 97010 with three independent biological replicates. Samples were prepared using Tru-Seq RNA Sample Preparation Kit v2 (Illumina Catalog No RS-122-2001) as per the manufacturer instructions. The RNA is fragmented into small pieces using divalent cations, 3' ends were adenylated to prevent the reads from self-ligation and both 5' and 3' adapters were ligated. PCR amplification was carried out to generate the final cDNA libraries. The cDNA libraries of all twelve samples were quantified using Qubit and Nanodrop and validated by Agilent 2200 TapeStation System. The samples with RNA integrity number more than seven were sequenced by using Illumina HiSeq2500 with 2 × 100 bp in paired-end format.

Bioinformatics analysis. Raw transcriptome sequences were quality checked using the tool FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Poor quality reads were removed using Trimmomatic¹⁰³ with a threshold Phred score of ≥ 30 (Q30) and a minimum length of ≥ 80 following the removal of the adapter sequences¹⁰⁴. SILVA database (release 132) (<https://www.arb-silva.de/>) was used to remove the rRNA present in the sample by alignment using bowtie2¹⁰⁵. For *de-novo* assembly, high-quality, adapter-free Q30 reads of all three biological replicates of four conditions were pooled together to generate four clustered *de novo* assemblies for both control and stressed samples of IND 99-907 and Co 97010 by using Trinity assembler v2.11.0 (<https://github.com/trinityrnaseq/trinityrnaseq/wiki>)¹⁰⁶.

The quality parameters of *de novo* assemblies were checked by using two approaches, (i) Quast 5.0.2¹⁰⁷ to obtain assembly statistics and (ii) BUSCO v4.1.2 (Benchmarking Universal Single-Copy Orthologs) analysis. Unigenes clustering was done based on 95% identity threshold using CD-HIT v4.8.1^{108–110} with parameters “cd-hit-est -c 0.95, -n 8” using the following ways, (i) for all four assemblies (containing isoforms) separately (ii) Control and Stressed unigene datasets were merged to generate a single dataset for every genotype (two unigene datasets for IND 99-907 and Co 97010) (iii) Further, unigenes clustering was performed to remove duplicate sequences to generate two reference unigene datasets for IND 99-907 and Co 97010 respectively. TransDecoder tool (<https://github.com/TransDecoder/TransDecoder/wiki>) was utilized to recognize the coding DNA sequences (CDS) and the long ORFs in all four unigene assemblies.

The transcript quantification by estimating the abundance of each gene was carried out using RSEM v1.2.25¹¹¹. Differential gene expression analysis (DGE) was carried out by DESeq2 tool in R with parameters, log₂FC > 2 (up-regulated) and log₂FC < -2 (down-regulated) with FDR < 0.01. DESeq2 requires estimated raw counts for performing DGE, hence the raw gene counts obtained from RSEM were used for performing the same^{112,113}. Annotations were performed with databases viz., NCBI-nr (1e-3) (<https://www.ncbi.nlm.nih.gov/refseq/about/nonredundantproteins>), Swissprot (<https://www.uniprot.org/statistics/Swiss-Prot>), Pfam (<http://pfam.xfam.org/>) and Uniprot Plants (<https://www.uniprot.org/help/plants>). The annotations of DEGs against NCBI-nr and Swissprot databases with an e-value of 1E-3 was done using Blast2GO v5.2.5³⁷, HMMer program was used against

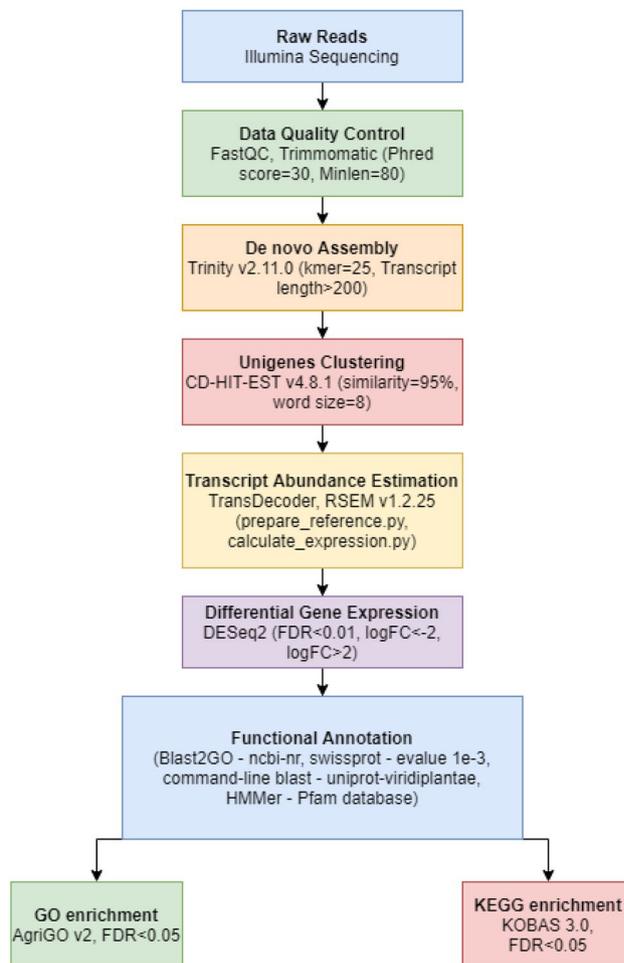


Figure 4. Complete Bioinformatic workflow of RNA-seq analysis involving two genotypes IND99-907 and Co 97010.

the Pfam database and Command-line Blast was used against Uniprot Plants to retrieve their respective protein annotations.

Annotation for Gene Ontology (GO mapping), InterProScan, Enzyme Code (EC) mapping was performed using Blast2GO v5.2.5³⁷. GO enrichment was done using Hypergeometric test and Benjamini–Hochberg correction through AgriGO v2⁷⁹. KOBAS 3.0 (<http://kobas.cbi.pku.edu.cn/kobas3/?t=1>).^{58,114–116} was used to perform KEGG pathway enrichment. The complete workflow followed for RNASeq data analysis is given in Fig. 4.

Validation using qRT-PCR. A total of 43 candidate DEGs were selected based on their involvement in the enriched stress-related pathways and biological processes in both stressed and control transcriptome data of IND99-907 and Co 97010 were used for validation by qRT-PCR (Supplementary Table S7). PrimerQuest tool¹¹⁷ was used to design the primers for conserved domains identified by Pfam domain search by SMART tool¹¹⁸. The total RNA was reverse transcribed by using PrimeScript 1st strand cDNA synthesis Kit (Takara, Catalog No.: 6110A). The 1.0 µg of total RNA from each sample taken from stressed and control samples was used for cDNA synthesis and total cDNA was synthesized as per the manufacturer instructions. The qRT-PCR validation was carried out using StepOne Real-Time PCR System (Applied Biosystems, Canada) and each gene was validated using three biological and two technical replicates. GAPDH gene expression was used as the internal control²³. The qRT-PCR Reaction Mix was prepared using TB Green® Premix Ex Taq™ (Tli RNase H Plus) Master Mix (Takara, Catalog No.: RR420A). The 20µL qRT-PCR mix consisted of SYBR green master mix (10µL), ROX reference dye (0.4µL), cDNA (1.0µL or 50 ng), forward and reverse primers (0.40 µl each) and molecular grade sterile water^{23,119}. The initial denaturation at 95 °C for 10 min, followed by 40 cycles of denaturation at 95 °C for 15 s and, annealing and extension at 60 °C for 60 s. Relative fold changes of gene expressions were determined using the $2^{-\Delta\Delta CT}$ method¹²⁰.

Statement of compliance. Experimental material involves the research between sugarcane wild relative *Erianthus arundinaceus* and modern sugarcane genotype (*Saccharum* spp) is in compliance with institutional, national, and international guidelines and legislation.

Data availability

The Illumina raw reads of samples were deposited in NCBI BIO PROJECT with accession number PRJNA716503.

Received: 6 May 2021; Accepted: 8 December 2021

Published online: 31 December 2021

References

- Singh, A. *et al.* Phytochemical profile of sugarcane and its potential health aspects. *Pharmacogn. Rev.* **9**, 45–54 (2015).
- Eggleston, G. Positive aspects of cane sugar and sugar cane derived products in food and nutrition. *J. Agric. Food Chem.* **66**, 4007–4012 (2018).
- FAO. Area, production and productivity of livestock, Food and Agriculture Organization of United Nations. *Food and Agriculture Organization of the United Nations, Database on Crops* <http://www.fao.org/faostat/en/#data/QC> (2019).
- Zaman, M., Shahid, S. A. & Heng, L. Soil Salinity: Historical Perspectives and a World Overview of the Problem. in *Guideline for Salinity Assessment, Mitigation and Adaptation Using Nuclear and Related Techniques* (eds. Zaman, M., Shahid, S. A. & Heng, L.) 43.53 (Springer Nature Switzerland AG, 2018). doi:<https://doi.org/10.1007/978-3-319-96190-3>.
- Munns, R. & Tester, M. Mechanisms of salinity tolerance. *Annu. Rev. Plant Biol.* **59**, 651–681 (2008).
- Wahid, A., Rao, A. U. R. & Rasul, E. Identification of salt tolerance traits in sugarcane lines. *F. Crop. Res.* **54**, 9–17 (1997).
- Lingle, S. E. & Wiegand, C. L. Soil salinity and sugarcane juice quality. *F. Crop. Res.* **54**, 259–268 (1997).
- Mahajan, S., Pandey, G. K. & Tuteja, N. Calcium- and salt-stress signaling in plants: Shedding light on SOS pathway. *Arch. Biochem. Biophys.* **471**, 146–158 (2008).
- Manishankar, P., Wang, N., Köster, P., Alatar, A. A. & Kudla, J. Calcium signaling during salt stress and in the regulation of ion homeostasis. *J. Exp. Bot.* **69**, 4215–4226 (2018).
- Lamers, J., Meer, T. V. & Der & Testerink, C.,. How plants sense and respond to stressful environments. *Plant Physiol.* **182**, 1624–1635 (2020).
- Liang, W., Ma, X., Wan, P. & Liu, L. Biochemical and biophysical research communications plant salt-tolerance mechanism: A review. *Biochem. Biophys. Res. Commun.* **495**, 286–291 (2018).
- Gupta, B. & Huang, B. Mechanism of salinity tolerance in plants: Physiological, biochemical, and molecular characterization. *Int. J. Genom.* **2014**, 2 (2014).
- Namba, C. I. N., Nag, A. & Sharma, R. K. Next generation sequencing technologies: The doorway to the unexplored genomics of non-model plants. *Front. Plant Sci.* **6**(1074), 1–16 (2015).
- Augustine, S. M., Cherian, A. V., Syamaladevi, D. P. & Subramonian, N. Erianthus arundinaceus HSP70 (EaHSP70) acts as a key regulator in the formation of anisotropic interdigitation in sugarcane (*Saccharum* spp. hybrid) in response to drought stress. *Plant Cell Physiol.* **56**, 2368–2380 (2015).
- Manoj, V. M. *et al.* Comparative analysis of glyoxalase pathway genes in *Erianthus arundinaceus* and commercial sugarcane hybrid under salinity and drought conditions. *BMC Genom.* **19**, 1–16 (2019).
- Wang, W. *et al.* Sweetcane (*Erianthus arundinaceus*) as a native bioenergy crop with environmental remediation potential in southern China: A review. *GCB Bioenergy* **11**, 1012–1025 (2019).
- Ram, B., Sreenivasan, T. V., Sahi, B. K. & Singh, N. Introgression of low temperature tolerance and red rot resistance from *Erianthus* in sugarcane. *Euphytica* **122**, 145–153 (2001).
- Cai, Q. *et al.* Verification of the introgression of *Erianthus arundinaceus* germplasm into sugarcane using molecular markers. *Plant Breed.* **124**, 322–328 (2005).
- Augustine, S. M. *et al.* *Erianthus arundinaceus* HSP70 (EaHSP70) overexpression increases drought and salinity tolerance in sugarcane (*Saccharum* spp. hybrid). *Plant Sci.* **232**, 23–34 (2015).
- Augustine, S. M. *et al.* Overexpression of EaDREB2 and pyramiding of EaDREB2 with the pea DNA helicase gene (PDH45) enhance drought and salinity tolerance in sugarcane (*Saccharum* spp. hybrid). *Plant Cell Rep.* **34**, 247–263 (2015).
- Narayan, J. A. *et al.* Isolation and characterization of water-deficit stress-responsive α -expansin 1 (EXPA1) gene from *Saccharum* complex. *3 Biotech* **9**, 186 (2019).
- Anunathini, P. *et al.* In silico characterisation and functional validation of chilling tolerant divergence 1 (COLD1) gene in monocots during abiotic stress. *Funct. Plant Biol.* **46**, 524–532 (2019).
- Dharshini, S. *et al.* Root transcriptome analysis of *Saccharum spontaneum* uncovers key genes and pathways in response to low-temperature stress. *Environ. Exp. Bot.* **171**, 103935 (2020).
- Thirugnanasambandam, P. P., Hoang, N. V., Furtado, A., Botha, F. C. & Henry, R. J. Association of variation in the sugarcane transcriptome with sugar content. *BMC Genom.* **18**, 1–22 (2017).
- Cardoso-Silva, C. B. *et al.* De novo assembly and transcriptome analysis of contrasting sugarcane varieties. *PLoS ONE* **9**, 1–10 (2014).
- Santa Brigida, A. B. *et al.* Sugarcane transcriptome analysis in response to infection caused by *Acidovorax avenae* subsp. *avenae*. *PLoS ONE* **11**, 1–30 (2016).
- Que, Y., Su, Y., Guo, J., Wu, Q. & Xu, L. A global view of transcriptome dynamics during *Sporisorium scitamineum* challenge in sugarcane by RNA-seq. *PLoS ONE* **9**, 1–18 (2014).
- Kasirajan, L., Hoang, N. V., Furtado, A., Botha, F. C. & Henry, R. J. Transcriptome analysis highlights key differentially expressed genes involved in cellulose and lignin biosynthesis of sugarcane genotypes varying in fiber content. *Sci. Rep.* **8**, 1–17 (2018).
- Nair, N. V. & Somarajan, K. G. Diversity of *Saccharum* germplasm in Kerala, India. *Plant Genet. Resour. Newsl.* **135**, 40–43 (2003).
- Vasanthi, S., Venkataramana, S., Gururaja Rao, P. N. & Gomathi, R. Long term salinity effect on growth, photosynthesis and osmotic characteristics in sugarcane. *Sugar Tech.* **12**, 5–8 (2010).
- Brindha, C., Srinivasavedantham, V., Arun, K. R. & Tayade, A. S. Characterization of the salt overly sensitive pathway genes in sugarcane under salinity stress. *Physiol. Plant.* **171**, 677–687 (2021).
- Parida, A. K., Veerabathini, S. K., Kumari, A. & Agarwal, P. K. Physiological, anatomical and metabolic implications of salt tolerance in the halophyte *Salvadora persica* under hydroponic culture condition. *Front. Plant Sci.* **7**, 1–18 (2016).
- Silva, B. R. S., Batista, B. L. & Lobato, A. K. S. Anatomical changes in stem and root of soybean plants submitted to salt stress. *Plant Biol.* **23**, 57–65 (2021).
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
- Waterhouse, R. M., Seppey, M., Sim, F. A. & Ioannidis, P. BUSCO applications from quality assessments to gene prediction and phylogenomics letter fast track. *Mol. Biol. Evol.* **35**, 543–548 (2017).
- Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).
- Götz, S. *et al.* High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* **36**, 3420–3435 (2008).

38. Liu, Y., Wang, L., Zhang, T., Yang, X. & Li, D. Functional characterization of KS-type dehydrin ZmDHN13 and its related conserved domains under oxidative stress. *Sci. Rep.* **7**, 1–10 (2017).
39. Arias, C. L. *et al.* NADP-dependent malic enzyme 1 participates in the abscisic acid response in *Arabidopsis thaliana*. *Front. Plant Sci.* **8**, 1–11 (2018).
40. Burlat, V., Kwon, M., Davin, L. B. & Lewis, N. G. Dirigent proteins and dirigent sites in lignifying tissues. *Phytochemistry* **57**, 883–897 (2001).
41. Luptovčíak, I., Komis, G., Takáč, T., Ovečka, M. & Šamaj, J. Katanin: A sword cutting microtubules for cellular, developmental, and physiological purposes. *Front. Plant Sci.* **8**(1982), 1–10 (2017).
42. Saxena, P., Srivastava, R. P. & Sharma, M. L. Studies on salinity stress tolerance in sugarcane varieties. *Sugar Tech* **12**, 59–63 (2010).
43. Chiconato, D. A. & da Silveira Sousa Junior, G., dos Santos, D. M. M. & Munns, R.,. Adaptation of sugarcane plants to saline soil. *Environ. Exp. Bot.* **162**, 201–211 (2019).
44. Neves-Piestun, B. G. & Bernstein, N. Salinity-induced inhibition of leaf elongation in maize is not mediated by changes in cell wall acidification capacity. *Plant Physiol.* **125**, 1419–1428 (2001).
45. de Lacerda, C. F., Cambraia, J., Oliva, M. A., Ruiz, H. A. & Prisco, J. T. Solute accumulation and distribution during shoot and leaf development in two sorghum genotypes under salt stress. *Environ. Exp. Bot.* **49**, 107–120 (2003).
46. Ashraf, M. & Foolad, M. R. Roles of glycine betaine and proline in improving plant abiotic stress resistance. *Environ. Exp. Bot.* **59**, 206–216 (2007).
47. Barzegargolchini, B., Movafeghi, A., Dehestani, A. & Mehrabanjoubani, P. Increased cell wall thickness of endodermis and protoxylem in *Aeluropus litoralis* roots under salinity: The role of LAC4 and PER64 genes. *J. Plant Physiol.* **218**, 127–134 (2017).
48. Sharma, S. K. Soil salinity effects on transpiration and net photosynthetic rates, stomatal conductance and Na⁺ and Cl⁻ contents in durum wheat. *Biol. Plant.* **38**, 519 (1996).
49. Khan, S., Li, M., Wang, S. & Yin, H. Revisiting the role of plant transcription factors in the battle against abiotic stress. *Int. J. Mol. Sci. J. Mol. Sci.* **19**(1634), 1–29 (2018).
50. Wen, W. *et al.* Switchgrass PvDREB1C plays opposite roles in plant cold and salt tolerance in transgenic tobacco. *He* **155**, 1–12 (2018).
51. de Cabral, G. A. L. *et al.* First expressed TFome of physic nut (*Jatropha curcas* L.) after salt stimulus. *Plant Mol. Biol. Rep.* **38**, 189–208 (2020).
52. Ramegowda, V. *et al.* GBF3 transcription factor imparts drought tolerance in *Arabidopsis thaliana*. *Sci. Rep.* **7**(9148), 1–13 (2017).
53. He, T. *et al.* Gene-set association and epistatic analyses reveal complex gene interaction networks affecting flowering time in a worldwide barley collection. *J. Exp. Bot.* **70**, 5603–5616 (2019).
54. Chano, V., de Heredia, U. L., Collada, C. & Alvaro, S. Transcriptomic analysis of juvenile wood formation during the growing season in *Pinus canariensis*. *Holzforchung* **71**, 919–937 (2017).
55. Badhan, S., Kole, P., Ball, A. & Mantri, N. Plant Physiology and Biochemistry RNA sequencing of leaf tissues from two contrasting chickpea genotypes reveals mechanisms for drought tolerance. *Plant Physiol. Biochem.* **129**, 295–304 (2018).
56. Zhou, F., Zhang, Y., Tang, W., Wang, M. & Gao, T. Transcriptomics analysis of the flowering regulatory genes involved in the herbicide resistance of Asia minor bluegrass (*Polypogon fugax*). *BMC Genom.* **18**(953), 1–13 (2017).
57. He, Y. *et al.* Jasmonic acid-mediated defense suppresses brassinosteroid-mediated susceptibility to Rice black streaked dwarf virus infection in rice. *New Phytol.* **214**, 388–399 (2017).
58. Xie, C. *et al.* KOBAS 2.0: A web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkr483> (2011).
59. Jonak, C. *et al.* Stress signaling in plants: A mitogen-activated protein kinase pathway is activated by cold and drought. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 11274–11279 (1996).
60. Sinha, A. K., Jaggi, M., Raghuram, B. & Tuteja, N. Mitogen-activated protein kinase signaling in plants under abiotic stress. *Plant Signal. Behav.* **6**, 196–203 (2011).
61. Ali, A. *et al.* Genome-wide analysis of mitogen-activated protein (MAP) kinase gene family expression in response to biotic and abiotic stresses in sugarcane. *Physiol. Plant.* **171**, 86–107 (2021).
62. Manohar, M. *et al.* Members of the abscisic acid co-receptor PP2C protein family mediate salicylic acid—abscisic acid crosstalk. *Plant Direct* **1**, 1–14 (2017).
63. Nakagami, H., Soukupová, H., Schikora, A., Žárský, V. & Hirt, H. A mitogen-activated protein kinase kinase kinase mediates reactive oxygen species homeostasis in *Arabidopsis*. *J. Biol. Chem.* **281**, 38697–38704 (2006).
64. Takahashi, Y. *et al.* MAP3 Kinase-dependent SnRK2-kinase activation is required for abscisic acid signal transduction and rapid osmotic stress response. *Nat. Commun.* **11**(12), 1–12 (2020).
65. Binder, B. M., Rodríguez, F. I. & Bleecker, A. B. The copper transporter RAN1 is essential for biogenesis of ethylene receptors in *Arabidopsis*. *J. Biol. Chem.* **285**, 37263–37270 (2010).
66. Tao, J. J. *et al.* The role of ethylene in plants under salinity stress. *Front. Plant Sci.* **6**, 1–12 (2015).
67. Lou, D., Wang, H. & Yu, D. The sucrose non-fermenting-1-related protein kinases SAPK1 and SAPK2 function collaboratively as positive regulators of salt stress tolerance in rice. *BMC Plant Biol.* **18**(203), 1–17 (2018).
68. Akbudak, M. A., Yildiz, S. & Filiz, E. Genomics Pathogenesis related protein-1 (PR-1) genes in tomato (*Solanum lycopersicum* L.): Bioinformatics analyses and expression profiles in response to drought stress. *Genomics* **112**, 4089–4099 (2020).
69. Isah, T. Stress and defense responses in plant secondary metabolites production. *Biol. Res.* **52**, 1–25 (2019).
70. Sytar, O., Mbarki, S., Zivcak, M. & Brestic, M. The Involvement of Different Secondary Metabolites in Salinity Tolerance of Crops. in *Salinity Responses and Tolerance in Plants, Volume 2* (eds. Kumar, V., Wani, S. H., Suprasanna, P. & Tran, L.-S. P.) vol. 2 21–49 (Springer International Publishing AG, 2018).
71. Marondedze, C., Thomas, L., Gehring, C. & Lilley, K. S. Changes in the *Arabidopsis* RNA-binding proteome reveal novel stress response mechanisms. *BMC Plant Biol.* **19**, 1–11 (2019).
72. Wu, X. *et al.* Gamma-aminobutyric acid (GABA) alleviates salt damage in tomato by modulating Na⁺ uptake, the GAD gene, amino acid synthesis and reactive oxygen species metabolism. *BMC Plant Biol.* **20**, 1–21 (2020).
73. Yang, Z. *et al.* Overexpression of β -ketoacyl-CoA synthase from *Vitis vinifera* L. Improves salt tolerance in *Arabidopsis thaliana*. *Front. Plant Sci.* **11**, 1–11 (2020).
74. Almaghamsi, A., Nosarzewski, M., Kanayama, Y. & Archbold, D. D. Effects of abiotic stresses on sorbitol biosynthesis and metabolism in tomato (*Solanum lycopersicum*). *Funct. Plant Biol.* **48**, 286–297 (2021).
75. Stepansky, A. & Galili, G. Synthesis of the *Arabidopsis* bifunctional lysine-ketoglutarate reductase/saccharopine dehydrogenase enzyme of lysine catabolism is concertedly regulated by metabolic and stress-associated signals. *Plant Physiol.* **133**, 1407–1415 (2003).
76. Pan, J. *et al.* Comparative proteomic investigation of drought responses in foxtail millet. *BMC Plant Biol.* **18**(35), 1–19 (2018).
77. Hirooka, K., Bamba, T., Fukusaki, E. & Kobayashi, A. Cloning and kinetic characterization of *Arabidopsis thaliana* solanesyl diphosphate diphosphate synthase. *Biochem. J.* **686**, 679–686 (2003).
78. Li, P. *et al.* The *Arabidopsis* UDP-glycosyltransferases UGT79B2 and UGT79B3, contribute to cold, salt and drought stress tolerance via modulating anthocyanin accumulation. *Plant J.* **89**, 85–103 (2017).

79. Tian, T. *et al.* AgriGO v2.0: A GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* **45**, W122–W129 (2017).
80. Shi, P. & Gu, M. Transcriptome analysis and differential gene expression profiling of two contrasting quinoa genotypes in response to salt stress. *BMC Plant Biol.* **20**(568), 1–15 (2020).
81. Liu, S. G., Zhu, D. Z., Chen, G. H., Gao, X. Q. & Zhang, X. S. Disrupted actin dynamics trigger an increment in the reactive oxygen species levels in the Arabidopsis root under salt stress. *Plant Cell Rep.* **31**, 1219–1226 (2012).
82. Frey, N. F. *et al.* Plasma membrane calcium ATPases are important components of receptor-mediated signaling in plant immune responses and development. *Plant Physiol.* **159**, 798–809 (2012).
83. Endler, A. *et al.* A mechanism for sustained cellulose synthesis during salt stress. *Cell* **162**, 1353–1364 (2015).
84. Fan, Y. *et al.* Over-expression of a plasma membrane H⁺-ATPase SpAHA1 conferred salt tolerance to transgenic Arabidopsis. *Protoplasma* **255**, 1827–1837 (2018).
85. Verissimo, A. F. & Daldal, F. Cytochrome c biogenesis system I: An intricate process catalyzed by a maturase supercomplex?. *Biochim. Biophys. Acta Bioenerg.* **1837**, 989–998 (2014).
86. Zhao, Z. & Assmann, S. M. The glycolytic enzyme, phosphoglycerate mutase, has critical roles in stomatal movement, vegetative growth, and pollen production in Arabidopsis thaliana. *J. Exp. Bot.* **62**, 5179–5189 (2011).
87. Grudkowska, M. & Zagdańska, B. Multifunctional role of plant cysteine proteinases. *Acta Biochim. Pol.* **51**, 609–624 (2004).
88. Yamada, H. *et al.* The Arabidopsis AHK4 histidine kinase is a cytokinin-binding receptor that transduces cytokinin signals across the membrane. *Plant Cell Physiol.* **42**, 1017–1023 (2001).
89. Wang, X. *et al.* The roles of endomembrane trafficking in plant abiotic stress responses. *J. Integr. Plant Biol.* **62**, 55–69 (2020).
90. Niopek-Witz, S., Deppe, J., Lemieux, M. J. & Möhlmann, T. Biochemical characterization and structure-function relationship of two plant NCS2 proteins, the nucleobase transporters NAT3 and NAT12 from Arabidopsis thaliana. *Biochim. Biophys. Acta Biomembr.* **1838**, 3025–3035 (2014).
91. Yang, H., Krebs, M., Stierhof, Y. D. & Ludewig, U. Characterization of the putative amino acid transporter genes AtCAT2, 3 & 4: The tonoplast localized AtCAT2 regulates soluble leaf amino acids. *J. Plant Physiol.* **171**, 594–601 (2014).
92. McMahon, H. T. & Boucrot, E. Molecular mechanism and physiological functions of clathrin-mediated endocytosis. *Nat. Rev. Mol. Cell Biol.* **12**, 517–533 (2011).
93. Shen, L. *et al.* CaCDPK15 positively regulates pepper responses to Ralstonia solanacearum inoculation and forms a positive-feedback loop with CaWRKY40 to amplify defense signaling. *Sci. Rep.* **6**, 1–12 (2016).
94. Nguyen, H. T., Leipner, J., Stamp, P. & Guerra-Peraza, O. Low temperature stress in maize (*Zea mays* L.) induces genes involved in photosynthesis and signal transduction as studied by suppression subtractive hybridization. *Plant Physiol. Biochem.* **47**, 116–122 (2009).
95. Ahn, C. S., Han, A. J. A. & Pai, H. S. Characterization of in vivo functions of Nicotiana benthamiana RabE1. *Planta* **237**, 161–172 (2013).
96. Geisler, D. A. *et al.* Downregulation of the δ -subunit reduces mitochondrial ATP synthase levels, alters respiration, and restricts growth and gametophyte development in Arabidopsis. *Plant Cell* **24**, 2792–2811 (2012).
97. Xu, N. *et al.* The Arabidopsis NRG2 protein mediates nitrate signaling and interacts with and regulates key nitrate regulators. **28**, 485–504 (2016).
98. Wang, W. *et al.* The Arabidopsis exocyst subunits EXO70B1 and EXO70B2 regulate FLS2 homeostasis at the plasma membrane. *New Phytol.* **227**, 529–544 (2020).
99. Ilk, N., Ding, J., Ichnatowicz, A., Koornneef, M. & Reymond, M. Natural variation for anthocyanin accumulation under high-light and low-temperature stress is attributable to the ENHANCER OF AG-4 2 (HUA2) locus in combination with production of anthocyanin pigment1 (PAP1) and PAP2. *New Phytol.* **206**, 422–435 (2015).
100. Shavrukov, Y. Salt stress or oceanic salt shock: Which genes are we studying? In Posidonia cadmium induces changes in DNA methylation and chromatin patterning. *J. Exp. Bot.* **64**, 119–127 (2013).
101. Soltys-Kalina, D., Plich, J., Strzelczyk-Żyta, D., Śliwka, J. & Marczewski, W. The effect of drought stress on the leaf relative water content and tuber yield of a half-sib family of ‘Katahdin’-derived potato cultivars. *Breed. Sci.* **66**, 328–331 (2016).
102. Bates, L. S., Waldren, R. P. & Teare, I. D. Rapid determination of free proline for water-stress studies. *Plant Soil* **39**, 205–207 (1973).
103. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
104. Pereira-Santana, A. *et al.* Transcriptional profiling of sugarcane leaves and roots under progressive osmotic stress reveals a regulated coordination of gene expression in a spatiotemporal manner. *PLoS ONE* **12**, 1–25 (2017).
105. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
106. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
107. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).
108. Li, W. & Godzik, A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
109. Geniza, M. & Jaiswal, P. Tools for building de novo transcriptome assembly. *Curr. Plant Biol.* **11–12**, 41–45 (2017).
110. Hoang, N. V., Furtado, A., Thirugnanasambandam, P. P., Botha, F. C. & De Henry, R. J. novo assembly and characterizing of the culm-derived meta-transcriptome from the polyploid sugarcane genome based on coding transcripts. *Heliyon* **4**, 1–36 (2018).
111. Li, B. & Dewey, C. N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **12**(323), 1–16 (2011).
112. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**(550), 1–21 (2014).
113. Love, M. I., Anders, S. & Huber, W. Analyzing RNA-seq data with DESeq2. <https://bioconductor.org/packages/release/bioc/vignettes/DESeq2/inst/doc/DESeq2.html#why-un-normalized-counts> (2021).
114. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
115. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951 (2019).
116. Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M. & Tanabe, M. KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Res.* **49**, D545–D551 (2021).
117. Owczarzy, R. *et al.* IDT SciTools: A suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Res.* **36**, 163–169 (2008).
118. Schultz, J., Milpetz, F., Bork, P. & Ponting, C. P. SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 5857–5864 (1998).
119. Selvarajan, D. *et al.* Differential gene expression profiling through transcriptome approach of *Saccharum spontaneum* L. under low temperature stress reveals genes potentially involved in cold acclimation. *Biotech* **8**, 1–18 (2018).
120. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* **25**, 402–408 (2001).

Acknowledgements

The authors are respectfully thanking and acknowledging the financial support from ‘Science and Engineering Research Board, Department of Science and Technology (DST-SERB)’ under Early Career Research Award (File No.: DST/SERB/ECR/2017/000349). We also thankfully acknowledge Mr. Selvamuthu K, Technical Officer, ICAR-Sugarcane Breeding Institute, Coimbatore for technical support for this work. We are also thankful to AgriGenome Labs Pvt Ltd, Cochin, India for transcriptome sequencing work. We thankfully acknowledge the Kanehisa (KEGG database project), www.kanehisa.jp for granting the permission to publish the enriched metabolic pathways of plant endocytosis (ko4144) and oxidative phosphorylation pathway (ko00190) in Scientific Reports of Springer Nature Ltd in print and digital under the CC BY 4.0 open access license.

Author contributions

P.V. & C.M.: Equal contribution to research work or shared first co-authors, performed genome assembly and annotation, differential gene expression, K.E.G.G. and G.O. enrichment analysis, submitted genome and raw sequence data to NCBI, prepared all tables, figures and wrote the manuscript. P.R., G.S., P.V., C.M.: Designed the experiment for performing genome assembly and differential gene expression. N.S.: Performed annotation. C.A., C.M., R.V., K.M., G.H.: Designed the experiment. V.R., C.M.: Root anatomical studies. P.V., S.D., S.S., G.S., C.M., V.S., H.K.M.: RNA isolation and performed qRT-PCR validation. C.M., C.A., G.H., B.R.: Overall supervised the experiment. All authors read the manuscript and approved.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-03735-5>.

Correspondence and requests for materials should be addressed to C.M. or C.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021