

# Unravelling wild carrot differentiation in Europe – preliminary data on a candidate gene approach

T. Nobre, C. Ragonezi and B. Arnholdt-Schmitt

EU Marie Curie Chair, ICAAM, Universidade de Évora, Núcleo da Mitra, Évora, Portugal.

## Abstract

Carrot is an outcrossing species and levels of gene flow between populations, and even between wild and domesticated relatives, are expected to be high. Cases of natural hybridization and introgression of crops and wild relatives have been reported. Have these events diluted any putative habitat-adapted genotypes? In other words, can we still find a correlation between wild carrot genotypes and regional/local environment? We have chosen to start addressing this question using a member of the alternative oxidase (AOX) gene family. AOX genes seem to be linked to all kinds of abiotic and biotic stress reactions. Wild carrots were sampled in an environmental gradient across Western Europe. This gradient included sampling points with more deviating conditions, such as Sierra de Guadarrama or the central Pyrenees and the French Massif Central. Phylogenetic reconstruction on this molecular marker is to be combined with geographic, climatic, and ecological evidence. So far, the preliminary results suggest the existence of a biogeographical barrier at the Pyrenees, and higher gene diversity than initially expected. From an applied point of view, diversity of functional traits is much more relevant than species diversity. Gene transfer from wild to cultivated plants has contributed to the evolution of crop species. Providing that deterioration of genetic resources and biodiversity loss have not been drastic, gene transfer from wild plants has the potential to further contribute to a (targeted) improvement of cultivars.

**Keywords:** *Daucus carota* L., AOX, habitat-adapted, wild crop relatives, hybridization

## INTRODUCTION

Crop plants typically include only a portion of the genetic diversity of their wild relatives. Since genetic variation is the raw material of evolution, low genetic diversity has as direct consequence a reduction on the ability of the species to evolve in response to changes in its environment. The extent of this loss of diversity depends on the population size during the domestication period, the mating system and the duration of that period, and it is not experienced equally by all genes in the genome (Wright et al., 2005). The narrowing down of genetic variability among modern cultivars has implications for survival from several biotic and abiotic stresses.

For many crops, this diversity loss due to domestication has made the plant totally dependent on humans such that it is no longer capable of propagating itself in nature; in carrot, however, the domestication process rendered more modest changes when compared to their progenitors – and they can even revert to the wild or become self-propagating weeds (Doebley et al., 2006).

Cultivated carrot (*Daucus carota* subsp. *sativus*) is a common vegetable with a well-known and widely distributed weedy ancestor, wild carrot (*D. carota*). Carrot is an outcrossing species and levels of gene flow between populations, and even between wild and domesticated relatives, are expected to be high. Cultivated carrots do hybridize with wild carrots (e.g. Brandenburg, 1981; Heywood, 1983), which grow in most regions of the world where carrots are cultivated. Cases of natural hybridization and introgression of crops and wild relatives have been reported (Magnussen and Hauser, 2007 and citations therein). However, other studies argue against extensive gene transfer between cultivated and wild populations (e.g. Small, 1984; Bradeen et al., 2002). Nonetheless, the outcrossing nature of



carrot without any clonal propagation, associated with the fact that open-pollinated seed production was (likely) used to propagate carrot during domestication, lead to no or small reduction in the genetic diversity of cultivated vs. wild carrot (Iorizzo et al., 2013). Do the carrot mating-system and consequent cases of natural hybridization and introgression have diluted any putative habitat-adapted genotypes? In other words, can we find a correlation between wild carrot genotypes and regional/local environment?

To start addressing this question, we have chosen a member of the alternative oxidase (AOX) gene family. AOX genes seem to be linked to all kinds of abiotic and biotic stress reactions in various organisms (e.g. Vanlerberghe and McIntosh, 1997; Liao et al., 2012; Li et al., 2012) and have the potential to be used in a model towards identification of crop wild relatives (CWR) hotspots of stress adapted diversity. The choice of this gene relates to the fact that AOX is not only part of the stress response in plants, but it also plays a central role in defining the stress response (Van Aken et al., 2009). The rendering proteins are active in mitochondria, organelles of crucial importance for environmental stress perception and stress signal transduction. This gene family is present in the respiratory chains of all plants, as well as in certain fungi, protists, animals and bacteria, and it is crucially involved in the adaptive regulation of metabolism. Within the same species, individual genotypes and/or groups of genotypes can be distinguished by polymorphic AOX gene sequences (e.g. Abe et al., 2002; Holtzapffel et al., 2003; Guerra Cardoso et al., 2009; Costa et al., 2009; Santos Macedo et al., 2009). Because diversity of functional traits is much more relevant than species diversity, scanning wild crop relatives with a stress related gene has the potential to highlight hot-spots of specific functional diversity.

## MATERIALS

Sampling was performed following an environmental gradient across Europe (Figure 1). This gradient accommodated some sampling points with more deviating climatic conditions, such as in Sierra de Guadarrama (considerable temperatures changes between summer and winter and a very dry summer; wild carrots could not be found above the 1100 m) or in the central Pyrenees and the French Massif Central (with a cold continental climate at equivalent height). The temperature variation, hours of sun and water availability varies considerable from the most southern-western sampling point (38.58252, -7.99804) to the most northern-eastern (54.07787, 12.11262). In general, the sampling was made on easily accessible non-cultivated fields (thus close to a road), with altitudes ranging from the referred 1100 m to sea level. In total, 12 places were sampled by collecting at each location 8 wild carrot roots. The samples were dried with silica gel and are stored at -80°C.

DNA from roots was extracted using an adaptation of the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) standard extraction protocol: 1) extra initial grinding step (with liquid nitrogen and using a tissue grinder) to further pulverize the root tissues; 2) addition of polyvinylpyrrolidone (PVP, 10 000 mol wt at 3%) to the extraction buffer to remove phenolic and other compounds that can inhibit PCR and 3) lysis was performed overnight at 60 rpm. DNA concentration of all samples was determined with the NanoDrop-2000C spectrophotometer. For the isolation of *AOX1* amplicon in wild carrots specific primers were designed based on the already available cDNA sequence at the NCBI GenBank (EU286573.2). A nested PCR approach was selected. For the first reaction, the primers used were located at the beginning of exon 1 as forward and at the end of exon 3 as reverse (more details in Nogales et al., in prep.). A standard 25 µL reaction at 2.5 mM of magnesium and BSA at 0.4 µg µL<sup>-1</sup> was run with an annealing temperature of 55°C. The PCR product was then diluted in 1:50 and used as template for the second reaction. The second reaction was done with the same forward primer and the degenerated primer P2 (Saisho et al., 1997), with a standard 25 µL reaction at 1.5 mM of magnesium and BSA at 0.4 µg µL<sup>-1</sup>. Annealing temperature was 60°C. Amplicons were purified from the agarose gel with GFX PCR DNA and Gel Band Purification Kit, directly cloned into pGEM®-T Easy vector (Promega, Madison, WI, USA), transformed into bacterial strain JM109 (Promega, USA) and bacterial colonies were tested using T7 and SP6 primers. Sequence was done from the PCR product, in sense and antisense strands.



Figure 1. Sampling locations according to a climatic SW-NE transect, distinguishing the populations for which *AOX1* gene amplifications were obtained (c).

Sequence assemblies at gDNA level were performed in CLC Main Workbench vs 6.8.1 software, and alignments in MAFFT (Kato et al., 2005) under the option L-INS-I (iterative refinement method incorporating local pairwise alignments; gap opening penalty: 1.5 and gap extension penalty 0.14; 1PAM/k  $\frac{1}{4}$  2 scoring matrix for nucleotide sequences). The option “leave gappy regions” was selected as it was observed that intron 1 is highly diverse and locally contaminated by unrelated segments. The optimal model of evolution was selected, separately for each of the intron and exon regions, using MrMODELTEST v. 2.2 (Posada and Crandall 1998). Bayesian inference was conducted using MrBayes version 3.0 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) according to the models selected and with MCMC runs being repeated twice as a safeguard against spurious results. The first  $10^3$  trees were discarded as burn-in, and the remaining trees were used to calculate a majority rule consensus tree. Stationarity was confirmed by analysis of the log likelihoods and the consistency between runs.

## RESULTS AND DISCUSSION

At the current stage, data was obtained only from 6 out of the 12 populations (Figure 1). The pattern suggests the existence of a biogeographical differentiation at the Pyrenees, as the chosen approach for amplification of the *AOX1* amplicon failed for all the samples from the Iberian Peninsula (while no changes were made to the methodology). The desired amplicon comprised part of exon 1, intron 1 and part of exon 2 (Campos et al., in prep). Overall, the data show much higher gene diversity than initially expected: the amplicon size varies considerably due to intron 1 size, which showed on average 677 bp but was found to be as long as 945 bp in one case. We hypothesized that the lack of results for the Iberian samples is due to several polymorphisms at the exon 1 region, impairing amplification using the primers designed on the basis of known sequences from cultivars and inbred lines. The hypothesis that the polymorphisms are on the other exons should also be acknowledge, however less likely as exon 2 and exon 3 are much more conserved (Nogales et al., submitted). This bottleneck is expected to be overcome with the gathering of more sequence

information from wild carrots, which will also allow verifying this hypothesis.

Nonetheless, the high variability found on intron 1 is worth notice. In particular, the longest intron found (945 bp) was on a wild carrot specimen collected in the vicinity of Port-la-Nouvelle (France), which showed an insertion of 300 bp at the end of intron. This insertion, when searched for on the databases through BLAST algorithms, shows high identity with clones of *Daucus sahariensis* and *Daucus syrticus* (Figure 2). Arbizu et al. (2014) found that even with combined molecular and morphological studies there are particular problems in distinguishing these two species. *Daucus syrticus*, and *D. sahariensis* together with *D. gracilis*, are probably the most closely related plants to *D. carota* (Lee and Park 2014). Thus, even though *D. sahariensis* is native in North Africa -and could eventually have been introduced in France-the insertion event observed at intron 1 of the referred AOX1 sequence might thus result from a hybridization event between wild *D. carota* and *D. sahariensis*.

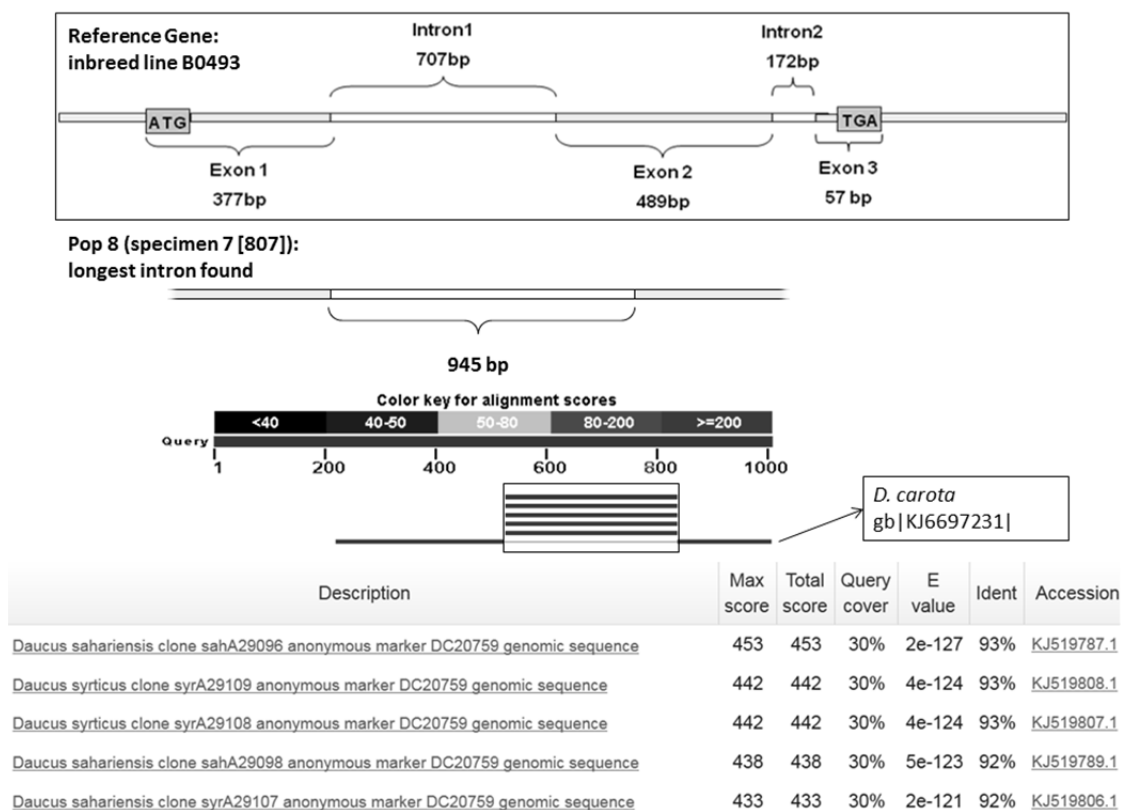


Figure 2. The longest insertion found at AOX1 intron 1 of *D. carota* wild representative, and alignment results with other *Daucus* species.

Because introns can have large influence in the control of gene expression in plants (Gianí et al., 2003; Fiume et al., 2004), they are also worth consideration. Exon and intron regions are obviously subjected to different evolutionary forces and sequence alignment and phylogeny reconstruction should take that into consideration. Indels for instance, are rather frequent in intron regions and can make alignment difficult and often subjective. Exon regions are per definition codifying regions, where mutations have higher costs. Therefore, prior to the phylogenetic analyses, each gene partition was subjected to alignment under more or less strict parameters (dependent on being an exon or intron region respectively) and the substitutional model of molecular evolution was assigned according to distance analyses performed using the log likelihood function for ML searches (Posada and Crandall, 1998). In fact, exons followed the nucleotide substitution model K2+G+I [exon 1 showed 21

variable codons (71 variable nucleotides) of which 9 codons were parsimony informative (with 2 rare codons) and exon 2 showed 25 variable codons (59 variable nucleotides) of which only 2 were parsimony informative] and the intron followed T92+G [358 variable nucleotide sites (putative hybridization event above described not included), of which 242 were parsimony informative]. Both models are based on the Kimura80 model (Kimura, 1980), which distinguishes between transitions and transversions, but in the case of the intron the model accommodates strong transition-transversion and G+C-content biases.

The reconstructed phylogeny of the available wild carrot data (ca. 35 sequences) based on AOX1 amplicon (Figure 3) is still preliminary and inconclusive. More resolution is expected when Iberian Peninsula specimens are included and populations are represented by a higher number of specimens (crucial due to the outcrossing potential of the species), and through the integration of the phylogenetic data with further geographic, climatic, and ecological evidence.

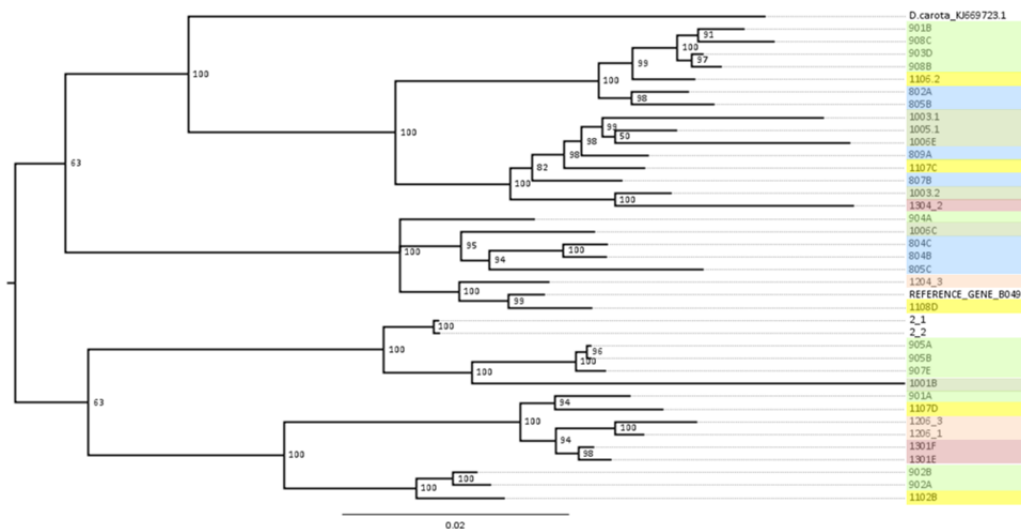


Figure 3. DcAOX1 sequences of analysed for this study, where the colors group individual plants in populations. No specific clades are apparent and the tree, although resolved, shows a relatively low support of the initial branching. The midpoint rooted tree corresponds to the majority rule consensus tree of trees sampled in a Bayesian analysis. The numbers at the nodes refer to the Bayesian posterior probability of the nodes (more than 50%) derived from 19500 Markov chain Monte Carlo-sampled trees.

Wild carrot is the most important gene pool for carrot breeding. Introducing genetic diversity from wild carrot can expand the genetic variation of cultivated carrot and therefore facilitate the improvement of cultivation traits (Simon et al., 2007), for example temperature extremes tolerance. Global climate change is another threat that agriculture is facing: higher temperatures likely reduce yields of crops, changes in precipitation increase the likelihood of short-run crop failures and long-run production declines (Nelson et al., 2009). The consequences on the European agricultural ecosystems will vary depending on the crop, the region and the likely climate changes (Bindi and Oleson, 2011) but it will certainly require a great deal of adaptation of the cropping systems to the stresses imposed by the new environmental conditions. A tool that can highlight putative hotspots of diversity in wild carrot populations in relation to a given environmental factor will allow directing and designing appropriate sampling strategies for the purpose of germplasm collections.

**ACKNOWLEDGMENTS**

TN was supported by a Marie Curie fellowship (FP7-PEOPLE-2012-CIG Project Reference 321725) and by the Portuguese Foundation for Science and Technology – FCT



(SFRH/BCC/52187/2013). BAS was supported by the Portuguese Foundation for Science and Technology under the program POPH – Programa Operacional Potencial Humano (Ciência 2007 and Ciência 2008). This work was aided by FEDER Funds through the Operational Program for Competitiveness Factors – COMPETE and National Funds through FCT – Foundation for Science and Technology under the Strategic Project Pest-C/AGR/UI0115/2011 and PEst-OE/AGR/UI0115/2014.

## Literature cited

Abe, F., Saito, K., Miura, K., and Toriyama, K. (2002). A single nucleotide polymorphism in the alternative oxidase gene among rice varieties differing in low temperature tolerance. *FEBS Lett.* *527* (1-3), 181–185.[http://dx.doi.org/10.1016/S0014-5793\(02\)03218-0](http://dx.doi.org/10.1016/S0014-5793(02)03218-0) PubMed

Arbizu, C., Reitsma, K.R., Simon, P.W., and Spooner, D.M. (2014). Morphometrics of *Daucus* (Apiaceae): a counterpart to a phylogenomic study. *Am. J. Bot.* *101* (11), 2005–2016.<http://dx.doi.org/10.3732/ajb.1400252> PubMed

Bindi, M., and Oleson, J.E. (2011). The responses of agriculture in Europe to climate change. *Reg. Environ. Change* *11* (S1), 151–158 <http://dx.doi.org/10.1007/s10113-010-0173-x>.

Bradeen, J.M., Bach, I.C., Briard, M., le Clerc, V., Grzebelus, D., Senalik, D.A., and Simon, P.W. (2002). Molecular diversity analysis of cultivated carrot (*Daucus carota* L.) and wild *Daucus* populations reveals a genetically nonstructured composition. *J. Am. Soc. Hortic. Sci.* *127*, 383–391.

Brandenburg, I.W.A. (1981). Possible relationships between wild and cultivated carrots (*Daucus carota* L.) in the Netherlands. *Kulturpflanze* *29* (1), 369–375 <http://dx.doi.org/10.1007/BF02014765>.

Costa, J.H., Cardoso, H.G., Campos, M.D., Zavattieri, A., Frederico, A.M., de Melo, D.F., and Arnholdt-Schmitt, B. (2009). *Daucus carota* L. - an old model for cell reprogramming gains new importance through a novel expansion pattern of AOX genes. *Plant Physiol. Biochem.* *47*, 753–759.<http://dx.doi.org/10.1016/j.plaphy.2009.03.011> PubMed

Doebley, J.F., Gaut, B.S., and Smith, B.D. (2006). The molecular genetics of crop domestication. *Cell* *127* (7), 1309–1321.<http://dx.doi.org/10.1016/j.cell.2006.12.006> PubMed

Fiume, E., Christou, P., Gianì, S., and Breviario, D. (2004). Introns are key regulatory elements of rice tubulin expression. *Planta* *218* (5), 693–703.<http://dx.doi.org/10.1007/s00425-003-1150-0> PubMed

Gianì, S., Morello, L., Bardini, M., and Breviario, D. (2003). Tubulin intron sequences: multi-functional tools. *Cell Biol. Int.* *27* (3), 203–205.[http://dx.doi.org/10.1016/S1065-6995\(02\)00301-3](http://dx.doi.org/10.1016/S1065-6995(02)00301-3) PubMed

Guerra Cardoso, H., Doroteia Campos, M., Rita Costa, A., Catarina Campos, M., Nothnagel, T., and Arnholdt-Schmitt, B. (2009). Carrot alternative oxidase gene AOX2a demonstrates allelic and genotypic polymorphisms in intron 3. *Physiol Plant* *137* (4), 592–608.<http://dx.doi.org/10.1111/j.1399-3054.2009.01299.x> PubMed

Heywood, V.H. (1983). Relationships and evolution in the *Daucus carota* complex. *Isr. J. Bot.* *32*, 51–65.

Holtzapffel, R.C., Castelli, J., Finnegan, P.M., Millar, A.H., Whelan, J., and Day, D.A. (2003). A tomato alternative oxidase protein with altered regulatory properties. *Biochim. Biophys. Acta* *1606* (1-3), 153–162.[http://dx.doi.org/10.1016/S0005-2728\(03\)00112-9](http://dx.doi.org/10.1016/S0005-2728(03)00112-9) PubMed

Huelsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: bayesian inference of phylogenetic trees. *Bioinformatics* *17* (8), 754–755.<http://dx.doi.org/10.1093/bioinformatics/17.8.754> PubMed

Iorizzo, M., Senalik, D.A., Ellison, S.L., Grzebelus, D., Cavagnaro, P.F., Allender, C., Brunet, J., Spooner, D.M., Van Deynze, A., and Simon, P.W. (2013). Genetic structure and domestication of carrot (*Daucus carota* subsp. *sativus*) (Apiaceae). *Am. J. Bot.* *100* (5), 930–938.<http://dx.doi.org/10.3732/ajb.1300055> PubMed

Katoh, K., Kuma, K., Toh, H., and Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* *33* (2), 511–518.<http://dx.doi.org/10.1093/nar/gki198> PubMed

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* *16* (2), 111–120.<http://dx.doi.org/10.1007/BF01731581> PubMed

Lee, B.Y., and Park, C.-W. (2014). Molecular phylogeny of *Daucus* (Apiaceae): evidence from nuclear ribosomal DNA ITS sequences. *J. Species Res.* *3* (1), 39–52 <http://dx.doi.org/10.12651/JSR.2014.3.1.039>.

Li, Y., Zhu, L., Xu, B., Yang, J., and Zhang, M. (2012). Identification of down-regulated genes modulated by an alternative oxidase pathway under cold stress conditions in watermelon plants. *Plant Mol. Biol. Report.* *30* (1), 214–224 <http://dx.doi.org/10.1007/s11105-011-0333-2>.

Liao, Y.W.K., Shi, K., Fu, L.J., Zhang, S., Li, X., Dong, D.K., Jiang, Y.P., Zhou, Y.H., Xia, X.J., Liang, W.S., and Yu, J.Q. (2012).

- The reduction of reactive oxygen species formation by mitochondrial alternative respiration in tomato basal defense against TMV infection. *Planta* 235 (2), 225–238.<http://dx.doi.org/10.1007/s00425-011-1483-z> PubMed
- Magnussen, L.S., and Hauser, T.P. (2007). Hybrids between cultivated and wild carrots in natural populations in Denmark. *Heredity* (Edinb) 99 (2), 185–192.<http://dx.doi.org/10.1038/sj.hdy.6800982> PubMed
- Nelson, G.C., Rosegrant, M.W., Koo, J., Robertson, R., Sulser, T., Zhu, T., and Lee, D. (2009). *Climate Change: Impact on Agriculture and Costs of Adaptation* (Intl. Food Policy Res. Inst.).
- Posada, D., and Crandall, K.A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14 (9), 817–818.<http://dx.doi.org/10.1093/bioinformatics/14.9.817> PubMed
- Ronquist, F., and Huelsenbeck, J.P. (2003). MrBayes 3: bayesian phylogenetic inference under mixed models. *Bioinformatics* 19 (12), 1572–1574.<http://dx.doi.org/10.1093/bioinformatics/btg180> PubMed
- Saisho, D., Nambara, E., Naito, S., Tsutsumi, N., Hirai, A., and Nakazono, M. (1997). Characterization of the gene family for alternative oxidase from *Arabidopsis thaliana*. *Plant Mol. Biol.* 35 (5), 585–596.<http://dx.doi.org/10.1023/A:1005818507743> PubMed
- Santos Macedo, E., Cardoso, H.G., Hernández, A., Peixe, A.A., Polidoros, A., Ferreira, A., Cordeiro, A., and Arnholdt-Schmitt, B. (2009). Physiologic responses and gene diversity indicate olive alternative oxidase as a potential source for markers involved in efficient adventitious root induction. *Physiol Plant* 137 (4), 532–552.<http://dx.doi.org/10.1111/j.1399-3054.2009.01302.x> PubMed
- Simon, P.W., Freeman, R.E., Vieira, J.V., et al. (2007). Carrot. In *Vegetables II Fabaceae, Liliaceae, Solanaceae, and Umbelliferae*. J. Prohens, and F. Nuez, ed. (New York: Springer).
- Small, E. (1984). Hybridization in the domesticated-weed-wild complex. In *Plant Biosystematics*, W.F. Grant (Toronto: Academic Press).
- Van Aken, O., Giraud, E., Clifton, R., and Whelan, J. (2009). Alternative oxidase: a target and regulator of stress responses. *Physiol Plant* 137 (4), 354–361.<http://dx.doi.org/10.1111/j.1399-3054.2009.01240.x> PubMed
- Vanlerberghe, G.C., and McIntosh, L. (1997). Alternative oxidase: from gene to function. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 48 (1), 703–734.<http://dx.doi.org/10.1146/annurev.arplant.48.1.703> PubMed
- Wright, S.I., Bi, I.V., Schroeder, S.G., Yamasaki, M., Doebley, J.F., McMullen, M.D., and Gaut, B.S. (2005). The effects of artificial selection on the maize genome. *Science* 308 (5726), 1310–1314.<http://dx.doi.org/10.1126/science.1107891> PubMed

