Phylogeny, function, and structure of rice oxalate oxidases

M. G. C. Carrillo,[1] P. H. Goodwin,[2] M. Variar,[3] S.S. Han,[4] J.H. Roh,[4] J. Cairns,[1] J. E. Leach,[5] H. Leung,[1] and C. M. VERA CRUZ[1]

[1]International Rice Research Institute, DAPO Box 7777, Metro Manila, Philippines; [2]University of Guelph, Guelph, ON, Canada N1G2W1; [3]Central Rainfed Upland Rice Research Station, Hazaribag, India; [4] Rural Development Administration, Suwon, South Korea; [5] Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO 80537-1177 USA

Oxalate oxidases (*OxO*) in cereals have been implicated to play a role in defense response to pathogen infection. Here, we analyzed the molecular changes in members of rice *OxO* (*OsOxO*) genes mapped to chromosome 3 associated with resistance to blast. *OxOs* are members of the cupin proteins with the conserved motifs ([G(X)5HXH-(X)3,4E(X)6G] and G(X)5PXG(X)2H(X)3N]) forming the active sites of the enzyme. There are four tandemly duplicated rice oxalate oxidases (*OsOxO*) exhibiting >90% similarity at the nucleotide and amino acid levels. Expression analysis using resistant and susceptible advanced backcross lines of Vandana/Moroberekan showed that only *OsOxO4* is expressed during rice-*Magnaporthe oryzae* interaction. Expression data from the Rice Gene Indices and Rice MPSS databases show that *OsOxO1* is expressed at the flowering stage, *OsOxO3* is in roots, and *OsOxO4* is in leaves; and these are expressed during biotic and abiotic stress conditions. No expression evidence was found for *OsOxO2*. Analyses of the overall *OsOxO* sequence from 62 rice cultivars show that purifying selection is the major factor maintaining *OxO* protein homogeneity although a 134-bp conserved region in *OsOxO4* was under diversifying selection. Phylogeny of the 62 cultivars using the *OsOxO* sequences show that majority of the cultivars fall into the expected subgroups –indica, aus, tropical japonica, and temperate japonica. Characterization of SNPs and InDels in each gene across cultivars using PolyPhen and SIFT analyses tools identified sites that are of functional importance. Plotting the SNPs against the different conserved domains of *OsOxO* shows that these are highly conserved, and that SNPs found in these regions were either silent mutations or which led to tolerated amino acid changes. No SNPs or InDels were found in motif 16, a unique motif in cereal *OxOs*. Overall, our data suggest that *OxO* in rice are conserved among different cultivars suggesting an important role in diverse biotic stress and developmental pathways.

Characterizing *OxOs* using diverse germplasm

Table 1. Rice germplasm representing different variety groups used for sequencing of four oxalate oxidases in rice

Germplasm	Origin	Subgroup¹	Isozyme Grp²	% DLA (Blast Nursery)³	SES score (0-9)⁴ PO6-6	CA89	M64-1-3-9-1
Apo (IR55423-01)	Philippines	Indica	I	10	4	2	2
ARC 10177	India	Aus/Indica	II	30	2	1	1
Aus Jota	Bangladesh	Aus	II	35	2	1	5
Azucena	Philippines	Aromatic/Tropical japonica	VI	33	5	3	1
Bala	Philippines	Indica	I	30	2	1	5
Basmati 370	India	Aromatic	V	12	0	0	2
Black Gora	India	Aus	II	12	0	0	3
Carolina Gold	USA	Tropical japonica	-	45	2	0	5
Chhote Dhan	Nepal	Aromatic	V	25	1	1	5
Chodongji	Korea	Temperate japonica	VI	15	2	1	1
CO39	India	Indica	I	80	7	9	8
CT9993-5-10-1-M	Thailand	Tropical japonica	VI	15	2	0	0
Dee-Geo-Woo-Gen	Taiwan	Indica	I	20	3	2	1
Donghae	Korea	Temperate japonica	-	25	3	5	2
Dongjin	Korea	Temperate japonica	-	35	5	3	0
Dular	India	Aromatic/Aus	II	15	0	0	3
Fandrapotsy 104	Madagascar	Indica	-	65	2	3	2
Gerdeh	Iran	Aromatic/Tropical Japonica	II	25	3	3	0
Gihobyeo	Korea	Temperate japonica	-	30	6	7	1
GL33	Korea	Temperate japonica	-	40	4	2	3
Hwacheong	Korea	Temperate japonica	-	25	2	4	1
Ilmibyeo	Korea	Temperate japonica	-	15	2	2	4
Ilpumbyeo	Korea	Temperate japonica	-	12	1	2	3
Iguape Cateto	Brazil	Tropical japonica	VI	10	1	0	1
IR36	Philippines	Aromatic/Indica	I	30	4	3	2
IR64	Philippines	Indica	I	15	3	4	0
IR60080-46A	Philippines	Indica/Tropical japonica	VI	30	3	1	1
IR62266-42-6-2	Philippines	Indica/Tropical japonica	I	10	1	0	1
Jing Xi 17	China	Aromatic/Aus	-	25	0	0	2
Junam	Korea	Temperate japonica	-	35	1	2	3
Kalukantha	Sri Lanka	Indica	I	70	5	4	6
Khao Dawk Mali 105	Thailand	Aromatic/Aus	I	45	1	3	5
Khao Kap Xang	Laos	Tropical japonica	I or VI	15	0	1	3
Kun Min Tsieh Hunan	Taiwan	Indica	I	50	4	5	2
Lemont	USA	Tropical japonica	VI	15	1	0	0
Leuang Pratew	Thailand	Indica	I	25	3	2	2
Matia-Aman 53-13	Bangladesh	Indica	IV	15	2	2	4
Milyang 23	Korea	Indica/Japonica	I	45	5	6	3
Milyang30	Korea	Indica/japonica	-	25	3	2	3
Moroberekan	Guinea	Temperate japonica	VI	5	1	0	2
N22	India	Aus	II	5	0	0	3
Nagdong	Korea	Temperate japonica	-	35	5	3	3
Nipponbare	Japan	Temperate japonica	VI	35	1	3	0
Pacholinha	Brazil	Indica/Tropical japonica	-	40	3	4	3
Palgong	Korea	Aromatic/Temperate japonica	-	15	2	3	5
Patbyeo	Korea	Temperate japonica	VI	10	1	1	2
Peta	Philippines	Aus/Indica	I	55	3	4	1
Pokkali	India	Aus/Indica	I	25	2	1	3
Rathal	Sri Lanka	Indica/Tropical japonica	VI	20	2	1	2
Rayada	Bangladesh	Aus/Deepwater 4	IV	20	2	2	1
RTS12	Taiwan	Indica	I	10	3	2	2
Seomjin	Korea	Temperate japonica	-	25	3	3	0
Sindongjin	Korea	Temperate japonica	-	25	2	3	2
Sintane Diofor	Burkina Faso	Aus/Indica	I	55	2	5	2
Suwon345	Korea	Temperate japonica	-	35	3	3	3
Taebaek	Korea	Indica/Temperate Japonica	-	10	2	3	2
Tchampa	Iran	Aromatic/Temperate Japonica	V	5	0	0	0
Teqing	China	Indica	I	10	1	0	0
Vandana	India	Aus/Indica	I	65	6	2	6
Vary Vato 462	Madagascar	Indica	-	15	2	2	2
Yancoussa	Ivory Coast	Indica/Tropical japonica	VI	15	1	0	0
Zenith	USA	Temperate japonica	-	30	3	4	3

¹Subgroup refers to the group determined by SSR markers and chloroplast loci (Garris et al.; 2005) and by IRRI (McNally, unpublished data).
² Isozyme grouping defined by Glaszman (1987). (-) indicates no data available.
³ Percent diseased leaf area (DLA) at the IRRI Blast Nursery.
⁴ Standard Evaluation System for rice (http://www.knowledgebank.irri.org/ses/)

To characterize rice *OXOs* we sequenced the four *OsOxO* genes (*OsOxO1*, *OsOxO2*, *OsOxO3* and *OsOxO4*) from 62 diverse rice cultivars. These cultivars were selected on the basis of their value in breeding and relative diversity with each other. These cultivars are classified based on isozymes (Grp 1-VI) and later by SSR - indica (I), aus/boro (II), photoperiod insensitive deepwater rice of Bangladesh (III), long-duration, photoperiod-sensitive floating rice of Bangladesh (IV), aromatic rice (V) and japonica (VI).

Three primer pairs were designed for each gene that span the 1000-bp upstream, gene-coding, and 3' UTR regions. These primers are currently being used to develop near-isogenic lines containing these genes.

The cultivars were evaluated for reaction to rice blast isolates in the greenhouse and lineages occurring in the blast nursery. The cultivars had varying degrees of resistance to rice blast isolates PO6-6, CA89, and M64-1-3-9-1 and to field blast. No significant association was observed between haplotypes and reaction to blast. This may be due to epistatic effects brought about by major resistance genes or a combination of other defense genes present in the cultivars.

Nucleotide diversity of rice *OxOs*

Table 2. Nucleotide variation in the coding region of oxalate oxidases from 62 rice cultivars.

Parameter	OsOxo1	OsOxo2	OsOxo3	OsOxo4
Gene-coding region (bp)	684	690	684	690
A. Total polymorphic and InDel sites	52	51	34	56
Number of synonymous sites	15	17	5	20
Number of nonsynonymous sites	27	28	10	18
B. Polymorphic Sites				
Total polymorphic sites	33	44	18	34
Single nucleotide polymorphism (SNP)	28	42	16	31
Dinucleotide polymorphism	5	2	1	1
Trinucleotide polymorphism	0	0	1	2
> 3 bases polymorphism	0	0	0	1
C. Insertions/Deletions (InDels)				
Total InDels	8	9	16	22
Single-base InDel	7	8	14	21
Dinucleotide InDel	0	0	1	1
Trinucleotide InDel	0	0	0	0
> 3 bases Indel	1	1	1	0

In general, there were more single nucleotide polymorphisms than dinucleotide or trinucleotide polymorphisms in all four genes. *OsOxO4* had the most counts for both nucleotide polymorphisms and InDels, while *OsOxO3* had the least variation. While a lot of variations can be found in *OsOxO4*, the ratio of synonymous substitutions (SS) to non-synonymous substitutions (NSS) is 1:1, while *OsOxO1-3* had 1:2 SS:NSS.

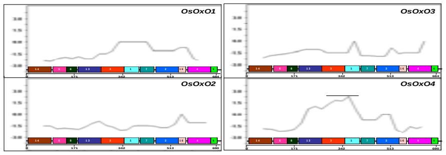

Fig. 1. Sliding window plots of Tajima's D estimates for *OsOxO* (coding region) genes in rice identified regions under diversifying selection (DNASP 4.0). The *OxO* motifs identified by MEME analyses is superimposed. A bar indicates regions under diversifying selection (P < 0.10). *OsOxO1, OsOxO2 and OsOxO3* were under purifying selection. For *OsOxO4*, most of the regions also have a negative values for Tajima's D, except for 134-bp window (indicated by a bar), which is significant for diversifying selection. These region span the motifs 3 and 1 which is the conserved domain for the active site for cupins such as *OxOs*.

SNPs and InDels in *OxOs*

• To assess whether SNPs and InDels in the OXO sequences potentially changethe protein structure of the resulting protein, two programs were used - SIFT (Sorting Intolerant F Tolerant) and PolyPhen.

• SIFT (http://blocks.fhcrc.org/sift/SIFT.html)) searches for similar protein sequences from different species in the database, obtains the multiple alignments of these sequences, and then calculates from the alignment the tolerance index (from 0 to 1) for all possible substitutions at each position. The higher a tolerance index, the less functional impact a particular amino acid substitution is likely to have.

• PolyPhen (http://genetics.bwh.harvard.edu/pph/) searches for 3D protein structures, performs multiple alignments of homologous sequences and amino acid contact information in several protein structure databases, calculates position-specific independent counts (PSIC) scores for each of two variants, and then computes the PSIC score difference of two variants.

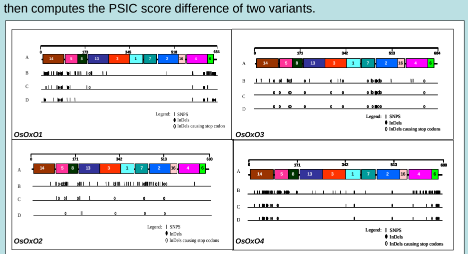

Fig. 2. Location of SNPs and InDels in the coding region of *OsOxOs* from 62 rice cultivars. a) The motifs are based on MEME analyses. b) All SNPs and InDels in the gene-coding region. c) SNPs and InDels that result to intolerant amino acid substitutions identified by SIFT. d) SNPs and InDels that result in potentially damaging amino acid substitutions identified by PolyPhen. Motif 14 is the 23-AA signal peptide sequence, while motif 16 is an eight-amino acid residue unique in *OxOs*. Motifs 1, 3, and 7 correspond to the conserved cupin domain active site. No SNPs or InDels were found in motif 16, a unique motif in cereal *OxOs*.

Analyses of the 1000 bp upstream regions of *OxOs*

Table 3. Copy number of *cis*-elements in the 1000 bp upstream region of rice *OXO* identified by the Plant *Cis*-Acting Elements database (PLACE).

Class	Cis-element	OsOxO1	OsOxO2	OsOxO3	OsOxO4
a. Tissue					
Roots	NODCON2GM	2	5	2	5
	OSE1ROOTNODULE	5	0	1	1
	OSE2ROOTNODULE	2	5	2	8
	RAV1AAT	3	4	1	2
	RHERPATEXPA7	0	0	3	4
	ROOTMOTIFTAPOX1	7	7	5	2
Pollen	POLLEN1LELAT52	8	6	7	16
	GTGANTG10	1	3	8	9
Seeds	EBOXBNNAPA	12	8	10	12
	RYREPEATBNNAPA	0	5	3	3
	RYREPEATLEGUMINBOX	0	4	3	2
b. Stress					
Disease	ARR1AT	13	7	7	12
	ELRECOREPCRP1	1	1	0	1
	GT1GMSCAM4	2	1	1	7
	SEBFCONSSTPR10A	0	2	0	2
	WRKY71OS	5	3	5	5
	WBOXES	2	8	10	9
Dehydration	MYB1AT	2	1	0	1
	MYB2CONSENSUSAT	0	1	1	2
	MYB2AT	0	1	1	0
	ABRELATERD1	0	0	4	3
	ACGTATERD1	6	2	8	6
	CBFHV	0	0	0	2
	MYBCORE	1	2	2	3
	MYCCONSENSUSAT	12	8	10	12
Metal	CURECORECR	2	6	6	2
Light	IBOXCORE	2	3	3	4
	GT1CONSENSUS	12	6	5	12
Salinity	GT1GMSCAM4	2	1	1	7
c. Transcription Factors	DOFCOREZM	27	7	15	28

To identify the *cis*-elements present in each promoter region, sequences were analyzed using PLACE (http://www.dna.affrc.go.jp/PLACE/). The types of *cis*-elements found can be categorized into tissue-specific, stress-related or transcription factor-related *cis*-elements. Tissue-specific *cis*-elements included those that are associated with expression in roots, pollen, and seeds and that occur in different locations in the 1000-bp upstream region as well as in different copy numbers. A 475-bp insertion similar to the Pong subclass of transposons was detected in *OsOxO2* 1000-bp upstream region of cultivars Rayada, Dular, Aus Jota, Black Gora, and ARC10177.

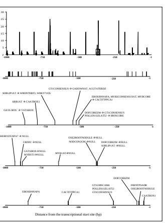

Fig. 4. Polymorphic sites and InDels in the 1000-bp upstream region of *OsOxO4*. A) Number of cultivars with polymorphisms and polymorphic sites in *OsOxO4*. Shaded boxes indicate four or more polymorphic or InDel sites. B) Polymorphic sites located between *cis*-elements. C) Polymorphisms which result in conversion of *cis*-element. D) Polymorphisms which result in a loss of *cis*-elements. E) Polymorphisms which result in silent mutations of *cis*-elements. A 26 bp deletion was observed in the 1000-bp upstream region of *OsOxO4* of cultivars Leuang Pratew, IR64, Vandana and Peta. These sequences contain *cis*-elements related to bacterial responsive nodulation genes active in the infected cells of the nitrogen-fixing zone of root nodules.

Expression of rice *OxOs*

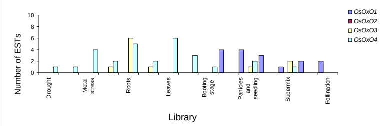

Fig. 5. Abundance of *OsOxO* expressed sequence tags (ESTs) from the DFCI *Oryza sativa* Gene Index (OsGI). No ESTs were detected for *OsOxO2*.

EST abundance of rice *OxOs* were retrieved from the Rice Gene Index (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=Rice). No ESTs were identified for *OsOxO2*. *OsOxO1* was expressed in the panicles and at flowering and pollination stages. *OsOxO3* ESTs were found in roots and seed libraries. *OsOxO4* ESTs were detected in drought, cold stress, and metal ($CuSO_4$) stress libraries as well as in root, shoot and leaf libraries. It was also expressed at tillering and booting stages.

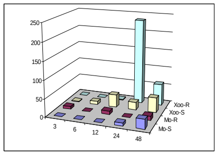

Fig. 6. MPSS expression profile of *OsOxO4* plotted in parts per million of transcripts in rice. Only *OsOxO4* was expressed in *Magnaporthe oryzae* (*Mo*) susceptible (*Mo*-S) and resistant (*Mo*-R) interactions and in rice-*Xanthomonas oryzae* pv *oryzae (Xoo)* susceptible (*Xoo*-S) and resistant interaction (*Xoo*-R). Leaf samples for rice-*Mg* interaction were collected from Nipponbare (susceptible) and Nipponbare-*Pi9* (resistant) inoculated with *Mo*. Leaf samples for rice-*Xoo* interaction were collected from Nipponbare (susceptible) and Nipponbare-*Xa21* (resistant) inoculated with *Xoo*.

Performance of *OsOxOs* in combination with defense genes

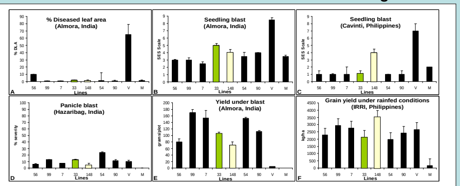

Fig. 7. Performance of promising Vandana/Moroberekan intermated advanced backcross lines with resistance to blast and tolerance for drought. We identified Vandana/Moroberekan lines with different candidate defense gene alleles from Moroberekan but with high similarity to Vandana. A-D) Resistance to blast was determined by screening advanced lines in blast hotspots in India and the Philippines (2004-05). E) Yield under blast was measured in Almora, India. F-I) Advanced lines were evaluated for phenotypic similarity to Vandana and drought tolerance at IRRI (2004 wet season, rainfed conditions, 5mm average rainfall at flowering stage). Lines 33 and 48 performed well under reproductive drought stress. These are also candidates for agronomic /testing in station trials in India. Lines: 56 - IR78221-19-6-56-B (Chi, OxO, POX, OXLP); 99 – IR78221-19-6-99-B (OxO); 7 – IR78221-6-7-B (OxO, HSP90, POX); 33 – IR78221-19-6-33-B (OxO, POX, OXLP); 148 – IR78222-7-148-2-B (Chi, OxO, OXLP, DHAP); IR78221-19-4-54-B (Chi, OxO, Thau, POX, OXLP); 90 – IR78221-19-6-90 (OxO, POX, OXLP); V – Vandana; and M – Moroberekan. Chi – chitinase; Pox – peroxidase; Thau – thaumatin; OXLP – oxalate oxidase like protein; DHAP - Deoxyphosphoheptonate aldolase).

Summary and conclusions

- The rice oxalate oxidases were sequenced from 62 rice cultivars belonging to different subgroups of rice. Cultivars in the same haplotype group did not necessarily have the same reaction to a blast isolate. This may be due to epistatic effects brought about by major resistance genes or combination of other defense genes present in the cultivars.
- Contrary to the divergent evolution of *R* genes, *OxOs* in rice are under purifying selection although regions in *OsOxO* were under diversifying selection. This observation is consistent with the reported roles of *OxOs* in general defense against pathogens.
- SNPs and InDels that were of functional importance were identified by both PolyPhen and SIFT analyses.
- The 1000-bp upstream region of the OxOs revealed several classes of *cis*-elements that point to their role in biotic and abiotic stress conditions as well as at different developmental stages in the rice plant.
- *OxOs*, in combination with other defense response genes, were successfully introgressed into Vandana, a drought-tolerant popular variety from India. Lines showing resistance to field blast in India and the Philippines and with comparable yield to Vandana under irrigated conditions have been identified.

IRRI INTERNATIONAL RICE RESEARCH INSTITUTE
Generation
RURAL DEVELOPMENT ADMINISTRATION