Exploratory QTL Analyses of some Pepper Physiological Traits in Two Environments

N.A. Alimi[1,2], M.C.A.M. Bink [2], A. Dieleman [3,4], A.M. Sage-Palloix[1], R.E. Voorrips [4], V. Lefebvre [1], A. Palloix [1], F.A. van Eeuwijk [2]

1 INRA- Avignon, GAFL UR 1052, BP 94, 84143 Montfavet Cedex France.
2 Wageningen UR, Biometris, P.O. Box 100, 6700AC, Wageningen, the Netherlands.
3 Wageningen UR Greenhouse Horticulture, P.O. Box 644, 6700 AP, Wageningen, Netherlands
4 Wageningen UR, Plant Research International, P.O. Box 16, 6700 AA, Wageningen, Netherlands.
Contact: Nurudeen.Alimi@avignon.inra.fr; nurudeenadeniyi.alimi@wur.nl

Abstract

behind phenotypic differences and led to selection of genotypes having favourable traits. Continuous monitoring of environmental conditions has also become an accessible option. Rather than single trait evaluation, we would prefer smarter approaches capable of evaluating multiple, often correlated and time dependent traits simultaneously as a function of genes (QTLs) and environmental inputs, where we would The use of molecular breeding techniques has increased insight into the genetics like to include intermediate genomic information as well. In this paper, an exploratory QTL analysis over two environments was undertaken using available genetic and phenotypic data from segregating recombinant inbred lines (RIL) of pepper (*Capsicum annuum*). We focused on vegetative traits, e.g. stem length, speed of stem development, number of internodes etc. We seek to improve the estimation of allelic values of these traits under the two environments and determine possible QTL x E interaction. Almost identical QTLs are detected for each trait under the two environments but with varying LOD scores. No clear evidence was found for presence of QTL by environment interactions, despite differences in phenotypes and in magnitude of QTLs expression. Within the EU project SPICY (Voorrips et al., 2010 this issue), a larger number of environments will be studied and more advanced statistical analysis tools will be considered. The correlation between the traits will also be modelled. The identification of markers for the important QTL (Nicolaï et al., 2010 this issue) will improve the speed and accuracy of genomic prediction of these complex phenotypes.

Introduction

The use of molecular breeding techniques has contributed considerably to the unraveling of crop traits that have impacted the quality and yield of plant products. It has increased insight into the genetics behind the genotypic differences and allows breeders to achieve earlier and more accurate selection of genotypes having favorable traits. Yield in agronomic and horticultural crops is a composite trait with many underlying traits and genetic factors that may mask or accentuate each other and also interact with environmental factors. Dealing with such a complex trait requires more advanced approaches capable of evaluating multiple traits simultaneously rather than single trait evaluation. This will enable breeders to investigate issues related to pleiotropy and genetic linkage that underlie commonly observed genetic correlations between traits. For such complex traits exhibiting considerable genotype by environment interaction, these QTLs have to be analyzed by considering their combination under different environment using the so called QTL x E analysis. The specific goal of this work is therefore to study the presence and magnitude of interaction between QTLs and environment.

Materials and methods

Data Sources and Description

Data from the first SPICY experiment at Wageningen University and Research Center (WUR), the Netherlands and the already published data from INRA, France (Barchi et al, 2009) are used. The genotypes are from the fifth generation of Recombinant Inbred Lines (RILs) of an intraspecific cross between large – fruited inbred cultivar 'Yolo Wonder' (YW) and the hot pepper cultivar 'Criollo de Morelos 334' (CM 334). There are a total of 297 RILs from the INRA experiment from which a subset of 149 lines was selected in the WUR experiment, using the MapPop software (Brown and Vision 2000), for selective phenotyping. The 149 most informative individuals were selected using the full linkage map as the input file, and the maximum bin length (eMBL) as the selection criterion. The genetic linkage map was constructed from genotypic data on a set of 587 markers (507 AFLPs, 40 SSRs, 19 RFLPs, 17 SSAPs and four STSs). A total of 489 markers were assembled into 49 linkage groups (LGs). Twenty-three of these LGs, composed of 69% of the markers and covering 1553 cM, were assigned to one of the 12 haploid pepper chromosomes, leaving 26 small LGs (304 cM) unassigned (Barchi et al., 2007).

The WUR data was obtained in a glasshouse experiment (*glasshouse trial*) in the Netherlands between December and May (winter/spring season). The plants were planted by randomizing genotypes in a designed but unbalanced way across four compartments in replicates of 4, 8 or 16 plants per genotypes. The replicates occurred within and between compartments. The data from INRA were measured in open field cultivation (*open field trial*) between July and August (summer season) in the south of France, in a randomized complete block design with 3 blocks of 3 individual plants (repeats) per genotype and block.

This paper concentrates on the following five traits that were in common in the two experiments:

1. The primary axis length (Axl) defined as the length (in cm) of the primary axis from the cotyledons to the first branch;
2. The number of leaves on the primary axis (Nle);
3. The mean internode length (Inl) given by the ratio Axl:Nle in cm;
4. The axis growth speed (Axs) given by the ratio Axl:(Flw-15 days), in $cm.day^{-1}$, in which Flw is the number of days from sowing to first flower anthesis from which the 15 days corresponding to the time of hypocotyl and cotyledons emergence after sowing were deducted to obtain the growth time of the axis; and
5. The mean internode growth time (Int) given by the ratio (Flw-15 days):Nle, in $day.internode^{-1}$.

The focus of this paper is the analysis of these common traits to discern if the same QTLs underlie identical traits in the two environments and possible interaction between QTL and environment.

Data Evaluation

Each trait was graphically explored for possible variation across blocks and presence of extreme observations (outliers). Further, multivariate analysis of variance (MANOVA) models were fitted to the traits simultaneously across blocks and genotypes. This model allows (a) calculation of trait heritability; (b) quantification of the effect of genotype and/or blocks on the traits and significance testing of these effects and (c) obtaining least square means per genotype after accounting for block and interaction effects. The magnitude and pattern of correlation between traits in each

experiment and across experiments are explored where correlation is expected between the original and derived traits.

Quantitative Trait Locus (QTL) Analysis

QTL detection based on interval mapping (Lander & Botstein, 1989) using the obtained least square means for all traits and the genetic map developed by Barchi et al. (2007), was done with MapQTL software (Van Ooijen, 2004). The significance thresholds for putative QTLs are derived via permutation (10000 runs) of marker genotype and trait phenotype data.

QTL x Environment (E) Interaction Analysis

Putative QTL by environment interactions were studied for the five common traits by considering for each genotype the difference (e.g. Axl_diff) and mean (e.g. Axl_ave) for each trait over the two environments. Identification of QTL for the trait mean would indicate that the QTL is expressed similarly in both environments, i.e., absence of interaction. Identification of QTL for the trait difference would indicate that the QTL is expressed differently, i.e.., presence of interaction. These pairs of derived traits are analyzed using interval mapping, similarly to the original traits. If a QTL is detected either for mean or difference, its effect size and the percentage of the effect size to the parental differences in the two trials are calculated and presented.

Result and discussion

Trait Evaluation

The variation between the three blocks in the *open field trial* (fig. 1) is negligible for all the traits as the difference in trait means across blocks is small. The variation across blocks in the *glasshouse trial* is slightly larger but not significant (fig. 2). Within each block however, there is prominent variation due to the presence of different genotypes, i.e., large genotypic variability. This genotypic variability is more clearly seen in the *glasshouse trial*. There are also indications for very few possible outlying or rather extreme observations. The influence of these outliers was not confirmed yet and they were left in the data. Mean values are comparable between trials, except for Internode length with values lower than 2 cm in *open field trial* and close to 3.5 cm in *glasshouse trial* and axis growth speed with a mean value of around 5 cm/day in *open field trial* and about 10 cm/day in *glasshouse trial*. The range of observations for traits in *glasshouse trial* is generally higher as compared to the same traits in *open field trial*. Some of the traits show very little skewness especially in the *glasshouse trial*.

Within the *open field trial*, the correlation among primary axis length (Axl), number of leaves (Nle) and axis growth speed (Axs) is high and positive (table 1). Internode growth time (Int) is negatively correlated with all other traits except internode length (Inl), with which it is weakly but positively correlated. Internode length (Inl) shows high correlation with axis length (Axl) and axis growth speed (Axs). This same trend is seen in the *glasshouse trial* but with generally lower magnitude. The orientation of measurements for a particular trait in the two trials (e.g. Axl1 and Axl2) coincides as revealed by their correlation coefficients. However, low correlations were observed between the trials for Internode length (Inl) and Axis growth speed (Axs).

The mean trait values for the two parents and estimated trait heritability from the MANOVA model are also listed in table 1. Genotype is consistently significant for all the traits, while block effect is seen in some of the traits especially in *glasshouse trial*, confirming what was observed from the graphical exploration. There is no interaction

between genotype and blocks. The sufficiency of this model to handle the unbalanced settings in the *glasshouse trial* is not guaranteed and the randomness created by genotype and blocks in the two trials deserve to be further explored. Also, the correlation within each trial is not explicitly modeled. The essence of using this model is to obtain least square means of the traits per genotype while accounting for possible block and interaction between genotype and block effects. Heritability is generally higher for traits in the *open field trial* except for axis length. However, our calculated heritability for the *open field trial* is lower than those reported in Barchi et al. (2009). This may be due to a combination of difference in sample size (here we studied a subset of 149 out of the original 297 RILs), the underlying model assumption and the correction for block effects. The parental lines display contrasting phenotypes with parent *Yolo Wonder* showing shorter axis length, fewer leaves, slower axis development but faster leaf development. This is consistent with what has been reported in the literature for these pepper cultivars. The *glasshouse trial* shows consistently higher rates of vegetative trait development, as is also revealed from the box plots (figures. 1 & 2).

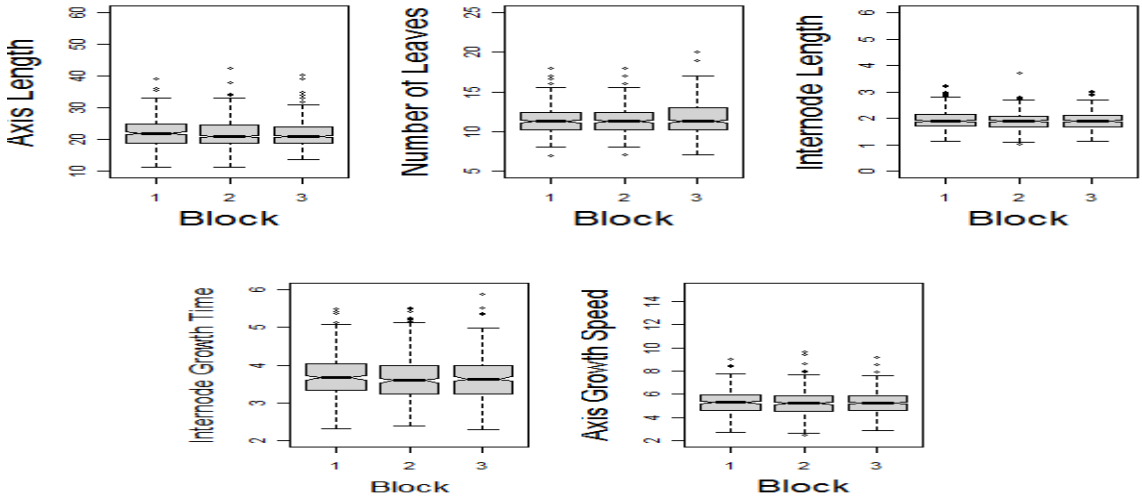

Figure 1. Box plots showing possible trait variation across blocks in the open field trial

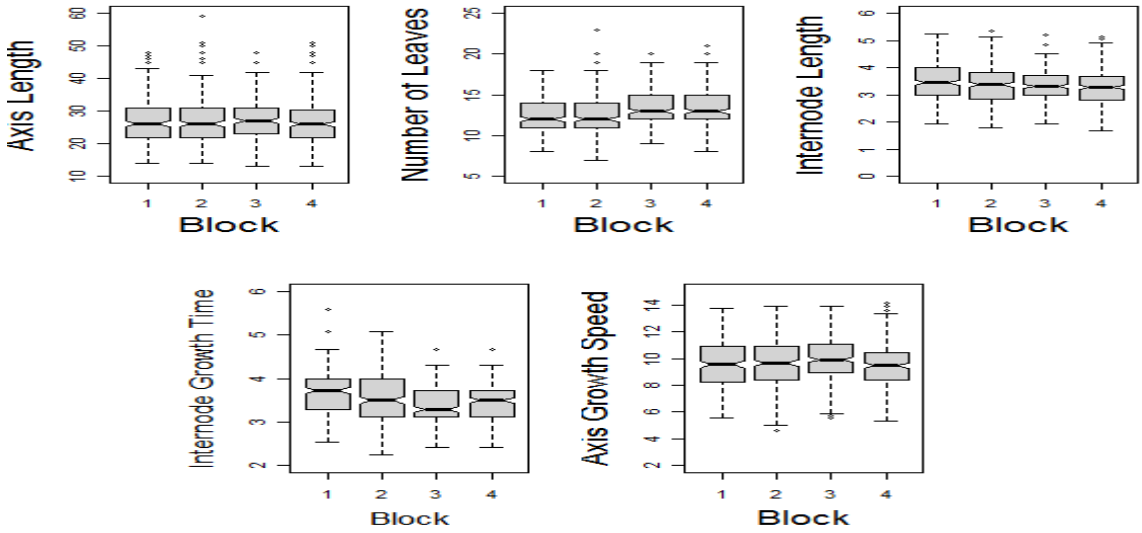

Figure 2. Box plots showing possible trait variation across blocks in the glasshouse trial

Table 1. Correlation coefficients, parent means and heritability for common traits in the two experiments

	Traits[a]	OPEN FIELD				GLASSHOUSE					
		Axl1	Nle1	Inl1	Int1	Axs1	Axl2	Nle2	Inl2	Int2	Axs2
		Correlation Matrix									
OPEN FIELD	Axl1										
	Nle1	0.61									
	Inl1	0.62	-0.23								
	Int1	-0.48	-0.88	0.28							
	Axs1	0.94	0.50	0.66	-0.52						
GLASS HOUS	Axl2	0.64	0.48	0.29	-0.36	0.55					
	Nle2	0.43	0.81	-0.27	-0.69	0.32	0.54				
	Inl2	0.10	-0.33	0.44	0.21	0.21	0.33	-0.47			
	Int2	-0.33	-0.76	0.35	0.74	-0.28	-0.45	-0.93	0.36		
	Axs2	0.35	0.20	0.22	-0.30	0.43	0.66	0.17	0.77	-0.31	
		Parental Means and Trait Heritability									
	Yolo Wonder	18.01	12.12	1.49	3.86	3.93	21.75	11.56	2.53	4.11	6.17
	Criollo de Morelos 334	22.92	12.50	1.85	3.09	6.01	38.75	15.75	3.25	3.06	10.64
Parental Differences		-4.92	-0.38	-0.36	0.77	-2.08	-17	-4.19	-0.72	1.05	-4.46
Heritability		0.78	0.80	0.51	0.62	0.86	0.97	0.19	0.42	0.16	0.94

[a] Axl1, Nle1, Inl1, Int1 and Axs1 stand for primary axis length, number of leaves on the primary axis, mean internode length, mean internode growth time and axis growth speed respectively in the *open field* trial; while Axl2, Nle2, Inl2, Int2 and Axs2 represent primary axis length, number of leaves on the primary axis, mean internode length, mean internode growth time and axis growth speed respectively in the *glasshouse* trial.

QTL Interval Mapping Analysis

The QTL test statistic (LOD score) profiles for significant linkage groups are presented in figure 3. In general, the patterns of the profiles for most linkage groups are consistent among the two trials; however, the magnitude of LOD scores can be different. The latter implies that a QTL may be significant in one trial but insignificant in the other trial. For example, such QTL are found for axis length (Axl) on chromosome 1, number of leaves (Nle) on chromosome 3 and internode growth time (Int) on chromosome 3. These might indicate that some QTLs are better expressed in certain environment though may be detected in various environments. Furthermore, some QTL are detected only in one trial. For example on chromosome 6, QTLs were found for internode length (Inl) and axis speed (Axs) in the *open field trial* but not in the *glasshouse trial*. There is also a possibility of QTLs for axis length (Axl) and axis speed (Axs) on chromosome 12 in the *open field trial*.

QTL x Environment interaction

Several QTL were detected for trait means between the two environments but no significant QTL was detected for trait differences (table 2). The effect sizes of the detected QTL are mostly in the direction of the parental differences in both trials though with varying magnitudes (fig 4). On chromosome 3, there are QTL for means across the two trials for all five vegetative traits. The effect sizes of QTL detected on chromosome 3 and LG 22 for internode length mean vary significantly between the two trials with the effect size greater in the *glasshouse trial*. There is however some QTL for trait means whose effect sizes are in opposite direction of parental differences in both trials. Such QTL for average could be seen for axis length (Axl) and axis speed (Axs) detected on chromosome 3, internode growth time (Int) and number of leaves (Nle) on chromosome 12 and axis speed (Axs) on LG 24.

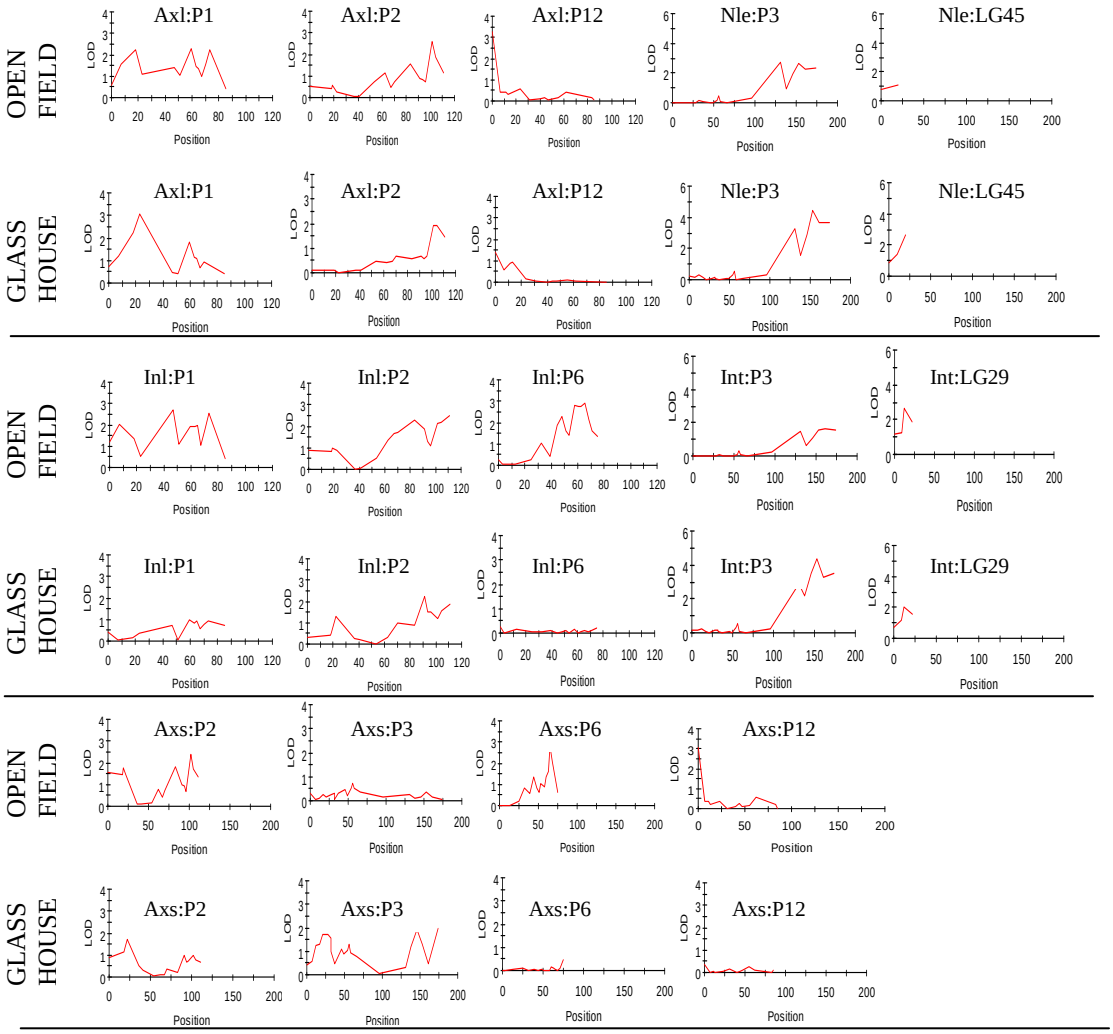

Figure 3: QTL profiles of significant chromosomes (P1, P2 etc) or unassigned linkage groups (LG29, LG45) in both trials. Abbreviated names of traits are explained in section Materials and Methods.

Table 2. Result of the QTL x E Analyses.

Trait	Locus	Position	LOD	Group	QTL Effect Size			95% GW Threshold
					QTLxE	INRA	WUR	
Axl_diff	EPMS_472	174.1	2.41$	P3	1.413	-0.105	-1.518	3
	e36/m52_190y	22.7	2.25$	P1	-1.324	0.645	1.968	
Axl_ave	e41/m48_159y	18.1	2.72$	P1	1.306	0.910	1.702	2.9
Nle_diff	p11/m49_196y	153	2.41$	P3	0.324	-0.407	-0.731	3
	e41/m54_412c	44	2.09$	P12	0.306	0.050	-0.256	
Nle_ave	p11/m49_196y	153	4.06	P3	-0.569	-0.407	-0.731	3.1
	c33/m54_221y	130.5	3.38	P3	-0.529	-0.415	-0.642	
	EPMS_472	174.1	3.38	P3	-0.539	-0.392	-0.687	
Inl_diff	e34/m53_181c	0	2.05$	LG22	0.134	-0.018	-0.152	3
	e31/m58_516y	11.7	1.89$	P3	-0.137	0.020	0.156	
Inl_ave	e44/m51_467c	5.8	3.06	LG28	0.117	0.073	0.161	2.9
	e44/m51_258c	91.1	2.78	P2	-0.109	-0.061	-0.157	
Int_diff	e38/m61_158y	111.5	2.25$	P4a	0.081	0.062	-0.019	3
	e41/m54_412c	44	1.89$	P12	-0.073	-0.016	0.056	
Int_ave	p11/m49_196y	153	3.21	P3	0.119	0.090	0.149	2.9
	EPMS_472	174.1	2.84	P3	0.116	0.094	0.139	
Axs_diff	EPMS_472	174.1	2.79$	P3	0.432	0.025	-0.407	3
	p11/m49_197y	18.7	2.15$	LG24	0.374	0.095	-0.279	
Axs_ave	p11/m49_343c	22.2	2.18$	P2	0.265	0.162	0.368	3

$ No significant QTLs found for these traits but the QTLs with the highest LOD scores are reported. Abbreviated names of traits are explained in section Materials and Methods.

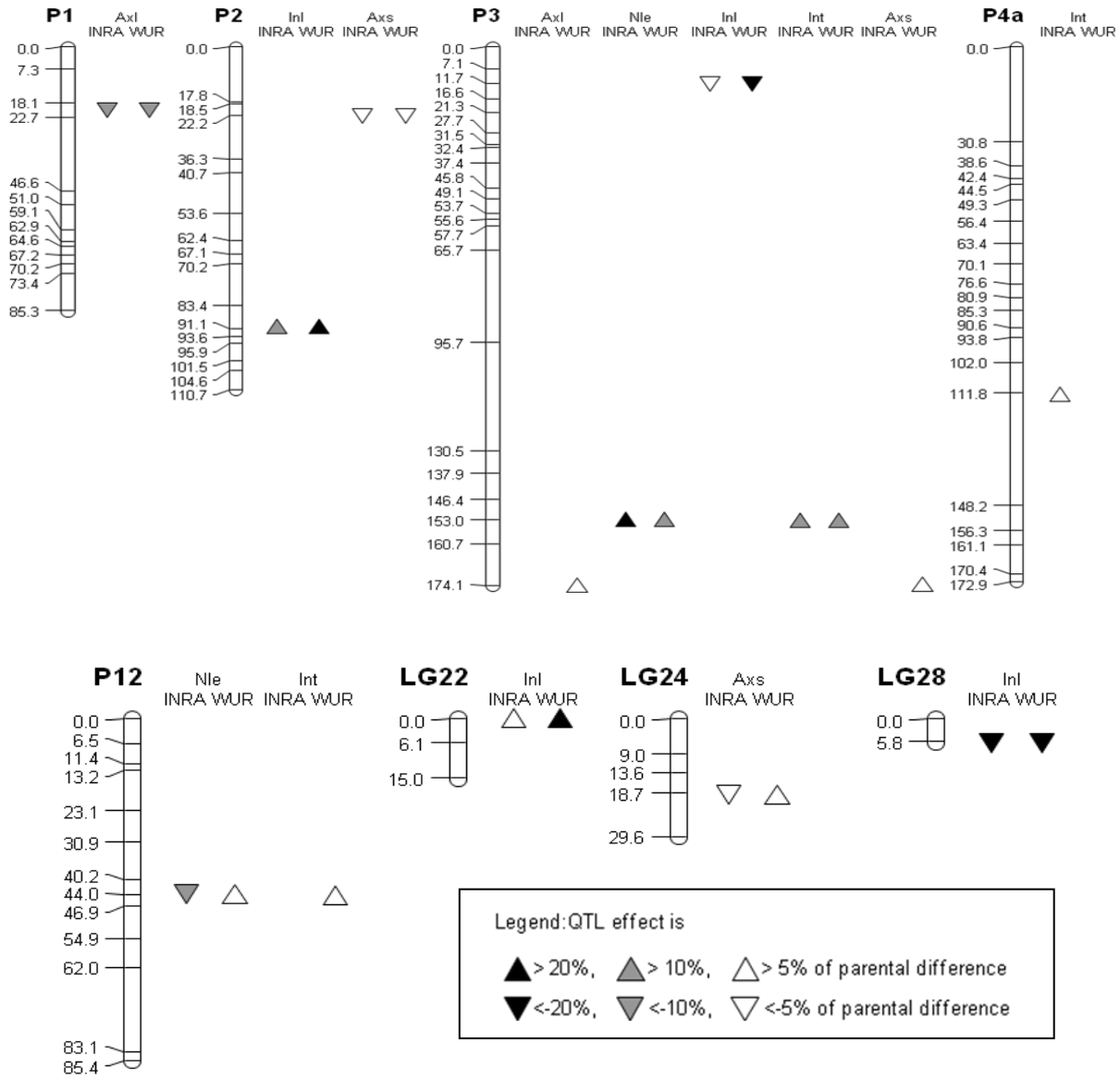

Figure 4. Charts showing positions on the chromosome or LG of QTLs with highest LOD scores for the traits considered in the QTL x E Analysis. Traits abbreviations are as discussed in methods section. INRA and WUR represent *open field* and *glasshouse* trials respectively.

Concluding Remarks

The vegetative development of pepper plant is more pronounced in the *glasshouse trial* than in the *open field trial*. The *glasshouse trial* showed higher length of internodes and faster rate of stem length development with more conspicuous genotypic variability indicating stronger parental differences or segregation. This is further confirmed from the parental means for each trait in both trials. Though parental differences exist for all traits in both trials, the magnitudes of these differences are much larger in the *glasshouse trial*. This resulted from a rather stable growth of '*Yolo Wonder*' in both environments but an environment dependent response of '*CM334*' which displayed a higher increase of vegetative growth in the winter glasshouse trial. Higher trait heritability seen in the *open field trial* could be linked to the higher block effect accounted for in the *glasshouse trial*.

About 17 putative QTL were detected for all traits in the two trials, 3 for axis length; 3 for number of leaves; 4 for internode length; 3 for internode growth time and 4 for axis speed. The test statistics scores for the significance of these QTL are generally low. Similar levels of low LOD scores were reported by Barchi et al. (2009) while analyzing two subpopulations (141 and 93 RILs) of the whole 297 genotypes in the INRA *open field trial*. They noted that LOD scores associated to detected QTL are usually much lower in the reduced sub populations than in the full RIL population, and only the QTL with the highest LOD scores remained significant. This is an indication that some QTL may not be detected in our analysis due to the size of the current dataset, giving room for possible false-negative QTL. It is known that the power to detect QTL increases as the population size is maximized (Charcosset and Gallais 1996) and the precision depends on the adopted sampling methods which can be random or based on selective genotyping/phenotyping. However, most often population size cannot be increased easily due to the large costs of phenotyping experiments.

Most of the 17 QTL are found in both trials but with different level of expression. Breeders know that most of the vegetative traits such as axis length and number of leaves, though genetically determined in constant environment, are strongly affected by environments. The detected QTL for axis length on chromosome 1, number of leaves on chromosome 3, internode growth time on chromosome 3 and axis speed also on chromosome 3 are better revealed in the *glasshouse trial*, while those detected for axis length on chromosome 2, internode length on chromosome 1 and 2 and axis speed on chromosome 2 are better expressed in the *open field trial*. A few of the QTL such as the one for axis growth speed on chromosome 6 and 12 were only expressed in one trial.

It was observed that co-localization occurs for many of these QTL i.e. most of the detected QTL affect more than a single trait. Axis length, internode length and axis growth speed are all affected by the same QTL on chromosome 2. On chromosome 3, number of leaves, internode growth time and axis growth speed are influenced by the same QTL; axis growth speed and internode length on chromosome 6, and axis length and axis growth speed on chromosome 12. This co-localization of trait QTL is in agreement with the established correlations between these traits. This may be an indication for linkage and/or pleiotropic effects of genes on the morphology (internode length, number of leaves) or growth speed of vegetative organs. Such linkage or pleiotropic effects can be more accurately studied by explicit modeling of the correlation mechanism and causal relationship among the traits. We will explore Bayesian QTL mapping approaches (such as Yandell et al. 2007 and Bink et al. 2008) that allow flexible models and also inclusion of prior knowledge on model parameters.

The result from our simple QTL by environment analysis does not reveal any significant QTL masked by environmental interaction since no QTL was detected for trait difference between the two environments. This result cannot be generalized yet as the number of environments considered is small and the sufficiency of the analysis is not guaranteed. Within the EU-SPICY project, phenotypic data on the same RIL population of 149 genotypes are being collected under 4 environments covering different seasons (winter and summer) and different geographical locations (Temperate and Mediterranean). A range of plant and fruit traits are being recorded and evaluated in these trials. Our model should incorporate analysis of these complex traits across a range of environmental conditions, considering the interaction between genotype and environment while accounting for the different developmental stages (time) for a given trait. We anticipate that the integration of QTL models and eco-

physiological models (Van Eeuwijk et al., 2010) to predict these complex traits in terms of their underlying QTLs will contribute to the genetic improvement of important crops across a range of environments.

Acknowledgements

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 211347.

References

Barchi et al. 2007. A high-resolution intraspecific linkage map of pepper (Capsicum annuum L.) and selection of reduced RIL subsets for fast mapping. Genome 50:51-60.

Barchi et al. 2009. QTL analysis of plant development and fruit traits in pepper and performance of selective phenotyping. Theor Appl Genet 118:1157-1171.

Bink, M.C.A.M.; Boer, M.P.; ter Braak, C.J.F.; Jansen, J.; Voorrips, R.E.; van de Weg, W.E. 2008. Bayesian analysis of complex traits in pedigreed plant populations. Euphytica 161:85–96. DOI: 10.1007/s10681-007-9516-1.

Brown, D.; Vision, T. 2000. MapPop 1.0: Software for selective mapping and bin mapping. http://www.bio.unc.edu/faculty/vision/lab/mappop/

Charcosset, A.; Gallais, A. 1996. Estimation of the contribution of quantitative trait loci (QTL) to the variance of a quantitative trait by means of genetic markers. Theor Appl Genet 93:1193-1201.

Lander, E.S.; Botstein, D. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185-199.

Nicolaï, M.; Sage-Palloix, A.M.; Nemouchi, G.; Savio, B.; Lefebvre, V.; Vuylsteke, M.; Palloix, A. 2010. Providing genomic tools to increase the efficiency of molecular breeding for complex traits in pepper. (this issue).

van Eeuwijk, F.A.; Bink, M.C.A.M.; Chenu, K.; Chapman, S.C. 2010. Detection and use of QTL for complex traits in multiple environments. Current Opinion in Plant Biology (online).

Van Ooijen, 2004. MapQTL® 5, Software for the mapping of quantitative trait loci in experimental populations. Kyazma B.V., Wageningen, Netherlands.

Voorrips, R.E.; Palloix, A.; Dieleman, A.; Bink, M.; Heuvelink, E.; van der Heijden, G.; Vuylsteke, M.; Glasbey, C.; Barócsi, A.; Magán, J.; van Eeuwijk, F. 2010. Crop Growth models for the –omics era: the EU-SPICY project. (this issue).

Yandell, B.S.; Mehta, T.; Banerjee, S.; Shriner, D.; Venkataraman, R.; Moon, J.Y.; Neely, W.W.; Wu, H.; von Smith, R.; Yi, N. 2007. R/qtlbim: QTL with Bayesian Interval Mapping in experimental crosses. Bioinformatics 23:641-643. DOI: 10.1093/bioinformatics/btm011.