

Supplement Figure S1 Setter et al. 2010

Supplement Figure 2. Alignment of amino acid sequences for aldehyde oxidases and xanthine dehydrogenases in maize (Zm), rice (Os) and Arabidopsis (At). Also shown are reference sequences from human and mouse. Maize sequences from www.maizesequence.org 4a.53. Three regions of the sequences are shown, around amino acids corresponding to mouse-1265, mouse-886, and mouse-806 (marked by arrows). See text for explanation.

		1390	1400	1410	1420	1430	144	40	1450
	osAO2_NP_001051518.1/1-1375	GLVINDGTWTYK			· · _ ·	, <b>1</b> ,			
	osAO3_NP_001051517.1/1-1356	GLVINDGTWTYKI							
	osAO1_NP_001064133.1/1-1387	GLVIHDGTWTYKI							
	zmAO2_bin5.01_NP_001105309.1/1-1349	GLVIHDGTWTYKI							
	zmAO4_bin1.11_GRMZM2G141473_/1-1357	GLVIHDGTWTYKI							
	zmAO1 bin1.11 GRMZM2G141535/1-1358	GLVIHDGTWTYKI							
	zmAO3_bin1.11_GRMZM2G124260_sn/1-128	3 <mark>g</mark> m v i h <mark>d g t</mark> wty <mark>k</mark> i	I P T V D T I P P	QFNVELINS	A <mark>g</mark> dr <mark>kr</mark> vl <mark>ss</mark>	K <mark>as</mark> gepplll	.AS <mark>SVH</mark> CAM <mark>b</mark>	EAIRAAR	KEFS
	atAAO3_NP_001077966.1/1-1332	GLVVQQ <mark>GT</mark> WDY <mark>K</mark> I		HENVEIVNT	G H H K N <mark>r</mark> v l s s	K <mark>as</mark> gepplll	.AASVHCAT <mark></mark>	SAIREARK	(H <mark>S</mark> L:
	atAAO4_NP_563711.1/1-1337	G L V N E E <mark>G T</mark> WD <mark>Y K</mark> I		QFNVQILNS	G H H K N <mark>r</mark> v l s s	K <mark>as</mark> geppllv	AASVHCAT <mark>F</mark>	SAIREARK	(QYL:
	atAAO1_NP_568407.2/1-1368	<mark>G</mark> LVVT <mark>D</mark> S <mark>TWTY</mark> KI	I <mark>P T</mark> V <mark>D T I P F</mark>	QFNVEILNS	GQHK <mark>nrvl</mark> ss	K <mark>as</mark> gepplll	.AA <mark>svh</mark> cav <mark>f</mark>	AAV <mark>k</mark> ea <mark>r</mark> i	( <mark>Q  </mark> L :
	atAAO2_NP_189946.1/1-1321	G L L L T <mark>D S T W T Y K</mark> I							
	osXDH_NP_001050420.1/1-1369	GHLFTC <mark>GPG<mark>Syk</mark>i</mark>							
	zmXDH_bin1.04_GRMZM2G050984/1-1384	G <mark>h</mark> lftcgpg <mark>syk</mark> i							
	atXDH1_NP_195215.2/1-1361	G <mark>S</mark> LLTC <mark>GPGNY</mark> KI							
1265	atXDH_NP_195216.2_chr04/1-1353	G <mark>SLL</mark> TC <mark>GPGSY</mark> KI							
	hsXOR_human_NP_000370_xanthine/1-1333								
mouse	mAOX1_mouse_NP_033806/1-1333	<mark>g t l y</mark> s r <mark>g p</mark> n q <mark>y k</mark> i	PALCD   P 1	EMHISFLPP	SEHSNT <mark>lyss</mark>	KGLGESGVFL	GCSVFFAIH	IDAV <mark>K</mark> AA <mark>RO</mark>	ERG
		980	9:	90 1	000	1010	1020	1030	10
	osAO2_NP_001051518.1/1-1375	<mark>d</mark> mima <mark>gg<mark>rh</mark>pm<mark>k</mark>v</mark>				1 I M			KVCK
	osAO3 NP 001051517.1/1-1356	DMIMAGGRHPMKV							
	osAO1_NP_001064133.1/1-1387	DMIMAGG <mark>RHP</mark> MKA							
	zmAO2_bin5.01_NP_001105309.1/1-1349	DMIMAGGRHPMKV							
	zmAO4 bin1.11 GRMZM2G141473 /1-1357	DMIMAGG <mark>RHP</mark> MKV							
	zmAO1_bin1.11_GRMZM2G141535/1-1358	DMIMAGG <mark>RH</mark> PM <mark>K</mark> A	KYSVGFKS	DGKITALHL	DLGINAGISP	DVSP-LMPRA	I I GALK <mark>KY</mark> N	W <mark>G</mark> TLEF <mark>D</mark> T	<mark>kvck</mark>
	zmAO3_bin1.11_GRMZM2G124260_sn/1-1283	8 <mark>d milagg<mark>rh</mark>pm<mark>kt</mark></mark>	KYSVGFKS	DG <mark>KIT</mark> AVHL <mark>I</mark>	DLGL <mark>N</mark> AGI <mark>T</mark> P	DLSA-ILPNT	I I <mark>g</mark> s f k <mark>ky</mark> n	W <mark>g</mark> alaf <mark>d</mark> i	KLCK
	atAAO3_NP_001077966.1/1-1332	DMIMAGG <mark>RH</mark> PMKI	NYNVGFRS	DG <mark>K</mark> L <mark>T</mark> ALEL1	TMLIDA <mark>gl</mark> e <mark>p</mark>	DV <mark>SP</mark> - IM <mark>P</mark> RN	IM <mark>g p l r k y</mark> d	WGALSFDV	K <mark>VCK</mark>
	atAAO4_NP_563711.1/1-1337	DMIMAGG <mark>RH</mark> PLKI	TYSVGFRS	DG <mark>K</mark> L <mark>T</mark> ALDLN	NLFIDA <mark>gs</mark> dv	DVSL-VMPQN	IMNSLR <mark>ky</mark> d	W <mark>G</mark> A L S F <mark>D</mark> I	K <mark>VCK</mark>
	atAAO1_NP_568407.2/1-1368	DMI <mark>T</mark> TGG <mark>RH</mark> PM <mark>K</mark> V							
	atAAO2_NP_189946.1/1-1321	DMIMTGG <mark>RHP</mark> M <mark>K</mark> I							
	osXDH_NP_001050420.1/1-1369	DMMTTGQ <mark>RHS</mark> FLG							
	zmXDH_bin1.04_GRMZM2G050984/1-1384	DMISTGQRHSFLG							
004	atXDH1_NP_195215.2/1-1361	DMM I T <mark>G H<mark>RHS</mark>FLG</mark>							
884	atXDH_NP_195216.2_chr04/1-1353	DMM I T <mark>G H<mark>RHS</mark>FVG</mark>							
mouse	hsXOR_human_NP_000370_xanthine/1-1333	DML I TGG <mark>RHPF</mark> LA							
	mAOX1_mouse_NP_033806/1-1333	DMLITGG <mark>RHPY</mark> LG	KYKAGEMN	DGRILALDVE			LLKLDNA <mark>y</mark> k	FPNLRCRG	WACR
			20	930	<u> </u>	950	960	970	980
	osAO2_NP_001051518.1/1-1375	EITQNTVARCLGI							
	osAO3_NP_001051517.1/1-1356	EITQNTVARCL <mark>G</mark> I							
	osAO1_NP_001064133.1/1-1387	<mark>evto</mark> divarci <mark>g</mark> v							
	zmAO2_bin5.01_NP_001105309.1/1-1349	ESTQNVVAKCV <mark>g</mark> i							
	zmAO4_bin1.11_GRMZM2G141473_/1-1357	EVTONVVAKCLGI							
	zmAO1_bin1.11_GRMZM2G141535/1-1358 zmAO3_bin1.11_GRMZM2G124260_sn/1-1283	EL TONLIARCEGI							
	atAAO3_NP_001077966.1/1-1332	EYVHSVIATCLGI							
	atAAO4_NP_563711.1/1-1337	EFTQTVIATCLGI							
	atAAO1_NP_568407.2/1-1368	EFVHQTIAGCLGV							
	atAAO2_NP_189946.1/1-1321	QYVQSSVAACLGI							
	osXDH_NP_001050420.1/1-1369	QKHQKYVANVLGL							
	zmXDH_bin1.04_GRMZM2G050984/1-1384	QKHQKYVASVLGL							
	atXDH1_NP_195215.2/1-1361	QKHQKYVSHVLGL							
	atXDH_NP_195216.2_chr04/1-1353	QQHQKYVSHVLGL							
806	hsXOR_human_NP_000370_xanthine/1-1333	MKTQSFVAKMLGV							
	mAOX1_mouse_NP_033806/1-1333	KYIQDIVAA <mark>T</mark> LKL							
mouse									

**Supplementary Figure S3.** Plot of squared correlations of allele frequencies (r<sup>2</sup>) against distance between polymorphic sites (SNPs) in the aldehyde oxidase ZmAO3 candidate gene. The curve shows a nonlinear regression (logarithmic) of r<sup>2</sup> on distance. Similar results were obtained with a Loess curve. The predicted value of r<sup>2</sup> rapidly declines to 0.2 by 800 bp and reaches background levels (0.1) by 3000 bp in ZmAO3. The SNPs used to generate this plot are from a 3500 bp region surrounding SNP position 285,274,032 bp on chromosome 1 in 27 diverse maize inbred lines. The investigated SNPs are an integration of HapMap SNPs (Gore et al., 2009) and loci genotyped in this study. The 27 maize inbred lines are parents of the Maize Nested Association (NAM) population and the Intermated B73 X Mo17 (IBM) population. Eight (CML103, CML228, CML247, CML277, CML322, CML333, CML52, and CML69) of the 27 lines are derived from the CIMMYT breeding program and were included in this association study.

