

RESEARCH ARTICLE

Open Access

# QTLs and candidate genes for desiccation and abscisic acid content in maize kernels

Valérie Capelle<sup>1,2</sup>, Carine Remoué<sup>2,3</sup>, Laurence Moreau<sup>4</sup>, Agnès Reyss<sup>1,2</sup>, Aline Mahé<sup>1,2</sup>, Agnès Massonneau<sup>5,6</sup>, Matthieu Falque<sup>4</sup>, Alain Charcosset<sup>4</sup>, Claudine Thévenot<sup>1,2</sup>, Peter Rogowsky<sup>5</sup>, Sylvie Coursol<sup>4</sup>, Jean-Louis Prioul<sup>1,2\*</sup>

## Abstract

**Background:** Kernel moisture at harvest is an important trait since a low value is required to prevent unexpected early germination and ensure seed preservation. It is also well known that early germination occurs in viviparous mutants, which are impaired in abscisic acid (ABA) biosynthesis. To provide some insight into the genetic determinism of kernel desiccation in maize, quantitative trait loci (QTLs) were detected for traits related to kernel moisture and ABA content in both embryo and endosperm during kernel desiccation. In parallel, the expression and mapping of genes involved in kernel desiccation and ABA biosynthesis, were examined to detect candidate genes.

**Results:** The use of an intermated recombinant inbred line population allowed for precise QTL mapping. For 29 traits examined in an unreplicated time course trial of days after pollination, a total of 78 QTLs were detected, 43 being related to kernel desiccation, 15 to kernel weight and 20 to ABA content. Multi QTL models explained 35 to 50% of the phenotypic variation for traits related to water status, indicating a large genetic control amenable to breeding. Ten of the 20 loci controlling ABA content colocalized with previously detected QTLs controlling water status and ABA content in water stressed leaves. Mapping of candidate genes associated with kernel desiccation and ABA biosynthesis revealed several colocalizations between genes with putative functions and QTLs. Parallel investigation via RT-PCR experiments showed that the expression patterns of the ABA-responsive *Rab17* and *Rab28* genes as well as the late embryogenesis abundant *Emb5* and aquaporin genes were related to desiccation rate and parental allele effect. Database searches led to the identification and mapping of two *zeaxanthin epoxidase* (*ZEP*) and five novel *9-cis-epoxycarotenoid dioxygenase* (*NCED*) related genes, both gene families being involved in ABA biosynthesis. The expression of these genes appeared independent in the embryo and endosperm and not correlated with ABA content in either tissue.

**Conclusions:** A high resolution QTL map for kernel desiccation and ABA content in embryo and endosperm showed several precise colocalizations between desiccation and ABA traits. Five new members of the maize *NCED* gene family and another maize *ZEP* gene were identified and mapped. Among all the identified candidates, aquaporins and members of the *Responsive to ABA* gene family appeared better candidates than *NCEDs* and *ZEPs*.

## Background

Maize (*Zea mays*) kernel moisture at harvest is an important trait in temperate regions because costly additional drying is needed to reach 15% water content, which is the level compatible with good seed preservation during storage. Although yield is correlated to maximum water content occurring approximately 40-60 days after pollination (DAP) [1] and to a lesser extent to

kernel moisture at maturity [2], the yield/moisture ratio at maturity is variable enough to allow selection for both higher yield and lower moisture at harvest [3,4]. For example, recurrent selection has been successfully applied for reduction of kernel moisture by the introduction of tropical germplasm into temperate-adapted germplasm [5]. Indirect inbred selection criteria to reduce gain moisture, based on husk senescence have been proposed [6]. The biochemical, biophysical and molecular phenomena controlling kernel moisture at harvest intervene mainly during the maturation phase,

\* Correspondence: [jean-louis.prioul@u-psud.fr](mailto:jean-louis.prioul@u-psud.fr)

<sup>1</sup>Univ Paris-Sud, Institut de Biotechnologie des Plantes, Bât 630, F-91405 Orsay, France

which corresponds to the last stage of seed development after the early and the grain filling phases. Physiological and genetic analyses of the maturation phase reveal a process, which both prevents early embryo germination and favours the synthesis of specialized proteins related to the acquisition of desiccation tolerance, enhancing embryo viability under strong dehydration. Accordingly, in many seed species including maize, embryos separated from endosperm at the early developmental phase can grow and germinate when placed in tissue culture, but their germination ability decreases as maturation proceeds [7].

Despite considerable progress in recent years in knowledge of maturation, the number of genes thought to be involved in regulation of kernel moisture remains extremely limited. The late embryogenesis abundant (LEA) proteins including the dehydrin family, are specifically produced during the maturation phase [8]. Although they have been assumed for a long time to protect cellular and molecular structures from the damaging effect of desiccation [8], only recent results shed some light on their precise action. Cytoplasmic LEA proteins prevent protein aggregation due to water loss *in vitro* [9] and mitochondria LEA proteins protect two matrix enzymes, fumarase and rhodanese [10]. Other proteins likely to be involved in kernel desiccation are the water channel aquaporins, which are present in nearly all organs. An extensive study of maize aquaporins described 31 full length cDNAs distributed into four groups comprising 13 plasma membrane (PIP) and 11 tonoplast (TIP) intrinsic proteins [11]. Among them, *ZmPIP1;1*, *ZmPIP1;2*, *ZmPIP1;3*, *ZmPIP2;1*, *ZmPIP2;2*, *ZmPIP2;3*, *TIP1;1* and *TIP2;1* were reported to be expressed in reproductive tissues. In rice (*Oryza sativa*), *OsTIP1* and *OsTIP3* are expressed in mature seeds in the embryo and the aleurone layer, respectively. Because members of the PIP2 and TIP1 families have much higher water transport capacities than those of the PIP1 family [12], aquaporins of the first two families may be of higher significance for desiccation.

The phytohormone abscisic acid (ABA) appears to play a central role in both the establishment of embryo dormancy and the synthesis of LEA proteins, as demonstrated by mutants impaired in ABA synthesis or sensitivity. ABA deficient maize mutants are viviparous, i.e. embryos germinate precociously on the ear [13,14], their vivipary being prevented by ABA addition. ABA synthesis mutants of Arabidopsis (*Arabidopsis thaliana*) and tomato (*Solanum lycopersicum*) have also impaired seed maturation and dormancy but are not viviparous [15]. Interestingly, maize plants with white endosperm (*yellow* mutants) have higher moisture than those with yellow endosperm [16]. This is due to the fact that the *y1* mutation causes a defect in phytoene synthase (PSY), an

enzyme involved in both carotenoid and ABA synthesis [17], highlighting the likely role of ABA in regulating kernel moisture. Furthermore, the expression of many *LEA* genes and more generally members of the *Responsive to ABA (Rab)* gene family, is induced by exogenous ABA [18].

The ABA biosynthetic and catabolic pathways are now well understood since almost all the biosynthetic genes have been identified through the isolation of auxotrophic mutants [19]. The enzymes downstream of the xanthophyll cycle are specific to ABA biosynthesis. The cloning and characterization of maize *Viviparous14 (Vp14)*, which encodes 9-*cis*-epoxycarotenoid dioxygenase 1 (NCED1) catalyzing the cleavage of the C40 neoxanthin chain into the C15 ABA skeleton xanthoxin [20], led to the identification of NCED as a rate controlling enzyme. Indeed, maize *nced1* mutants have a strongly reduced kernel ABA content [21] and in Arabidopsis, *NCED1* overexpression confers a significant increase in ABA accumulation in the plant [22]. In Arabidopsis, nine *NCED*-related sequences have been identified and phylogenetic analysis has indicated that five of these clustered with functionally characterized NCED proteins from other species [23]. Aside from this main regulatory step in the ABA biosynthesis pathway, other metabolic steps of ABA metabolism also contribute to determining ABA level. One is the conversion of zeaxanthin into violaxanthin catalyzed by zeaxanthin epoxidase (ZEP), which is encoded by a single-copy gene in Arabidopsis and tobacco (*Nicotiana plumbaginifolia*) and whose overexpression causes an enhanced accumulation of ABA in seeds [24]. The cloning and characterization of the maize *Viviparous10/Viviparous13 (Vp10/Vp13)* and *Viviparous15 (Vp15)* genes further demonstrated that ABA biosynthesis is also dependent on a molybdenum cofactor involved in the abscisic aldehyde oxidase reaction, the last step of ABA biosynthesis [25,26]. In addition, the recent discovery that *Viviparous8 (Vp8)* encodes a putative peptidase, together with the predominant effect of *vp8* mutant on ABA synthesis and turnover in maize embryos, indicate that ABA level is also controlled indirectly through regulation of seed-specific factors [27].

The mutant approach is powerful in identifying the mandatory steps (genes) in a pathway, but it does not provide any insight into the relative impact of each step on the quantitative value of the final product (trait) of the pathway. The genetic variability of quantitative traits is controlled by one or generally several loci named quantitative trait loci (QTLs) which may be mapped using appropriate segregating populations. In addition, the relative contribution of each locus to the trait genetic variation and the allelic effect at each locus may be estimated. More than 20 years ago, Robertson proposed a very fruitful hypothesis bridging mutation and

QTL approaches by simply saying that the “qualitative and quantitative traits may be the result of different types of variation of genic DNA at the loci involved”; in other words minor allelic effects produce quantitative variations, while major variations (null alleles) produce qualitative variations (mutations) [28]. This opened the way for research aiming at the identification of the genes underlying QTLs. The considerable international efforts in mapping known function genes in maize now provide rather precise genetic maps that can be used to identify candidate genes from their map common location with detected QTLs [29]. This comparison is easier when dealing with physiological and biochemical traits since the number of possible candidate genes may be restricted to those encoding enzymes or cofactors acting in relevant pathways [30]. However, one limitation is the confidence interval of the QTL position which may reach more than 10 centimorgans (cM) in classical recombinant inbred lines (RILs) as illustrated in one of the few reports on QTL for kernel moisture and drying rate [31]. A way to reduce this interval is to increase the number of recombination events by the inclusion of four generations of random intermating after the second generation and before the single seed descent, thus providing intermated recombinant inbred lines in which the QTL confidence interval is substantially reduced by a factor of two to three [32]. The candidate gene selection is thus facilitated. A useful criterion to validate the identified candidate genes is to examine the corresponding gene expression during the process under investigation [33]. Differences in the transcription level related to the trait variation may support the role of the functional difference of the parental alleles.

Much remains to be learned about the genes explaining the variability in the desiccation rate and the genetic relationship between this process and ABA content. Here, we describe a QTL-candidate gene analysis of the desiccation process in maize using an intermated recombinant inbred line population (LHRF<sub>F<sub>3:4</sub></sub>) derived from the cross between the maize inbred lines F2 and F252 differing in desiccation rate. First, QTLs were searched for traits related to kernel moisture and ABA content in the endosperm and the embryo during kernel desiccation. Second, an extensive data mining of the genes mapped in the confidence interval of the QTLs was performed in order to short list candidate genes with annotations related to desiccation rate and/or ABA content. In addition to these *in silico* studies, six members of the *NCED* gene family and two members of the *ZEP* gene family were identified and mapped by PCR amplification and sequencing. Third, expression profiles of the candidate genes during desic-

cation were examined by RT-PCR experiments for correlations with kernel desiccation rates or changes in ABA content.

## Results

### Genetic variability in desiccation rate and ABA content

Kernel water content relative to dry weight (%DW) was evaluated in the two parental inbred lines and the segregating LHRF<sub>F<sub>3:4</sub></sub> population. This trait continuously declined from 12 DAP when the kernel was still in the filling stage and long before the onset of the maturation stage at 30 DAP (Fig. 1A, Table 1). However, kernel water content (g/kernel) reached a maximum between 30 to 40 DAP (data not shown). This maximum corresponded to the end of the intensive starch accumulation and indicated the beginning of the desiccation-maturation process. Thus, further data presentation was limited to the 30-80 DAP period. The two parental lines had different desiccation rates, especially after 40 DAP, F252 line having approximately 9% less moisture ((FW - DW)/FW\*100) than F2 line at harvest (Fig. 1A, Table 1).

Mean ABA concentration in the LHRF<sub>F<sub>3:4</sub></sub> population, when expressed on DW basis, declined drastically from 12 to 30 DAP and then, increased slightly up to 60 DAP (Fig. 1B). Similar kinetics albeit with a much lower amplitude between 12 and 30 DAP was observed when ABA was expressed on FW basis, whereas the ABA content per mg of water increased continuously and markedly after 40 DAP (Fig. 1C). Because ABA is water soluble, the later mode of expression was likely the most physiologically relevant, but also the most difficult to obtain when working with lyophilized powder as in the present experiments. The interpretation of global changes at the kernel level was further complicated by large differences between the different kernel parts, ABA being 5 to 60 times more concentrated in the embryo than in the endosperm (Table 1). In addition, the kinetics in the two tissues was also clearly different. In the embryo, the bell-shaped ABA concentration curve peaked at 60 DAP and remained high at 80 DAP. In contrast, ABA concentration continuously declined in the endosperm (Fig. 1D). Large genotypic variability among the 153 LHRF<sub>F<sub>3:4</sub></sub> lines was noted in the general trend as shown by the large standard deviations. Principal component analysis and Pearson coefficient tables with all the measured variables showed that FW was highly correlated to DW or kernel water content ( $r > 0.69$  to  $0.94$ ) at a given DAP stage, but not between stages (Additional file 1). By contrast, regardless of the considered stage, the correlation between endosperm ABA content and embryo ABA content was not significant or very low ( $r = 0.23$ ). This was also true across DAP stages

**Table 1 Characteristics of the parental lines and their offspring for desiccation rate and ABA content**

Trait	LHRF_F <sub>3:4</sub> mean ± SD	LHRF_F <sub>3:4</sub> min-max	F2 mean ± SD	F252 mean ± SD
FW30	188 ± 33	123-395	183 ± 14	180 ± 20
FW40	247 ± 36	156-351	260 ± 17	249 ± 19
FW60	316 ± 44	217-445	332 ± 36	319 ± 13
FW80	306 ± 50	189-450	314 ± 38	295 ± 33
DW30	55 ± 9	34-87	52 ± 7	55 ± 4
DW40	103 ± 14	67-146	111 ± 7	108 ± 9
DW60	183 ± 24	122-259	190 ± 20	190 ± 9
DW80	208 ± 30	138-288	206 ± 25	220 ± 18
%DW30	30 ± 4	17-39	29 ± 2	31 ± 1
%DW40	42 ± 3	32-49	43 ± 2	43 ± 1
%DW60	58 ± 3	46-66	57 ± 2	60 ± 1
%DW80	68 ± 4	58-78	66 ± 2	74 ± 3
Water30	133 ± 28	87-308	131 ± 8	124 ± 16
Water40	143 ± 25	85-222	149 ± 13	142 ± 10
Water60	132 ± 23	82-193	142 ± 19	128 ± 5
Water80	98 ± 24	43-165	108 ± 16	75 ± 16
ABAend30	87 ± 25	41-200	93 ± 8	114 ± 19
ABAend40	67 ± 14	38-121	72 ± 33	76 ± 5
ABAend60	47 ± 14	17-92	75 ± 27	85 ± 15
ABAend80	8 ± 4	2-21	79 ± 28	31
ABAemb30	278 ± 175	42-1004	273	212 ± 3
ABAemb40	352 ± 106	158-714	489 ± 56	408 ± 44
ABAemb60	630 ± 184	290-1356	619 ± 152	380 ± 0.2
ABAemb80	505 ± 178	115-1091	544 ± 105	350 ± 60
Slope	-0.083 ± 0.006	-0.100-0.070	-0.071	-0.085
Rate30_80	0.082 ± 0.006	0.060-0.096	0.074	0.086
Rate30_40	0.124 ± 0.045	0.020-0.240	0.142	0.119
Rate40_60	0.082 ± 0.020	0.010-0.130	0.072	0.084
Rate60_80	0.050 ± 0.018	0.010-0.110	0.042	0.073

The LHRF\_F<sub>3:4</sub> segregating population derived from the cross between the F2 and the F252 parental inbred lines differing in desiccation rate. FW: kernel fresh matter weight (mg/kernel); DW: kernel dry matter weight (mg/kernel); %DW = DW/FW × 100; Water: kernel water content (mg/kernel); ABAend: ABA in endosperm (pg/DW); ABAemb: ABA in embryo (pg/DW); Slope: regression line slope of (Water/FW × 100) as a function of thermal time; Rate = (Water/FW × 100)/(thermal time interval).

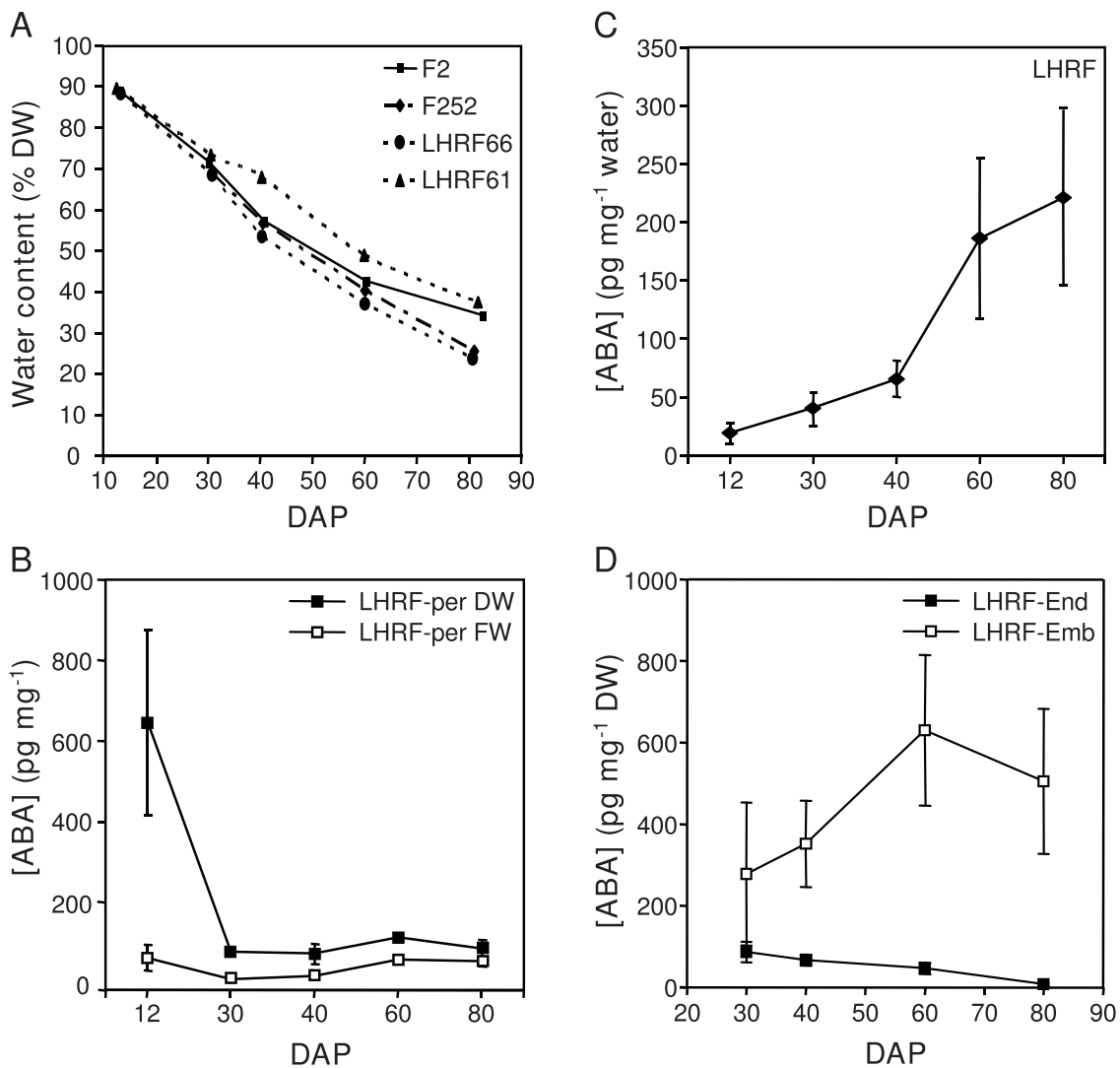
(Additional file 1). Some low but significant correlations were noted between some water-related variables and ABA content in the endosperm or the embryo at a same DAP stage (e.g. ABAemb80 and FW80, Additional file 1).

The trait variability was higher in the inbred lines than in parental lines, which is usual with a complex trait, the best line having lower moisture than F252 line and the worst line having higher moisture than F2 line (Fig. 1A for water content; Table 1 for others traits), illustrating the so-called transgression effect. The existence of positive and negative allelic effects for each trait is in favour of a genotypic origin of the transgressions observed at the phenotypic level (Table 2).

#### QTLs for kernel desiccation and ABA content

QTLs were searched at 30, 40, 60 and 80 DAP for three traits related to desiccation (Water, FW and %DW, blue

in Fig. 2), for one related to growth (DW, black in Fig. 2) and for ABA content either in embryo or endosperm (pink in Fig. 2). In addition, desiccation rate was evaluated by five traits, the slope of the FW decrease from 30 to 80 DAP vs. thermal time and the slope between each of the sampling dates (30 to 40, 40 to 60, 60 to 80 and 30 to 80 DAP, green in Fig. 2). A total of 78 QTLs were detected for 25 of the 29 examined traits (Table 2). Out of the 29 traits analyzed, 13 displayed at least one QTL with a genomewide *P* value below 5%, confirming unambiguously that their variation was unlikely due to environmental effect, but rather to a genetic effect (Table 2). On average, two to three QTLs were detected for each trait, their sum explaining from 6.1 (ABA\_embryo\_60\_1) to 52.8% (DW\_60) of the phenotypic variation albeit rather low LOD scores (2 to 4.6). The fact



**Figure 1 Time course of mean water status and ABA content in kernel in parents and inbreds.** The LHRF<sub>F<sub>3:4</sub></sub> segregating population derived from the cross between the F2 and the F252 parental inbred lines differing in desiccation rate (four intermated cycle were performed after the second generation and before single descent). A, water content expressed as a percentage of dry weight (% DW) in the two parental inbred lines and the best (LHRF61) and worst (LHRF66) LHRF<sub>F<sub>3:4</sub></sub> families. B, ABA concentration per kernel dry matter weight (DW) or per kernel fresh matter weight (FW) in the LHRF<sub>F<sub>3:4</sub></sub> population (LHRF). C, ABA concentration per kernel water in the LHRF<sub>F<sub>3:4</sub></sub> population (LHRF). D, ABA concentration per dry matter weight (DW) in endosperm (End) and embryo (Emb) of the LHRF<sub>F<sub>3:4</sub></sub> population (LHRF). Data from LHRF are means  $\pm$  SD;  $n = 153$ .

that each trait was measured on two plants per F<sub>3:4</sub> family from the same plot (unreplicated design) might explain these moderate effects. The experimental design did not allow for a correct estimation of heritabilities but one may assume they were low. In order to evaluate the uncertainties due to low heritability, genomewide risk was calculated for each QTL (see asterisks in Tables 2 and 3). As classically observed, the QTLs for different traits tend to be grouped in clusters that were not evenly distributed in the genome (Fig. 2). In nearly each cluster one or several QTLs were detected with a

genomewide risk below 5% (e.g. bins 1.04, 1.08, 2.05, 3.02, 4.04, 5.07, 7.03, 9.07). Forty three QTLs were related to kernel desiccation, 15 to kernel weight and 20 to ABA concentration (Fig. 2, Table 2).

The largest cluster on bin 4.04 only consisted of desiccation traits, whereas in the other clusters desiccation and growth traits were intermixed, which was expected due to the observed correlation between FW and DW at each DAP stage (Additional file 1). A focus on the 80 DAP stage at which maximum differences in drying were observed between both genotypes (Fig. 1A),

**Table 2 QTLs detected for kernel desiccation and ABA content at 30, 40, 60 and 80 DAP**

Number	QTL name	Chr. bin	Position	ProxMarker	DistMarker	Confidence interv.+/- LOD = 1	LOD value	R <sup>2</sup>	Additive value/ F252
1	FW_30_1	4.03	88	nc004	umc2039	76-104	2.54*	7.5	-13.203
2	FW_40_1	5.05	396	umc1822	umc54	372-424	2.04	6.1	15.444
3	FW_60_1	1.05	356	bnlg2295	umc1124	332-380	2.14	6.5	-15.819
4	FW_60_2	4.04	144	gsy4	mmc0471	140-160	2.02	6.1	-11.272
5	FW_60_3	8.09	504	Umc2052	umc1384	488-504	2.26	6.9	14.543
<i>cumulated</i>								<b>19.5</b>	
6	FW_80_1	1.04	336	Bnlg1811	dupssr26	308-364	2.58*	7.5	-16.785
7	FW_80_2	1.08	584	Umc83C	mmc0041	576-592	3.36***	9.7	-18.124
8	FW_80_3	2.02	80	mmc0111	gsy53b	48-92	2.17	6.4	15.758
9	FW_80_4	4.04	156	mmc0471	umc1088	144-188	2.29	6.7	-13.621
10	FW_80_5	8.09	496	Umc2052	umc1384	484-504	2.73*	8.1	21.382
<i>cumulated</i>								<b>38.4</b>	
11	DW_30_1	2.08	484	Umc1536	umc1049	468-496	2.2	6.5	-2.717
12	DW_30_2	5.01	60	mmc0151	umc147a	44-76	2.36	7	2.939
13	DW_30_3	9.07	480	gsy330	bnlg1129	460-492	4.09****	11.8	4.186
<i>cumulated</i>								<b>25.3</b>	
14	DW_40_1	5.05	392	Umc1822	umc54	372-420	2.1	<b>6.3</b>	5.466
15	DW_60_1	1.04	340	dupssr26	bnlg2295	332-364	4.52****	13.1	-9.324
16	DW_60_2	2.08	536	dupssr24	umc1464	512-548	2.97**	8.8	-6.879
17	DW_60_3	5.04	344	gsy34	umc1221	324-364	2.14	6.4	-7.937
18	DW_60_4	5.07	536	Umc1537	umc68	508-552	2.34	7	7.869
19	DW_60_5	5.07	580	Umc68	bnlg118	568-592	3.33****	9.9	-7.540
20	DW_60_6	8.09	500	Umc2052	umc1384	484-504	2.5*	7.6	8.503
<i>cumulated</i>								<b>52.8</b>	
21	DW_80_1	1.01	0	Bnlg1124	umc1354	0-20	2.08	6.2	7.824
22	DW_80_2	1.04	336	Bnlg1811	dupssr26	328-356	3.6***	10.3	-11.965
23	DW_80_3	1.08	584	Umc83C	mmc0041	572-596	2.03	6	-8.734
24	DW_80_4	5.04	324	Umc1110	bnl7_71	284-344	2.41	7	-9.851
25	DW_80_5	8.09	504	Umc2052	umc1384	488-504	2.46*	7.3	10.686
<i>cumulated</i>								<b>36.8</b>	
26	%DW_30_1	4.04	136	phi096	gsy4	132-148	2.23	6.6	1.025
27	%DW_30_2	4.06	264	Bnlg1621	umc1329	236-292	2.01	6	-1.134
28	%DW_30_3	5.01	12	Bnlg1006	mmc0151	0-32	2.47*	7.3	-1.175
<i>cumulated</i>								<b>19.9</b>	
29	%DW_60_1	1.01	32	Umc1354	bnlg1014	0-64	2	6	1.318
30	%DW_60_2	1.06	440	Umc1590	umc1335	420-484	2.13	6.4	1.197
31	%DW_60_3	3.02	52	Umc1892	bnlg1325	40-72	4.65****	13.5	1.750
32	%DW_60_4	5.02	128	Umc90	csu108	104-152	2.61*	7.8	1.573
33	%DW_60_5	6.05	308	Umc1020	bnlg1702	296-332	2.33	7	-0.944
<i>cumulated</i>								<b>40.7</b>	
34	%DW_80_1	4.05	180	Bnlg1217	umc1511	172-208	2.24	6.6	1.051
35	%DW_80_2	5.01	52	mmc0151	umc147a	32-72	2.62*	7.6	1.294
36	%DW_80_3	5.07	528	Umc1537	umc68	508-544	2.35	6.9	-0.991
37	%DW_80_4	8.01	48	Bnlg1194	bnl9_11	32-76	2.85*	8.3	-1.221
38	%DW_80_5	9.04	220	csu147	umc38c	196-244	3.04***	8.8	1.625
39	%DW_80_6	9.07	464	Umc1137	gsy330	440-488	4.38****	12.4	1.438
<i>cumulated</i>								<b>50.6</b>	
40	Water_30_1	4.03	116	Umc2039	phi096	76-144	2.25	6.7	-10.951



**Table 2: QTLs detected for kernel desiccation and ABA content at 30, 40, 60 and 80 DAP (Continued)**

78	Rate_40_60_1	3.03	64	Bnlgl325	bnlgl523	48-72	3.51***	10.5	0.008
----	--------------	------	----	----------	----------	-------	---------	------	-------

Position in pcM on the LHRF\_F<sub>3,4</sub>\_1201 map of the segregating population. ABA\_embryo: ABA in embryo (pg/DW); ABA\_endosperm: ABA in endosperm (pg/DW); DW: kernel dry matter weight (mg/kernel); FW: kernel fresh matter weight (mg/kernel); %DW = DW/FW × 100; Water: kernel water content (mg/kernel); Rate = (Water/FW × 100)/(thermal time interval); Slope: regression line slope of (Water/FW × 100) as a function of thermal time. The genomewide risk, *P*, (absence of any QTL in the genome) was computed by classical permutation for a given LOD score (see Methods) and reported after the LOD value: \* *P* < 0.25, \*\* *P* < 0.10, \*\*\* *P* < 0.05 and \*\*\*\* *P* < 0.01. The individual significance level (risk of absence of a QTL at a specified locus, not taking in account the multiple comparisons at other loci) is 0.001 for a 2.45 LOD score.

**Table 3 Colocation between QTLs and candidate genes related to kernel desiccation and ABA biosynthesis**

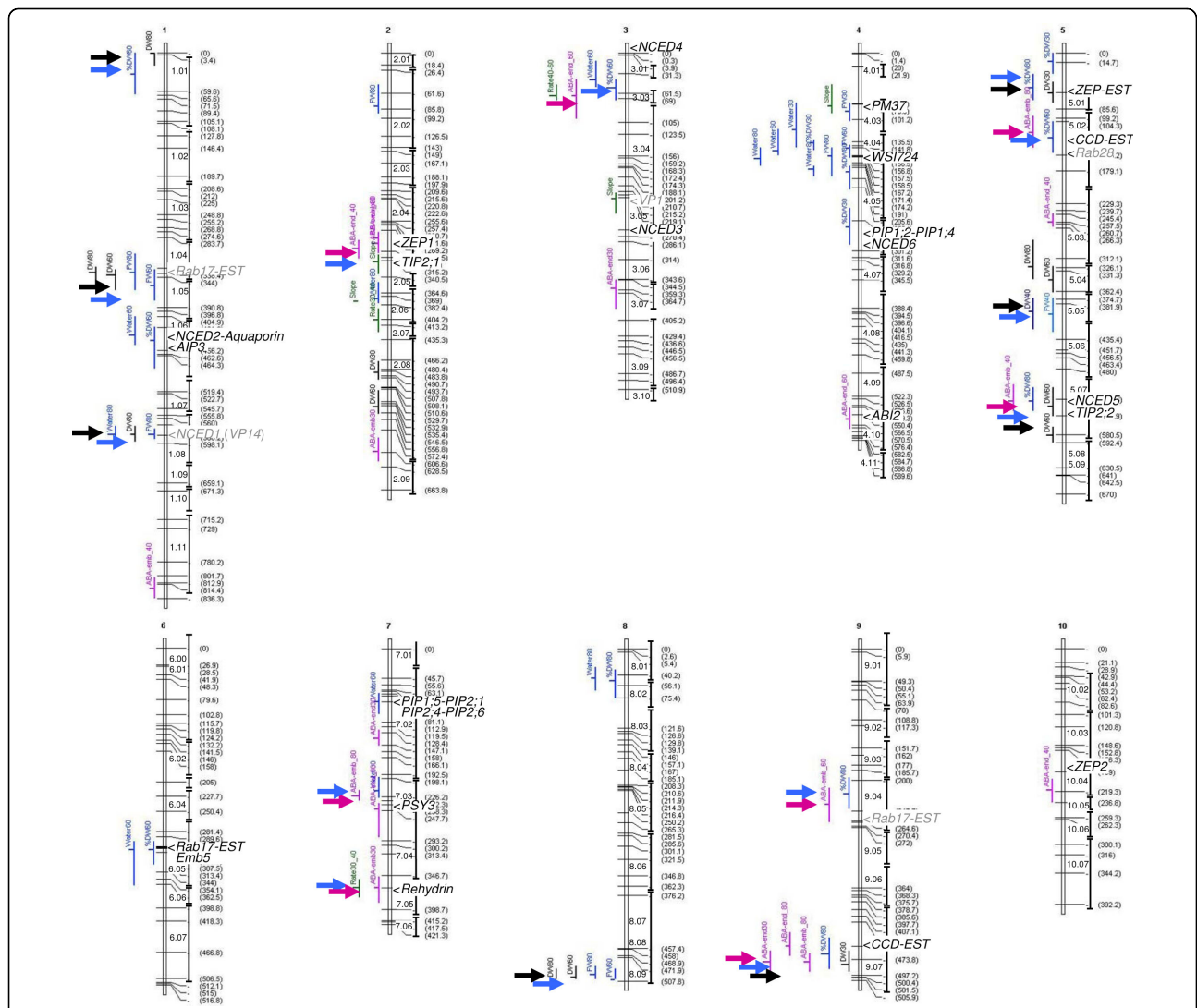
QTL name	Class	Bin	Position	From-to	Candidate gene
DW_80_2 ***	Growth	1.04		250-273	<i>Rab17-EST</i> (241)
FW_80_1 *	Desiccation		255	245-277	
Water_60_1 *	Desiccation		333	305-344	<i>NCED2</i> (310)
%DW_60_2	Desiccation	1.06	338	323-386	<i>Aquaporin</i> (314) <i>AIP3</i> (323)
DW_80_3	Desiccation		460	455-464	
FW_80_2 ***	Desiccation	1.08-1.09	460	457-463	<i>NCED1 = Vp14</i> (465)
Water_80_1 ***	Desiccation		460	455-463	
ABA_embryo_30_1	ABA content		222	215-232	<i>ZEPI</i> (227)
ABA_embryo_40_2 ***	ABA content	2.04	226	219-234	
ABA_endosperm_40_1 ***	ABA content		230	224-237	<i>TIP2;1</i> (236)
Slope_1 **	Desiccation rate		237	235-241	
Slope_3	Desiccation rate	3.05	220	212-244	<i>Vp1</i> (209)
Slope_4	Desiccation rate		82	66-97	<i>PM37</i> (82)
FW_30_1 *	Desiccation	4.03	92	79-109	
FW_80_4	Desiccation	4.04-4.05	130	124-142	<i>WSI724</i> (132)
Water_80_3 ***	Desiccation		132	124-132	
%DW_30_2	Desiccation	4.06	173	161-189	<i>PIP1;2 PIP1;4</i> (161)
ABA_endosperm_60_2	ABA content	4.10	340	329-346	<i>ABI2</i> (334)
%DW_80_2 *	Desiccation		44	33-58	<i>ZEP-EST</i> (41)
DW_30_2	Growth	5.01	48	40-61	
ABA_embryo_80_1	ABA content		96	77-115	<i>CCD-EST</i> (81)
%DW_60_4 *	Desiccation	5.02	102	83-122	<i>Rab28</i> (132)
%DW_80_3	Desiccation		312	304-322	
ABA_embryo_40_3 *	ABA content	5.07		302-322	<i>NCED5</i> (311)
DW_60_4	Growth		317	304-327	<i>TIP2;2</i> (320)
%DW_60_5	Desiccation	6.05	148	136-175	<i>Rab17-EST</i> (138)
Water_60_4 **	Desiccation			136-201	<i>Emb5</i> (138)
Water_60_5	Desiccation		167	163-173	<i>PIP1;5</i> (164)
ABA_endosperm_30_2 **	ABA content	7.02	183	179-189	<i>PIP2;1</i> (164) <i>PIP2;4</i> (164) <i>PIP2;6</i> (164)
ABA_endosperm_80_1	ABA content	7.03	244	236-288	<i>PSY3</i> (241)
ABA_embryo_30_3	ABA content	7.05	383	365-402	<i>Rehydrin</i> (380)
Rate_30_40_2 ***	Desiccation rate			370-396	
%DW_80_5 **	Desiccation	9.04	125	108-140	<i>Rab17-EST</i> (150)
ABA_embryo_60_1	ABA content		135	120-149	
ABA_endosperm_80_2	ABA content	9.07	279	262-294	<i>CCD-EST</i> (271)
%DW_80_6 ****	Desiccation		291	269-314	

To compare with gene positions, QTLs mapped on the LHRF\_F<sub>3,4</sub> segregating population (Additional file 1) were projected on the REFMAP050110 map [32] using Biomercator [34]. Distances on REFMAP050110 are expressed in pcM and are shown in parentheses for gene candidates. Genes in bold were mapped by PCR in this study (Tables 4 and Additional file 1). Genes involved in kernel drying and located in the vicinity of the QTL confidence interval, are indicated in grey. ABI2: ABA-insensitive protein phosphatase 2C 2; AIP3: ABI3-interacting protein 3; Emb5: Embryogenic-ABA-inducible LEA 5; PM37: seed maturation protein; PSY3: phytoene synthase 3; Vp1: Viviparous1; WSI724: dehydrin. The asterisks behind each QTL represent the genomewide risk *P* (for definition see legend of Table 2) with \* *P* < 0.25, \*\* *P* < 0.10, \*\*\* *P* < 0.05 and \*\*\*\* *P* < 0.01.

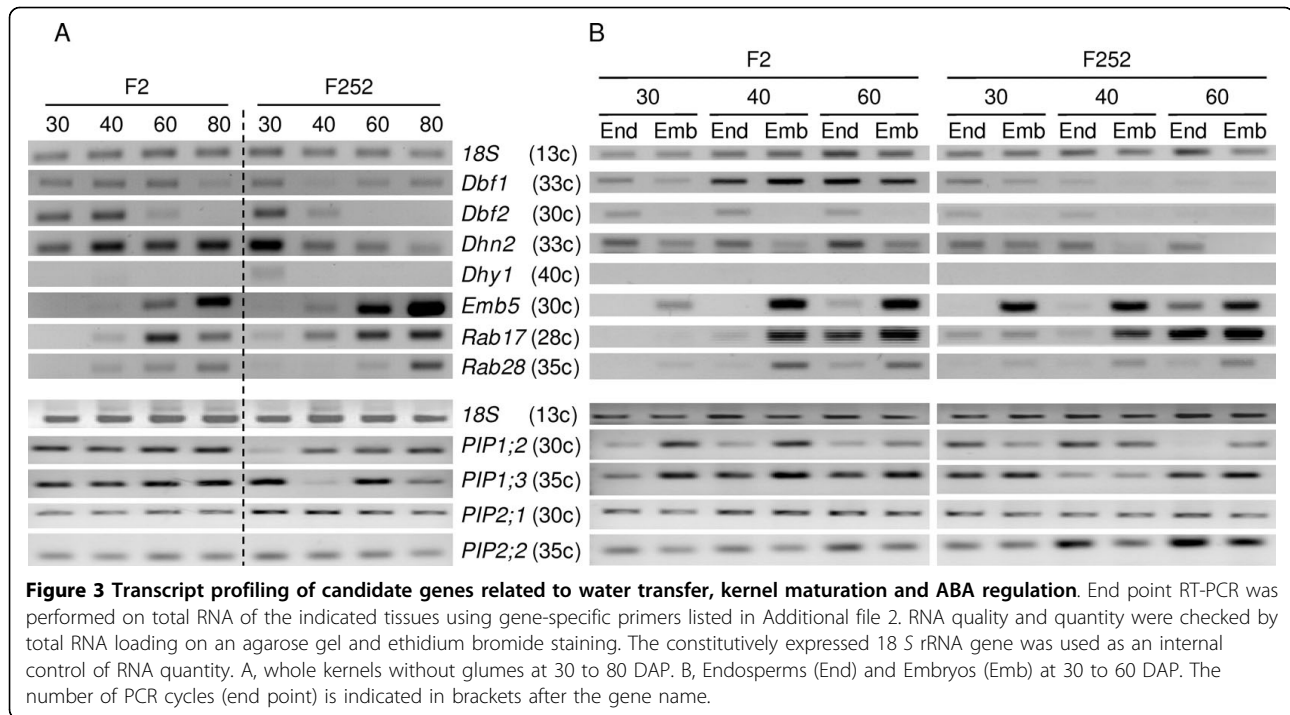


showed that among the six %DW80 QTLs, a variable which mirrored moisture content, two collocated with FW80 and/or Water80 (bins 4.04 and 8.02) and three with QTLs for ABA content (bins 5.07, 9.04 and 9.07). The allele effect was consistent with the better drying performance of the F252 parent since four of the six % DW80 QTLs with a cumulated  $R^2$  of 35.4% (total  $R^2$ : 50.6%) presented a positive allele effect originating from F252.

QTLs for ABA content in embryo and endosperm were rarely collocated, which was somewhat predictable due to the poor correlation between ABA contents in the two tissues. By contrast, numerous collocations were observed between desiccation traits and ABA traits at eight loci (bins 2.04, 3.03, 5.02, 5.07, 7.03, 7.05, 9.04 and 9.07), illustrated by overlaps of their QTL confidence intervals (compare blue and pink arrows in Fig. 2).



**Figure 2 QTLs and candidate gene map.** The LHRF\_ F<sub>3,4</sub> segregating population was used. Bins are shown on the right of chromosomes. Distances are in pcM in the LHRF\_ F<sub>3,4</sub> map (LHREF3\_1201). QTLs for desiccation (blue), desiccation rate (green), growth (black) and ABA content (pink) are on the left of the chromosomes (Tables 2 and 3). The confidence intervals of the QTLs are indicated by vertical bars. Arrows highlight collocations between desiccation QTLs (blue) and growth QTLs (black) and between desiccation QTLs and ABA QTLs (pink). Gene locations in the QTL confidence interval are indicated close to their corresponding QTLs. Gene codes are detailed in Tables 3 and 4. Genes in bold were mapped by PCR in this study (Additional file 2). The others were mapped by Génoplante <http://urgi.versailles.inra.fr/GnpMap/>. Genes involved in kernel drying and located in the vicinity of the QTL confidence interval, are indicated in grey. ABA-emb: ABA in embryo (pg/DW); ABA-end: ABA in endosperm (pg/DW); DW: kernel dry matter weight (mg/kernel); FW: kernel fresh matter weight (mg/kernel); %DW = DW/FW x 100; Water: kernel water content (mg/kernel); Rate = (Water/FW x 100)/(thermal time interval); Slope: regression line slope of (Water/FW x 100) as a function of thermal time.



### QTL and gene colocation

When clusters are composed of traits of different classes (desiccation, desiccation rate, growth, ABA content), QTL colocations raise the classical question of the existence of common genes which may control genetic variability of two or more classes at a single locus. A way to find candidate genes is to examine the list of reported genes which have been mapped in the QTL region. For this purpose, the QTLs mapped on the LHRF\_F<sub>3;4</sub> population (LHREF3\_1201 map) were projected with Biomecator [34] on a reference map (REFMAP050110), which is based on the internationally used IBM population. The list of known or putative cDNAs provided by the data base in each QTL confidence interval was manually scanned to select functions related to water transfer (e.g. aquaporin), kernel maturation (e.g. LEA proteins) and ABA metabolism or regulation (e.g. ABA biosynthesis, ABA-responsive proteins and related transcription factors).

Relevant colocations (i.e. genes in the QTL confidence interval and functionally related to the trait) were observed on chromosomes 1, 2, 3, 4, 5, 6, 7 and 9 (Table 3). It has to be noted that most colocations of interest (Table 3) involved at least one QTL with a LOD score higher than 2.45, corresponding to an individual significance level below 0.001 and a genomewide level below 25%, including eight QTLs with a risk below 5% (see Methods). Desiccation trait clusters collocated with aquaporin ESTs (bins 1.06 and 4.06), maturation proteins (bins 1.04 and 4.03) and/or ABA-related genes

(bins 1.06, 1.08-1.09, 3.05, 4.04-4.05, 5.01 and 6.05). Surprisingly, ABA biosynthetic *NCED1* (*Vp14*; bin 1.08-1.09) and *NCED2* (bin 1.06) genes did not collocate with QTLs for ABA content but rather with QTLs for desiccation. Nevertheless, colocations were identified for the ABA biosynthetic *ZEP1* (bin 2.04) and *NCED5* (bin 5.07) genes and QTLs for both desiccation and ABA content. Colocation of two *carotenoid cleavage dioxygenase* (*CCD*) ESTs was noted with clusters comprising desiccation and ABA traits (bins 5.02 and 9.07). However, although very close to *NCED* genes, *CCD* genes did not appear to be involved in ABA biosynthesis, the encoded enzymes being able to cleave carotenoids at 9,10 (9',10') bonds to generate multiple apocarotenoid products [35], whereas *NCEDs* cleaved carotenoids asymmetrically at positions 11-12 [36]. Another colocation involved the ABA\_endosperm\_80\_1 QTL (bin 7.03) and the maize *phytoene synthase 3* (*PSY3*) gene whose expression influences abiotic stress-induced root carotenogenesis [37]. Three other colocations involved QTLs for both desiccation and ABA content with genes encoding aquaporins (*TIP2;1* on bin 2.04, *TIP2;2* on bin 5.07 and *PIP1;5*, *PIP2;1*, *PIP2;4* and *PIP2;6* on bin 7.02) and LEA proteins (*Rab28* on bin 5.02, *Rehydrin* on bin 7.05 and *Rab17* on bin 9.04).

### Transcript expression of candidate genes related to water status, kernel maturation and ABA regulation during kernel desiccation

As a first step to validate candidate genes, RT-PCR analysis was performed during kernel desiccation in both

**Table 4 Maize *NCED* and *ZEP* gene mapping and colocation with QTLs for desiccation and/or ABA content**

Gene Name	EST	IGR ID <sup>a</sup> HTGS ID <sup>b</sup> yrGATE ID <sup>c</sup>	Map	Bin	MM coord. <sup>d</sup>	Proj. coord. <sup>e</sup>	Flanking markers	QTL colocation
<i>NCED1</i>	[GenBank:ZMU95953] E746	<i>Vp14</i> [GenBank: AC230016.2]	LHFR_Gnp2004	1.08- 1.09	504	466	bnlg1643- phi055	DW_80_3 FW_80_2 Water_80_1
<i>NCED2</i>	E746	AZM4_115740 [GenBank: AC217286.3]	LHFR_Gnp2004	1.05- 1.06	328	306	umc 1906- umc 67	Water_60_1 %DW_60_2
<i>NCED3</i>	No	AZM4_115695 [GenBank: AC199036.2]	LHFR_Gnp2004	3.05- 3.06	193	207	umc1501- dupssr23	No
<i>NCED4</i>	No	AZM4_114127 [GenBank: AC212820.3]	LHFR_Gnp2004	3.00	-108.7	-125	Before umc1746	No
<i>NCED5</i>	[GenBank:QCD5h12]	AZM4_50254 [GenBank: AC194862.3]	LHFR_Gnp2004	5.06- 5.07	364	298	phi087- umc1537	%DW_80_3 ABA_embryo_40_3 DW_60_4
<i>NCED6</i>	No	AZM4_50252 [GenBank: AC190588.3]	REFMAP050110	4.06	177	177	bnlg 1621- umc 66a	%DW_30_2
<i>ZEP1</i>	QAG5c10 [GenBank:AI977858]	AZM5_13314 [GenBank: AC194845.3] yrGATE_Zm2 gZEP1	REFMAP050110	2.04	221	221	bnlg1018- bnlg166	ABA_embryo_30_1 ABA_embryo_40_2 ABA_endosperm_40_1 Slope_1
<i>ZEP2</i>	QAG5c10	AZM5_13312 [GenBank: AC206194.3] yrGATE_Zm10 gZEP2	REFMAP050110	10.04- 10.05	94.5	94.5	bnlg1526- umc259	No

<sup>a</sup>The Institute for Genomic Research.

<sup>b</sup>HTGS: the High Throughput Genomic Sequences.

<sup>c</sup>yrGATE: Your Gene Structure Annotation Tool for Eukaryotes.

<sup>d</sup>Map coordinate computed with MapMaker using 'RI self Haldane' options.

<sup>e</sup>Map coordinate on REFMAP050110 obtained by homothetic projection with BioMercator [30].

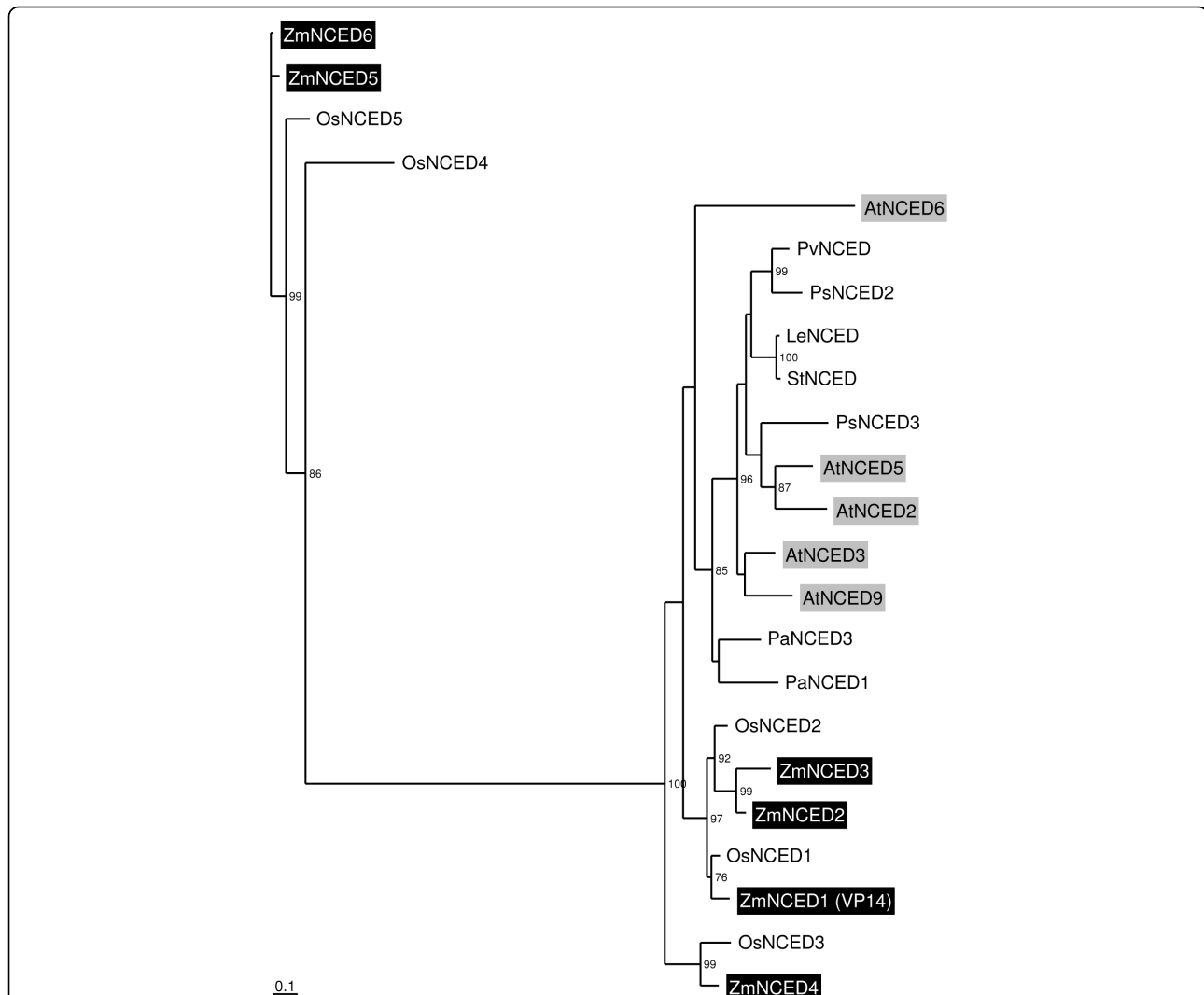
Mapping of candidate genes was performed on the LHFR\_Gnp2004 population derived from F2xF252 crossing, except for maize *NCED6*, *ZEP1* and *ZEP2* genes which were mapped on the IBM population (REFMAP050110 map) because of no polymorphism between F2 and F252 lines. QTL codes are detailed in Table 3.

parental inbred lines for genes encoding LEA proteins, ABA-responsive transcription factors and aquaporins. In addition to parental differences, responses to desiccation might be classified into three categories: up-regulation, down-regulation and up-and-down-regulation (Fig. 3). Expression of LEA *Emb5* and *Rab17* (*Dhn1*) genes and dehydrin *Rab28* gene increased during desiccation, especially in F252 at 60 and 80 DAP (Fig. 3A). Expression of *Dbf1* and *Dbf2* genes encoding transcription factors regulating the LEA *Rab17* gene [38] diverged. *Dbf2* expression clearly decreased over the time, whereas *Dbf1* expression was still substantial at 80 DAP (Fig. 3A). The dehydrin *Dhy1* gene was expressed at very low level. Nevertheless, its transcript levels clearly decreased in F252 genotype only, a pattern clearly established for the dehydrin *Dhn2* gene (Fig. 3A). Furthermore, *Emb5*, *Rab17* and *Rab28* genes were more strongly expressed in embryo, whereas *Dhn2* and *Dbf2* genes were mostly expressed in endosperm (Fig. 3B). This pattern was conserved in both genotypes, although the magnitude of the expression was frequently different as previously noted in Fig. 3A.

Among the 12 analyzed *PIP* genes (Table 1), only *PIP1;2*, *PIP1;3*, *PIP2;1* and *PIP2;2* were repeatedly expressed in kernels (Fig. 3). *PIP1;2* and *PIP1;3* transcripts levels generally increased at later stages (Fig. 3A), but the time course was gene and genotype-dependent. Both genes showed stronger expression in embryo than in endosperm in F2 genotype, while no clear preference was seen in F252 genotype (Fig. 3B). In contrast, *PIP2;1* and *PIP2;2* expression was more or less stable during grain desiccation (Fig. 3A), with a preferential expression of *PIP2;2* in older endosperm (Fig. 3B). The three TIPs tested (Additional file 2) had expression maximum in leaf and root tissue and only *TIP1;1* was weakly expressed in kernels (data not shown).

#### Molecular analysis of maize *NCED* and *ZEP* genes

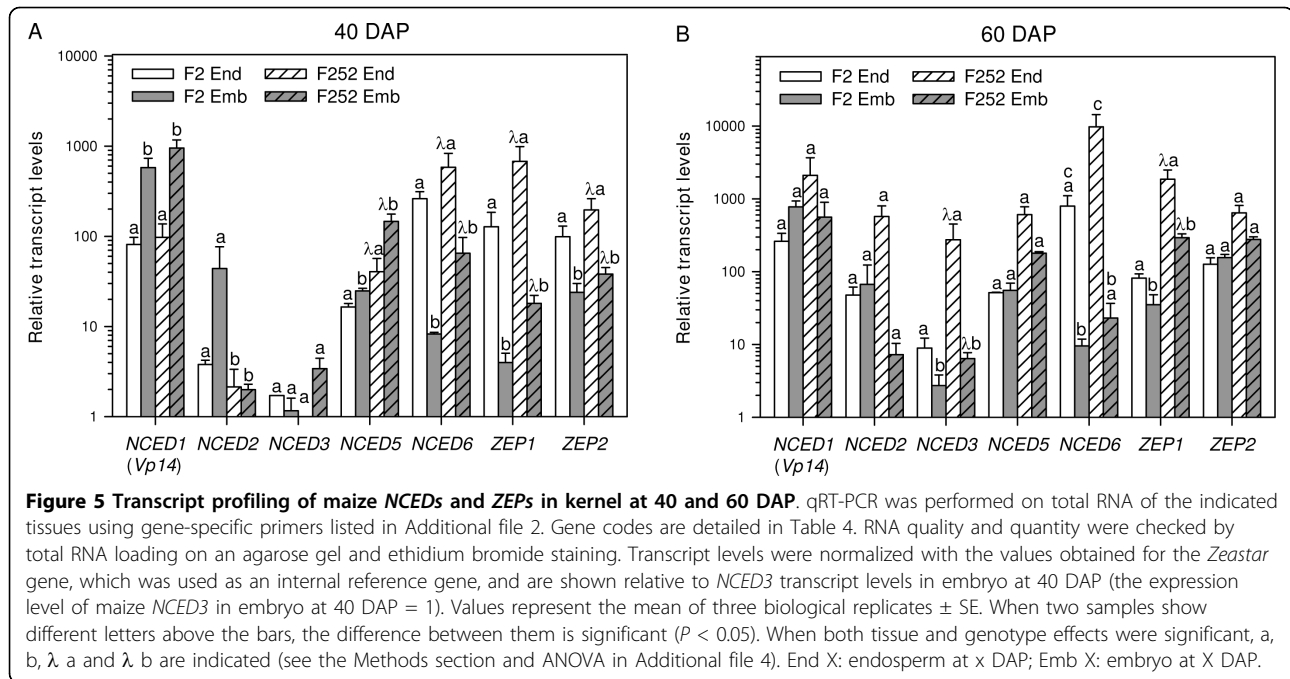
One of the problems in analyzing accurately *NCED* and *ZEP* expression was the design of specific primers mainly because of insufficient knowledge of the actual number and sequence of *NCED* and *ZEP* genes in maize. Therefore, we first analyzed the public databases for the presence of putative *NCED* and *ZEP* coding sequences in maize. Nine different *NCED* loci had



**Figure 4 Phylogenetic tree of NCED proteins.** The five novel maize NCED protein sequences were aligned with known NCED protein sequences [39,44,45]. In addition, four other NCED protein sequences were included in the alignment: OsNCED4 [GenBank:AAW21318.1] and OsNCED5 [GenBank:AAW21317.1] from rice and PsNCED2 [GenBank:BAC10550] and PsNCED3 [GenBank:BAC10551] from *Pisum sativum*. An unrooted tree of the NCED protein sequences was obtained using the maximum Likelihood method. Branch lengths are scaled proportional to substitution rate. Bootstrap values (percent) of 500 bootstrap replicates exceeding 70% are indicated above the supported branches. The Arabidopsis and maize NCEDs are highlighted by grey and dark boxes, respectively. Maize accession numbers are indicated in Table 4. Other NCEDs include: AtNCED2 [GenBank:NP\_193569], AtNCED3 [GenBank:NP\_188062], AtNCED5 [GenBank:NP\_174302], AtNCED6 [GenBank:NP\_189064] and AtNCED9 [GenBank:NP\_177960] from Arabidopsis; LeNCED [GenBank:CAB10168] from *Lycopersicon esculentum*; OsNCED1 [GenBank:AAW21319.1], OsNCED2 [GenBank:AAW21321.1] and OsNCED3 [GenBank:AAW21320.1] from rice; PaNCED1 [GenBank:AF224672] and PaNCED3 [GenBank:AF224671] from *Phaseolus vulgaris*; StNCED [GenBank:CAB76920] from *Solanum tuberosum*.

previously been identified with *NCED*-like EST sequences in the Génoplante programme, some ESTs mapping to several loci (e.g. E746, Table 4). To identify a maximum of *NCED* genes in maize, the search was extended to maize genomic contigs (The Institute for Genomic Research) and high throughput genomic sequences, which were matched with the EST sequences and used to design specific primers (Additional file 2) for DNA and cDNA amplification in lines F2 and F252. Focus on longest ESTs in the mixed cDNA/genomic

DNA contigs allowed us to discard contigs with introns, which belonged to the *CCD* gene family [23] and to determine six maize *NCED* genes that could potentially encode NCED proteins (Table 4). Phylogenetic analysis performed with the deduced NCED amino acid sequences showed that the six maize proteins fell into four monocot clusters that were distinguishable from the eudicot cluster (Fig. 4). In contrast, close homologs for each of the maize *NCED* genes were found in rice. The two putative maize NCED5 and NCED6 proteins



fell into a highly divergent cluster, including two maize and two rice *NCED* proteins. The orthologs of maize *NCED1* (*VP14*) and *NCED4* proteins were rice *NCED1* and *NCED3* proteins, respectively. Additionally, maize *NCED2* and *NCED3* proteins appeared as a pair of paralogs that was related to rice *NCED2*, whereas maize *NCED5* and *NCED6* proteins appeared as a pair of paralogs that was related to rice *NCED5*. *NCED1* (*Vp14*), *NCED2*, *NCED3*, *NCED4* and *NCED5* genes were mapped on the LHRF mapping panel at five different loci (bins 1.08-1.09, 1.05-1.06, 3.05-3.06, 3.00 and 5.06-5.07) (Fig. 2, Table 4). *NCED6* gene was mapped on the IBM mapping panel because of no polymorphism between F2 and F252 (bin 4.06; Fig. 2, Table 4). Only *NCED1*, *NCED2* and *NCED5* corresponded to previously mapped *NCED*-annotated ESTs, whereas no EST was detected for maize *NCED3*, *NCED4* and *NCED6* gene loci. Consistently, none of the identified *NCED* genes mapped at the *CCD* loci (bins 5.02 and 9.07; Table 3).

Four *ZEP*-annotated ESTs had previously been mapped at bins 2.02-2.03, 4.09-4.11 and 5.01 (Géno-plante data base, <http://urgi.versailles.inra.fr/GnpMap>, data not shown) in maize. Nevertheless, only two maize *ZEP* genes, *ZEP1* (*yrGATE\_Zm2 gZEP1*) and *ZEP2* (*yrGATE\_Zm10 gZEP2*), were ascertained from our data mining and PCR fragment joining (Additional file 3), in agreement with a recent report [37], and mapped on IBM mapping panel at bins 2.04 and 10.04-10.05, respectively (Fig. 2, Table 4).

#### Transcript expression of maize *NCED* and *ZEP* genes during kernel desiccation

Colocations were identified for *NCED5* and *ZEP1* loci and QTLs for both desiccation and ABA traits, suggesting that these ABA biosynthetic genes are potential candidates (Fig. 2, Tables 3 and 4). Interestingly, *NCED1* (*Vp14*), *NCED2* and *NCED6* genes mapped at loci colocalizing with QTLs for desiccation only (Fig. 2, Tables 3 and 4). Therefore, transcript levels of maize *NCED* and *ZEP* genes were quantified by quantitative RT-PCR (qRT-PCR) in embryo and endosperm from both parental lines at 40 and 60 DAP, ABA content being at the highest level in embryo at these stages (Fig. 1D). *NCED4* expression was not analyzed because it was not possible to design *NCED4*-specific primers of sufficient quality for qRT-PCR.

*NCED1*, *NCED2*, *NCED3*, *NCED5* and *NCED6* transcripts were detected in all tissues analyzed, *NCED2* and *NCED3* being frequently expressed at a level close to the detection limit in most of the samples (Fig. 5). At 40 DAP, differential expression between embryo and endosperm was significant for the two mainly expressed *NCED* genes, *NCED1* and *NCED6* (Fig. 5A, Additional file 4) but the response was opposite since, in both genotypes, *NCED1* transcript was present at the highest level in the embryo, while *NCED6* was the most highly expressed transcript in the endosperm. *NCED5* had a similar pattern to that of *NCED1*, although the magnitude of the embryo-endosperm difference in transcript level was lower than for *NCED1* (Fig. 5A, Additional file

4). There was also a genotype effect since differential expression between F2 and F252 genotypes was significant for *NCED2*, *NCED5* and *NCED6* genes (Fig. 5A, Additional file 4). At 60 DAP, *NCED1*, *NCED2*, *NCED3* and *NCED6* transcript levels were higher in both genotypes than at 40 DAP (Fig. 5B). The difference between embryo and endosperm was no longer significant for *NCED1* and *NCED5* genes (Fig. 5B, Additional file 4). In contrast, expression was significantly higher in endosperm than in embryo for *NCED3* and *NCED6* genes (Fig. 5B, Additional file 3). Additionally, the genotype effect was significant for *NCED3*, its transcript level being lower in F2 genotype than in F252 genotype in both tissues (Fig. 5B, Additional file 4).

Regarding the *ZEP* genes, at 40 DAP, they were both mainly expressed in the endosperm rather than in the embryo (Fig. 5A, Additional file 4). There was also a genotype effect since expression of both *ZEP* genes was higher in the F252 genotype than in the F2 genotype in both tissues. At 60 DAP, *ZEP1* preferential expression in the endosperm and higher expression in F252 were significant as observed at 40 DAP but magnitude of the effects was lower (Fig. 5B, Additional file 4). In contrast, there was no significant difference for *ZEP2* expression between embryo and endosperm whatever the genotypes (Fig. 5B, Additional file 4).

## Discussion

### Relationship between kernel moisture and kernel weight

The observed genetic variability in the desiccation rate of maize inbred lines F2 and F252 and their offspring means that this trait is amenable for breeding purposes. The use of intermated recombinant inbred lines allows the reduction of QTL confidence interval to only a few pseudo cM (pcM, see Methods for equivalence between pcM and cM). On the basis of a map of roughly 6000 pcM and a genome size of 30 000 to 40 000 genes per maize genome, an estimate of 5 to 7 genes per pcM is reached. With the large progress in maize sequencing <http://www.maizesequence.org>, in maize genetic mapping, and syntenic relationships between maize and rice, a reasonable discrimination among candidate genes is possible in a given QTL confidence interval of a few pcM. In the case of physiological traits, candidate genes may be *a priori* selected for validation due to the knowledge of the process under study [30]. Although such a correlative approach does not establish direct causal relationships, it must be emphasized that the chance to find a random collocation between a QTL for the rather simple traits considered in this manuscript and a gene functionally related to the trait should be low. Nevertheless, the situation is different for more complex traits involving multiple developmental or metabolic pathways

and consequently a much larger number of candidate genes. Using this approach, several candidates [39] were identified and then validated through molecular and/or genetic studies [40-42]. In the present case, obvious candidates for desiccation QTLs were LEA proteins and aquaporins, whereas obvious candidates for ABA QTLs were key enzymes of the ABA biosynthetic pathway, such as *NCED* and *ZEP*, ABA responsive genes and corresponding transcription factors.

Kernel moisture and kernel weight are non independent variables. It was recently shown that maximum kernel weight (MKW) may be predicted from maximum water content (MWC) occurring at 40 to 60 DAP in maize hybrids grown at three densities [1]. It was interpreted as meaning that maximum water content is an essential determinant of kernel volume, e.g. of sink capacity, which depends on early developmental events taking place during the lag phase. This MKW/MWC relationship was verified in the present parental lines and the LHRF\_F<sub>3:4</sub> population (data not shown), explaining the high correlations observed between FW and Water content and to a lesser extent FW and DW at each DAP stage. However, it has no influence on kernel drying as shown by the absence of correlation in the LHRF\_F<sub>3:4</sub> population between moisture content at harvest (%DW80) and final DW. The large observed variability in %DW, allowed detection of six QTLs explaining *ca.* 50% of the variance with most fast desiccation alleles contributed by parent 252. This was consistent with the higher desiccation rate of F252 parental line. Interrogation of Maize GDB <http://www.maizegdb.org> showed that the %DW60\_4 and %DW80\_5 QTLs mapped very close to earlier reported QTLs for kernel moisture on bins 5.02 and 9.04, respectively [43]. Comparison with the more recently reported grain moisture and drying rate QTLs [31] was difficult to assess because the map used was not anchored to any reference map and the QTL confidence intervals were very large. However, striking similarities were observed at bins 1.04, 1.08 and 5.01 when positions were estimated proportionally to the total length of each chromosome.

### Candidate gene mapping and collocations with QTLs

To the best of our knowledge, before the present study, QTLs for ABA content and water status have only been searched in leaves of plants submitted to water stress [39,44,45]. These QTLs were compared to those for ABA content and water related traits in kernel by projecting all of the QTLs on the reference map. The major *root-ABA1* QTL identified near the RFLP marker *csu133* on chromosome 2 (bin 2.04) [39,44,45] and affecting root architecture and grain yield in maize [46], did not collocate with the grain ABA QTLs detected in the same chromosome region (Fig. 2, Table 2). Nevertheless,

**Table 5 Common QTL locations for traits related to desiccation and ABA content in leaf and kernel**

Bin	Locus or marker	QTL position (pcM) on REFMAP050110	Candidate gene	QTLs from kernel	QTLs from leaf		
					Lebreton et al [39,44,45].	Tuberosa et al [39,44,45].	Pelleschi et al [39,44,45].
1.06	bnl5.59	305-344	<i>NCED2</i> <i>Aquaporin</i>	Water_60_1	ABAL		RWC, ABAX
1.08-1.09	umc39c	455-463	<i>NCED1</i>	DW_80_3, FW_80_2 Water_80_1	ΨL	ABAL	RWC, ABAX
3.05	umc10 gsy406	212-244	<i>Vp1</i>	Slope_3	ABAX		RWC ΨL
3.06	Gsy224	210-220	<i>NCED3</i>				ABAL
4.03-4.04	Gsy431	89 (66-97)	<i>PM37</i>	Slope_4			RWC
4.05-4.06	umc66	161-189	<i>PIP1;2, PIP1;4</i> <i>NCED6</i>	%DW_30	ABAL	ABAL	ABAL, ΨL ABAX
7.01-7.02	gsy113 umc116	163-173 179-189	<i>PIP1;5, PIP2;1,</i> <i>PIP2;4, PIP2;6</i>	Water_60_5 ABA_endosperm_30_2	stomatal conductance	ABAL	ABAL ABAX
7.03	umc110	234-244 253-312	<i>PSY3</i>	ABA_embryo_80_2 ABA_endosperm_80_1	stomatal conductance	ABAL	RWC
9.04	umc114	108-149	<i>Rab17-EST</i>	%DW_80_5 ABA_embryo_60_1	ABAL		ΨL
9.07	gsy330 umc94	262-314	<i>CCD-EST</i>	%DW80_6 ABA_endosperm_80_2			ABAX

For kernel, the LHRF\_F<sub>3,4</sub> segregating population derived from F<sub>2</sub> × F<sub>252</sub> crossing. For leaf, populations derived from Polj17 × F<sub>2</sub> crossing [394445], high ABA × low ABA lines [394445] and F<sub>2</sub> × MBS847 RILs [394445]. Plants were submitted to water deprivation at comparable young stage in [394445] and [394445]. In [394445], ABA concentration was determined in leaf samples from stressed plants irrigated with 50% normal rainfall during stem elongation and before anthesis. Gene codes are detailed in Tables 3 and 4. Genes in bold were mapped by PCR in this study (Table 4 and Additional file 2). Genes involved in kernel drying and located in the vicinity of the QTL confidence interval, are indicated in grey. ABA\_embryo: [ABA] in embryo; ABA\_endosperm: [ABA] in endosperm; ABAL: leaf [ABA] ABAX: xylem [ABA]; DW: kernel dry matter weight (mg/kernel); FW: kernel fresh matter weight (mg/kernel); %DW = DW/FW × 100; Slope: regression line slope of (Water/FW × 100) as a function of thermal time; RWC: relative water content; Water: kernel water content (mg/kernel); ΨL: leaf water potential.

striking similarities were observed in ten other regions (Table 5). Interestingly, among these ten regions, four (bins 1.08-1.09, 4.06, 7.01-7.02 and 7.03) were detected in all studies (one for kernel and three for leaves), the common candidates being *PIP* (in three regions out of four) and *NCED* (in two regions out of four) genes. The coincidence of QTL locations between a source (leaf) and a sink organ (kernel) strongly suggests that common genetic factors (genes) may control the trait at each locus. In this respect, our finding that *PIPs* are the most frequent candidate genes at the common loci is important as the data indicate that they are good candidates for the regulation of kernel moisture. Of particular interest is the observation that *PIP1;2* and *PIP2;1* are highly expressed isoforms, which are associated with water relations in both leaf [47] and kernel. By contrast, *NCED* candidates only appeared in relation to QTLs for leaf ABA content and not to QTLs for kernel ABA content.

In addition to the above mentioned *PIP* and *NCED* genes, our kernel specific QTLs revealed other candidates such as the *ABI3-interacting protein 3* (*AIP3*) on chromosome 1; *ZEP1* and *TIP2;1* on chromosome 2; *Viviparous1* (*Vp1*) on chromosome 3; the *seed maturation protein 37* (*PM37*), the dehydrin *WSI724* and the *ABA-insensitive protein phosphatase 2C 2* (*ABI2*) on

chromosome 4; *Rab28*, *NCED5* and *TIP2;2* on chromosome 5; *Rehydrin* and *PSY3* on chromosome 7; *Rab17* and *CCD-EST* on chromosome 9 (Tables 3 and 4). The transcript patterns of *LEA*, *aquaporin*, *NCED* and *ZEP* genes during kernel maturation and desiccation provided further help for the identification of potential candidates. Regarding the desiccation traits, the consistency of *Rab17*, *Rab28* and *Emb5* genes as candidates was supported by the fact that their transcript levels increased from 40 to 80 DAP and was stronger in F<sub>252</sub> than in F<sub>2</sub>, which was consistent with the positive effect of F<sub>252</sub> allele on desiccation. It is also notable that ABA-responsive complex 3 found in *Rab28* promoter sequence [48] is transactivated by VP1, which colocalizes with Slope\_3 QTL. The situation was not as clear for the aquaporins, although some *PIP1* and *PIP2* were very frequently associated with QTLs for desiccation and ABA content and *PIP1* gene expression increased by the end of the maturation phase. We suspect that specificity problem linked to the high sequence similarity between aquaporins led to cross-hybridization during RFLP-based mapping and/or non-specific amplifications in RT-PCR experiments, because primers were designed from the known 3' UTR B73 sequences, which likely differ from that of F<sub>2</sub> and F<sub>252</sub>. We are confident that the emerging high throughput sequencing applied to

individual lines like F2 and F252 will give us the necessary tools to obtain clear-cut answers in the future.

#### **Mapping, colocation and expression studies of the maize *NCED* and *ZEP* gene families**

Effort was made to determine the number of *NCED* and *ZEP* genes in the maize genome, to map the genes on the LHRF mapping panel and to determine their expression patterns in the kernel of both maize parental inbred lines. The available data at the beginning of this work were (1) the *NCED1* (*Vp14*) genomic sequence from B73, (2) the map positions of nine loci identified with a *NCED1* RFLP probe or similarly annotated probes and (3) the mapping of four partial *ZEP* cDNAs at six loci. *In silico* analysis of genomic sequences coupled with PCR verification using genomic DNA and cDNA to confirm the contigs' physical existence and expression of sequence contigs, led to the identification of six *NCED* and two *ZEP* genes in maize. These figures remain minimum estimates based on the available data, but they are consistent with other species, since five *NCED* genes and a single copy *ZEP* gene have been found in *Arabidopsis* and rice. The tree topology of *NCED1*-related proteins suggests that at least four maize duplications might have occurred, one of which followed by high divergence in the branch supporting more recent duplicate *NCED5* and *NCED6* proteins. Additionally, blast analysis of rice high throughput genomic sequences revealed that maize *ZEP1* and *ZEP2* genes were closer to each other than to the rice *ZEP* gene, suggesting a maize specific duplication (data not shown). Mapping of maize *NCED* genes led to the identification of six different loci, only three being common to the nine previously reported loci. This highlights the caution needed in the interpretation of colocations between a QTL and a locus when the candidate gene was mapped using RFLP probes. Similar situations were found with *ZEP* genes since among the six loci identified with four partial *ZEP*-like cDNAs, only two were confirmed in agreement with a recent report [37].

The comparison of the transcript expression profiles of five of the six identified maize *NCED* genes shed new light on the relative autonomy of the embryo and the endosperm compartments. Accordingly, the expression patterns were independent in the endosperm and the embryo whatever the gene tested, suggesting different regulatory mechanisms. This was consistent with the absence of any correlation between ABA content or its time course in endosperm and embryo. The most strongly expressed maize *NCED* transcripts were *NCED1* (*Vp14*) and *NCED6* in embryo and endosperm, respectively. The higher expression of *NCED1* in the embryo may explain why the *vp14* mutation produced both a viviparous phenotype (vivipary being a property of the embryo) and a kernel ABA deficiency (the

embryo being far much richer in ABA than the endosperm). However, the *NCED1* locus did not collocate with embryo ABA QTLs, although it collocated with leaf ABA QTLs. A way to interpret this inconsistency in the kernel would be to assume that *NCED1* is not limiting or provides a coarse control. The plausible existence of such a coarse control is consistent with the coincidence between maximum ABA accumulation and maximum *NCED1* transcript level in the embryo at 60 DAP. An additional finer control might explain the genotype differences. It is noticeable that only *NCED5* collocated with a QTL for ABA (*ABAembryo40\_3*), this collocation being consistent with the fact that *NCED5* transcript level was higher in the embryo than in the endosperm. Nevertheless, the low transcript level and the moderate  $R^2$  associated with the QTL value might undermine the significance of this finding. Similarly to *NCED5*, and *NCED6*, *ZEP* transcript levels were significantly different in both genotypes and both tissues at 40 DAP. The expression of both maize *ZEP* genes was prevalent in the endosperm and higher in F252 line than in F2 line, *ZEP1* being more strongly expressed than *ZEP2*. Interestingly, *ZEP1* collocated with two QTLs for ABA content in the embryo at 30 and 40 DAP and one QTL for ABA content in endosperm at 40 DAP.

#### **Effect of ABA content on the genetic control of kernel desiccation**

The potential role of ABA content on the genetic control of kernel desiccation was supported by numerous colocations between QTLs for both traits. The colocation of QTL for desiccation with *ABA-responsive* genes and *LEA* genes also known to be controlled by ABA provided further support, although it did not establish a triggering role of ABA on drying. To prove such a link, one would need to be able to ectopically manipulate ABA content in different parts of the kernel and examine the effect on desiccation. To this end, the genes controlling ABA content in grain must first be identified. *NCED* and *ZEP* genes which were the first to be shown to have a key role in ABA biosynthesis seemed good candidates prior to this study. However, among the 15 regions grouping the 20 ABA QTLs, only four QTLs collocated with *NCED5* and *ZEP1* loci. As previously mentioned, *NCED5* was expressed at a lower level than *NCED1*, which did not collocate with any ABA QTLs, and both *ZEP1* and *ZEP2* genes were mainly expressed in endosperm in which ABA levels was much lower than in embryo. This can be interpreted as an indication that the control of ABA levels is unlikely to be determined by transcript expression. Post-transcriptional regulation should be checked by measuring *NCED* and *ZEP* protein levels as well as their enzyme activities. Other biosynthetic and/or other regulatory may also be considered. However, none of the newly reported



regulator *Vp8*, *Vp10/Vp13* and *Vp15* genes [25-27] colocalized with the presently detected ABA QTLs. It is noticeable that *PSY3* gene colocalized with ABA\_endosperm\_80\_1 QTL on chromosome 7. Nevertheless, the low levels of *PSY3* transcripts in maize endosperm [37] might undermine the significance of this finding. On the other hand, the genetic control of ABA content in leaf seems to present similarities to that in kernel as shown by the striking number of ABA QTLs of common loci between the two organs. Therefore, shared QTL loci and candidates merit special attention in future work, although the control is seemingly more complex in grain than in leaf.

## Conclusions

This study showed that kernel drying in field condition is genetically controlled. QTLs for kernel water status were frequently associated with QTLs for ABA content in embryo and/or endosperm, providing tools for marker assisted selection. In addition, striking colocalizations were found between the presently mapped QTLs for water status and ABA concentration in kernel and those previously reported in leaf for the same traits. Five novel *NCED* genes were identified and mapped. Phylogenetic analysis established homologies with rice genes. Unexpectedly, it was difficult to establish a causal relationship between the expression of individual members of the maize *NCED* and *ZEP* gene family and ABA QTL effect, although a few colocalizations between these genes and QTLs for ABA were detected. Among the other candidate genes mapped, colocalizations occurred for *aquaporins*, *LEA* and *ABA-responsive* genes and QTLs for kernel desiccation, indicating the potential interest of these genes for breeding and highlighting the necessity to validate them through transgenesis and association genetics.

## Methods

### Plant population and growing conditions

The population used was obtained by performing four generations of random intermating in the F<sub>2</sub> population derived from the maize F<sub>2</sub> × F252 hybrid. Random mating was performed by crossing two plants taken at random. Each plant being used only once, between 80 and 100 crosses were done at each generation to produce the next one. This was followed by one generation of selfing to produce a generation equivalent to F<sub>3</sub> (LHRF\_F<sub>3</sub>) in terms of genotypic frequencies. Each plant was again selfed to derive LHRF\_F<sub>3:4</sub> families (Y.F. Huang, D. Madur, V. Combes, C. L. Ky, P. Bertin, A. Charcosset, L. Moreau, in preparation). The parental lines and the whole LHRF\_F<sub>3:4</sub> population (322 families) were grown in the field at Gif-sur-Yvette, France, in April 2001, in rows (25 plants per rows, one row per family and two rows per parental line). Among the 322

grown families, 153 were sampled 120, 300, 400, 600 and 800 degree × day (cumulated average day temperature above a 6°C basis) after flowering for analyzing time course of desiccation and ABA content. At each stage, two plants per row were sampled for the LHRF\_F<sub>3:4</sub> families. For the parents, three plants in each of the two rows were sampled. We used a thermal scale to compensate for the difference in the dates of plant pollination between genotypes (from July 10th to August 1st, 2001). This scale corresponded to 12, 30, 40, 60 and 80 DAP for the year 2001, thus for simplification all the results were presented as DAP on a 2001 basis. At each date, 20 grains were sampled in the median part of the ear from two plants in each line. They were rapidly frozen in liquid nitrogen and stored in liquid nitrogen.

Statistical analysis of the data (mean, standard error and principal component analysis PCA) were performed using StatBox-Pro package (Grimmersoft, version 6.6, Issy Les Moulineaux, France)

### Kernel desiccation measurement

FW (mg/kernel) and DW (mg/kernel) were measured on five kernels and the four following variables were calculated: i) Water: absolute kernel water content = FW-DW (mg/kernel); ii) %DW: = DW/FW × 100, this variable mirrored relative water content; iii) Slope: regression line slope of (Water/FW × 100) as a function of thermal time between 30 and 80 DAP; iv) Rate = (Water/FW × 100)/(thermal time interval) between two sampling dates.

### ABA content measurement

Embryo and endosperm were separated from the pericarp and treated with two slightly different extraction protocols to optimise grinding efficiency. The embryos were ground with two stainless beads (3 mm diameter) at 20 Hz during 2 fold 30 s and the powder was lyophilized for 24 h. Then, 15 mg were placed in 200 µL distilled water and the slurry was agitated for 48 h at 4°C in the dark. This procedure prevented the formation of a lipid-powder mixture impeding ABA extraction. The whole kernels or the endosperms were lyophilized and ground with 2 stainless beads (7 mm diameter) at 20 Hz during 2 fold 30 s. Two or three aliquots of 250 mg were suspended in 700 µL distilled water for 48 h at 4°C in the dark. All samples were then centrifuged (12,000 g, 10 min) for clarification and the supernatants were used directly for radioimmuno assay (RIA) measurement as previously described. The monoclonal antibody used is only directed against the non-glycosylated ABA form [49]. Values were expressed as ABA content in pg per mg of FW, pg per mg of DW or pg per mg kernel water.

### QTL mapping

The map used for QTL detection was established using the whole population of 322 LHRF\_F<sub>3:4</sub> families (Y.F.

Huang, D. Madur, V. Combes, C. L. Ky, P. Bertin, A. Charcosset, L. Moreau, in preparation). Briefly, leaves of each LHRF\_F<sub>3:4</sub> family (pool of 5 to 20 plants) were harvested to perform genomic DNA isolation, digestion and hybridization for RFLPs and amplified fragment length polymorphisms as previously described [50]. All the lines were characterized for 78 RFLP and 280 SSR markers. Genetic map was constructed using MAPMAKER 3.0b [51]. The genetic distances were calculated with Haldane mapping function. Because the LHRF\_F<sub>3:4</sub> population was regarded as a conventional F<sub>3</sub> population (ignoring the generations of random intermating), those distances were not expressed in "true" cM and were referred to pcM. The linkage map included 358 markers with a total length of 5568.3 pcM and an average distance of 16 pcM between adjacent markers. By comparison with the map of a conventional F<sub>3</sub> population derived from the same parental lines, the map expansion factor was estimated to be 2.63 on average over the whole genome (Y.F. Huang, D. Madur, V. Combes, C. L. Ky, P. Bertin, A. Charcosset, L. Moreau, in preparation).

The QTL detection was performed using composite interval mapping method implemented in the Plabqtl software [52]. Stepwise method was used for covariable identification using the "cov Select" option of the software. The LOD curve shape obtained with the "cov Select" option is strongly affected by the choice of cofactors, leading to artificially narrow peaks in the vicinity of covariate position, which is not appropriate to estimate QTL confidence intervals based on the one-LOD unit fall method. Therefore, confidence intervals, allele effect and individual R<sup>2</sup> of each QTL were estimated in a second step using the "cov-" option, which removes the markers located on the chromosome being analysed from the set of covariables initially selected by the "cov Select" option. The QTL detection threshold and the entry threshold for covariables were set respectively at LOD = 2 and P = 0.005.

To evaluate the risks associated with these threshold values, classical permutations were performed to determine the probability of the maximum of the test statistics under the null hypothesis (absence of any QTL in the genome). Maximum LOD values of 2.0, 2.36, 2.45, 2.88, 3.21 and 3.98 correspond to a genomewide type I risks of 0.5, 0.3, 0.25, 0.10, 0.05 and 0.01, respectively. The use of intermated recombinant inbred population increases the equivalent number of independent tests (over the genome) compared to conventional segregating populations and therefore increases the genomewide type I risk associated to a given LOD value. As such, a LOD 2 value can be considered as being permissive since it corresponds to a genomewide type I error risk of 0.5. Nevertheless, it corresponds to a low individual type I risk of 0.0025. Because we were looking for

colocation between QTLs and candidate genes, even a colocation between a sub-significant QTL and a gene might be of interest especially if several QTL related to similar traits were also colocating. Therefore, we considered QTL with LOD > 2 in colocation analysis and carefully checked that each QTL region displayed at least one QTL with lower individual P values. The genome-wide significance level of each QTL was indicated in Tables 2 and 3. The presently used unreplicated design was chosen to optimize the power of the QTL detection given the number of phenotypical observations, which are very labour intensive for the physiological traits of interest such as ABA content.

#### Colocation between QTLs and candidate genes

Colocations between QTLs and candidate genes were based on manual search of the genes related to kernel desiccation and ABA biosynthesis in the confidence interval of each QTL. In order to have the most exhaustive list of mapped genes, the consensus REFMAP050110 (IBM\_Gnp2004) framework map <http://urgi.versailles.inra.fr/GnpMap> was used [50]. It contains ESTs, genes and QTLs mapped in the French plant genomics programme 'GénoPlante'. The QTLs detected in the LHRF\_F<sub>3:4</sub> population (LHREF3\_1201 map) were projected on the REFMAP050110 using BioMercator software [34]. Data mining was performed using the tools set up by 'GénoPlante' <http://urgi.versailles.inra.fr/>, which yields the list and genetic coordinates of the genes or ESTs reported in the vicinity of the QTLs. Because of the use of intermated recombinant inbred lines, appropriate conversion was performed to compute actual cM from map coordinates obtained with MapMaker [32].

#### Maize *NCED* and *ZEP* gene mapping

Based on Arabidopsis and rice sequences, putative *NCED* and *ZEP* coding sequences in maize were identified by Blast searches of EST assemblies in The Gene Index data base <http://compbio.dfci.harvard.edu/tgi/plant.html>, assemblies of genomic DNA fragments at The Institute for Genomic Research <http://www.tigr.org/plantProjects.shtml> and genomic sequences from BAC clones <http://www.ncbi.nlm.nih.gov/HTGS>. The members of the maize *NCED* and *ZEP* gene families suggested from these combined *in silico* analyses were confirmed by sequencing of amplified cDNA and gDNA which specific primers. Maize *ZEP* genes were annotated using yrGATE at PlantGDB <http://www.plantgdb.org/prj/yrGATE/>, yrGATE\_Zm2 gZEP1 is substantially the same as GRMZM2G127139\_T01). The yrGATE\_Zm10 gZEP2 annotation is incomplete since the last version of the maize HTGS database is not currently included. Therefore, both maize *ZEP1* and *ZEP2* genes are also shown in Additional file 3, which was based on the alignment of rice and maize *ZEP* genes and constructed

using gbioseq [http://www.bioinformatics.org/project/?group\\_id=94](http://www.bioinformatics.org/project/?group_id=94). The maize *NCED* and *ZEP* genes were then mapped on the LHRF or IBM mapping panels using the LHRF\_Gnp2004 or REFMAP050110 (IBM\_Gnp2004) framework map, respectively [50]. LHRF is an intermated recombinant inbred line population derived from the cross between maize F2 and F252 inbred lines. IBM is an intermated recombinant inbred line population derived from the cross between maize B73 and Mo17 inbred lines <http://www.maizegdb.org>. It was used in the absence of polymorphism between F2 and F252 [53]. Genotyping was PCR-based as previously described [50]. Primers used for mapping maize *ZEP* and *NCED* genes are shown in Additional file 2. Gene positions obtained on LHRF\_Gnp2004 map were then projected on the REFMAP050110 map using BioMercator software [34].

#### RT-PCR and qRT-PCR

Gene expression patterns using RT-PCR were determined as previously described [54]. In these end point analyses, the cycle number was adjusted to yield non-saturated bands for samples with low expression. Primers used for RT-PCR analyses are shown in Additional file 2. They were designed based on Gene Bank data <http://www.ncbi.nlm.nih.gov/Genbank/> for aquaporins using the accession numbers provided in [11] and based on the maize EST assembly produced by Genoplante-Info and now housed at the URGI <http://urgi.versailles.inra.fr/GnpSeq/> for *LEA*, *RAB* and other maturation-related genes.

For qRT-PCR, RNA isolation was carried out essentially as previously described [54]. Briefly, tissues were ground to powder under liquid nitrogen and transferred to a tube containing equal volumes of extraction buffer (200 mM Tris-HCl pH 9, 400 mM KCl, 200 mM sucrose, 35 mM MgCl<sub>2</sub>, 25 mM EGTA) and phenol/chloroform/isoamyl alcohol (pH 8) and vortexed for 30 s. The resulting supernatant collected after centrifugation (15,000 g, 10 min at 15°C) was re-extracted twice with phenol/chloroform/isoamyl alcohol and then precipitated 20 min at -80°C by addition of 1 M acetic acid (0.1 volume) and 100% ethanol (2.5 volumes). The RNA pellet collected after centrifugation (15,000 g, 30 min at 4°C) was washed with 3 M Na-acetate (pH 5.2) and re-suspended in water. A second acetic acid/ethanol precipitation was performed before final re-suspension in water. RNA was treated with RNase free DNase and the DNase inactivated according to the instructions of the supplier (Ambion, Applied Biosystems, Foster City, CA, USA). RNA was quantified in a spectrophotometer at 260 nm and its quality was assessed by gel analysis and 260/280 and 260/230 nm spectrophotometric ratios. Total RNA (5 µg) was used as a template to synthesize cDNA using 0.5 µg oligo dT primers (Invitrogen, Breda,

The Netherlands), 100 units of SuperScript II (Invitrogen) and 40 units of recombinant Rnasin ribonuclease inhibitor (Promega, Madison, Wisconsin, USA) in a final volume of 20 µl.

Quantitative real-time PCR was performed with 10 µl of 1:50 (v/v) dilution of the first cDNA strands synthesized as described above using the 7500HT real-time PCR system (Applied Biosystems, Foster City, CA, USA) and the SYBR green PCR master mix (Applied Biosystems) according to the manufacturer's instructions. Each reaction was performed in duplicate and the real-time experiment was repeated two times. The absence of genomic DNA and primer dimers was confirmed by analysis of RT-minus and water control samples and by examination of dissociation curves. The change in fluorescence for each sample was analyzed using 7500 real-time PCR system sequence detection software 1.3.1 (Applied Biosystems). Primer sequences were designed using Primer Express 2.0 (Applied Biosystems). PCR products for candidate genes were sequenced to confirm their identity. Genes and primers used for qRT-PCR analyses are shown in Tables 4 and Additional file 2. Transcript levels were normalized with the values obtained for the housekeeping *Zeastar* gene (accession number AC196679), which was used as an internal reference gene. The fold change of transcript abundance of candidate genes was first calculated as  $2^{-\Delta C_t}$ , where  $\Delta C_t$  is the number of PCR cycles required to reach the log phase of amplification for the target gene minus the same measure for *Zeastar*. Transcript abundance of maize *NCED3* in embryo at 40 DAP was adjusted to 100% and fold changes of transcripts from other genes were normalized via comparison with that of maize *NCED3*. A Log transformation of the expression data was then performed because the standard error values showed that the standard error increased in proportion to the treatment. Log-transformed values were used for two-way ANOVA analyses, using R software version 2.8.1. A F-test for interactions was performed to determine whether the additive model could be retained [55,56]. This additive model corresponds to a multiplicative model for the initial non-transformed values, with  $a$ ,  $b$ ,  $\lambda a$  and  $\lambda b$  being the four values of the multiplicative model. When the additive model could not be retained, the Bonferroni method was applied for pairwise comparisons (see Additional file 3 for calculations). The significance was placed at a 0.05 level.

#### Sequence alignment and phylogenetic analysis

Amino acid sequences of open reading frames were initially aligned using ClustalW [57] with the BioEdit Sequence Alignment Editor 4.8.8 [58]. For the phylogenetic analysis, the N-terminal portions of the proteins were omitted because of the difficulty of confidently assessing primary homology among these sequences.

After the N-terminal regions, the *NCED* sequences share high degrees of similarity. These regions were therefore included in the phylogenetic analyses which were performed using the maximum Likelihood method with PHYML v2.4.4 [59]. JTT was used as the substitution model with four substitution rate categories, the gamma parameter being estimated from the data. Node support was assessed through 500 bootstraps. A dendrogram was constructed using TreeView 1.5.2 <http://taxonomy.zoology.gla.ac.uk/rod/rod.html>.

**Additional file 1: Pearson correlation coefficient between variables for significant values ( $P \leq 0.05$ ).**

Abbreviations: ABAemb: ABA in embryo (pg/DW); ABAend: ABA in endosperm (pg/DW); ABAgrain: ABA in whole grain (pg/DW); DW: kernel dry matter weight (mg/kernel); FW: kernel fresh matter weight (mg/kernel); %DW = DW/FW  $\times$  100; Water: kernel water content (mg/kernel); Rate = (Water/FW  $\times$  100)/(thermal time interval); Slope: regression line slope of (Water/FW  $\times$  100) as a function of thermal time.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2229-10-2-S1.DOC>]

**Additional file 2: Primers used for mapping, RT-PCR and qRT-PCR.**

displaying primers used for mapping, RT-PCR and qRT-PCR.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2229-10-2-S2.DOC>]

**Additional file 3: Identification of two *ZEP* genes in maize.** The rice *ZEP* gene (Os04 g0448900; black) was used in BLAST analysis to identify five putative homologs from maize (The Institute for Genomic Research ID OGAGIC79TC, AZM5\_13314, AZM5\_13316, AZM5\_24223, AZM\_13312 and AZM5\_13315; light brown). Maize ESTs were found for all this maize genomic contigs, except for AZM5\_13315. Sequencing of amplified gDNA and cDNA with specific primers showed that *ZEP* is encoded by only two genes in maize (deduced CDS indicated in blue). This was confirmed using the maize HTGS database (maize *ZEP1*, [GenBank: AC194845.3]; maize *ZEP2*, [GenBank:AC206194.3]; brown) and in agreement with recent reports [394445]. We used yrGATE at PlantGDB <http://www.plantgdb.org/prj/yrGATE> to annotate both genes (maize *ZEP1*: yrGATE\_Zm2 gZEP1; maize *ZEP2*: yrGATE\_Zm10 gZEP2; yrGATE\_Zm10 gZEP2 gene annotation is incomplete since the last version of the maize HTGS database is not currently included in the software).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2229-10-2-S3.DOC>]

**Additional file 4: Analysis of variance table and Bonferroni method.**

containing analysis of variance table and Bonferroni method.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2229-10-2-S4.XLS>]

#### Acknowledgements

We thank Valérie Morin and Benjamin Pouvreau for technical assistance with RT-PCR experiments, Catherine Damerval for guiding us through the phylogenetic analysis and Jean Coursol for guiding us through the statistical analysis. We also thank Carl K.-Y. Ng for carefully reading of the manuscript. This work was supported by grants from the French plant genomics programme 'Génoplatte' (ZMS3P1, B01, B06, GNP05055C and GNP05075C) to AC, SC, VC, AM<sup>1,2</sup>, LM, JLP, AR, CT, CR and PR, and by a grant from the IFR87 La Plante et son Environnement to SC.

#### Author details

<sup>1</sup>Univ Paris-Sud, Institut de Biotechnologie des Plantes, Bât 630, F-91405 Orsay, France. <sup>2</sup>CNRS, UMR 8618, F-91405 Orsay, France. <sup>3</sup>CNRS, UMR 0320/

UMR 8120 Génétique Végétale, F-91190 Gif-sur-Yvette, France. <sup>4</sup>INRA, UMR 0320/UMR 8120 Génétique Végétale, F-91190 Gif-sur-Yvette, France. <sup>5</sup>INRA, Reproduction et Développement des Plantes, UMR 879 INRA-CNRS-ENSL-UCBL, IFR128 Biosciences Lyon-Gerland, F-69364 Lyon Cedex 07, France. <sup>6</sup>52, Av de la Marjolaine, 34110 Frontignan, France.

#### Authors' contributions

VC planned and carried out the sample collection from the field, set up and performed RIA for ABA quantification, carried out QTL detection, managed interaction with the Génoplatte data base, and did preliminary molecular analysis for qRT-PCR; CR, performed *in silico* search of maize *ZEPs*, primer design and sequencing for confirming *in silico* analyses, applying to qRT-PCR and mapping analyses of *NCED* and *ZEP* gene families; LM, provided expertise in the use of intermated recombinant inbred lines and performed the statistical analysis for QTL detection with VC; AR participated in sample collection, in setting up and carrying out high throughput RIA for ABA quantification and in molecular analysis; AM<sup>1,2</sup> contributed to RIA and provided her expertise in aquaporin by designing specific primers; AM<sup>5,6</sup> carried out the RT-PCR; MF performed *NCED* and *ZEP* gene mapping by PCR; AC conceived the intermated recombinant inbred line population and participated in the experimental design; CT participated in the experimental design, in planning and collecting the samples, she largely contributed to manuscript editing; PR managed the RT-PCR measurements and contributed to editing; SC performed *in silico* search of maize *NCED* genes, carried out the phylogenetic analysis, managed the qRT-PCR experiment and the corresponding statistical analysis, then, critically revised the manuscript; JLP conceived and coordinated the study, managed its design and drafted the manuscript. All authors read and approved the final manuscript.

Received: 3 August 2009

Accepted: 4 January 2010 Published: 4 January 2010

#### References

1. Borrás L, Westgate ME: **Predicting maize kernel sink capacity early in development.** *Field Crops Res* 2006, **95**:223-233.
2. Moreau L, Charcosset A, Gallais A: **Use of trial clustering to study QTL  $\times$  environment effects for grain yield and related traits in maize.** *Theor Appl Genet* 2004, **110**:92-105.
3. Dwyer LM, Ma BL, Evenson L, Hamilton RI: **Maize physiological traits related to grain yield and harvest moisture in mid- to short-season environments.** *Crop Sci* 1994, **34**(4):985-992.
4. Eta-Ndu JT, Openshaw SJ: **Selection criteria for grain yield and moisture in maize trials.** *Crop Sci* 1992, **32**(2):332-335.
5. Hawbaker MS, Hill WH, Goodman MM: **Application of recurrent selection for low grain moisture content at harvest in tropical maize.** *Crop Sci* 1997, **37**(5):1650-1655.
6. Sweeney PM, St Martin SK, Clucas CP: **Indirect inbred selection to reduce grain moisture in maize hybrids.** *Crop Sci* 1994, **34**(2):391-396.
7. Bewley JD, Black M: **Physiology of development and germination.** *Seeds* New-York: Plenum Press Publ 1994, 71-88.
8. Battaglia M, Olvera-Carrillo Y, Garciarubio A, Campos F, Covarrubias AA: **The Enigmatic LEA Proteins and Other Hydrophilins.** *Plant Physiol* 2008, **148**(1):6-24.
9. Goyal K, Walton LJ, Tunnacliffe A: **LEA proteins prevent protein aggregation due to water stress.** *Biochem J* 2005, **388**(Pt 1):151-157.
10. Grelet J, Benamar A, Teyssier E, Avelange-Macherel MH, Grunwald D, Macherel D: **Identification in pea seed mitochondria of a late-embryogenesis abundant protein able to protect enzymes from drying.** *Plant Physiol* 2005, **137**(1):157-167.
11. Chaumont F, Barrieu F, Wojcik E, Chrispeels MJ, Jung R: **Aquaporins constitute a large and highly divergent protein family in maize.** *Plant Physiol* 2001, **125**(3):1206-1215.
12. Takahashi H, Rai M, Kitagawa T, Morita S, Masumura T, Tanaka K: **Differential localization of tonoplast intrinsic proteins on the membrane of protein body Type II and aleurone grain in rice seeds.** *Biosci Biotechnol Biochem* 2004, **68**(8):1728-1736.
13. Durantini D, Giulini A, Malgioglio A, Pilu R, Tuberosa R, Sanguineti C, Gavazzi G: **Vivipary as a tool to analyze late embryogenic events in maize.** *Heredity* 2008, **101**(5):465-470.
14. Robertson DS: **The Genetics of viviparity in maize.** *Genetics* 1955, **40**(5):745-760.

15. Koornneef M, Jorna ML, Brinkhorst-van der Swan DLC, Karssen CM: **The isolation of abscisic acid (ABA) deficient mutants by selection of induced revertants in non-germinating gibberellin sensitive lines of *Arabidopsis thaliana* (L.) Heynh.** *Theor Appl Genet* 1982, **61**(4):385-393.
16. Kang MS, Zuber MS: **Combining ability for grain moisture, husk moisture, and maturity in maize with yellow and white endosperms.** *Crop Sci* 1989, **29**(3):689-692.
17. Buckner B, Miguel PS, Janick-Buckner D, Bennetzen JL: **The *y1* gene of maize codes for phytoene synthase.** *Genetics* 1996, **143**(1):479-488.
18. Shinozaki K, Yamaguchi-Shinozaki K: **Gene expression and signal transduction in water-stress response.** *Plant Physiol* 1997, **115**(2):327-334.
19. Nambara E, Marion-Poll A: **Abscisic acid biosynthesis and catabolism.** *Annu Rev Plant Biol* 2005, **56**:165-185.
20. Schwartz SH, Tan BC, Gage DA, Zeevaert JA, McCarty DR: **Specific oxidative cleavage of carotenoids by VP14 of maize.** *Science* 1997, **276**(5320):1872-1874.
21. Tan BC, Schwartz SH, Zeevaert JA, McCarty DR: **Genetic control of abscisic acid biosynthesis in maize.** *Proc Natl Acad Sci USA* 1997, **94**(22):12235-12240.
22. Iuchi S, Kobayashi M, Tajiri T, Naramoto M, Seki M, Kato T, Tabata S, Kakubari Y, Yamaguchi-Shinozaki K, Shinozaki K: **Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in *Arabidopsis*.** *Plant J* 2001, **27**(4):325-333.
23. Tan BC, Joseph LM, Deng WT, Liu L, Li QB, Cline K, McCarty DR: **Molecular characterization of the *Arabidopsis* 9-cis epoxycarotenoid dioxygenase gene family.** *Plant J* 2003, **35**(1):44-56.
24. Frey A, Audran C, Marin E, Sotta B, Marion-Poll A: **Engineering seed dormancy by the modification of zeaxanthin epoxidase gene expression.** *Plant Mol Biol* 1999, **39**(6):1267-1274.
25. Porch TG, Tseung CW, Schmelz EA, Settles AM: **The maize *viviparous10/viviparous13* locus encodes the *Cnx1* gene required for molybdenum cofactor biosynthesis.** *Plant J* 2006, **45**(2):250-263.
26. Suzuki M, Settles AM, Tseung CW, Li QB, Latshaw S, Wu S, Porch TG, Schmelz EA, James MG, McCarty DR: **The maize *viviparous15* locus encodes the molybdopterin synthase small subunit.** *Plant J* 2006, **45**(2):264-274.
27. Suzuki M, Latshaw S, Sato Y, Settles AM, Koch KE, Hannah LC, Kojima M, Sakakibara H, McCarty DR: **The Maize *viviparous8* locus, encoding a putative ALTERED MERISTEM PROGRAM1-like peptidase, regulates abscisic acid accumulation and coordinates embryo and endosperm development.** *Plant Physiol* 2008, **146**(3):1193-1206.
28. Robertson DS: **A possible technique for isolating genic DNA for quantitative traits in plants.** *J Theor Biol* 1985, **117**(1):1-10.
29. Tuberosa R, Salvi S: **QTL for Agronomic Traits in Maize Production.** *Handbook of Maize: Its Biology* New-York: SpringerBennetzen JL, Hake SC 2009, 501-541.
30. Prioul JL, Quarrie S, Causse M, deVienne D: **Dissecting complex physiological functions through the use of molecular quantitative genetics.** *J Exp Bot* 1997, **48**(311):1151-1163.
31. Sala RG, Andrade FH, Camadro EL, Ceroni JC: **Quantitative trait loci for grain moisture at harvest and field grain drying rate in maize (*Zea mays*, L.).** *Theor Appl Genet* 2006, **112**(3):462-471.
32. Falque M: **IRILmap: linkage map distance correction for intermated recombinant inbred lines/advanced recombinant inbred strains.** *Bioinformatics* 2005, **21**(16):3441-3442.
33. Prioul JL, Pelleschi S, Sene M, Thévenot C, Causse M, deVienne D, Leonardi A: **From QTLs for enzyme activity to candidate genes in maize.** *J Exp Bot* 1999, **50**(337):1281-1288.
34. Arcade A, Labourdette A, Falque M, Mangin B, Chardon F, Charcosset A, Joets J: **BioMercator: integrating genetic maps and QTL towards discovery of candidate genes.** *Bioinformatics* 2004, **20**(14):2324-2326.
35. Vogel JT, Tan BC, McCarty DR, Klee HJ: **The carotenoid cleavage dioxygenase 1 enzyme has broad substrate specificity, cleaving multiple carotenoids at two different bond positions.** *J Biol Chem* 2008, **283**(17):11364-11373.
36. Schwartz SH, Tan BC, McCarty DR, Welch W, Zeevaert JA: **Substrate specificity and kinetics for VP14, a carotenoid cleavage dioxygenase in the ABA biosynthetic pathway.** *Biochim Biophys Acta* 2003, **1619**(1):9-14.
37. Li F, Vallabhaneni R, Wurtzel ET: ***PSY3*, a new member of the phytoene synthase gene family conserved in the poaceae and regulator of abiotic stress-induced root carotenogenesis.** *Plant Physiol* 2008, **146**(3):1333-1345.
38. Kizis D, Pages M: **Maize DRE-binding proteins DBF1 and DBF2 are involved in rab17 regulation through the drought-responsive element in an ABA-dependent pathway.** *Plant J* 2002, **30**(6):679-689.
39. Pelleschi S, Leonardi A, Rocher JP, Cornic G, deVienne D, Thévenot C, Prioul JL: **Analysis of the Relationships between Growth, Photosynthesis and Carbohydrate Metabolism Using Quantitative Trait Loci (QTLs) in Young Maize Plants Subjected to Water Deprivation.** *Mol Breed* 2006, **17**(1):21-39.
40. Manicacci D, Camus-Kulandaivelu L, Fourmann M, Arar C, Barrault S, Rousselet A, Feminias N, Consoli L, Frances L, Mechin V, Murigneux A, Prioul JL, Charcosset A, Damerval C: **Epistatic interactions between *Opaque2* transcriptional activator and its target gene *CyPPDK1* control kernel trait variation in maize.** *Plant Physiol* 2009, **150**(1):506-520.
41. Pelleschi S, Guy S, Kim JY, Pointe C, Mahe A, Barthes L, Leonardi A, Prioul JL: ***lvr2*, a candidate gene for a QTL of vacuolar invertase activity in maize leaves. Gene-specific expression under water stress.** *Plant Mol Biol* 1999, **39**(2):373-380.
42. Thévenot C, Simond-Cote E, Reyss A, Manicacci D, Trouverie J, Le Guilloux M, Ginhoux V, Sidicina F, Prioul JL: **QTLs for enzyme activities and soluble carbohydrates involved in starch accumulation during grain filling in maize.** *J Exp Bot* 2005, **56**(413):945-958.
43. Beavis WD, Smith OS, Grant D, Fincher RR: **Identification of quantitative trait loci using a small sample of topcrossed and F4 progeny from maize.** *Crop Sci* 1994, **34**:882-896.
44. Lebreton C, Lazic-Jancic V, Steed A, Pekic S, Quarrie SA: **Identification of QTL for drought responses in maize and their use in testing causal relationships between traits.** *J Exp Bot* 1995, **46**:853-865.
45. Tuberosa R, Sanguineti MC, Landi P, Salvi S, Casarini E, Conti S: **RFLP mapping of quantitative trait loci controlling abscisic acid concentration in leaves of drought-stressed maize (*Zea mays* L.).** *Theor Appl Genet* 1998, **97**(5-6):744-755.
46. Landi P, Sanguineti MC, Liu C, Li Y, Wang TY, Giuliani S, Bellotti M, Salvi S, Tuberosa R: **Root-ABA1 QTL affects root lodging, grain yield, and other agronomic traits in maize grown under well-watered and water-stressed conditions.** *J Exp Bot* 2007, **58**(2):319-326.
47. Hachez C, Heinen RB, Draye X, Chaumont F: **The expression pattern of plasma membrane aquaporins in maize leaf highlights their role in hydraulic regulation.** *Plant Mol Biol* 2008, **68**(4-5):337-353.
48. Busk PK, Pages M: **Protein binding to the abscisic acid-responsive element is independent of VIVIPAROUS1 in vivo.** *Plant Cell* 1997, **9**(12):2261-2270.
49. Quarrie SA, Whitford PN, Appleford MEJ, Wang TL, Cook SK, Henson IE, Loveys BR: **A monoclonal antibody to (S)-abscisic acid: its characterization and use in a radioimmunoassay for measuring abscisic acid in crude extracts of cereals and lupin leaves.** *Planta* 1988, **173**:330-339.
50. Falque M, Decousset L, Dervins D, Jacob AM, Joets J, Martinant JP, Raffoux X, Ribiere N, Ridet C, Samson D, Charcosset A, Murigneux A: **Linkage mapping of 1454 new maize candidate gene loci.** *Genetics* 2005, **170**(4):1957-1966.
51. Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L: **MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations.** *Genomics* 1987, **1**:174-181.
52. Utz HF, Melchinger AE: **A program for composite interval mapping of QTL.** *J Quant Trait Loci* 1996, **2**(1), electronic journal.
53. Lee M, Sharopova N, Beavis WD, Grant D, Katt M, Blair D, Hallauer A: **Expanding the genetic map of maize with the intermated B73 x Mo17 (IBM) population.** *Plant Mol Biol* 2002, **48**(5-6):453-461.
54. Massonneau A, Condamine P, Wisniewski JP, Zivy M, Rogowsky PM: **Maize cystatins respond to developmental cues, cold stress and drought.** *Biochim Biophys Acta* 2005, **1729**(3):186-199.
55. Rao CR: **Linear Statistical Inference and its Applications.** Wiley, John & Sons, New-York 2002.
56. Scheffé H: **The analysis of Variance.** Wiley, John & Sons, New-York 1999.
57. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through**

sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994, **22**(22):4673-4680.

58. Hall T: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucleic Acids Symp* 1999, **41**:95-98.
59. Guindon S, Lethiec F, Duroux P, Gascuel O: **PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference.** *Nucleic Acids Res* 2005, **33** Web Server: W557-559.

doi:10.1186/1471-2229-10-2

**Cite this article as:** Capelle *et al.*: QTLs and candidate genes for desiccation and abscisic acid content in maize kernels. *BMC Plant Biology* 2010 **10**:2.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

