The Genetic Architecture of Gene Expression in *Caenorhabditis elegans*

Ana Viñuela Rodríguez

Thesis committee

Thesis supervisor:
Prof. dr. ir. J. Bakker
Professor of Nematology
Wageningen University

Thesis co-supervisor:
Dr. ir. J.E. Kammenga
Associate professor at the Laboratory of Nematology
Wageningen University

Other members:
Dr. M.A. Herman, Kansas State University, Kansas, USA
Prof. dr. R.C. Jansen, University of Groningen, Groningen, The Netherlands
Prof. dr. D.J. de Koning, Swedish University of Agricultural Sciences, Uppsala, Sweden
Prof. dr. B.J. Zwaan, Wageningen University, Wageningen, The Netherlands

This research was conducted under the auspices of the Graduate School PE&RC

The Genetic Architecture of Gene Expression in *Caenorhabditis elegans*

Ana Viñuela Rodríguez

Thesis
submittted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 4 March 2011
at 11 a.m. in the Aula.

Viñuela Rodríguez, A. (2011)

The Genetics Architecture of Gene Expression in *Caenorhabditis elegans*

PhD Thesis Wageningen University, Wageningen, the Netherlands, with references, with summary in English, Spanish and Dutch.

ISBN: 978-90-8585-838-6

- 'But I don't want to go among mad people,' Alice remarked.
- 'Oh, you can't help that,' said the Cat: 'we're all mad here. I'm mad. You're mad.'
- 'How do you know I'm mad?' said Alice.
- 'You must be,' said the Cat, 'or you wouldn't have come here.'

Alice's adventures in Wonderland by Lewis Carroll

Wanderer, there is no road; you make your path as you walk

Antonio Machado (1875-1936)

Contents

Chapter: 1

General introduction

Ana Viñuela

From genome to phenome

The genetic architecture of any trait includes the genetic elements and their effects that contribute to the phenotype and the phenotypic variation (Le Rouzic and Carlborg, 2008). For a given trait the architecture consists of genes, alleles and their dominance, epistasis, or pleitropy interactions (Mackay, 2001; Hansen, 2006). But the process that translates a genotype into a phenotype involves many other levels, from expressed genes through proteins, metabolites and the many levels of interactions between these components that lead to the produced phenotype. Because of this, the whole process is usually described as an intricate web of the many influential factors acting at each level, whose structure defines the observed phenotype (Bard, 2010). However, on a basic level of organization, phenotypes and their variation may be explained by variation in gene expression. To better understand the genetic architecture of any given trait we therefore need to unravel the genetic architecture of gene expression.

The genetic architecture of gene expression is complex. The elements and the effects that contribute to the expression of a gene and the variation in gene expression are multiple and diverse. In other words, gene expression is a phenotypic trait that displays variation. One advantage of the gene expression phenotypes are associations with their location in the genome itself (Romanoski *et al.*, 2010). The number of influential variables affecting gene expression may be enormous, but the number of interactions and effects is limited. In this thesis I exploit this advantage, where the architecture of gene expression was explored considering some of the elements that contribute to gene expression as a phenotype and its variation.

Gene expression in Caenorhabditis elegans

Since Sydney Brenner introduced *Caenorhabditis elegans* (Nematoda, Rhabditidae) in research (Brenner, 1974), this little worm has become a model organism in genetics and genomics. It is easy, fast and cheap to grow; and it can produce an entire new generation of hundreds of individuals in less than three days. In addition, laboratory protocols are standardized all around the world and information is made publicly available online by *the worm* community members (www.wormbook.org, www.wormbase.org). An additional

advantage of *C. elegans* as a model organism for genomics is the ability of self-fertilization. Yet, occasionally occurring males allow to sexually cross different strains (Gutteling *et al.*, 2007) or even closely related species for more complex genomic studies (Woodruff *et al.*, 2010). This, together with its small genome has been an advantage for studies in evolution, cancer, toxicology, and ecology (Saito and Van den Heuvel, 2002; Kammenga *et al.*, 2007; Leung *et al.*, 2008; Cutter *et al.*, 2009). Furthermore, the full sequence of the *C. elegans* genome which became available in 1998 (The C. elegans Sequencing Consortium, 1998) was a starting point for a prolific work in gene expression. To illustrate this point it is enough to consider the approx. 400 papers published per year with the keywords *C.elegans* and *gene expression* over the last ten years (www.ncbi.nlm.nih.gov/pubmed). Overall, *C. elegans* has become an excellent tool for genomic research and for unraveling the structure and dynamics of eukaryotes genomes.

Genotype-by-environment interactions

Phenotypic responses to environmental changes are genetically dependent. Therefore, gene expression responses to modifications in the environment depend on genetic elements and the way they interact with the environment. The estimation of the genetic and environmental contribution to a phenotype is a necessary step to understand how a genotype is translated to a phenotype. However, both factors are not necessarily additive (Lewontin, 2006). They may interact and produce a differential phenotypic response between genotypes exposed to identical environments. This phenomenon is known as genotype-by-environment interaction (GxE).

With the development of high throughput technologies (e.g.: microarrays, proteomics, metabolomics, etc) many studies have addressed the contribution of the genome, the environment and their interaction to the resultant phenotype. Gene expression studies have been particularly useful to quantify the influence of genotypic and environmental factors (Gibson, 2008). The influence of the environmental factors is often measured as variation in gene expression in genetically identical individuals exposed to different environments (Sørensen *et al.*, 2005; Cui *et al.*, 2007; Nota *et al.*, 2010); while the influence of the genotype is measured as variation in gene expression in genetically different individuals in the same environments (Brem and Kruglyak, 2005; Keurentjes *et al.*, 2007; Capra *et al.*,

2008). More complex studies have been able to address GxE interactions in gene expression by combining limited number of genotypes exposed to different environments (Li *et al.*, 2006; Smith and Kruglyak, 2008). Yet, we are far from fully understanding the influence of the environment on gene expression changes (Gibson, 2008; Hodgins-Davis and Townsend, 2009).

Perturbations affecting gene expression

A common method to study gene expression is by means of perturbation. The responses to environmental perturbations are usually characterized in expression profiles, networks or responsive gene pathways. Perturbations help to uncover hidden phenotypic variability and specific elements (e.g.: allelic variants) involved in the response to different environments. This technique has been particularly successful with single gene perturbations (Murphy *et al.*, 2003), where the lack of a gene activity (e.g.: *daf-2*) is associated to a phenotype (extended lifespan), and the study of some of the elements involved (a gene pathway) allowed to partially describe the gene expression pattern responsible for the phenotype. This approach often attributes a single gene as the ultimate cause of a phenotype, which ignores the many intermediate levels involved in the generation of the phenotype. Still, those studies help to establish the relative importance of genetic elements, since the absence of just one gene can produce extreme phenotypes such as lethality, extended life span or lack of offspring.

Single gene approaches; however are oversimplifications of nature. Organisms in their natural environments are confronted by severe and multiple stresses. We can study their effects on gene expression by employing multiple perturbations instead of one at the time (Jansen, 2003; Rockman, 2008; Holmstrup *et al.*, 2010). These perturbations or stressors may be introduced sequentially or simultaneously. With sequential perturbation the individual effect of each stressor and the interacting effects of all of them can be distinguished. With the exposure to simultaneous perturbations, non-obvious connections between the components influencing the phenotype may be unraveled, since we are exploring a much larger space of responses and elements (Rockman, 2008).

This thesis explores the transcriptional responses to multiple environmental stressors as a tool to unravel the genetic architecture of gene expression.

Environmental stressors with specific molecular targets

Pesticides

Over their lifetime, organisms are often faced with multiple environmental stresses. Soil related animals such as *C. elegans* are occasionally exposed to anthropogenic stressor, such as pesticides and other noxious chemicals. Those chemicals are designed to target a range of species via specific molecular receptors. One of the most widely employed family of pesticides in household and pest control are Organophosphorus Pesticides (OP). They were designed to target the enzyme acetylcholinesterase (AChE) and inhibit its activity. This inhibition leads to a cholinergic hyper stimulation affecting the nervous system of the pest organism (Pope, 1999). However, studies in human, rats and worms have shown crucial differences in the mode of action of different OP at enzymatic level (Slotkin and Seidler, 2007; Slotkin *et al.*, 2008; Lewis *et al.*, 2009). The differences suggest that their secondary targets may be more relevant for the toxic outcome than expected in previous studies, what is especially relevant considering how often organisms are exposed to combined toxicants. Also, it highlights the relevance of studies with mixtures of pesticides because most of the studies are based in single treatment experiments and little is known about toxicants interactions. Moreover, the few studies that have investigated combined toxicants treatments conclude that exposure to mixtures of pesticides induce a specific gene transcription response in the mixtures compared with the single treatments (Mumtaz *et al.*, 2002; Hook *et al.*, 2008).

Environmental stressors with unspecific molecular targets

Temperature

Temperature affects almost all biological processes. This is particularly evident in ectothermic organisms such as *C. elegans*, which cannot regulate their body temperature. Changes in temperature do not target a specific molecular receptor but modifies the whole metabolism (Gillooly *et al.*, 2001). As a consequence, temperature can change the gene responses to other environmental cues. For instance, temperature modifies the effects to

environmental toxicants exposure (Leon, 2008; Holmstrup *et al.*, 2010). Increased temperature elevates the toxic consequences of OP exposure and other toxicants (Osterauer and Köhler, 2008; De Silva *et al.*, 2009). This increase in toxicity is usually attributed to an elevated uptake and accumulation of the toxicants in the organisms, but it does not explain how such a clear connection between both environmental stressors was not found for zebrafish exposed to pesticides (Scheil and Köhler, 2009; Holmstrup *et al.*, 2010). Moreover, temperature also influences *C. elegans* life history traits as lifespan or offspring, induces temperature specific proteins and regulates gene expression (Klass, 1977; Mádi *et al.*, 2003; Li *et al.*, 2006; Gutteling *et al.*, 2007). However, some genetically determined processes have temperature independent responses, which suggest a compensation mechanism to environmental changes (Kammenga *et al.*, 2007). In other words, the ability to respond to temperature changes differs between genotypes and it is modulated by GxE interactions.

Other factors affecting gene expression

Age

Organisms age since they were born and the rate of aging is determined by the ability of an organism to survive. Aging is not a regulated process, while survival is. Thus, it is important to distinguish between aging and survival (Lithgow, 2006). How well (or poorly) survival is regulated determines the rate of aging, and defines the lifespan of an organism. Survival is a genetically determined process modulated by GxE interactions, which influences how an organism ages. Likewise, the genetic basis of plasticity in aging lie in the gene expression responses to environmental cues (Flatt and Schmidt, 2009). For instance, Coolon *et al*. (2009) analyzed transcription profiles of *C. elegans* in four bacterial growth media to identify genes that function in environmental interactions. They found a correlation between differences in selected genes transcription profiles and differences on survival in different bacterial environments (Coolon *et al.*, 2009). Moreover, gene expression changes with age (Hill *et al.*, 2000; Jin *et al.*, 2001; McCarroll *et al.*, 2004). Also, the influence of environmental factors on gene expression differs with age, as many

stress responses are attenuated with age like heat shock, oxidative stress or responses to toxicants (Darr and Fridovich, 1995; Olsen *et al.*, 2006; Przybysz *et al.*, 2009). In order to understand the relationship of the variables affecting gene expression we need to consider the age of the organism, since the interactions between the different factors are not the same at different ages.

Genotype

The genotype is a source of multiple and simultaneous perturbations. Genotypic perturbations can be introduced with genetic crosses to produce a segregating population of recombinant individuals with different combinations of allelic variants (Jansen, 2003). The gene expression variation between segregants may be used in the analysis of quantitative trait loci (QTL) to identify the causal genomic regions of this phenotypic variation in a method called *genetical genomics* (Jansen and Nap, 2001). An important advantage of this method is the possibility to distinguish between local and distant regulation of gene expression. The QTL for expression differences (eQTLs) identify genomic regions with regulatory elements influencing the expression of a gene. If the eQTL maps to the physical position of the gene, it is inferred to be involved in local regulation, also called *cis*-eQTL. If the eQTL maps to a different genomic region, it is likely a distant regulator, and called *trans*-eQTL. Genetical genomics has been successfully used to describe the genetics of gene transcription regulation in many model organisms (Brem *et al.*, 2002; Schadt *et al.*, 2003; Li *et al.*, 2006; Keurentjes *et al.*, 2007). Also, it has provided information about the nature of gene expression regulation in different environments, tissues and cell types (Li *et al.*, 2006; Petretto *et al.*, 2006; Gerrits *et al.*, 2009).

In Chapter 4, we used a recombinant inbred line (RIL) population generated from a cross between the *C. elegans* canonical strain N2 and the natural isolate CB4856 in a genetical genomic experiment. The two strains are among the most diverse natural isolates of *C. elegans* (Denver *et al.*, 2003; Rockman and Kruglyak, 2009). This diversity is well documented and included large deletions relative to N2 in chromosome *II*, high SNPs density, divergences in life history traits, gene expression and genotype-by-environment interactions in complex traits (Wicks *et al.*, 2001; Li *et al.*, 2006; Gutteling *et al.*, 2007;

Gutteling *et al.*, 2007; Maydan *et al.*, 2007; Capra *et al.*, 2008; Palopoli *et al.*, 2008). The study of natural variation, especially in gene expression, is an expanding field in *C. elegans* and in combination with other gene expression studies it will help to reveal the genetic architecture of gene expression in *C. elegans*.

Outline of the thesis

First, in **Chapter 2** we describe how a single genotype, the canonical strain N2, responds to environmental stressors with a specific mode of action (pesticides) by studying their effects on gene expression. The chapter provides a detailed insight into the transcriptional responses following exposure to two widely used organophosphorus pesticides, chlorpyrifos and diazinon. We produced genome-wide transcription profiles of *C. elegans* exposed to both pesticides. Furthermore, we studied the effect of interacting pesticides by exposing nematodes to a mixture of them. **Chapter 3** incorporates the study of temperature as a non-target environmental factor influencing gene expression. Expression profiles from different temperatures and pesticide exposures are combined to enlighten the complexity of gene-by-environment interactions.

Second, in **Chapter 4** we elaborate on the influence of different genotypes as multiple perturbations on gene expression using genetical genomics. Simultaneously, we investigate how the genotype-phenotype relationship progresses with age. Recombinant inbred lines were employed to explore the dynamics of regulatory loci affecting genome-wide gene expression. We performed a genetic mapping strategy of gene transcription variation (eQTL) at three different ages, juvenile, reproductive and old ages. Also, we investigated the influence of age in gene expression to reveal the presence of genes with a genotype-by-age effect ($_{gxa}$eQTL). Moreover, we explored variation in gene expression with age and its implication for the aging process. **Chapter 5** addresses questions about transcriptional consequences of genomic recombination in aging *C. elegans*. Gene expression profiles of the recombinant inbred lines parental strains (N2 and CB4856) were generated. Then, we explored gene expression heritability and transgression as genetic parameters for the analysis of gene expression divergence in different genotypes. Moreover, we investigated the effect of age over those parameters.

In the final **Chapter 6,** I present a general discussion on the genetic architecture of gene expression in *C. elegans*. Also, I discuss the transcriptional responses to perturbation within the scope of the robustness theory to clarify how the genetic architecture of gene expression affects the genotype-phenotype map in *C. elegans*.

References

Bard, J. (2010). A systems biology view of evolutionary genetics. *BioEssays* **32**(7): 559-563.

Brem, R. B. and L. Kruglyak (2005). The landscape of genetic complexity across 5,700 gene expression traits in yeast. *Proceedings of the National Academy of Sciences of the United States of America* **102**(5): 1572-1577.

Brem, R. B., G. Yvert, et al. (2002). Genetic Dissection of Transcriptional Regulation in Budding Yeast. *Science* **296**(5568): 752-755.

Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* **77**(1): 71-94.

Capra, E. J., S. M. Skrovanek, et al. (2008). Comparative Developmental Expression Profiling of Two *C. elegans* Isolates. *PLoS ONE* **3**(12): e4055.

Coolon, J. D., K. L. Jones, et al. (2009). *Caenorhabditis elegans* Genomic Response to Soil Bacteria Predicts Environment-Specific Genetic Effects on Life History Traits. *PLoS Genet* **5**(6): e1000503.

Cui, Y., S. McBride, et al. (2007). Toxicogenomic analysis of *Caenorhabditis elegans* reveals novel genes and pathways involved in the resistance to cadmium toxicity. *Genome Biology* **8**(6): R122.

Cutter, A. D., A. Dey, et al. (2009). Evolution of the Caenorhabditis elegans Genome. *Mol Biol Evol* **26**(6): 1199-1234.

Darr, D. and I. Fridovich (1995). Adaptation to oxidative stress in young, but not in mature or old, caenorhabditis elegans. *Free Radical Biology and Medicine* **18**(2): 195-201.

De Silva, P. M. C. S., A. Pathiratne, et al. (2009). Influence of temperature and soil type on the toxicity of three pesticides to Eisenia andrei. *Chemosphere* **76**(10): 1410-1415.

Denver, D. R., K. Morris, et al. (2003). Phylogenetics in Caenorhabditis elegans: An Analysis of Divergence and Outcrossing. *Molecular Biology and Evolution* **20**(3): 393-400.

Flatt, T. and P. S. Schmidt (2009). Integrating evolutionary and molecular genetics of aging. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1790**(10): 951-962.

Gerrits, A., Y. Li, et al. (2009). Expression Quantitative Trait Loci Are Highly Sensitive to Cellular Differentiation State. *PLoS Genet* **5**(10): e1000692.

Gibson, G. (2008). The environmental contribution to gene expression profiles. *Nat Rev Genet* **9**(8): 575-581.

Gillooly, J. F., J. H. Brown, et al. (2001). Effects of Size and Temperature on Metabolic Rate. *Science* **293**(5538): 2248-2251.

Gutteling, E. W., A. Doroszuk, et al. (2007). Environmental influence on the genetic correlations between life-history traits in Caenorhabditis elegans. *Heredity* **98**(4): 206-213.

Gutteling, E. W., J. A. G. Riksen, et al. (2007). Mapping phenotypic plasticity and genotype-environment interactions affecting life-history traits in Caenorhabditis elegans. *Heredity* **98**(1): 28-37.

Hansen, T. F. (2006). The Evolution of Genetic Architecture. *Annual Review of Ecology, Evolution, and Systematics* **37**(1): 123-157.

Hill, A. A., C. P. Hunter, et al. (2000). Genomic Analysis of Gene Expression in C. elegans. *Science* **290**(5492): 809-812.

Hodgins-Davis, A. and J. P. Townsend (2009). Evolving gene expression: from G to E to G × E. *Trends in Ecology & Evolution* **24**(12): 649-658.

Holmstrup, M., A.-M. Bindesbøl, et al. (2010). Interactions between effects of environmental chemicals and natural stressors: A review. *Science of The Total Environment* **408**(18): 3746-3762.

Hook, S. E., A. D. Skillman, et al. (2008). Gene Expression Profiles in Rainbow Trout, Onchorynchus mykiss, Exposed to a Simple Chemical Mixture. *Toxicol. Sci.* **102**(1): 42-60.

Jansen, R. C. (2003). Studying complex biological systems using multifactorial perturbation. *Nat Rev Genet* **4**(2): 145-151.

Jansen, R. C. and J.-P. Nap (2001). Genetical genomics: the added value from segregation. *Trends in Genetics* **17**(7): 388-391.

Jin, W., R. M. Riley, et al. (2001). The contributions of sex, genotype and age to transcriptional variance in Drosophila melanogaster. *Nat Genet* **29**(4): 389-395.

Kammenga, J. E., A. Doroszuk, et al. (2007). A Caenorhabditis elegans wild type defies the temperature-size rule owing to a single nucleotide polymorphism in tra-3. *Plos Genetics* **3**(3): 358-366.

Kammenga, J. E., M. A. Herman, et al. (2007). Microarray challenges in ecology. *Trends in Ecology & Evolution* **22**(5): 273-279.

Keurentjes, J. J. B., J. Y. Fu, et al. (2007). Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. *Proceedings of the National Academy of Sciences of the United States of America* **104**(5): 1708-1713.

Klass, M. R. (1977). Aging in the nematode Caenorhabditis elegans: Major biological and environmental factors influencing life span. *Mechanisms of Ageing and Development* **6**: 413-429.

Le Rouzic, A. and Ö. Carlborg (2008). Evolutionary potential of hidden genetic variation. **23**(1): 33-37.

Leon, L. R. (2008). Thermoregulatory responses to environmental toxicants: The interaction of thermal stress and toxicant exposure. *Toxicology and Applied Pharmacology* **233**(1): 146-161.

Leung, M. C. K., P. L. Williams, et al. (2008). Caenorhabditis elegans: An Emerging Model in Biomedical and Environmental Toxicology. *Toxicol. Sci.* **106**(1): 5-28.

Lewis, J., M. Szilagyi, et al. (2009). Distinct patterns of gene and protein expression elicited by organophosphorus pesticides in *Caenorhabditis elegans*. *BMC Genomics* **10**(1): 202.

Lewontin, R. C. (2006). The analysis of variance and the analysis of causes. *Int. J. Epidemiol.* **35**(3): 520-525.

Li, Y., O. A. A. Alvarez, et al. (2006). Mapping determinants of gene expression plasticity by genetical genomics in C-elegans. *Plos Genetics* **2**(12): 2155-2161.

Lithgow, G. J. (2006). Why aging isn't regulated: A lamentation on the use of language in aging literature. *Experimental Gerontology* **41**(10): 890-893.

Mackay, T. F. C. (2001). Quantitative trait loci in Drosophila. *Nat Rev Genet* **2**(1): 11-20.

Mádi, A., S. Mikkat, et al. (2003). Mass spectrometric proteome analysis for profiling temperature-dependent changes of protein expression in wild-type *Caenorhabditis elegans*. *PROTEOMICS* **3**(8): 1526-1534.

Maydan, J. S., S. Flibotte, et al. (2007). Efficient high-resolution deletion discovery in Caenorhabditis elegans by array comparative genomic hybridization. *Genome Research* **17**(3): 337-347.

McCarroll, S. A., C. T. Murphy, et al. (2004). Comparing genomic expression patterns across species identifies shared transcriptional profile in aging. *Nat Genet* **36**: 197 - 204.

Mumtaz, M. M., D. B. Tully, et al. (2002). Gene induction studies and toxicity of chemical mixtures. *Environmental Health Perspectives* **110**(SUPPL. 6): 947-956.

Murphy, C. T., S. A. McCarroll, et al. (2003). Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. *Nature* **424**(6946): 277-283.

Nota, B., N. M. Van Straalen, et al. (2010). Gene expression microarray analysis of heat stress in the soil invertebrate Folsomia candida. *Insect Molecular Biology* **19**(3): 315-322.

Olsen, A., M. Vantipalli, et al. (2006). Lifespan extension of *Caenorhabditis elegans* following repeated mild hormetic heat treatments. *Biogerontology* **7**(4): 221-230.

Osterauer, R. and H.-R. Köhler (2008). Temperature-dependent effects of the pesticides thiacloprid and diazinon on the embryonic development of zebrafish (Danio rerio). *Aquatic Toxicology* **86**(4): 485-494.

Palopoli, M. F., M. V. Rockman, et al. (2008). Molecular basis of the copulatory plug polymorphism in Caenorhabditis elegans. *Nature* **454**(7207): 1019-1022.

Petretto, E., J. Mangion, et al. (2006). Heritability and Tissue Specificity of Expression Quantitative Trait Loci. *PLoS Genet* **2**(10): e172.

Pope, C. N. (1999). Organophosphorus Pesticides: They all have the same mechanism of toxicity? *Journal of Toxicology and Environmental Health Part B: Critical Reviews* **2**: 161-181.

Przybysz, A. J., K. P. Choe, et al. (2009). Increased age reduces DAF-16 and SKN-1 signaling and the hormetic response of Caenorhabditis elegans to the xenobiotic juglone. *Mechanisms of Ageing and Development* **130**(6): 357-369.

Rockman, M. V. (2008). Reverse engineering the genotype-phenotype map with natural genetic variation. *Nature* **456**(7223): 738-744.

Rockman, M. V. and L. Kruglyak (2009). Recombinational Landscape and Population Genomics of *Caenorhabditis elegans*. *PLoS Genet* **5**(3): e1000419.

Romanoski, C. E., S. Lee, et al. (2010). Systems Genetics Analysis of Gene-by-Environment Interactions in Human Cells. **86**(3): 399-410.

Saito, R. M. and S. Van den Heuvel (2002). Malignant worms: What cancer research can learn from C. elegans. *Cancer Investigation* **20**(2): 264-275.

Schadt, E. E., S. A. Monks, et al. (2003). Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422**(6929): 297-302.

Scheil, V. and H.-R. Köhler (2009). Influence of Nickel Chloride, Chlorpyrifos, and Imidacloprid in Combination with Different Temperatures on the Embryogenesis of the Zebrafish <i>Danio rerio</i&gt. *Archives of Environmental Contamination and Toxicology* **56**(2): 238-243.

Slotkin, T. A. and F. J. Seidler (2007). Comparative developmental neurotoxicity of organophosphates in vivo: Transcriptional responses of pathways for brain cell development, cell signaling, cytotoxicity and neurotransmitter systems. *Brain Research Bulletin* **72**(4-6): 232-274.

Slotkin, T. A., F. J. Seidler, et al. (2008). Targeting of neurotrophic factors, their receptors, and signaling pathways in the developmental neurotoxicity of organophosphates in vivo and in vitro. *Brain Research Bulletin* **76**(4): 424-438.

Smith, E. N. and L. Kruglyak (2008). Gene-Environment Interaction in Yeast Gene Expression. *PLoS Biol* **6**(4): e83.

Sørensen, J. G., M. M. Nielsen, et al. (2005). Full genome gene expression analysis of the heat stress response in *Drosophila melanogaster*. *Cell Stress and Chaperones* **10**(4): 312-328.

The C. elegans Sequencing Consortium (1998). Genome Sequence of the Nematode *C.elegans*: A Platform for Investigating Biology. *Science* **282**(5396): 2012-2018.

Wicks, S. R., R. T. Yeh, et al. (2001). Rapid gene mapping in Caenorhabditis elegans using a high density polymorphism map. *Nature Genetics* **28**(2): 160-164.

Woodruff, G. C., O. Eke, et al. (2010). Insights Into Species Divergence and the Evolution of Hermaphroditism from Fertile Interspecies Hybrids of Caenorhabditis Nematodes. *Genetics*: genetics.110.120550.

Chapter: 2

Transcriptional responses to environmental factors with specific mode of action

Ana Viñuela, L. Basten Snoek, Joost A.G. Riksen, Jan E. Kammenga

Published as:

Ana Viñuela, L. Basten Snoek, Joost A.G. Riksen, Jan E. Kammenga.

Genome-wide gene expression analysis in response to Organophosphorus pesticides Chlorpyrifos and Diazinon in C. elegans.

PLoS ONE 5(8): e12145. doi:10.1371/journal.pone.0012145.(2010)

Abstract

Organophosphorus pesticides (OP) were originally designed to affect the nervous system by inhibiting the enzyme acetylcholinesterase, an important regulator of the neurotransmitter acetylcholine. Over the past years evidence is mounting that these compounds affect many other processes. Little is known, however, about gene expression responses against OP in the nematode *Caenorhabditis elegans*. This is surprising because *C. elegans* is extensively used as a model species in toxicity studies. To address this question we performed a microarray study in *C. elegans* which was exposed for 72 hrs to two widely used OP, chlorpyrifos and diazinon, and a low dose mixture of these two compounds. Our analysis revealed transcriptional responses related to detoxification, stress, innate immunity, and transport and metabolism of lipids in all treatments. We found that for both compounds as well as in the mixture, these processes were regulated by different gene transcripts. Our results illustrate intense and unexpected crosstalk between gene pathways in response to chlorpyrifos and diazinon in *C. elegans*.

Introduction

Organophosphorus pesticides (OP) are widely used to control agricultural and household pests, and consequently these compounds have been well studied for their impact on gene function and physiology. Initially, OP were designed to affect the nervous system of pest organisms by inhibiting acetylcholinesterase (AChE). This leads to a cholinergic hyper stimulation as a common mode of action (Pope, 1999). However, mounting evidence suggests that this is not the only process that is being affected. For instance, mice lacking AChE are hypersensitive to OP toxic effects, indicating that OP inhibit targets other than AChE (Duysen *et al.*, 2001). In addition, neurotransmitter receptors, proteases and several other enzymes interact with OP to modify the consequences of AChE inhibition (Pope *et al.*, 2005). Moreover, many OP alter immune functions in mammals by oxidative damage, metabolism modifications and stress-related immunosupression (Li, 2007).

Of the hundreds of OP, chlorpyrifos (CPF) and diazinon (DZN) are intensively used as insecticides and acaricides. CPF is the most thoroughly studied OP and, like DZN and most OP, is desulfurated by cytochrome P450 enzymes (CYPs). CPF and DZN share detoxification pathways and molecular targets as they are structurally similar (Poet *et al.*, 2003). DZN, however, induces a different inhibition ratio of AChE (Slotkin *et al.*, 2008). Some studies also investigated the response of AChE to a combination of CPF and DZN (Timchalk *et al.*, 2002; Timchalk *et al.*, 2005). Their results suggest a greater inhibition of cholinesterase by CPF than by DZN and a greater influence of CPF in the combination. Likewise, kinetic models of humans and rats exposed to CPF and DZN indicate an interaction between both compounds at the enzymatic level. In particular, CPF and DZN seem to compete for common CYPs (Timchalk and Poet, 2008).

Despite our knowledge about OP responses in different organisms, little is known about genome-wide gene expression responses of OP in the nematode *Caenorhabditis elegans*. This is surprising because *C. elegans* has been well established as a model for understanding human toxicology, especially for studying neurotoxic compounds like OP (Cole *et al.*, 2004). The effects of CPF on *C. elegans* have also been investigated in contrast to other chemical effects (Ruan *et al.*, 2009). These studies showed that neurotoxic

compounds affect behaviour and movement in *C. elegans*. Additionally, gene transcription changes by CPF were investigated by Roh & Choi (2008) (Roh and Choi, 2008) who found that CPF regulates stress related genes. Still, a comparative analysis of the genome-wide transcriptional responses of *C. elegans* to DZN and CPF has not been conducted.

Here we measured genome wide gene transcription profiles of *C. elegans* exposed to CPF and DZN and to a low dose mixture (LDM) of both compounds. Gene Ontology (GO) and domain enrichment analysis illustrates the complexity with novel and known pathways associated to OP response.

Results and Discussion

Effects of CPF and DZN on transcript levels in C. elegans

We analyzed global gene transcription profiles of *C. elegans* treated with CPF or DZN. These profiles were used to reveal transcriptional responses of each OP. Compared to control treatment, 551 and 245 genes were significantly regulated in CPF and DZN treated worms, respectively. Of these genes, 126 were regulated by either compound (Figure 1 and Table S1).

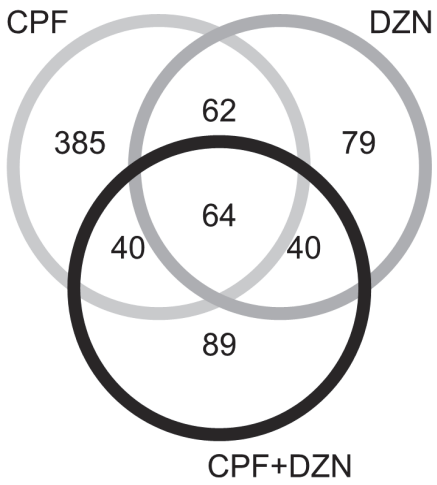


Figure 1.– Transcriptional response to CPF, DZN and combination. Venn diagram showing significantly regulated genes by CPF (light grey), DZN (dark grey), a combination of both (CPF + DZN, black) and their overlap.

To gain insight into the biological processes associated with the regulated genes, we determined which GO annotation terms were over-represented. In both treatments, significantly enriched GO terms (P-value < 0.01, hypergeometric test) were related with detoxification (monooxygenase activity), metabolic process, and transport of lipids (Figure 2, Table 1 and Supplementary Table ST2). To add meaningful information to the results from the GO term enrichment we extended our investigation by using a similar analysis with protein domains associated to the regulated genes as categories (Figure 3 and Supplementary Table ST2). The significantly overrepresented groups (P-value < 0.01, hypergeometric test) also included domains related to detoxification, stress-response, and transport of fatty acids. In addition, we also found significantly enriched domains associated with other metabolic pathways and immune response.

Detoxification of CPF and DZN

Enzymes and functional domains associated with detoxification in *C. elegans* are mainly cytochrome P450 (CYP) and short-chain dehydrogenase (SDR) in phase *I* of the xenobiotic metabolism, and UDP-glucuronosil/transferase (UGT) and glutathione-S-transferase (GST) in phase *II* (Tim H. Lindblom, 2006). Their implication in the detoxification of CPF and DZN has been well characterized in different organisms (Timchalk *et al.*, 2002; Poet *et al.*, 2004). In our experiment, 19 genes that code for domains related to phase *I* responded to the CPF treatment, whereas 14 genes with these domains responded to the DZN treatment. Of those genes, 10 were shared between CPF and DZN treatment, of which nine were CYPs. Five of those shared CYPs belong to the *cyp-35* subfamily, which has previously been identified as strongly inducible by a range of xenobiotics in *C. elegans* (Menzel *et al.*, 2005). In addition, only one gene (F25D1.5) coding for an SDR domain was regulated in both treatments.

Also genes coding for proteins with domains involved in phase *II* of detoxification were affected by CPF (21) and DZN (14). From those genes, 8 were affected by both treatments. One of these common genes was *cdr-1* (F35E8.1), a member of the cadmium responsive genes family. Two other genes from this family were affected only by CPF, *cdr-4* (K01D12.11) and *cdr-6* (K01D12.12), and *cdr-5* (K01D12.14) only by DZN. Whereas cadmium modifies the expression levels of *cdr-4* and *cdr-1*, the level of expression of *cdr-5*

and *cdr-6* are impervious to this metal's presence (Dong *et al.*, 2005). Moreover, other stressors are capable of modulating *cdr-4* expression, but until now, *cdr-1* transcription has only been induced by cadmium.

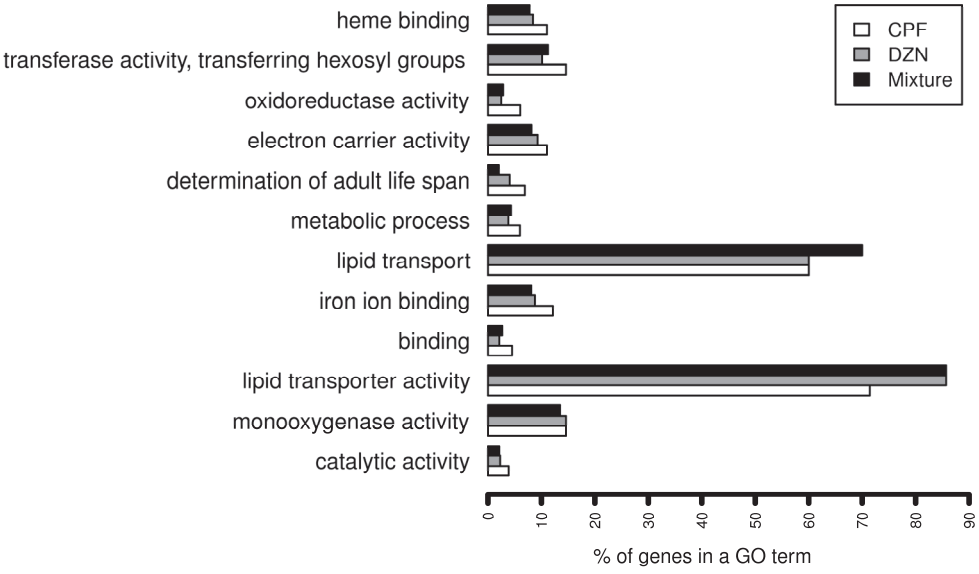


Figure 2.– Differentially expressed genes within enriched GO terms in CPF, DZN and combination. Percentage of genes within each GO terms significantly regulated in each treatment: CPF (white), DZN (grey) and a combination of both (black). Only significantly enriched GO terms for all the treatments are shown (P-value < 0.01 using hypergeometric test). Full list of GO terms can be found in Table S2.

Role of the daf-16 pathway in response to OP

Another gene involved in phase *II* of detoxification and shared between treatments was *gst-10* (Y45G12C.2). This gene acts downstream of *daf-16* (R13H8.1), a FOXO class transcription factor involved in the insulin-like growth factor signalling in *C. elegans* (Murphy *et al.*, 2003). Detoxification genes are known to constitute a big group among the identified up-regulated genes by *daf-16* (McElwee *et al.*, 2003; Murphy *et al.*, 2003). Other genes acting downstream *daf-16* have also been associated with toxic stress response. We found from those groups metallothioneins (*mtl*) and vitellogenins (*vit*) to be affected by CPF or DZN. DZN decreased the transcript levels of *mtl-2* (T08G5.10), while CPF did not

Table 1.- Major Gene Ontology (GO) terms represented in the different treatments

Definition	GO terms	Genes in CPF	Genes in DZN	Genes in mixture	Common genes in all three treatments
Catalytic activity	GO:0003824	22	13	12	2
Monooxygenase activity	GO:0004497	13	13	12	6
Lipid transporter activity	GO:0005319	5	6	6	5
Binding	GO:0005488	32	15	19	4
Iron ion binding	GO:0005506	18	13	12	6
Lipid transport	GO:0006869	6	6	7	5
Metabolic process	GO:0008152	36	23	26	8
Determination of adult life span	GO:0008340	17	10	5	3
Electron carrier activity	GO:0009055	19	16	14	6
Oxidoreductase activity	GO:0016491	17	7	8	1
Transferase activity, transferring hexosyl groups	GO:0016758	13	9	10	6
Heme binding	GO:0020037	17	13	12	6

Table 1.- A summary of common significant (P-value < 0.01, hypergeometric test) GO terms in each treatments are shown. The final column show the number of common genes within a GO term in all the three treatments. Only GO terms significant in all the three treatments are shown. A complete list with all the GO terms and its correspondent P-values can be found in Supporting Table ST1.

regulate any of the two *mtl* genes present in *C. elegans*. Furthermore, vitellogenin genes, lipid-binding proteins involved in lipid mobilization, were down-regulated in both treatments. CPF modified the expression of 5 out of 6 members of the *vit* gene family (*vit-1* to *vit-3*, *vit-5*, and *vit-6*) while DZN regulated all 6 *vit* genes. Furthermore, both OP regulated collagen genes (34 by CPF, 16 by DZN) and major sperm proteins type genes (33 by CPF, 3 by DZN) acting downstream of *daf-16*. Many of these have been associated with stress responses, although their specific connection to these responses remains unresolved (Roh *et al.*, 2006; Cui *et al.*, 2007). The overlap in genes affected by CPF and/or DZN and the *daf-16* pathway suggest that this pathway is a primary response pathway following exposure to OP.

OP affect innate immunity genes

The insulin-like signalling pathway regulated by *daf-16* is known to induce the immune response as part of a general stress response (Ewbank, 2006). On the other hand, Lewis *et al.* (2009) showed that OP regulate *daf-16* in *C. elegans* as a key modulator of physiological responses other than stress, including innate immune response (Lewis *et al.*, 2009). Therefore, it is not surprising to find that *lys-7* (C02A12.4), a gene coding for a well defined antimicrobial lysozyme, was regulated by both toxicants. This gene is presumed to promote resistance to infection (O'Rourke *et al.*, 2006). Interestingly, while *lys-7* is up-regulated under both OP treatments and suggests an immune response, *spp-1* (T07C4.4) was down-regulated by CPF, but not by DZN. This saposin is also a well known pathogen responsive gene. However, suppression of *spp-1* during *Pseudomonas aeruginosa* infection seems to be under regulation of a different pathway than *daf-16* (Evans *et al.*, 2008). Therefore, other gene pathways may participate in parallel to *daf-16* insulin-like pathway to promote stress and infection responses in *C. elegans* as a result of OP exposure.

To further investigate an innate immunity response to OP, we compared the regulated genes in our experiments with genes associated with immune response and stress in *C. elegans*. An indication that different genes are involved in stress and infection response was found by Shapira *et al.* 2006 (Shapira *et al.*, 2006). The authors compared genes regulated in response to *P. aeruginosa* with those affected by cadmium. They reported several overlapping genes among these two responses, but they concluded that a reaction to infection is largely distinct from stress response. Based on this, we considered that an innate immune reaction signature may be distinguishable from OP stress response in *C. elegans*. Subsequently, we also compared our set of genes with the same set of cadmium responsive genes and pathogens responsive genes Shapira *et al*. used (Table 2). We found more genes overlapping between the OP affected genes and the cadmium responsive genes. The similarities between genes modulated by OP and by cadmium suggest the implication of a common response mechanism of gene expression responses to toxic compounds. Nevertheless, the larger number of genes affected by OP as well as *P. aeruginosa* infection supports the idea of an innate immune response to OP. For example, a CUB-like domain (also known as DUF141) containing gene (F55G11.2) is induced during infection but repressed under cadmium exposure. This gene was up-regulated in both OP treatments.

Another gene (F08G5.6) up-regulated in all OP treatments is strongly induced by infection but not by cadmium.

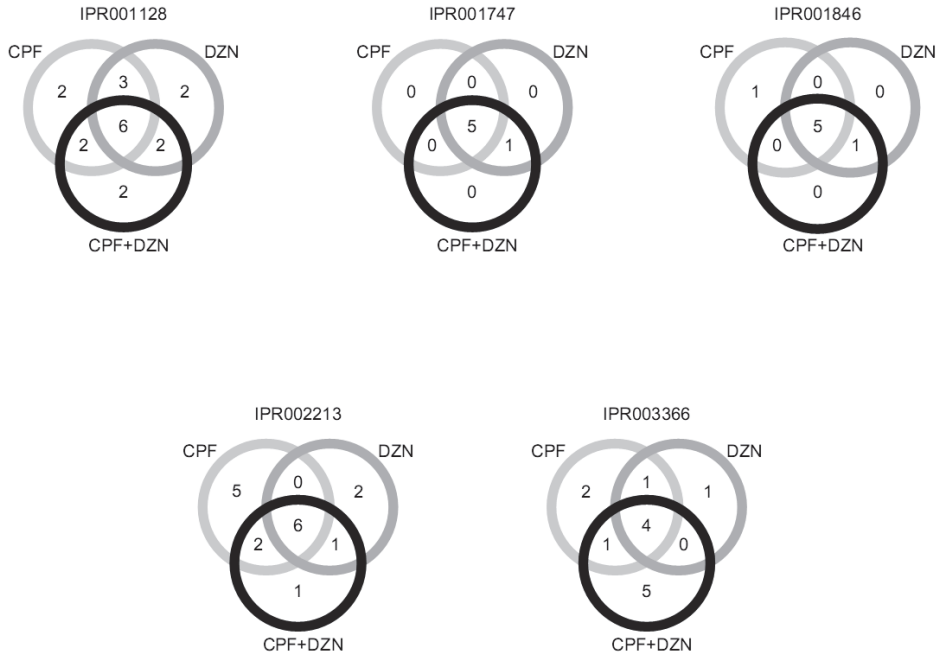


Figure 3.– Differentially expressed genes within enriched domains in CPF, DZN and combination. Major functional groups represented in all the treatments: CPF (light grey), DZN (dark grey) and a combination of both (black). The number of genes that show a functional domain are shown in each case, indicating the common regulated genes. Only functional domains common and significant (P-value < 0.01, hypergeometric test) within the three treatments are shown. Functional domain categories were extracted from defined proteins domains in Wormbase release 190 using BioMart.

Domains description: IPR001128 = Cytochrome P450 ; IPR001747 = Lipid transport protein, N-terminal; IPR001846 = von Willebrand factor, type D; IPR002213 = UDP-glucuronosyl/UDP-glucosyltransferase; IPR003366 = CUB-like region.

The same study from Shapira and colleagues identified ELT-2, a GATA transcription factor, as key regulator for transcriptional responses to infection. ELT-2 together with SKN-1, also a transcription factor, have been proposed as complementary regulators of *spp-1* during infection (Evans *et al.*, 2008). Since *spp-1* is also regulated by the *daf-16* pathway, genes acting downstream of *elt-2* and *skn-1* may act in parallel to induce a combined stress-

immune response to CPF and DZN (Figure 4). In that sense, SKN-1 is known to mediate an expression response of genes involved in phase II of detoxification. To induce these genes, SKN-1 is phosphorylated through the PMK-1 p38 mitogen-activated protein kinase pathway (p38 MAPK) (An *et al.*, 2005). Genes regulated by *pmk-1* are essential for immune and stress responses in C. elegans. These genes code for antimicrobials proteins with CUB-like domains, C-type lectins domains, ShK domains, and DUF274 domains (Troemel *et al.*, 2006). Our domain enrichment analysis showed that OP regulated genes involved in phase *II* of detoxification. It also showed an overrepresentation of domains associated to innate immune response through the *pmk-1*/*skn-1* pathway (Supplementary Table ST2). Therefore, we compared our data sets with the genes downstream *pmk-1* identified by Troemel *et al*. (Table 2). We found significant overlapping between genes belonging to this pathway and the genes that were regulated in each treatment (13 genes, P-value < 0.01) for CPF response, but not for DZN (5 genes, P-value = 0.012). This strengthens the implication of this pathway in response to CPF and indicates differences in immunomodulation between CPF and DZN. Those differences have been reported in *in vitro* studies with human cells, where CPF showed a stronger effect in cell viability and some immune parameters (Oostingh *et al.*, 2009). Altogether, CPF and DZN modified innate immunity related genes in *C. elegans* in a different manner and as a part of complex stress response.

Transcriptional effects of a low dose mixture (LDM)

In the LDM, 233 genes were significantly regulated (Supplementary Table ST1), of which 64 were affected by all, the combination and the both single compound treatments (Figure 1 and Supplementary Table ST1). More genes (89) were affected by the LDM only. The GO terms enrichment analysis (Figure 2, Table 1 and Supplementary Table ST2) showed a significant modulation of detoxification pathways, stress response, and transport and metabolism of lipids, which is highly similar to the single compound treatments. Moreover, the domain enrichment analysis identified differentially expressed genes that code for CYP, SDR, GST, UDP, C-type lectin, and CUB-like domains (Figure 3 and Supplementary Table ST2).

Table 2. -Pathways involved in OP response.

Pathway /mechanism	num. genes	CPF		DZN		CPF + DZN	
		Overlap. genes	P-value	Overlap. genes	P-value	Overlap. genes	P-value
DAF-16	541	47	1.679e-10	36	2.920e-16	29	3.242e-11
ELT-2	255	30	1.779e-10	18	1.723e-09	21	3.562e-12
MDT-15	305	57	1.346e-28	29	1.319e-17	29	8.589e-18
DBL-1/ SMA-9	320	31	1.224e-08	15	7.045e-06	22	2.829e-11
PMK-1	143	13	0.00019	5	0.01267	5	0.01190
Cd	290	44	9.917e-19	44	8.403e-35	40	2.447e-30

Table 2. - Pathway/mechanism investigated were: genes downstream *daf-16 (*DAF-16) (Murphy *et al.*, 2003), genes regulated by the GATA transcription factor ELT-2 (ELT-2) (Shapira *et al.*, 2006), genes regulated by the Mediator subunit MDT-15 (MDT-15) (Taubert *et al.*, 2008), genes downstream the transcription factors DBL-1 and SMA-9 (DBL-1/SMA-9) (Liang *et al.*, 2007), genes downstream *pmk-1* (PMK-1) (Troemel *et al.*, 2006) and cadmium responsive genes (Cd) (Cui *et al.*, 2007). Numbers of genes associated with each pathway or mechanism are shown. Overlapping genes are the number of genes belonging to a pathway that are regulated in each treatment. P-values were calculated using hypergeometric test.

Regulation of CYPs genes is required for detoxification of OP in the single and combined treatments. Our findings showed that the OP combination regulated 12 CYPs (IPR001128), 6 of which were also affected by CPF and DZN in single treatments. On the other hand, 2 CYPs were regulated only in the combined treatment. Differences in the regulated CYPs were expected, because OP seem to compete for binding sites of CYPs (Timchalk and Poet, 2008). Moreover, *in vitro* analysis of CYPs metabolism of CPF and DZN indicates that CPF inhibits the metabolism of DZN to IMHP (the inactive pyrimidol derived from DZN), with no effect of DZN over CPF metabolism. This inhibition seems to promote oxon synthesis from DZN, and therefore, increases the total levels of oxon in the organism after a treatment of more than one compound. Moreover, the toxicity of OP is determined by a balance between bioactivation (production of oxon metabolite) and detoxification (inactivation of OP and its intermediate oxon) (Poet *et al.*, 2003). Therefore, higher levels of DZN bioactivation, due to modification of this balance, may be a reason for a greater-than-additive inhibition of cholinesterase enzymes in mixtures of OP (Moser *et al.*, 2006;

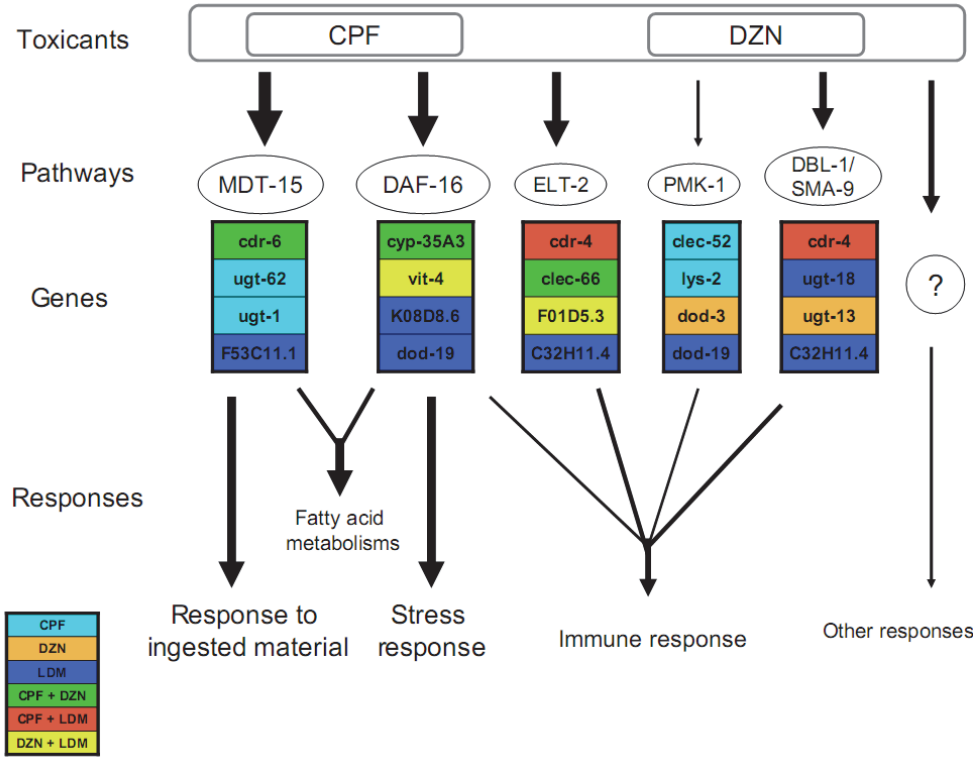


Figure 4 .– A model for the transcriptional responses to CPF and DZN in *C. elegans*. Exposure to the toxicants CPF, DZN or a low dose mixture (LDM) of both organophosphates modulates the expression of genes in common pathways. The most regulated pathways are shown with the thicker arrows: genes regulated by MDT-15 (MDT-15), genes downstream *daf-16* (DAF-16), and genes downstream *elt-2* (ELT-2). For details, see Table 2. Despite these shared regulated pathways, the three treatments regulated different genes to induce common responses such as stress response, response to ingested material or immune response. Here we show some regulated genes by chlorpyrifos (CPF), diazinon (DZN) and a low dose mixture (LDM). The different colors designate shared regulated genes by chlorpyrifos and diazinon (CPF + DZN), by chlorpyrifos and low dose mixture (CPF + LDM), and by diazinon and low dose mixture (DZN + LDM). An intense crosstalk between gene pathways in response to the toxicants may be explained by the genes shared between pathways. For example, *cdr-4* was regulated by CPF and LDM but not by DZN. Previous studies showed that its expression is modulated by ELT-2 and DBL-1, two transcription factors implicated in responses to infection. Exposure to the toxicants modulated the expression of many genes from those pathways that, acting in parallel, induced the detected responses.

Svendsen *et al.*, 2010). At the transcriptional level however, the number of CYPs regulated by the combined treatment do not suggest a higher bioactivation. Even though the combined treatment contained a higher concentration of toxicants, a similar number of genes involved in detoxification was regulated in the combined and the single treatments. These results were not contradictory to the previous studies, since a direct comparison between transcriptional and physiological studies is difficult. Nevertheless, our results indicated that dissimilar genes were involved in the regulation of detoxification of the different OP treatments and may explain the differences between the LDM and the single treatments.

Other processes were modulated by all treatments besides detoxification, these include: stress, fat metabolism and innate immunity. The overlap between the analyzed pathways for the single treatments and the regulated genes in the combinations was significant (Table 2). Interestingly, as it is shown in the last column of table 1, the number of common genes associated to some of these responses was low. For example, the modulation of stress and innate immune genes was observed for the combined treatment. Genes involved in response to infections were up-regulated such as F55G11.2, F08G5.6 and genes with the CUB-like functional domain (Figure 3 and Supplementary Table ST2). This again suggests an innate immune response to OP. Furthermore, the combined treatment showed up-regulation of *sma-9* (T05A10.1), a gene coding for a transcriptional factor involved in body size regulation. This transcriptional factor modulates a subset of genes in the TGF-ß signaling pathway which is associated with innate immunity, aging and germ line regulation (Savage-Dunn, 2005). Of the identified genes in the TGF-ß signaling pathway (Liang *et al.*, 2007), the combined treatment regulated 22 genes (Table 2). A closer look to the regulated genes by CPF and DZN showed also the implication of this pathway in their transcriptional responses. Once more, as in response to the *pmk-1* pathway, CPF seems to regulate more genes implicated in innate immunity, but the LDM did not show a stronger influence. Likewise, genes downstream *daf-16* and *elt-2* were regulated in the combined treatment.

Overall, as shown in Figure 4, the ingestion of CPF, DZN or a low dose mixture of both induced many common responses like stress and innate immunity pathways. The most regulated pathways included genes modulated by MDT-15 (MDT-15), genes downstream *daf-16* (DAF-16), and genes downstream *elt-2* (ELT-2). The three treatments regulated

different genes within the same pathways to induce the common responses. The analysis of the regulated genes showed an intense crosstalk between gene pathways involved in the toxicants response. This suggested that the structural differences between CPF and DZN induced the expression of different genes. These differences, however, are not large enough to induce a specific different response to the exposure of CPF or DZN.

Conclusion

We present the analysis and comparison of whole genome transcription analysis of a model organism exposed to two toxicants and a combination of both. Our results revealed that, CPF and DZN induced dissimilar genes, even though they have a similar chemical structure. The toxicants induced genes related to detoxification, stress, innate immunity and response to ingested material in single and combined treatments. The differences between transcript responses in the combined treatment suggest that the effect of a mix of low doses of CPF and DZN is not a summed effect of the single components. But at the same time, the similarities in the induced pathways (e.g.: *daf-16*, *elt-2*, *skn-1/pmk-1*, *sma-9*) indicate the regulation of similar responses to them.

Materials and Methods

C. elegans culturing

The Bristol N2 strain was cultured on standard nematode growth medium (NGM) with *E. coli* OP50 as food source. Nematodes were bleached (0.5 M NaOH, 1% hypochlorite) to collect eggs which were inoculated in 9 cm dishes for toxicity experiments. After 72 hours, nematodes were collected in the L3-L4 stage, frozen in liquid nitrogen and kept at -80C until the RNA extraction procedure.

Toxicant exposures

We analyzed gene expression in response to the toxicants at concentrations below the EC_{50} values for different fitness traits as reproduction (CPF: EC50 = 3.5 mg/L (Martin *et al.*, 2009) DZN: EC_{50} = 30 mg/L (Svendsen *et al.*, 2010)) or growth (CPF: EC_{50} = 14 mg/L

(Boyd *et al.*)), Nematodes were exposed to 0.5 mg/L of CPF (Cyren®/Nufos®, Cheminova A/S [Lemvig, Denmark]) and 1 mg/l of DZN (Supelco [Bellefonte, Pennsylvania 16823, USA]). The low dose mixture (LDM) of the two OP contained the sum of both single concentrations (DZN [1 mg/l] and CPF [0.5 mg/l]).

The experiment started with eggs placed on NGM dishes with the OP-treatments and *E. coli* OP50 as food source. After 72 hours, worms from 4 petri dishes were collected as one sample. A total of 6 replicates per treatment were collected (24 petri dishes), and immediately frozen in liquid nitrogen until RNA extraction. All OP were dissolved in acetone and added to 10 ml of NGM poured in each 9 cm petri dish used for the culture. Nematodes without treatments were grown simultaneously with the same concentrations of acetone in a control culture.

Microarray experiments

RNA from nematodes was extracted following the Trizol method, and the RNeasy Micro kit (Qiagen, Valencia, CA, USA) was used to clean up the samples. Labeled cDNA was produced with the kit Array 900 HS from Genisphere and Superscript II from Invitrogen. The 60-mers arrays were purchased from Washington University (http://genomeold.wustl.edu) and they were hybridized following the Genisphere Array 900 HS protocol with modifications. Extracts from CPF, DZN and the CPF/DZN combination exposures were hybridized with the control samples in each array. Six independent biological replicates were used per treatment to produce six replicate microarrays per experiment in a dye-swap design.

Microarray Analysis

A Perking & Elmer scanner was used to extract the raw intensities from the microarrays. Normalization within arrays and normalization between arrays of raw intensities was done using loess method (Smyth and Speed, 2003) and aquantile method (Yang, 2003), respectively. Both methods are included in the Limma package (Smyth, 2005) from R software (www.r-project.org/). The Rank Product package (Hong *et al.*, 2006) was used to identify the differentially expressed genes between controls and treatment in each experiment. Briefly, genes were ranked based on up- or downregulation by the treatment in each experiment. Then, for each gene a combined probability was calculated as a rank product (RP). The RP values were used to rank the genes based on how likely it was to observe them by chance at that particular position on the list of differentially expressed genes. The RP can be interpreted as a P-value. To determine significance levels, the RP

method uses a permutation-based estimation procedure to transform the *P-value* into an *E-value* that addresses the multiple testing problem derived from testing many genes simultaneously. Genes with a percentage of false-positives (PFP) < 0.05 were considered differentially expressed between treatments and control in each experiment. This method has the advantage to identify genes with a response to the toxicants even when the absolute effect of the response was low. Because we used sub-lethal concentrations of the toxicants, methods that use thresholds based on absolute fold change would not identify small changes in gene expression. Moreover, RP has proved to be a robust method for comparing microarray data from different sources and experiments (Hong and Breitling, 2008).

Gene Ontology (GO) data and functional domain data were extracted from Wormbase release WB195 using the R package BioMart (Durinck *et al.*, 2005). GO terms and domains with less than 4 genes were discarded. Over-represented groups of GO terms and functional domains were identified using a hypergeometric test, with a threshold of P-value <0.01. The hypergeometric test compared a group of 396 GO terms, with 16,947 annotated genes, with the GO terms associated with the significantly regulated genes in each treatment (551,245 and 233 for CPF, DZN and the LDM). For functional domains analysis, 1003 InterPro and Pfam domain terms were used, with 8682 annotated genes.

The same hypergeometric test was used to determine significant regulation of the different pathways analyzed. The lists of regulated genes, in this case, were extracted from the original publications. Group sizes were 541 genes for the *daf-16* pathway, 255 for the *etl-2* pathway, 305 for the *mdt-5* pathway, 320 for the *dbl-1* pathway, 143 for the *pmk-1* pathway, and 290 genes for Cd responsive genes. Annotated genes were all the unique genes on the microarray, 18893.

Microarray data have been deposited in Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo/), accession number GSE16719.

Supplementary Files

All the supplementary tables and figures are available from the author on http://www.anavinuela.com/thesis/thesis.html

Supplementary Table ST1. Differentially expressed genes in nematodes treated with CPF (0.5 mg/l), DZN (1 mg/l), CPF (0.5 mg/l) + DZN (1 mg/l). RankProduct (RP) was used to identify regulated genes in each of the treatments. Each worksheet contains the output for RP in one experiment. Sequence names for each regulated gene are in gene.names column. Gene.index are the microarray index for the original data. RP/sum are the rank product sum calculated per each gene. FC:(class1/class2) are the expression fold change of class 1

(treatment) / class 2 (control). pfp: estimated percentage of false positive for each gene. P.value: estimated P-value for each gene.

Supplementary Table ST2. Significantly enriched Gene Ontology (GO) terms and functional domains.

GO terms worksheets contain full GO enriched analysis. Genes in GO database are the number of genes belonging to each GO term. GO number are the GO term. Regulated genes refer to number of significantly regulated genes in each treatment that belong to a GO term. *P-values* were calculated using hypergeometric test, values lower that 0.01 and more than two genes in a category were considered significant.

Functional domains worksheets: Genes with the domain are the number of genes that code for a specific functional domain. The number of significantly regulated genes in each treatment that code for a specific functional domain is listed in the regulated genes column. The functional domains terms were extracted from Wormbase 195 using BioMart.

References

An, J. H., K. Vranas, et al. (2005). Regulation of the *Caenorhabditis elegans* oxidative stress defense protein SKN-1 by glycogen synthase kinase-3. *Proceedings of the National Academy of Sciences of the United States of America* **102**(45): 16275-16280.

Boyd, W. A., M. V. Smith, et al. Medium- and high-throughput screening of neurotoxicants using *C. elegans*. *Neurotoxicology and Teratology* **32**(1): 68-73.

Cole, R. D., G. L. Anderson, et al. (2004). The nematode *Caenorhabditis elegans* as a model of organophosphate-induced mammalian neurotoxicity. *Toxicology and Applied Pharmacology* **194**(3): 248-256.

Cui, Y., S. McBride, et al. (2007). Toxicogenomic analysis of *Caenorhabditis elegans* reveals novel genes and pathways involved in the resistance to cadmium toxicity. *Genome Biology* **8**(6): R122.

Dong, J., M. O. Song, et al. (2005). Identification and characterization of a family of Caenorhabditis elegans genes that is homologous to the cadmium-responsive gene *cdr-1*. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1727**(1): 16-26.

Durinck, S., Y. Moreau, et al. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* **21**(16): 3439-3440.

Duysen, E. G., B. Li, et al. (2001). Evidence for Nonacetylcholinesterase Targets of Organophosphorus Nerve Agent: Supersensitivity of Acetylcholinesterase Knockout Mouse to VX Lethality. *J Pharmacol Exp Ther* **299**(2): 528-535.

Evans, E. A., T. Kawli, et al. (2008). *Pseudomonas aeruginosa* Suppresses Host Immunity by Activating the DAF-2 Insulin-Like Signaling Pathway in *Caenorhabditis elegans*. *PLoS Pathog* **4**(10): e1000175.

Ewbank, J. J. Signaling in the immune response (January 23, 2006), WormBook, ed. The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.83.1, http://www.wormbook.org

Hong, F. and R. Breitling (2008). A comparison of meta-analysis methods for detecting differentially expressed genes in microarray experiments. *Bioinformatics* **24**(3): 374-382.

Hong, F., R. Breitling, et al. (2006). RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. *Bioinformatics* **22**(22): 2825-2827.

Lewis, J., M. Szilagyi, et al. (2009). Distinct patterns of gene and protein expression elicited by organophosphorus pesticides in *Caenorhabditis elegans*. *BMC Genomics* **10**(1): 202.

Li, Q. (2007). New mechanism of organophosphorous pesticide-induced immunotoxicity. *Journal of Nippon Medical School* **74**(2): 92-105.

Liang, J., L. Yu, et al. (2007). Transcriptional repressor and activator activities of SMA-9 contribute differentially to BMP-related signaling outputs. *Developmental Biology* **305**(2): 714-725.

Martin, H. L., C. Svendsen, et al. (2009). Measurement and modelling of the toxicity of binary mixtures in the nematode *Caenorhabditis elegans* - A test of independent action. . *Environmental Toxicology and Chemistry* **28**(1): 97-104.

McElwee, J., K. Bubb, et al. (2003). Transcriptional outputs of the Caenorhabditis elegans forkhead protein DAF-16. *Aging Cell* **2**: 111 - 121.

Menzel, R., M. Rödel, et al. (2005). CYP35: Xenobiotically induced gene expression in the nematode *Caenorhabditis elegans*. *Archives of Biochemistry and Biophysics* **438**(1): 93-102.

Moser, V. C., J. E. Simmons, et al. (2006). Neurotoxicological Interactions of a Five-Pesticide Mixture in Preweanling Rats. *Toxicol. Sci.* **92**(1): 235-245.

Murphy, C. T., S. A. McCarroll, et al. (2003). Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. *Nature* **424**(6946): 277-283.

O'Rourke, D., D. Baban, et al. (2006). Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of *C. elegans* with *M. nematophilum*. *Genome Res* **16**: 1005 - 1016.

Oostingh, G. J., G. Wichmann, et al. (2009). The cytotoxic effects of the organophosphates chlorpyrifos and diazinon differ from their immunomodulating effects. *Journal of Immunotoxicology* **6**(2): 136-145.

Poet, T. S., A. A. Kousba, et al. (2004). Physiologically Based Pharmacokinetic/Pharmacodynamic Model for the Organophosphorus Pesticide Diazinon. *NeuroToxicology* **25**(6): 1013-1030.

Poet, T. S., H. Wu, et al. (2003). In Vitro Rat Hepatic and Intestinal Metabolism of the Organophosphate Pesticides Chlorpyrifos and Diazinon. *Toxicol. Sci.* **72**(2): 193-200.

Pope, C., S. Karanth, et al. (2005). Pharmacology and toxicology of cholinesterase inhibitors: uses and misuses of a common mechanism of action. *Environmental Toxicology and Pharmacology* **19**(3): 433-446.

Pope, C. N. (1999). Organophosphorus Pesticides: They all have the same mechanism of toxicity? *Journal of Toxicology and Environmental Health Part B: Critical Reviews* **2**: 161-181.

Roh, J.-Y. and J. Choi (2008). Ecotoxicological evaluation of chlorpyrifos exposure on the nematode *Caenorhabditis elegans*. *Ecotoxicology and Environmental Safety* **71**(2): 483-489.

Roh, J. Y., J. Lee, et al. (2006). Assessment of stress-related gene expression in the heavy metal-exposed nematode Caenorhabditis elegans: A potential biomarker for metal-induced toxicity monitoring and environmental risk assessment. *Environmental Toxicology and Chemistry* **25**(11): 2946-2956.

Ruan, Q.-L., J.-J. Ju, et al. (2009). Evaluation of Pesticide Toxicities with Differing Mechanisms Using *Caenorhabditis elegans*. *Journal of Toxicology and Environmental Health, Part A: Current Issues* **72**(11): 746 - 751.

Savage-Dunn, C. TGF-ß signaling (September 9, 2005), WormBook, ed. The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.22.1, http://www.wormbook.org

Shapira, M., B. J. Hamlin, et al. (2006). A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. *Proc Natl Acad Sci USA* **103**: 14086 - 14091.

Slotkin, T. A., F. J. Seidler, et al. (2008). Targeting of neurotrophic factors, their receptors, and signaling pathways in the developmental neurotoxicity of organophosphates in vivo and in vitro. *Brain Research Bulletin* **76**(4): 424-438.

Smyth, G. K. (2005). Limma: Linear models for microarray data. *Bioinformatics and Computational Biology Solution Using R and Bioconductor*. New York, Springer**:** 397-420.

Smyth, G. K. and T. Speed (2003). Normalization of cDNA microarray data. *Methods* **31**(4): 265-273.

Svendsen, C., P. Siang, et al. (2010). Similarity, independence, or interaction for binary mixture effects of nerve toxicants for the nematode *Caenorhabditis elegans*. *Environmental Toxicology and Chemistry* **29**(5): 1182-1191.

Taubert, S., M. Hansen, et al. (2008). The Mediator Subunit MDT-15 Confers Metabolic Adaptation to Ingested Material. *PLoS Genetics* **4**(2): e1000021.

Tim H. Lindblom, A. K. D. (2006). Xenobiotic detoxification in the nematode *Caenorhabditis elegans*. *Journal of Experimental Zoology Part A: Comparative Experimental Biology* **305A**(9): 720-730.

Timchalk, C., R. J. Nolan, et al. (2002). A physiologically based pharmacokinetic and pharmacodynamic (PBPK/PD) model for the organophosphate insecticide chlorpyrifos in rats and humans. *Toxicological Sciences* **66**(1): 34-53.

Timchalk, C. and T. S. Poet (2008). Development of a physiologically based pharmacokinetic and pharmacodynamic model to determine dosimetry and cholinesterase inhibition for a binary mixture of chlorpyrifos and diazinon in the rat. *NeuroToxicology* **29**(3): 428-443.

Timchalk, C., T. S. Poet, et al. (2005). Pharmacokinetic and pharmacodynamic interaction for a binary mixture of chlorpyrifos and diazinon in the rat. *Toxicology and Applied Pharmacology* **205**(1): 31-42.

Troemel, E. R., S. W. Chu, et al. (2006). p38 MAPK Regulates Expression of Immune Response Genes and Contributes to Longevity in C. elegans. *PLoS Genetics* **2**(11): e183.

Yang, Y. H., and Thorne, N. P. (2003). Normalization for two-color cDNA microarray data. *D. R. Goldstein (ed.), Science and Statistics: A Festschrift for Terry Speed* **40**: 403-418.

Chapter: 3

Transcriptional responses to interacting environmental factors

Ana Viñuela, L. Basten Snoek, Joost A.G. Riksen, Jan E. Kammenga

Manuscript under revision

Abstract

Gene expression plays an important role in the response to environmental stressors. The number and type of genes depends on the stressors and their possible interactions. A common method to study transcriptional responses to stressors is to analyze expression profiles of model organisms under the influence of single factors. However, organisms rarely encounter single stress factors. Here, we analyzed gene expression levels of the model organism *Caenorhabditis elegans* exposed to two organophosphorus pesticides (OP) in a single and mixture treatment at two different temperatures. Furthermore, we investigated the transcriptional effect of the interactions between those stressors. Our results show that the expression of detoxification genes may be affected by OP in a concentration-related manner. Moreover, the transcriptional responses to the combination of OP and high temperature affect detoxification genes and increased the toxic levels of the pesticides.

Introduction

Organisms are exposed to changing environmental stresses throughout their life. Transcriptional profiling with microarrays has been extensively used to identify those genes differentially expressed in different single stressors and to understand how organisms respond to them. However, although those transcriptional profiles have provided new insights in genotype-by-environment interactions; organisms rarely encounter single stressors. More often they are exposed to multiple environmental factors which effects can be larger or smaller than the single factors. Understanding the responses to combined stressors can be especially relevant for risk assessment of chemicals in the environment. The concentration-response relationships and toxicity levels of many chemicals, including pesticides, are usually based on single exposure studies and optimal environmental conditions (Spurgeon *et al.*, 2010). This approach leads to under- or overestimation of risk, with serious consequences for human health and economy. However, over the last few decades there have been an increased interest in multiple stressors studies, although their number is still limited (Holmstrup *et al.*, 2010).

Pesticides are widely used to eliminate pests. Due to his common domestic and agricultural use, organisms are often exposed to pesticide combinations. Usually, pesticides are classified by their mode-of-action. For instance, all organophosphorus pesticides (OP) inhibit acetylcholinsterase (AChE), leading to a cholinergic hyper stimulation (Pope, 1999). However, the effect of OP exposure includes secondary targets that are different among members of the group (Pope, 1999; Slotkin *et al.*, 2007; Slotkin and Seidler, 2007; Lewis *et al.*, 2009). In addition, it has been shown that treatments with other mixtures of pesticides induce specific gene transcription responses compare to the single treatments (Mumtaz *et al.*, 2002; Hook *et al.*, 2008). In other words, toxicants with a similar mode-of-action can induce a different molecular response and their interactions in mixture exposures may affect the toxic outcome. Examples of interactions between chemicals are numerous in literature (Spurgeon *et al.*, 2010), but studies of gene transcriptional responses to mixtures are quite limited. Furthermore, other environmental parameters can also interact with the toxicants and the organism in single or mixture treatment and modify their toxic outcome.

For example, increased temperature elevates the toxicity of OP such as diazinon in zebrafish, or chlorpyrifos in earthworms (Osterauer and Köhler, 2008; De Silva *et al.*, 2009). But very low temperatures do not show significant interactions with pesticides like abamectin and carbendazim in earthworms (Bindesbøl *et al.*, 2009).

In general, it is assumed that temperature positively correlates with toxicant effects. This effect has been attributed mainly to increased uptake and increased accumulation of the toxicant at elevated temperatures (Holmstrup *et al.*, 2010). Nevertheless, some studies found a decreased toxicity at higher temperatures in aquatic organisms (Scheil and Köhler, 2009). This indicates that the metabolic disturbance of a toxicant depends on the temperature. A reason may be that temperature modifies the metabolic rate and therefore can have a strong effect on the whole organism (Gillooly *et al.*, 2001). Transcriptional responses to high temperatures have been characterized in many model organisms such as fruitflies, springtails and worms, among others (Sørensen *et al.*, 2005; Lee and Kenyon, 2009; Nota *et al.*, 2010). Moreover, Li *et al.* (2006) mapped genetic determinants for gene expression at different temperatures and for gene-environment interactions in *C. elegans* (Li *et al.*, 2006). Their results indicated that gene expression regulation differs with temperature, and therefore it strongly suggests that the interaction between toxicants and temperature may affect also transcriptional responses.

To investigate possible transcriptional responses to multiple interacting environmental factors we analyzed genome-wide gene expression profiles using microarrays of the *C. elegans* canonical strain N2 in different environments. Nematodes were treated with two OP, chlorpyrifos (CPF) and diazinon (DZN) in a single and combined treatment at 24ºC. Then, those gene expression data were combined with previously published expression profiles with identical toxicant treatments at 16ºC (Chapter 2). Our analysis focused on the identification of responsive genes to the different factors under study: CPF, DZN and temperature. The microarray profiles from the mixture of toxicants allowed us to identify those genes and biological functions whose expression was affected by interactions between environmental factors.

Results & Discussion

To study the effect of different interacting stressors we analyzed gene expression profiles of *C. elegans* treated with two widely used OP, chlorpyrifos (CPF) and diazinon (DZN) and a mixture of them, at two different temperatures, 16°C and 24°C. We used a linear model to investigate the relationship between toxicants and temperature and the observed gene expression levels. The model included as variables: CPF, DZN, Temperature (Temp), and the interactions: CPF*Temp, DZN*Temp, CPF*DZN and CPF*DZN*Temp. Genes with a –log10 P-value above the threshold were considered significantly influenced by the corresponding variable or interaction (Supplementary Table ST1). Table 1 shows the significant number of genes per variable. It also shows the number of overlapping genes between them, since the changes in transcript abundance can be influenced by more than one environmental factor. In summary, we tested 18893 genes from which we found the largest number of genes to be significantly affected by variables considering the interaction of the toxicants (CFP*DZN, 15.8%) and the interaction of the toxicants with temperature (CPF*Temp, 15.5% and DZN*Temp, 11.8%). The temperature as environmental factor, alone and interacting with other factors, affected 87.2% of the differentially expressed genes (6623), while the variables including toxicants factors influenced 72.8% of them. Therefore, the overall influence of temperature seems to be larger than the toxicants as it affected more genes.

We also investigated the biological functions that were affected. A general description of the processes affected by each variable may be obtained by Gene Ontology (GO) analysis of the significant genes as well as by significant enrichment of predicted domain sequences. The complete results can be viewed in the supplemental material (Supplementary Tables ST2, and Supplementary Figure SF1 and SF2). The analysis indicated regulation of monooxygenase activity, lipid transport and metabolic process, among others. Some GO terms were significantly enriched for only one variable. For example, the GO term associated to cellular components of the presynaptic active zone was significant for genes influenced by DZN but not in interaction with any other factor. More terms were affected only by temperature such as protein folding, dephosphorylation or ATPase activity. Likewise, the domain enrichment analysis indicated overrepresentation of domains related

with detoxification and metabolic transport. Among domains enriched only by interacting variables, we identified SMAD-like domains associated with genes influenced by the interaction of CFP*DZN. This domain is present in the SMAD proteins which are highly conserved mediators of the TGF-ß signal pathway involved in growth and stress response in *C. elegans* (Savage-Dunn, 2005). In summary, low concentration exposure to OP at different temperatures affected many biological processes known to be involved in toxicant responses (Chapter 2). However, as a result of interactions among environments specific biological functions were affected at each treatment.

Table 1.- Significantly expressed genes per environmental variable

	CPF	**DZN**	**Temp**	**CPF* Temp**	**DZN* Temp**	**CPF* DZN**	**CPF*DZN**
CPF	**1342**						
DZN	821	**1249**					
Temp	1180	997	**1933**				
CPF*Temp	249	256	319	**2933**			
DZN*Temp	185	174	233	1570	**2231**		
CPF*DZN	252	265	327	1970	1710	**2997**	
CPF*DZN*Temp	733	715	809	182	130	189	**1062**
Unique	69	96	371	655	256	574	98

Table 1.- We analyzed the transcriptional effect of three environmental factors: two toxicants, chlorpyrifos and diazinon, and temperature. Genes which expression level was affected by those three environmental factors were identified with a linear model that considered also interaction between factors. In this way, we indentified genes significantly affected (-log P-value > 2) by seven variables: chlorpyrifos (CPF), diazinon (DZN), temperature (Temp) and any interaction between them (CPF*Temp, DZN*Temp, CPF*DZN*Temp). Numbers of genes per variable are indicated in bold (diagonal). The number of genes commonly affected by two variables are also shown. The last row shows the number of genes significantly affected by only one variable in the model.

Influence on detoxification genes

Transcriptional regulation of detoxification genes is required in response to many environmental changes, including exposure to single and combined OP (Chapter 2). Moreover, toxicity of OP is determined by a balance between bioactivation and detoxification (Poet *et al.*, 2003). Bioactivation of CPF and DZN occurs in the initial phase

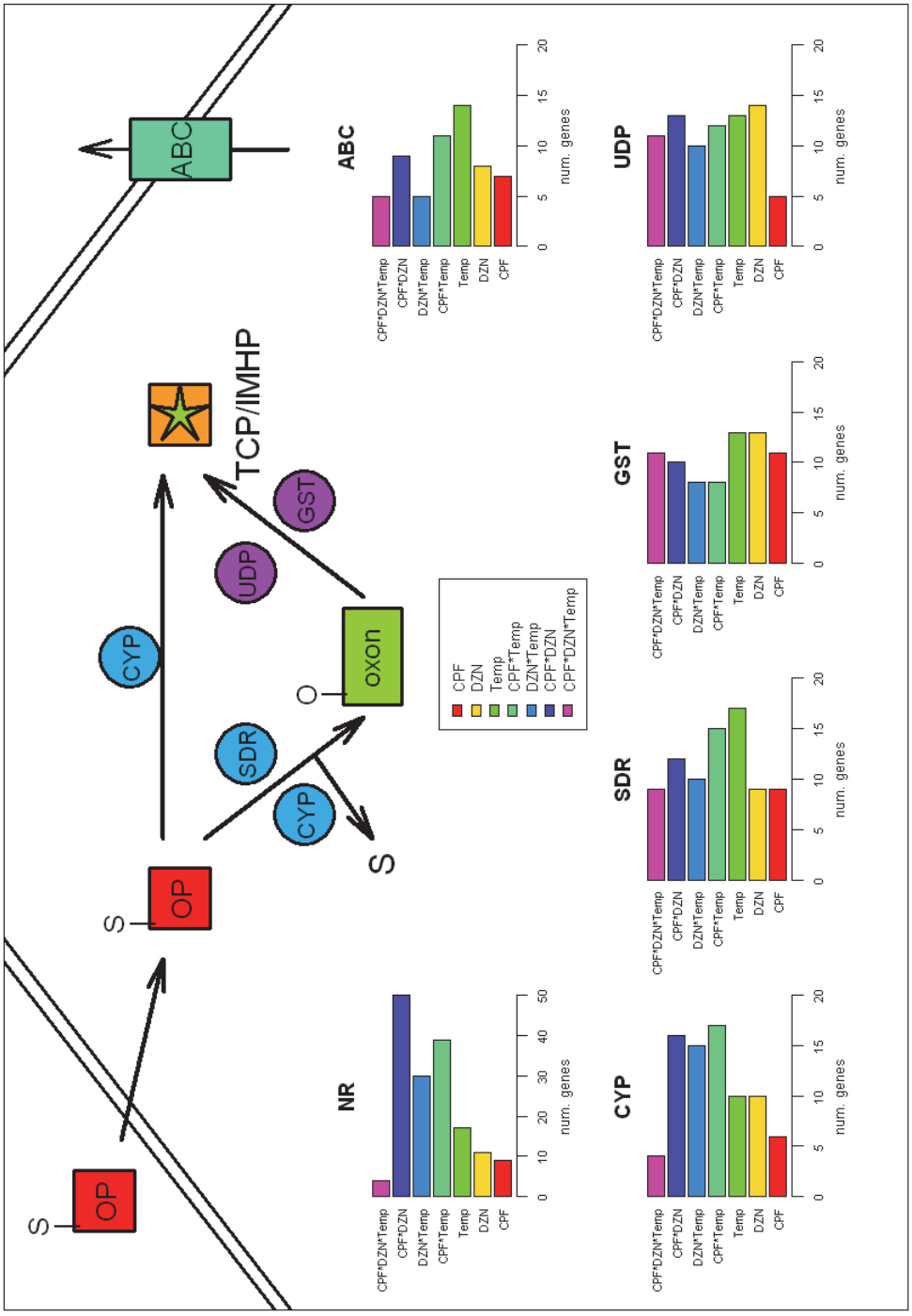
of detoxification, when cytochrome P450 enzymes (CYP) and short chain dehydrogenases (SDR) enzymes transform the toxicants into an oxygenated and highly toxic form called oxon-OP (Eaton *et al.*, 2008). The process usually follows with the effective detoxification (hydrolysis) of the oxon intermediates by UDP-glucuronosyl transferases (UGT) and glutathione-S-transferases (GST) enzymes to a final inactive compound (Figure 1). Direct dearylation of CPF and DZN to this final compound may also be mediated by CYPs in a direct detoxification reaction (Poet *et al.*, 2003). To better understand the differences in toxicity between highly similar toxicants we focused on detoxification genes and the analysis of interacting environmental factors affecting them.

We identified a large number of detoxification genes significantly affected by the different treatments (Figure 1). CPF influenced the expression of a smaller number of genes than DZN. This suggests that the transcriptional regulation of detoxification genes may be concentration-related, since the CPF treatment had a lower concentration of the toxicant. It also strengthens the idea that secondary targets play a more important role in CPF toxicity, because CPF-oxon bioactivation depended on CYP and SDR and the toxicity level of CPF and DZN were equivalent despite the concentration differences. On the other hand, temperature had an important effect as an interacting variable. Variation in temperature seemed to modify the expression of many genes affected by the toxicants (CPF*Temp and DZN*Temp interactions). This was in agreement with toxicological studies in other species showing increased effect of both toxicants with an increase in temperature (Lydy *et al.*, 1999; Osterauer and Köhler, 2008; De Silva *et al.*, 2009). However, only a small number of the genes were affected by the simultaneous interacting effect of both toxicants and temperature, while the interaction between toxicants seemed to have the strongest influence on detoxification genes. In the particular case of CYP, both CPF and DZN compete for common enzymes (Timchalk and Poet, 2008) and regulate other common detoxification genes (Chapter 2). This might explain the elevated number of genes affected by interacting toxicants. Nevertheless, a group of genes seemed to have specificity in their responses to either CPF or DZN, with weakly or no influence by temperature.

Next to the main detoxification genes mentioned above, ATP-binding cassette transporters (ABC) and nuclear receptors (NR) are known to be involved in activation of detoxification

genes and transport of detoxification products (Tim H. Lindblom, 2006). ABC transporter genes are implicated in transport of metabolites during detoxification and deletion of some ABC transporter genes alter nematode sensitivity to many environmental toxicants (Mahajan-Miklos *et al.*, 1999). *C. elegans* has 60 ABC transporters (Sheps *et al.*, 2004), of which 26 were affected by our treatments. One of the ABC genes affected, encode for a P-glycoprotein related protein 3 (*pgp-3*). This *pgp* gene is required for the defense against *Pseudomonas aeruginosa* and response to other toxicants (Zhao *et al.*, 2004). It also belongs to the *pgp* family of transmembrane proteins involved in ATP transport. Other members of the family were also found to be affected. For example, *pgp-5*, *pgp-6*, *pgp-7* and *pgp-8* are all located in a four gene cluster on chromosome *X* as a result of tandem duplication (Zhao *et al.*, 2004). All of them, except *pgp-8,* had expression levels affected by some of the variables tested. Only *pgp-7* was significantly influenced by the interaction of toxicants and temperature. That is, in an exposure of both toxicants, temperature modified the expression response of *pgp-7*, but not the other genes in the genomic cluster. A reason for this different regulation might be in a tissue specific expression pattern of the *pgp* genes. *pgp-6* and *pgp-5* are known to be expressed in the intestine, a main responsive organ to toxicant exposure, while *pgp-8* and *pgp-7* are expressed in neurons and male tail rays (Zhao *et al.*, 2004). Since our experiment included only hermaphrodites and no males, *pgp-7* was most likely expressed in other tissues not yet identified. In summary, the expression of ABC transporter genes was more influenced by temperature than by toxicants.

Figure 1.- (next page) Effects of organophosphate pesticides (OP) on detoxification genes. Transport of OP inside the cell activates transcription of nuclear receptor (NR) genes and regulates expression of genes involved in detoxification. An initial phase of CPF and DZN detoxification starts with cytochrome P450 enzymes (CYP) and short chain dehydrogenases (SDR) enzymes transforming the toxicants into an oxygenated form called oxon-OP (oxon). This highly toxic intermediate metabolite is effectively detoxified by hydrolysis mediated by UDP-glucuronosyl transferases (UGT) and glutathione-S-transferases (GST) enzymes. The final inactive compound of CPF and DZN are 3,5,6-trichloro-2-pyridinol (TCP) and 2-isopropyl-4-methyl-6-hydroxypyrimidine (IMHP), respectively. ABC-transporters (ABC) transport metabolites during the detoxification. The graphs show the number of detoxification genes affected by each variable (Timchalk *et al.*, 2005; Tim H. Lindblom, 2006).


S
OP
S
OP
CYP
SDR
CYP
S
O
oxon
UDP
GST
TCP/IMHP
ABC
CPF
DZN
Temp
CPF*Temp
DZN*Temp
CPF*DZN
CPF*DZN*Temp
NR
SDR
CYP
ABC
GST
UDP
num. genes
CPF*DZN*Temp
CPF*DZN
DZN*Temp
CPF*Temp
Temp
DZN
CPF

The transcriptional induction of CYP and other detoxification genes is mediated by nuclear receptors (NR). Indeed, it was suggested that NR are an important mechanism in environmental changes response in *C. elegans* (Sluder *et al.*, 1999). We investigated the expression of 274 NR, of which 76 were significantly affected by at least one variable. For this group of genes, the most relevant variable affecting gene expression was the interaction between toxicants (CPF*DZN). Like other detoxification genes, this regulation may be a sign for a concentration-related response, and higher toxicant concentration might have promoted regulation of more pathways via NR activity. As an example, we highlight three NR, *daf-12*, *nhr-8* and *nhr-48*, all members of a conserved subfamily of NR (Gissendanner *et al.*, 2004). The first one, *daf-12* (abnormal DAuer Formation-12) is an essential NR for dauer arrest, a resistant form of *C. elegans* under stressful environmental conditions (Hu, 2007). Like many NR (50), its expression level was influenced by the interaction of both toxicants. On the other hand, *nhr-8* and *nhr-48* (Nuclear Hormone Receptors) had a different response to CPF treatment in the interaction with temperature (CFP*Temp). Besides, *nhr-48* was also affected by the interaction between DZN and temperature. Unfortunately, the specific function of *nhr-48* is unknown, but the function of *nhr-8* has already been associated to xenobiotics exposure (Lindblom *et al.*, 2001).

Temperature as variable regulating gene expression

It is widely known that temperature affects gene expression. As shown above, it also interacts with other environmental factors, like toxicants. To better understand the role of the temperature on the expression profiles of OP we extracted data from reference samples analyzed in Chapter 2 to compare them with the reference samples at 24ºC. Since both experiment used identical experimental conditions with only different temperatures, their comparison could provide information about temperature influence in gene expression. As a result of the analysis we identified 22.9% (4343) genes to be differentially expressed between temperatures (-log10 P-value = 3, 0.0073 FDR, Supplementary Table ST3). Compared to the genes affected by temperature in the previous analysis, 37.2% (720) genes overlap between groups. In other words, these genes were strongly influenced by temperature despite the toxicants treatments. A full GO analysis and domain enrichment analysis can be found in Supplementary Table ST4. In summary, significant GO terms

affected by temperature were related with development and growth like epidermis development, oogenesis, body morphogenesis, lipid transporter activity, monooxygenase activity and ATP binding, among others. Significantly enriched domains included lipid transport protein, Serine/Threonin protein kinase, protein kinase, and ABC transporters. Since time of growth and development in *C. elegans* is negatively correlated with temperature, it is not surprising that we found these functions affected by the treatments. Moreover, the identified GO terms related with detoxification, strengthen the idea that temperature is an important environmental factor influencing OP induced gene expression changes.

In previous sections, we highlighted NR as part of the detoxification genes. Nematodes have a surprisingly large number of genes coding for NR compared to other species (Sluder *et al.*, 1999). Yet, they do not seem to have a highly specific response to toxicants and bacteria (Antebi, 2006; Ewbank, 2006). For this reason, it has been suggested that NR play an important role in diversifying the responses to environmental factors. Here again, we selected two NR, *nhr-8* and *daf-12*, as examples to describe the complexity of gene-by-environment interactions. The first example is *nhr-8*, a xenobiotic responsive gene (Gissendanner *et al.*, 2004) with a significant interaction between CPF and the temperature. It also had a significant temperature effect when we analyzed only the temperature influence. One of its transcriptional targets is the ABC transporter *pgp-3* (Lindblom *et al.*, 2001). Although direct evidence for this relation is lacking, their similar pattern of expression suggested it (Tim H. Lindblom, 2006). We also found similarities in response to xenobiotics, suggesting as Lindblom *et al.* (2001) that both genes participate together in some environmental responses, but not in all of them. The spectra of environmental signaling triggering transcriptional regulation of both genes overlap but they were different. That is, *nhr-8* might not be the only regulator of *pgp-3* activity. In this sense, Li *et al*. employed a genetical genomics approach to identify polymorphic regulatory elements affected by temperature (Li *et al.*, 2006). They identified eQTLs at 16ºC and 24ºC that describe genotype-by-environment interactions in complex traits. Interestingly, they indentified a distant eQTL on chromosome *I* for *pgp-3* at 16ºC, but not for *nhr-8*. Since *nhr-8* is located on chromosome *IV*, this NR cannot be the polymorphic regulatory element of *pgp-3* that Li *et al*. identified at low temperature. Unfortunately, the genomic region

associated with this eQTL is too large to speculate about possible regulators of *pgp-3*. The second example is *daf-12*, a NR involved in dauer formation (Antebi, 2006). This gene was affected by the toxicants interactions, and by CPF and temperature interaction (-log10 P-value = 1.44, 0.0355 FDR). It also had a significant temperature influence in the independent analysis, but in combination of toxicant exposure the temperature influence was weak. The study from Li *et al*. detected an eQTL at 16ºC in chromosome *IV*. Again, it is difficult to speculate about the gene or genes responsible for *daf-12* regulation, but the accumulated data indicated that previously identified regulatory elements active at 16ºC were modified by the exposure to the toxicants.

Conclusions

We analyzed the transcriptional responses to interacting environmental stressors. Three factors were considered: the toxicants CPF and DZN, and temperature. We focused our analysis on detoxification genes because toxicity on OP can be partially explained by bioactivation of OP to highly toxic oxon forms mediated by detoxification enzymes. Our results indicated that detoxification genes may be influenced in a concentration-dependent manner, since more genes were regulated by DZN and by the mixture of toxicants than by CFP itself. On the other hand, temperature showed a strong effect on transcript abundance as influenced the expression of many genes, included those involved in detoxification. Moreover, a larger number of detoxification genes were significantly regulated by the interactions between temperature and the toxicants than by the OP alone. Indeed, it has been shown that toxicity of OP increased with temperatures (Osterauer and Köhler, 2008; De Silva *et al.*, 2009); accordingly, our results suggested that higher toxicity of OP with temperature might be a consequence of gene-environment interactions on detoxification genes. Those interactions are probably specific to the particular environmental factors here studied; because they may change their effect on gene expression under the influence of different concentrations of toxicants, different pesticides combinations or temperatures. However, they highlight the significance of transcriptional studies that focus on more environmental factors and aiming to better understand the full process behind gene-environment interactions.

Methods & Materials

C. elegans culturing

The Bristol N2 strain was cultured on standard nematode growth medium (NGM) with *E. coli* OP50 as food source. Nematodes were bleached (0.5 M NaOH, 1% hypochlorite) and eggs were collected and inoculated in (9 cm diameter) dishes with the toxicants. After 40 hours in 24 °C, L3-L4 stage nematodes were collected, frozen in liquid nitrogen and kept at -80C until the RNA extraction procedure.

Toxicant exposures

We applied concentration of the toxicants below the EC_{50} values as described in Chapter 2. Concentrations were 0.5 mg/l of CPF (Cyren®/Nufos®, Cheminova A/S [Lemvig, Denmark]) and 1.0 mg/l of DZN (Supelco [Bellefonte, Pennsylvania 16823, USA]). The combination of the two OP contained the exact sum of both single concentrations (DZN [1mg/l] and CPF [0.5 mg/l]).

The experiment started with eggs placed on NGM dishes with the toxicants and *E. coli* OP50 as food source. After 72 hours at 24 °C, worms from 4 petri-dishes were collected as one sample. A total of 6 replicates per treatment were collected (24 petri dishes), and immediately froze in liquid nitrogen until RNA extraction. All the OP were dissolved in acetone and added to 10 ml of NGM poured in each 9 cm petri dish used for the culture. Nematodes without treatments were grown simultaneously with similar concentrations of acetone in a control culture.

Microarrays

RNA from nematodes was extracted following the Trizol method, and the RNeasy Micro kit (Quiagen, Valencia, CA, USA) was used to clean up the samples. Labeled cDNA was produced with the kit Array 900 HS from Genisphere and Superscript II from Invitrogen. The 60-mers arrays were purchased by Washington University (http://genomeold.wustl.edu) and they were hybridized following the Genisphere Array 900 HS protocol with modifications. Extracts from CPF, DZN and the CPF/DZN combination exposures were hybridized with the control samples in each array. Six independent biological replicates were used per treatment to produce six replicate microarrays per experiment in a dye-swap design.

Microarray data have been deposited in Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo/), accession number GSE24257. Expression profiles of *C. elegans* exposure at 16°C were downloaded from GEO (GSE16719).

Microarray Analysis

A Perking & Elmer scanner was used to extract the raw intensities from the microarrays. The R software was used for preprocessing and normalization (www.r-project.org/) using Limma package (Smyth, 2005). The Loess method (Smyth and Speed, 2003) was used for normalization within arrays and normalization between arrays was done using aquantile method (Yang, 2003), both of them are included in the Limma package for R.

To identify the differentially expressed genes for each treatment we used a linear model (gene expression = CPF (effect) * DZN (effect) * Temperature (effect) + error). Therefore, we were able to estimate the influence of each factor on the measured variation in gene expression as well as the effect of any possible interaction between them. For threshold determination we used a permutation approach. For each of the 23,232 permutations used we randomly picked a transcript, which could only be picked once. We combined all the expression values of this transcript and randomly distributed them over the replicates (and experimental conditions) and used for the analysis. In this way we obtained a threshold for each of the explanatory variables. We used a –log10 P-value 2 as common threshold, which resembles to the following FDR per variable: 0.0096 for CPF, 0.0098 for DZN, 0.0098 for Temp, 0.0098 for CPF*Temp, 0.0098 for DZN*Temp, 0.0102 for CPF*DZN, 0.0099 for CPF*DZN*Temp.

Differentially expressed genes between temperatures were identified using single channel data from all the microarrays profiles. A total of 18 biological replicates per temperature were combined. We used a linear model (gene expression = Temperature (effect) + error) to identify differentially expressed genes. Level of significance was determined by permutation, as describe above. The genome-wide threshold was set to –log10 P-value = 2, which resembles to a FDR of 0.0108.

Detoxification genes

We identified genes involved in detoxification based on protein domains. For each category of enzymes or proteins referred in the text (CYP, SDR, UDP, GST, ABC) we selected functional domains related to their function in INTERPRO (www.ebi.ac.uk/interpro/). Information of functional domains in *C. elegans* genes was downloaded from Wormbase (www.wormbase.org) release WB195. In this way, we assumed a gene with a domain,

predicted or confirmed, related to detoxification is a candidate gene to be involved in the OP metabolism.
Domain IDs were: CYP (PF00067, IPR001128, IPR002397, IPR002401, IPR002402, IPR002403); SDR (PF00106, IPR002198); UDP (PF00201, IPR002213); GST (IPR010987, IPR004045, IPR004046, PF02798, IPR005442, IPR003082, PF00043); ABC (IPR003439, IPR011527, PF00005, IPR010509, PF06472).
Nuclear Receptors were selected from literature (Antebi, 2006).

Enrichment analysis

Gene Ontology (GO) data and functional domain data were extracted from Wormbase release WB195. GO terms and domains with less than 4 genes were discarded. Over-represented groups of GO terms and domains were identified using a hypergeometric test (P-value < 0.01). In this way we analyzed 396 unique GO terms and 1003 unique INTERPRO id numbers, form 16,947 and 8682 annotated genes, respectively.

Supplemental material

All the supplementary tables and figures are available from the author on http://www.anavinuela.com/thesis/thesis.html

Supplementary Table ST1: Significantly affected genes by each of the analyzed variables. Each worksheet contains the list of affected genes by each factor: chlorpyrifos (CPF), diazinon (DZN), temperature (Temp); and the interactions of them: CPF*Temp, DZN*Temp, CPF*DZN, CPF*DZN*Temp.

Supplementary Table ST2: List of enriched Gene Ontology (GO) terms and functional domain in the affected genes. Data were extracted from Wormbase release WB195. GO terms worksheets: Genes in GO database are the number of genes belonging to each GO term. GO number are the GO term. Regulated genes refer to number of significantly regulated genes in each treatment that belong to a GO term.
Functional domains worksheets: Genes with the domain are the number of genes that code for a specific functional domain. The number of significantly regulated genes in each treatment that code for a specific functional domain is listed in the regulated genes column.

Supplementary Table ST3: Significantly affected genes by temperature.

Supplementary Table ST4: List of enriched GO terms and functional domains in the genes affected by temperature.

References

Antebi, A. Nuclear hormone receptors in C. elegans (January 03, 2006), WormBook, ed. The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.64.1, http://www.wormbook.org

Bindesbøl, A. M., M. Bayley, et al. (2009). Impacts of heavy metals, polyaromatic hydrocarbons, and pesticides on freeze tolerance of the earthworm dendrobaena octaedra. *Environmental Toxicology and Chemistry* **28**(11): 2341-2347.

De Silva, P. M. C. S., A. Pathiratne, et al. (2009). Influence of temperature and soil type on the toxicity of three pesticides to Eisenia andrei. *Chemosphere* **76**(10): 1410-1415.

Eaton, D. L., R. B. Daroff, et al. (2008). Review of the Toxicology of Chlorpyrifos With an Emphasis on Human Exposure and Neurodevelopment. *Critical Reviews in Toxicology* **38**(1 supp 2): 1 - 125.

Ewbank, J. J. Signaling in the immune response (January 23, 2006), WormBook, ed. The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.83.1, http://www.wormbook.org

Gillooly, J. F., J. H. Brown, et al. (2001). Effects of Size and Temperature on Metabolic Rate. *Science* **293**(5538): 2248-2251.

Gissendanner, C. R., K. Crossgrove, et al. (2004). Expression and function of conserved nuclear receptor genes in Caenorhabditis elegans. *Developmental Biology* **266**(2): 399-416.

Holmstrup, M., A.-M. Bindesbøl, et al. (2010). Interactions between effects of environmental chemicals and natural stressors: A review. *Science of The Total Environment* **408**(18): 3746-3762.

Hook, S. E., A. D. Skillman, et al. (2008). Gene Expression Profiles in Rainbow Trout, Onchorynchus mykiss, Exposed to a Simple Chemical Mixture. *Toxicol. Sci.* **102**(1): 42-60.

Hu, P. J. Dauer (August 08, 2007), WormBook, ed. The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.144.1, http://www.wormbook.org

Lee, S.-J. and C. Kenyon (2009). Regulation of the Longevity Response to Temperature by Thermosensory Neurons in Caenorhabditis elegans. *Current biology : CB* **19**(9): 715-722.

Lewis, J., M. Szilagyi, et al. (2009). Distinct patterns of gene and protein expression elicited by organophosphorus pesticides in *Caenorhabditis elegans*. *BMC Genomics* **10**(1): 202.

Li, Y., O. A. A. Alvarez, et al. (2006). Mapping determinants of gene expression plasticity by genetical genomics in C-elegans. *Plos Genetics* **2**(12): 2155-2161.

Lindblom, T. H., G. J. Pierce, et al. (2001). A C. elegans orphan nuclear receptor contributes to xenobiotic resistance. *Current Biology* **11**(11): 864-868.

Lydy, M. J., J. B. Belden, et al. (1999). Effects of Temperature on the Toxicity of M-Parathion, Chlorpyrifos, and Pentachlorobenzene to *Chironomus tentans*. *Archives of Environmental Contamination and Toxicology* **37**(4): 542-547.

Mahajan-Miklos, S., M.-W. Tan, et al. (1999). Molecular Mechanisms of Bacterial Virulence Elucidated Using a Pseudomonas aeruginosa Caenorhabditis elegans Pathogenesis Model. **96**(1): 47-56.

Mumtaz, M. M., D. B. Tully, et al. (2002). Gene induction studies and toxicity of chemical mixtures. *Environmental Health Perspectives* **110**(SUPPL. 6): 947-956.

Nota, B., N. M. Van Straalen, et al. (2010). Gene expression microarray analysis of heat stress in the soil invertebrate Folsomia candida. *Insect Molecular Biology* **19**(3): 315-322.

Osterauer, R. and H.-R. Köhler (2008). Temperature-dependent effects of the pesticides thiacloprid and diazinon on the embryonic development of zebrafish (Danio rerio). *Aquatic Toxicology* **86**(4): 485-494.

Poet, T. S., H. Wu, et al. (2003). In Vitro Rat Hepatic and Intestinal Metabolism of the Organophosphate Pesticides Chlorpyrifos and Diazinon. *Toxicol. Sci.* **72**(2): 193-200.

Pope, C. N. (1999). Organophosphorus Pesticides: They all have the same mechanism of toxicity? *Journal of Toxicology and Environmental Health Part B: Critical Reviews* **2**: 161-181.

Savage-Dunn, C. TGF-ß signaling (September 9, 2005), WormBook, ed. The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.22.1, http://www.wormbook.org

Scheil, V. and H.-R. Köhler (2009). Influence of Nickel Chloride, Chlorpyrifos, and Imidacloprid in Combination with Different Temperatures on the Embryogenesis of the Zebrafish <i>Danio rerio</i&gt. *Archives of Environmental Contamination and Toxicology* **56**(2): 238-243.

Sheps, J., S. Ralph, et al. (2004). The ABC transporter gene family of Caenorhabditis elegans has implications for the evolutionary dynamics of multidrug resistance in eukaryotes. *Genome Biology* **5**(3): R15.

Slotkin, T. A., E. A. MacKillop, et al. (2007). Screening for developmental neurotoxicity using PC12 cells: Comparisons of organophosphates with a carbamate, an organochlorine, and divalent nickel. *Environmental Health Perspectives* **115**(1): 93-101.

Slotkin, T. A. and F. J. Seidler (2007). Comparative developmental neurotoxicity of organophosphates in vivo: Transcriptional responses of pathways for brain cell development, cell signaling, cytotoxicity and neurotransmitter systems. *Brain Research Bulletin* **72**(4-6): 232-274.

Sluder, A. E., S. W. Mathews, et al. (1999). The Nuclear Receptor Superfamily Has Undergone Extensive Proliferation and Diversification in Nematodes. *Genome Research* **9**(2): 103-120.

Smyth, G. K. (2005). Limma: Linear models for microarray data. *Bioinformatics and Computational Biology Solution Using R and Bioconductor*. New York, Springer**:** 397-420.

Smyth, G. K. and T. Speed (2003). Normalization of cDNA microarray data. *Methods* **31**(4): 265-273.

Sørensen, J. G., M. M. Nielsen, et al. (2005). Full genome gene expression analysis of the heat stress response in *Drosophila melanogaster*. *Cell Stress and Chaperones* **10**(4): 312-328.

Spurgeon, D. J., O. A. H. Jones, et al. (2010). Systems toxicology approaches for understanding the joint effects of environmental chemical mixtures. *Science of The Total Environment* **408**(18): 3725-3734.

Tim H. Lindblom, A. K. D. (2006). Xenobiotic detoxification in the nematode *Caenorhabditis elegans*. *Journal of Experimental Zoology Part A: Comparative Experimental Biology* **305A**(9): 720-730.

Timchalk, C. and T. S. Poet (2008). Development of a physiologically based pharmacokinetic and pharmacodynamic model to determine dosimetry and cholinesterase inhibition for a binary mixture of chlorpyrifos and diazinon in the rat. *NeuroToxicology* **29**(3): 428-443.

Timchalk, C., T. S. Poet, et al. (2005). Pharmacokinetic and pharmacodynamic interaction for a binary mixture of chlorpyrifos and diazinon in the rat. *Toxicology and Applied Pharmacology* **205**(1): 31-42.

Yang, Y. H., and Thorne, N. P. (2003). Normalization for two-color cDNA microarray data. *D. R. Goldstein (ed.), Science and Statistics: A Festschrift for Terry Speed* **40**: 403-418.

Zhao, Z., L. L. Fang, et al. (2004). ATP-binding cassette protein E is involved in gene transcription and translation in Caenorhabditis elegans. *Biochemical and Biophysical Research Communications* **323**(1): 104-111.

Chapter: 4

Gene expression regulation as a function of genotype and age

Ana Viñuela & L. Basten Snoek, Joost A.G. Riksen, Jan E. Kammenga

Published as:

Ana Viñuela[1], L. Basten Snoek[1], Joost A.G. Riksen, Jan E. Kammenga.

Genome-wide gene expression regulation as a function of genotype and age in C. elegans.

Genome Research (20): 929-937. (2010)

[1]These authors contributed equally to this work

Abstract

Gene expression becomes more variable with age and it is widely assumed that this is due to a decrease in expression regulation. But currently there is no understanding how gene expression regulatory patterns progress with age. Here we explored genome-wide gene expression variation and regulatory loci (eQTL) in a population of developing and aging *Caenorhabditis elegans* recombinant inbred worms. We found almost 900 genes with an eQTL of which almost half were found to have a genotype-by-age effect ($_{gxa}$eQTL). The total number of eQTL decreased with age whereas the variation in expression increased. In developing worms, the number of genes with increased expression variation (1282) was similar to the ones with decreased expression variation (1328). In aging worms the number of genes with increased variation (1772) was nearly 5 times higher than the number of genes with a decreased expression variation (373). The number of *cis*-acting eQTL in juveniles decreased by almost 50% in old worms whereas the number of *trans*-acting loci decreased by ~27%, indicating that *cis*-regulation becomes relatively less frequent than *trans*-regulation in aging worms. Of the 373 genes with decreased expression level variation in aging worms, ~39% had an eQTL compared to ~14% in developing worms. $_{gxa}$eQTL were found for ~21% of these genes in aging worms compared to only ~6% in developing worms. We highlight three examples of linkages: in young worms (*pgp-6*), in old worms (*daf-16*) and throughout life (*lips-16*). Our findings demonstrate that eQTL patterns are strongly affected by age and suggest that gene network integrity declines with age.

Introduction

The expression level of most genes changes with age in all species studied so far. Some genes are expressed during development whereas others are expressed during or after the reproductive period (Hill *et al.*, 2000; Lund *et al.*, 2002; Ramoni *et al.*, 2002; Geigl *et al.*, 2004; Zhan *et al.*, 2007). For example, McCarroll *et al*. (2004) studied the temporal dynamics of gene expression throughout the life-cycle in *C. elegans* and *D. melanogaster* (McCarroll *et al.*, 2004). They reported that genes involved in mitochondrial metabolism, DNA repair, catabolism, peptidolysis and cellular transport are expressed during adulthood, whereas other transcriptional signatures were observed during larval development, embryogenesis and gametogenesis.

But next to the expression level, the *variation* in gene expression also changes with age. For instance Somel *et al*. (2006) reported that gene expression becomes heterogeneous with age in humans and rats. The age-correlated heterogeneity of gene expression had a minor effect on individual genes, but was widespread throughout the transcriptome (Somel *et al.*, 2006). These, and other studies, have led to the assumption that regulation of gene expression decreases with age leading to more variation in gene expression (Lu *et al.*, 2004; McCarroll *et al.*, 2004). But currently there is no understanding how genome-wide gene expression regulatory patterns progress with age.

Genetic mapping of gene expression regulatory loci (eQTL) provides a means to address this problem (Jansen and Nap, 2001; Rockman and Kruglyak, 2006). eQTL can be *cis*- or *trans*-acting, reflecting local and distant regulation of gene expression respectively. They have been thoroughly studied in humans, yeast, plants, rats, mice and worms (Brem *et al.*, 2002; Schadt *et al.*, 2003; Li *et al.*, 2006; Petretto *et al.*, 2006; Keurentjes *et al.*, 2007; Farber *et al.*, 2009). But, as far as we are aware of, all eQTL studies only focused on one time point or age group whereas most complex traits and diseases such as cancer are age-related. Surprisingly, we have no understanding of the dynamics of eQTL throughout the life of an organism, although it is long known that gene expression itself is strongly affected by age.

Here we present the first genome-wide study of heritable gene expression regulation as a function of age using a set of 36 single nucleotide polymorphisms (SNP)-genotyped

recombinant inbred lines (RILs) (Supplemental Table ST1) derived from *C. elegans* wildtypes N2 and CB4856 (Li *et al.*, 2006). The CB4856 x N2 RIL population displays large genetic and phenotypic differences for a wide range of traits such as reproduction, growth (Gutteling *et al.*, 2007; Kammenga *et al.*, 2007), gene expression (Li *et al.*, 2006) and copulatory plug formation (Palopoli *et al.*, 2008), to name a few. We used microarrays to measured genome-wide gene expression from the RILs reared at 24°C at three different ages: young worms, t1: age was 40 hrs; reproductive worms, t2: age was 96 hrs, and old worms, t3: age was 214 hrs. In this way we obtained per gene 108 observations perturbed by age and genotype. First we investigated whether genome-wide gene expression was affected by age across all recombinant genotypes. Then we asked if age-specific gene expression can be attributed to regulatory loci (eQTL) and whether eQTL patterns change during aging.

Results

Variation in gene expression level increases in aging worms

First, we assessed the influence of age on gene expression levels across all RILs. We found that more than 14,000 gene transcript levels, ~75% of all genes, were affected by age (P-value ≤ 0.0001). This proportion is congruent with findings by Jiang *et al*. (2001) who reported ~69% of all genes being dynamic in the larval stages of wildtype N2 worms. Yet, we demonstrate that dynamic patterns of gene expression proceed at older ages for many different genotypes. We could identify 8 different dynamic expression patterns, each containing ~1200 or more genes (See Figure 1 and Table 1, and Supplementary Table ST2). Most genes (2688) were found in the pattern in which expression increases during reproduction and was significantly higher during the reproductive and senescent stage compared to juveniles. Together these results confirm that age is a most influential factor affecting gene transcript levels in *C. elegans*.

Gene ontology (GO) analysis of the 8 expression patterns are significantly enriched with different GO terms (Supplemental Table ST3). Patterns B, F and H are enriched for genes involved in embryonic development. These patterns have in common that they all have

genes of which the expression goes up after the juvenile stage. The expression of the genes in these patterns remains high or even goes up after reproduction. This suggests that these genes might have different functions later in life, or that their expression is no longer suppressed, i.e. less regulated in old worms compared to young worms. However, this assumes the existence of a mechanism for suppression. An alternative explanation is that these genes are not suppressed, but induced during reproduction and there is no efficient way to turn off the induction.

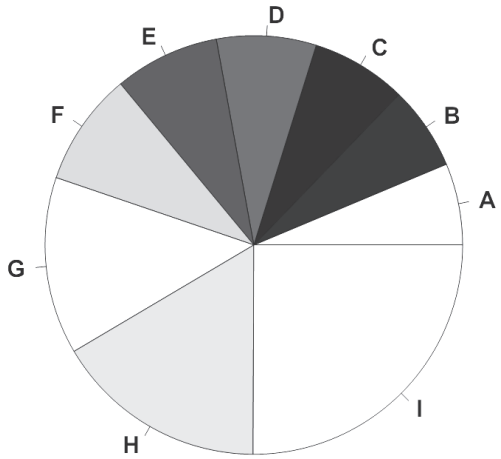


Figure 1.- Pie graph of age affected genes. The pie graph show the number of genes with an age dependent expression (P-value ≤ 0.0001; FDR see methods and materials)

Table 1.- Number of genes with an age dependent expression.

Group	Direction in developing worms	Direction of expression in aging worms	Number of genes	Percentage of total
A	Down	Down	1188	6.30
B	-	Up	1217	6.46
C	Down	Up	1394	7.40
D	Up	Down	1454	7.72
E	-	Down	1534	8.14
F	Up	Up	1651	8.76
G	Down	-	2580	13.68
H	Up	-	3101	16.46
I	-	-	4720	25.05

Table 1.- Number of genes with an age dependent expression. Categories were determined according to direction of change in gene expression: increased expression (up), decreased expression (down) or no change (-).

Patterns C and D are enriched for genes involved in metabolic and catabolic processes, respectively; E for dephosphorylation, I for transport, and G for locomotion. Moreover, we compared our list of genes in these expression patterns to a list of innate immunity genes described by Capra *et al*. (2008). They investigated gene expression in worms during development until young reproductive stage in the two parental wildtypes, N2 and CB4856, and reported that differentially expressed genes were enriched with genes involved in innate immune responses (Capra *et al.*, 2008). A significant enrichment of innate immuntiy genes in patterns C, D and I was found (P-value < 0.0001). Those patterns have in common that they represent genes with lower expression in old worms than in reproductive ones.

Second, we investigated which genes had increased or decreased *variation* in expression levels during aging. This was done for both the differences between young and reproductive worms (called developing worms) and between the reproductive and old worms (called aging worms). In developing worms the transcript level variation of a similar number of genes increased and decreased (1282 and 1328, respectively). However, after reproduction, in aging worms the number of genes with an increased variation was almost 5 times larger than the group with decreased variation (1772 and 373, respectively). GO term analysis revealed that this group of genes are mainly related with protein kinase activity while genes with decreased variation showed enrichment of monooxygenase activity and transferase activity (Supplemental Table ST4).

Regulatory loci of gene expression depend on age

To investigate if regulatory loci affecting gene expression also depended on age, we first mapped eQTL for each of the three age classes separately. In Figure 2 the upper panels show the peak position of the eQTL (*x-axis*) against the genomic position of genes having an eQTL (*y-axis*) at each age group. Overall, the number of eQTL decreased with age; we detected 978, 864, 577 eQTL in young, reproductive and old worms, respectively. The lower panels of Figure 2 show the frequency of eQTL per marker for each age. A distinction was made between *cis*- and *trans*-acting loci to show that both types of heritable regulation are very dynamic throughout the life of *C. elegans*. We detected 685, 642 and 373 *cis*- and 293, 222 and 204 *trans*-acting for the young, reproductive and old worms,

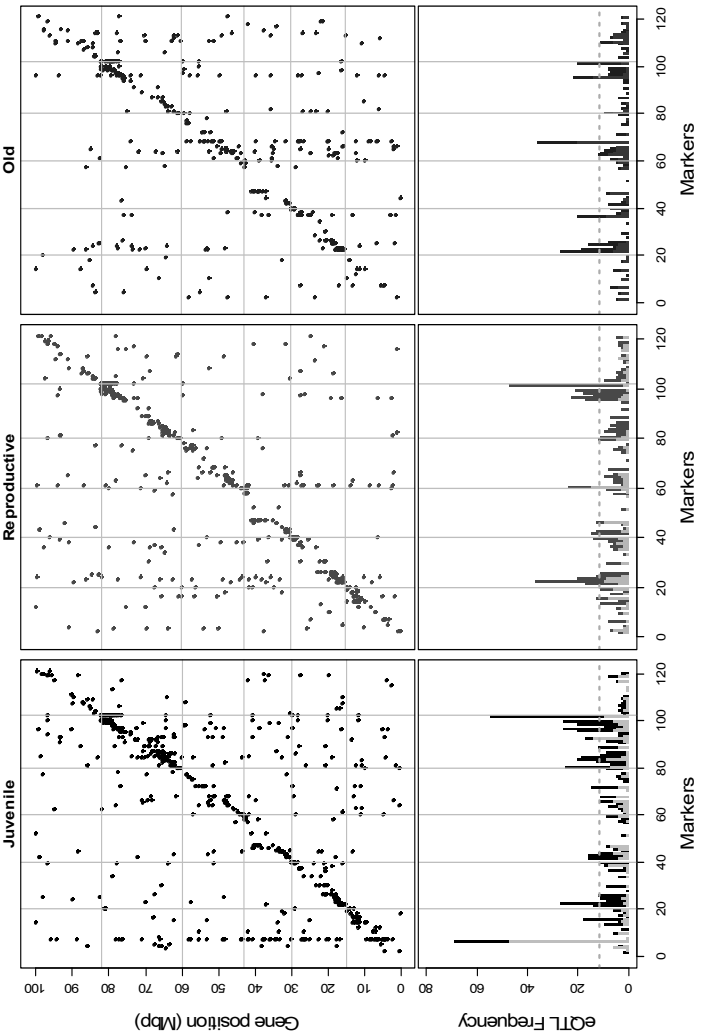


Figure 2.- Heritable transcript-level differences at three different time-points. The time-points are juvenile, reproductive and old worms. The upper graphs how the position of the genes with an eQTL (-log10 P-value > 3.8 ; FDR=0.01) on the *y-axis* and on the *x-axis* the marker at which the peak of the eQTL was found. From left to right: young worms in black, late reproductive worms in green and old worms in red. The lower graphs show the distribution of eQTL per marker. Height of the bar is the total number of eQTL at that marker, *cis*-acting eQTL are indicated in dark and *trans*-acting in the lighter color. The constitutive and age specific hotspots can easily be identified by comparing the three panels. Any deleted genes in CB4856 were left out of the analysis. The horizontal yellow lines are the *trans*-band thresholds to 0.01 as determined by permutation.

respectively. At all age-groups a major part of the *cis*-eQTLs were caused by genes deleted from the CB4856 genome (Maydan *et al.*, 2007), ~24%, ~30% and ~31% for the young, reproductive and old worms, respectively. To analyze the *cis*- and *trans*-eQTL frequency at the three age-groups we left out the CB4856 deleted genes, because of their high probability to generate a *cis*-eQTL. In this case, the number of *cis*-acting eQTL decreased in old worms with ~50% (compared to juveniles) whereas the number of *trans*-acting loci decreased with ~27%. This indicates that *cis*-regulation becomes relatively less frequent than *trans*-regulation in aging worms.

A two-time-point model identifies genotype-by-age interaction eQTL

Previously we found that post-reproductive lifespan varied strongly among N2 x CB4856 recombinant lines with some lines living almost twice as long than others (Doroszuk *et al.*, 2009). We found similar differences in lifespan among the RILs used in this study (Figure 3).

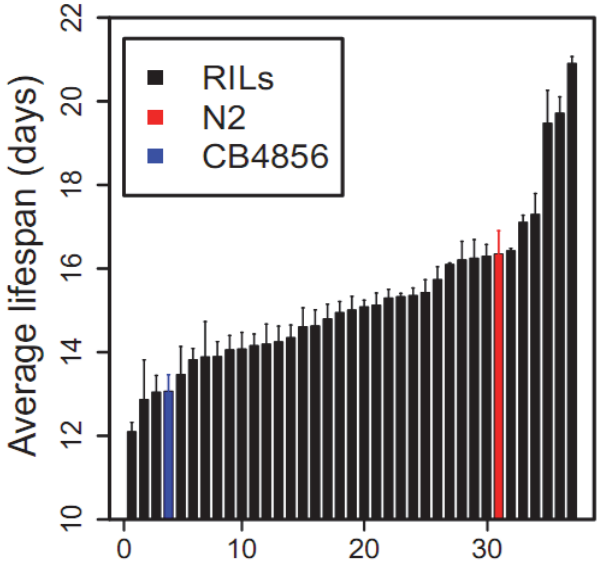


Figure 3.- Rils lifespan.. Average lifespan in days of the wildtypes N2 x CB4856 and the derived RILs grown at 24°C. Bars are SD.

This variation in lifespan across the RILs, ranged from a minimum of approx. ~13 days to a maximum of ~20 days. The average lifespan for some of the RILs transgressed beyond the mean value of either parental strain (N2: ~16 days; CB4856: ~13 days). This is often observed for many phenotypic traits in recombinants derived from genetically divergent

parental strains (Kammenga *et al.*, 2007). The differences among RILs means that, after reproduction, the physiological age differs between the RILs and that sampling mRNA at an older post-reproductive age, the third time point t3, would implicitly result in different physiological stages being sampled. Therefore we analyzed gene expression in relation to relative age, which we defined for each RIL as the age at the time of mRNA extraction divided by the average lifespan of that RIL. In this way the age-physiological differences among the RILs were taken into account when comparing the gene expression profiles.

In order to find the causal loci for heritable differences in transcript levels and possible interactions between age and genotype, we applied a two-time-point model. In this model we used three factors, i) relative age, ii) genotype (marker) and iii) the interaction between i) and ii), to explain the differences in gene expression between RILs and age-groups. With this mapping procedure we found almost 900 genes that had an eQTL or $_{gxa}$eQTL in developing and/or aging worms (P-value < 0.0001; Figure 4) Almost half of these genes with heritable transcript differences were found to have a genotype-by-age effect (396 at P-value < 0.0001; Table 2) allocated to a specific marker, which we coined genotype-by-age expression-QTL ($_{gxa}$eQTL). One specific hotspot (*trans*-band) for $_{gxa}$eQTL was found on chromosome *IV* for aging worms and a *trans*-band for eQTL on chromosome *I* was detected in developing worms (Figure 4).

Table 2.- Number of eQTLs

	Developing worms	**Aging worms**	**Overlap**
eQTL	460	362	83
$_{gxa}$eQTL	267	175	46
Overlap	134	108	

Table 2.- Type of eQTL (P-value < 0.0001 = -log10 P-value ≥ 4) at different ages calculated with the two-time point model. Significant linkage to either marker (eQTL) or the genotype-by-age interaction ($_{gxa}$eQTL) is indicated as well as the overlap between groups. In total we identified 893 unique genes with at least one significant feature.

Increased gene expression variation in old worms co-occurs with a lack of regulatory loci

Increased variation of gene expression during aging is often assumed to be the result of decreased gene expression regulation (Lu *et al.*, 2004; McCarroll *et al.*, 2004). We found

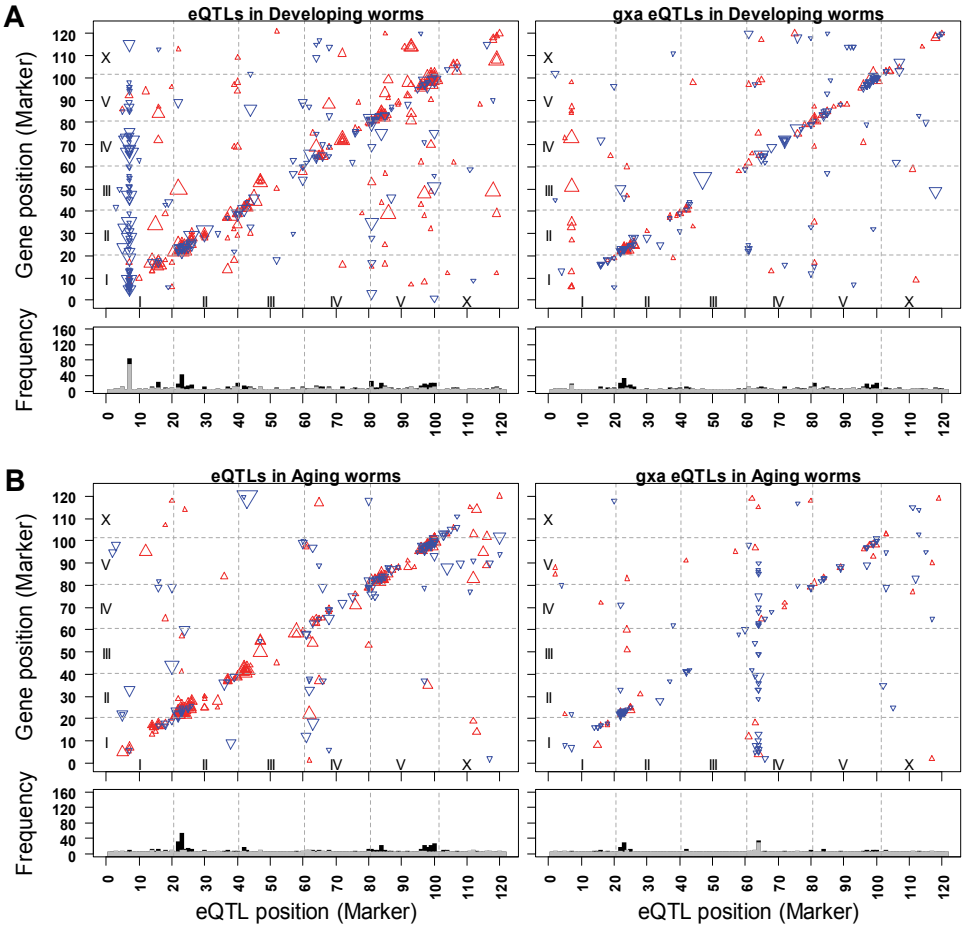


Figure 4 .- eQTLs location. Gene position (*y-axis*) plotted against the marker (*x-axis*) of the peak of the expression-Quantitative Trait Locus (eQTL) (P-value ≤ 0.0001) of genotype-by-age eQTL ($_{gxa}$eQTL) for that gene (4**A**: developing worms, 4**B**: aging worms). The size of the triangle is relative to the positive effect of the allele, in red for N2 effect, in blue for CB4856 effect. Chromosomes are separated by the thin grey lines, chromosome *I* is located at the left/bottom side of the panels and chromosome *X* at the right/top of the panels. **Developing; eQTL** (left upper panel) shows the significant (independent from age) marker (eQTLs) for the developing worms (the juvenile and reproductive worms in one model). **Aging; eQTL** (right upper panel) shows the significant (independent from age) marker (eQTLs) for the aging worms (the reproductive and old worms in one model). **Developing; $_{gxa}$eQTL** (left lower panel) shows the significant genotype (marker) by age effects ($_{gxa}$eQTL) for the developing worms (the juvenile and reproductive worms in one model). **Aging; $_{gxa}$eQTL** (right lower panel) shows the significant genotype (marker) by age effects ($_{gxa}$eQTL) for the aging worms (the reproductive and old worms in one model). Two distinct *trans*-bands (hotspots) can be observed in the **developing; eQTL** and in the **aging; $_{gxa}$eQTL** showing that hotspots can be found for genotype independent eQTL as well as genotype-by-age interaction $_{gxa}$eQTLs.

that genes with decreased expression level variation had a much higher percentage of eQTL (i.e. regulatory loci) and $_{gxa}$eQTL in aging worms compared to developing worms (39.1% and 13.5%, respectively for eQTLs and 20.6% and 6.4%, respectively for $_{gxa}$eQTLs). We found hardly any eQTLs or $_{gxa}$eQTL in the group of genes that had increased expression level variation in aging worms (Table 3; Supplemental Figure SF2). Although we found a small but significant group that showed less transcript variation and increased heritable regulation at old age (Supplemental Table ST5 and Supplemental Figure SF2), these results show that a large group of genes are more variably expressed at old-age and had a lack of eQTL which indicates a reduction of heritable regulation.

Table 3.- Relation between transcript variation and number of eQTLs at different ages

	Developing worms			**Aging worms**		
	Total genes	**Genes with an eQTL**	**Genes with a $_{gxa}$eQTL**	**Total genes**	**Genes with an eQTL**	**Genes with a $_{gxa}$eQTL**
Total		460	267		362	175
Genes with an **increase** in transcript variation	1282	1 (0.08%) P-value = 1 (under-rep P-value < $3*10^{-13}$)	102 (7.96%) P-value < $3*10^{-51}$	1772	0 (0.00%) P-value = 1 (under-rep P-value < $3*10^{-16}$)	21 (1.18%) P-value ~ 0.10
Genes with a **decrease** in transcript variation	1328	179 (13.5%) P-value < $8*10^{-89}$	85 (6.40%) P-value < $1*10^{-34}$	373	146 (39.1%) P-value < $7*10^{-162}$	77 (20.6%) P-value < $5*10^{-87}$

Table 3.- Genes with a significantly changing transcript level variance in developing and/or aging worms. The number of genes having an eQTL or an $_{gxa}$eQTL are indicated per group (number as well as the percentage of the total genes with an increased or decreased of variation). Also the significance of the over- (or under) representation of (gxa) eQTLs in those groups. The total number of eQTL and $_{gxa}$eQTL can be found in the row indicated with Total (also in Table 2)

Examples of genes having a $_{gxa}$eQTL

We show three examples of genes which have characteristic heritable expression patterns (Figure 5). These examples show i) differential gene expression in young worms, ii)

differential gene expression in old worms and iii) differential gene expression throughout the worm's life.

Example i) was found for *pgp-6* for which we detected a *trans*-acting $_{gxa}$eQTL at chromosome *V* in developing worms. The $_{gxa}$eQTL disappeared in the aging worms. The CB4856 allele had a much lower expression in young worms. In reproductive worms the expression of the CB4856 allele was strongly up-regulated and climbed to "N2"-levels. After reproduction, in old worms, the expression of both alleles decreased. A smaller, yet significant eQTL was also detected on chromosome *II* in aging worms. Example ii) was found for the forkhead family transcription factor encoding gene *daf-16* which had an $_{gxa}$eQTL on chromosome *IV* in aging worms. Example iii) was found for *lips-16*: this *cis*-acting $_{gxa}$eQTL was found at all age-groups although the difference between N2 and CB4856 was largest by the end of the reproductive phase. Both alleles showed a similar pattern but the CB4856 allele of this gene had a much higher expression throughout the worm's life. The $_{gxa}$eQTL in this case is caused by a much steeper increase in developing worms and a steeper decrease in aging worms of the CB4856 allele expression.

Discussion

We showed that variation in genome-wide gene expression increases in aging worms. This agrees with findings by Bahar *et al*. (2006) who reported increased cell-to-cell variation in gene expression in aging mouse heart (Bahar *et al.*, 2006). Somel *et al*. (2006) used a number of microarray gene expression tissue data sets from humans and rats, and found that gene expression becomes more variable with advancing age. Recently, a meta-analysis of age-related gene expression profiles with microarrays including 27 datasets from mice, rats and humans revealed evidence of up-regulation of gene transcription repression and negative regulation of transcription suggesting that transcriptional activity decreases with age (de Magalhaes *et al.*, 2009). Southworth *et al*. (2009) used a co-expression network approach to detect age-dependent differences in mice based on AGEMAP, a large DNA microarray study of gene expression as a function of age (Southworth *et al.*, 2009). They reported a decrease in gene co-expression in old mice and identified potential transcriptional mechanisms as a possible cause underlying the correlation decline.

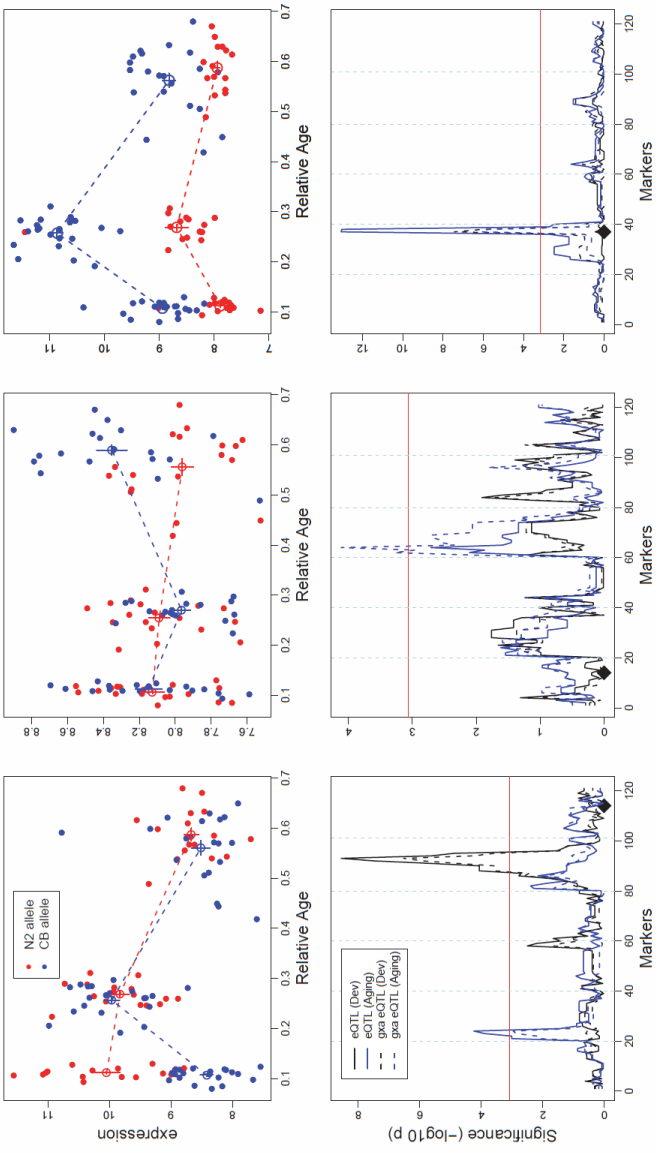


Figure 5.- Three gene examples where gene expression depends on age, genotype and their interaction. Upper panels shows the relative expression (*y-axis*) of the N2 allele (in red) and the CB4856 allele (in blue) of all the RILs at the peak of the major eQTL average and standard errors. Relative age (*x-axis*) is the time at which the specific RIL was sampled divided by the average age of that RIL. **Lower panels** shows the eQTL profiles of the different factors affecting gene expression. Gene position indicated by a black diamond. The eQTL profile obtained with juvenile/reproductive age-groups are in black. Those found with the reproductive/old age-groups are in blue. Independent (from age) genotype effects are in solid lines. Genotype-by-age interaction effects are in dotted lines. Chromosomes are separated by vertical dotted lines. X-axis: marker positions, chromosome *I* is on the left, chromosome *X* on the right. Y-axis: significance level (-log10 P). **Left)** *pgp-6*, **Middle)** *daf-16*, **Right)** *lips-16*. Thresholds are designated by the red horizontal line.

We provide for the first time insight into the regulatory loci of gene expression as a function of age and show that the number of eQTL decreases with age in *C. elegans*. As an extension of the three single age-group models we used a two-time-point model in which we tested if a time and/or marker (genotype) and/or an interaction between the two is present. If an eQTL is present at both time points, a significant marker effect will be found. If the eQTL is only present at one time point we will find a significant marker-by-time (GxA) interaction effect. In fact this approach increases the detection power in respect to eQTL appearing or disappearing at the different ages. Therefore we could work with a very strict significance threshold (P-value < 0.0001; FDR=0.006). Furthermore, the same conclusion could be drawn when we repeated our analysis at P-value < 0.001. This will not guarantee that eQTLs at all age groups are detected for every individual gene but our general conclusions will not depend on the thresholds used.

When juveniles were compared to old worms, the decline in number of *cis*-acting eQTL was stronger than the decline of *trans*-acting loci. Previous studies have shown that *cis*-acting linkages are less prone to variation induced by different environments (Li *et al.*, 2006) and tissues (Petretto *et al.*, 2006). We demonstrate that the proportion of *cis*-acting linkages is not stable during aging but decreases. Apparently aging influences the efficacy of *cis*-regulatory sites for controlling local gene expression. If we assume that the *trans*-acting loci influence gene expression levels via protein intermediates, as suggested by Smith and Kruglyak (2008), than these protein regulators become relatively more influential during aging (Smith and Kruglyak, 2008). In contrast to *cis*-regulation, *trans*-acting intermediates can affect the expression of many genes, thus allowing a coordinated gene expression response. Our results show that, although variation in genes expression increases at old ages and the number of eQTL decreases, there is a substantial contribution of *trans*-acting loci at old ages. This suggests that studies focusing on regulatory patterns underlying complex traits at older ages, such as progressive diseases, should focus on *trans*-acting loci late in life rather than *cis*-acting.

The *trans*-band for eQTL in developing worms could be due to segregation distortion which results in unequal marker distribution at this locus on chromosome *I* (Li *et al.*, 2006). This distortion is caused by a partial genome incompatibility as described in Seidel *et al.*, 2008. However, if this was a false positive caused by population structure it would also be

present in aging worms, were it was absent. Although one needs to be cautious with the conclusions about this locus, an important/influential polymorphic regulator in developing worms could be located at this genomic region on chromosome *I*. The *trans*-band for $_{gxa}$eQTL in aging worms on chromosome *IV* shows the genomic position of a polymorphic regulator that becomes more influential as the worms age. Li *et al*. (2006) found a temperature dependent *trans*-band on chromosome *V* when they mapped eQTLs at two different temperatures, 16°C and 24°C. The worms grown at 24°C are physiologically older than those grown at 16°C, although their gene expression profiles were measured at the same stage. A re-examination by scoring the frequency of eQTLs per marker in this experiment revealed a second *trans*-band at the top of chromosome *IV* in 24°C grown worms, but not in the 16°C grown worms. This *trans*-band co-locates with our $_{gxa}$eQTL *trans*-band in aging worms. This suggests that specifically physiological age dependent gene expression is affected by a polymorphic gene on the top of chromosome *IV*. Together, our results show that the increase in transcript variation in aging worms is due to a lack of large-effect regulatory loci. Alternatively, this increased expressionvariation could also be the result of an increase of small-effect eQTL. To get an indication for multiple small effect regulators we investigated if the number of genes with two eQTLs would be larger in aging worms. This was not the case (Supplemental Table ST6) and suggests that a lack of heritable regulation instead of more regulators is causal to the larger variation in a gene expression levels at older age.

The results allow us to infer the genetic architecture of heritable gene regulation with age. We highlighted three genes, *pgp-6, daf-16* and *lips-16*, each displaying different differential expression patterns during development and aging. Figure 5 illustrates that *pgp-6* is under the control of different factors at different ages. We detected a major eQTL on chromosome *V* in developing worms. We also detected a small but significant eQTL at chromosome *II* in aging worms. This indicates that the expression of a gene can be regulated by different regulators depending on the age of a worm. In *C. elegans, pgp-6,* an ATP-binding cassette (ABC) transporter is expressed in the larval and adult intestine and adult amphids in N2 (Zhao *et al.*, 2004). ABC transporters are one of the largest protein families in almost all species. They bind ATP and use its energy for transporting various molecules across membranes of organelles, such as the endoplasmic reticulum, peroxisomes, and

mitochondria (Dean and Allikmets, 1995). Mutation of ABC transporters can result in various diseases in humans or in hyper-sensitivity to drugs. Moreover, *pgp*-genes have been shown to be involved in multi-drug resistance (Mahajan-Miklos *et al.*, 1999) and resistance against *Pseudomonas aeruginosa* (Kurz *et al.*, 2007). In our experiment the expression level of *pgp-6* is much higher in N2 compared to CB4856 in young worms and decreases in older worms. Our results imply to study the regulatory control of *pgp-6* conferring differential resistance between N2 and CB4856 in young worms and its *trans*-eQTL on chromosome *V*.

In contrast to *pgp-6*, expression of *daf-16* is differentially controlled in old worms and not in younger ones. Its $_{gxa}$eQTL maps to the top of chromosome *IV*, were a *trans*-regulatory band for $_{gxa}$eQTL was found in aging worms. *daf-16* controls the insulin/IGF-1 signaling pathway underlying lifespan determination in *C. elegans*. This pathway operates during adulthood affecting lifespan and not during the juvenile period (Dillin *et al.*, 2002). Nevertheless, *daf*-16 cannot be the causal gene for this *trans*-regulatory band in aging worms since it is not physically located at the genomic position of the *trans*-regulatory band. Still, it is intriguing to think that a polymorphic regulator operating through the *daf*-16 pathway must be present at this locus.

The third example was provided for *lips-16*, encoding a tri-acylglycerol lipase. Lipases break down lipids and possible mobilization of stored tri-acylglycerides. Our results show that in both wildtypes N2 and CB4856, *lips-16* is more expressed in reproductive worms compared to juveniles and old worms. But, CB4856 displays a higher *lips-16* expression throughout its life than N2. This suggests that lipid catabolism mediated by *lips-16* is more intense in CB4856 than in N2.

Although it is tempting to search the detected eQTL for candidate regulators of these three transcripts, the current study was not designed for this purpose, but this would be a likely next step. The primary goal was to explore the landscape of regulatory loci during aging rather than pinpointing potential causal genes. Future mapping studies, using N2 x CB4856 introgression lines (Doroszuk *et al.*, 2009) and complementation studies (Kammenga *et al.*, 2007) would be most suitable for finding the causal polymorphic genes at the regulatory loci.

In our two-point-model we corrected for physiological differences among the RILs during aging because we assumed different physiological states associated with post/reproductive ages. This raises the question as to whether the explanatory variable here is age in a continuous sense or physiological state in a categorical sense. As pointed out by Collins *et al*. (2008) it is important to make a distinction between age-related changes that represent development and growth and age-related changes that represent senescence. A criteria they put forward is the life-cycle stage during which the change occurs. But development and senescence are arbitrary definitions of two periods within the same continuous process, namely aging. The lack of heritable regulation later in life is an age-related change, since we can associate this process to old worms. The physiological age is related with genotype-by-environment interactions, and the absolute age (in days) is an arbitrary measurement, as it is time. We measured gene expression in juveniles, late reproductive and post-reproductive old worms. Clearly these are defined stages, each with their own physiology. Correcting for these physiological differences allows for a more precise measurement related to age rather than stage. Yet, there is no clear moment during the life-cycle that clearly separates development from aging, indeed it seems very likely that developmental changes and aging changes occur simultaneously (Collins *et al.,* 2008).

Our results show that the heritable regulation of genome-wide gene expression decreases with age. What could be the reason why the regulation is no longer genetically determined in old worms for most genes but is stronger for a small group of genes? One explanation might be found in our GO analysis which revealed that monooxygenase activity was highly significant within the group of genes that displayed decreased variation in old worms. This points at increased regulation of stress pathways. Regulation of the functions of monooxygenase enzymes and their catalytic reactions has important significance for cell homeostasis. Halaschek-Wiener *et al*. (2005) identified longevity-associated genes in a long-lived *C. elegans daf-2* (insulin/IGF receptor) mutant using serial analysis of gene expression and reported that expression of metabolism-associated genes diminished and stress response genes increased during aging (Halaschek-Wiener *et al.*, 2005).

Our work provides new insights into the regulatory control of gene expression as a function of age. So far, only a few mapping studies of regulatory loci of gene expression have focused on the dynamics of eQTL. For example, our study of the environmental

influence on eQTL distribution (Li *et al.*, 2006) and the studies of Schadt *et al*. (2008) and Gerrits *et al*. (2009) who focus on the spatial distribution of eQTL among different tissues and cell types (Schadt *et al.*, 2008; Gerrits *et al.*, 2009). With this study, we have shown that age influences the distribution of eQTLs and that the temporal dynamics of regulatory loci are heritable and can become more prominent at older ages. We believe our results have important implications for the identification of loci that are linked to complex progressive and geriatric disease phenotypes.

Methods & Materials

C. elegans culturing

A set of 36 recombinant inbred lines (RILs) derived from a cross between wildtypes N2 and CB4856 (Li *et al.*, 2006) was used to generate genetical genomics data (Supplemental Table ST1). The RILs were cultured on standard nematode growth medium (NGM) with *E. coli* OP50 as food source and a constant temperature of 24ºC. Populations were started with non-mated hermaphrodites and screened regularly to remove any occurring males. Reproductive nematodes from all RILs were bleached (0.5 M NaOH, 1% hypochlorite) in order to collect age synchronized eggs (Emmons *et al.*, 1979) which were inoculated (t0) in 24 dishes per RIL. After 40 hours (t1) nematodes from 6 dishes were collected in late L4 stage, frozen in liquid nitrogen and stored at -80ºC until RNA extraction. The remaining 18 dishes were kept in culture until hour 41 when the nematodes were transferred to fresh NGM dishes (with *E. coli* OP50) treated with 0.05-0.01 mg/ml of FUDR (fluorodeoxyuridine) in order to avoid hatching of the eggs. FUDR is widely used in *C. elegans* and does not affect lifespan (Wu *et al.*, 2009). 30 hours after that the nematodes with FUDR were transferred to fresh dishes (without FUDR) to prevent starvation and remove the FUDR. After 23 hours, 96 hours of total culture time (t2), nematodes from 6 dishes were collected and frozen in liquid nitrogen prior to RNA extraction. The remaining 12 dishes were kept at constant temperature until 214 hours of culture (t3), when they were harvested and frozen in liquid nitrogen. All the dishes were visually inspected before the harvest. Any population with infection, more than one generation (reproduction) or starving nematodes (lacking bacteria) were discarded.

Life-span assays in the RILs

The RILs were grown at 24°C on NGM plates containing *E. coli* (OP50) as a food source. Eggs from the RILs were placed on NGM plates containing *E. coli* (OP50) as a food source. After hatching, L4 worms were transferred and kept on fresh plates with FUDR [0.05-0.01 mg/ml] for their whole life. We measured the lifespan of 30 - 80 individual worms per RIL from the time the eggs were placed in the NGM dishes (Supplemental Table ST1). Worms were scored every day and those that crawled off the dishes or displayed bagging were censored. Worms were considered dead when they failed to respond to a gentle prodding. Average lifespan per RIL was calculated and used for further analysis.

Microarray experiments

RNA from nematodes was extracted following the Trizol method, followed by the RNeasy Micro kit (Qiagen, Valencia, CA, USA) to clean up the samples. Labeled cDNA was produced with an Array 900 HS kit from Genisphere and Superscript II from Invitrogen. The Nucleospin kit (Bioké, Leiden, NL) was used to clean the cDNA samples in order to reduce unspecific binding to the arrays. The 60-mers arrays were obtained from Washington University. Microarrays were hybridized following the Genisphere Array 900 HS protocol. The differential hybridization due to SNP differences is low in these 60-mers arrays (Li *et al.*, 2006).

We aimed to hybridize samples following the distant pair design proposed by Fu and Jansen (2006). (Supplemental Table ST1) A total of 54 microarrays were hybridized, 18 microarrays per time point. All microarray data have been deposited in Gene Expression Omnibus (GEO), microarray data with the common accession number GSE17071.

Microarray normalization

A Perking & Elmer scanner was used to extract the raw intensities. The R software was used for preprocessing and normalization (www.r-project.org) using the Limma package (Smyth, 2005). The Loess method (Smyth and Speed, 2003) was used for normalization within arrays and normalization between arrays was done using the quantile method, both methods are included in the Limma package.

Gene Ontology analysis

Gene Ontology (GO) data were extracted using BioMart (Durinck *et al.*, 2005) and the *C. elegans* annotation package, both packages for the R programming environment. GO terms

with less than 4 genes in GO database were discarded from the analysis. Over-represented groups of GO terms were identified using a hyper-geometric test, with a threshold of P-value < 0.01. Therefore, the hypergeometric test uses a group size of 396 GO terms, with 16,947 annotated genes. Over-represented GO terms in all the expression patterns for age affected genes (Supplemental Figure SF1) are accessible in Supplemental Table ST3. The hyper-geometric test was also used to determine significant over-representation of innate immunity genes in genes modulated by age. The list of genes was extracted from the original paper (Capra *et al.*, 2008). To identify biological functions associated to genes with increased or decreased variation in gene expression four groups of genes were investigated: genes with increased variance in developing worms, genes with increased variance in aging worms, genes with decreased variance in developing worms and genes with decreased variance in aging worms. GO terms were considered significant with a P-value < 0.01.

Assigning cis- or trans-acting regulation to eQTLs

We assigned *trans*-acting to any eQTL of which the peak was found at least 5 markers from the position of the gene the eQTL was calculated for. The rest was assigned *cis*-acting.

eQTL determination for daf-16

A closer inspection of *daf-16* was needed because it is represented with 7 different probes on the micro-array (Washington University; Array ID 3998, 5934, 7870, 8690, 9806, 17494 and 20336). To (re-)determine the quality of the probes we blasted them against the *C. elegans* genome. Two groups were found, three probes (3998, 7870 and 17494) were perfectly aligned to genomic clone R13H8 and the four other to F55A3. The four probes that aligned to genomic clone F55A3 showed possible cross-hybridization and were discarded from eQTL mapping. Furthermore one of the probes (17494) had a mean log2 intensity P-value < 7.5 and was also discarded. The two remaining probes (3998 and 7870) showed a similar eQTL profile and had a co-locating eQTL on chr *IV*. This eQTL had in both cases a –log10 P-value significance of ≥ 4. For clarity we took the mean log2 intensity of these two probes per RIL as a measure for *daf-16* expression and subsequently calculated the eQTL profile as shown in Figure 3.

Threshold determination

We used a permutation approach to determine the thresholds for the different mapping strategies (as in Li *et al*. 2006). For each of the used models for eQTL mapping we used

23,000 permutations. For each permutation we randomly picked a spot, each spot could only be picked once. The gene expression and relative lifespan values were than randomly distributed over the RILs (and time points) and used for mapping. In this way we obtained a threshold for each of the explaining factors. For the single time-points we used a false discovery rate (FDR) of 0.01 to adjust for multiple testing (Li *et al.*, 2006). The genome-wide threshold for this FDR is –log10 P-valuep of 3.8 for each of the three time-points. For the combined models (t1 to t2 and t2 to t3) we used a genome-wide threshold of –log10 P-value of 4, which resembles an FDR of 0.006, 0.001 and 0.006 for marker, age and the interaction between marker and age respectively.

To determine the threshold for the single gene examples we used 1000 permutations as in the genome-wide threshold. The difference is that we use the gene under study in all of the permutations. The P-values for the gene specific thresholds were determined at FDR 0.05.

Genome wide eQTL mapping

Prior to eQTL mapping we made sure that the absolute correlation of all marker and the dyes was less than 0.35. Furthermore the absolute correlation between markers on different chromosomes was less than 0.55 for all markers. Also, before calculating the eQTL profiles the outliers (> mean +2*SD or < mean -2*SD) per spot per stage were removed, to prevent the signals from dust etc.

First, we used a linear model to calculate the linkage of each marker with the measured expression levels for each of the time points separately (to make Figure 1). We used the log2 single channel normalized intensities as a measure for gene expression. We did not use the ratios because they can be notoriously problematic when it comes to inflation of variance. The model used for each of the three age-groups was: gene expression = marker (effect) + error. In this way we obtained the genome-wide eQTL profiles for all genes for the three age-groups.

Secondly, to quantify the heritable differences in gene expression that are age dependent we extended our linear model by analyzing two age-groups at once and including the relative age of the RILs as an explanatory factor. The model used for both combinations, t1 (=juvenile) and t2 (=reproductive), and t2 and t3 (=old worms) was: gene expression = marker (effect) + relative age + interaction (marker*relative age) + error. The age-groups t1 and t2 in one model are revered to as *developing worms*. The age-groups t2 and t3 in one model are revered to as *aging worms*. The significance and effect of each marker, relative age (time) and for the interaction between the two was obtained for all genes on the array.

Determination of temporal expression patterns

The significance (P-value ≤ 0.0001; see threshold determination) of the relative age (time) was used to determine if a gene was differentially expressed between the three age (time) groups. The effect of this factor explaining gene expression differences was used to determine if the expression went up or down during the two age/time periods (t1-t2 and t2-t3). So, we could assign each gene affected by age to one of the 8 possible classes (see Figure 1).

Determination of variance differences

We determined for all genes if a difference in variance exists between the juvenile and reproducing worms and between reproducing an old worms. A two-sided *F*-test was used to determine the significance of the difference in variance. As a threshold we used a P-value < 0.0001. Genes that had a significant difference in transcript variance where then assigned to increasing or decreasing in developing and aging worms. Over (or possible under-) representation of genes with an eQTL in the different significantly increased or decreased transcript level variance groups was determined by a hyper-geometric test. As example, the group of 1289 genes with significantly increased variance in developing worms contained 113 genes with an $_{gxa}$eQTL. We asked what is the likelihood to find this by chance when the total number of genes (18893) contains 280 genes with a $_{gxa}$eQTL.

To investigate if a possible relation between mean expression and an increase or decrease in variance influenced our results we used three methods. First we calculated the correlation (Pearson) between the variance and mean expression level. For all three stages these were low but positive (0.30, 0.32, and 0.40 resp.). So a possible increase in variation of the low expressed genes by normalization is very unlikely to have influenced our results. Secondly when we divide the expression levels per stage by the mean expression level per stage we found an even larger difference between genes with increased variance (3403) and decreased variance (641), both at P-value < 0.0001, which is confirming and strengthening our initial result. Thirdly, when we use a permutation approach we only found small difference between genes that have an increased or decreased variance when comparing stage 2 and 3 (979 and 743). In this permutation approach we made two random groups per gene from the expression levels of stage 2 and 3. Then we calculated the variances, means and P-values. Again this result shows that there is a large difference between the number of genes that have and increased variance and those that have a decreased variance when stage 2 and 3 are compared.

Supplemental material

All the supplementary tables and figures are available from the author on http://www.anavinuela.com/thesis/thesis.html

Supplementary Figure SF1: Plot bar with the average lifespan in days of the wildtypes N2 x CB4856 and the derived RILs grown at 24°C.

Supplementary Figure SF2: Relation between an in- or de-crease in variance during development and aging. On the *y-axis* the –log 10 P value of the most significant marker is shown. On the *x-axis* the log2 of the variance of gene expression in juvenile worm divided by the variance in reproducing worms for developing worms and the log2 of the variance of gene expression in reproducing worms divided by the variance in old worms is shown. A negative number (on the x-axis) means an increase in variation and a positive number a decrease. The red line indicates the eQTL threshold (-log10 P-value >= 4; P-value <= 0.0001). Type of eQTL (marker (eQTL) or genotype-by-age ($_{gxa}$eQTL) are indicated above the panels. The blue marked dots indicate the genes that had a significantly different variance (-log10 P-value >= 4; P-value <= 0.0001). The difference between developing and aging worms is most visible in the relatively small number of genes with a $_{gxa}$eQTL and increase in variation in aging worms and those with a $_{gxa}$eQTL and an increase in variation in developing worms.

Supplementary Table ST1: Recombinant inbred lines used for the arrays (including what label was used) for each age-group (juvenile, (late) reproductive and senescent) and their average lifespan at 24°C. Lifespan of N2 and CB4856 was (16.3 ± SE 0.6) (n=70), (13.1 ± SE 0.4) (n=36), respectively.

Supplementary Table ST2: List of genes in each pattern of expression described in figure 1 and table 1.

Supplementary Table ST3: List of significant GO terms in each pattern of expression described in figure 1 and table 1.

Supplementary Table ST4: List of significant GO terms in genes with a significant increased or decreased in transcript variance.

Supplemental Table ST5: Genes with a $_{gxa}$eQTL in aging worms (not included are the genes that are deleted in CB4856 compared to N2).

Supplemental Table ST7: Number of genes with two eQTLs at three different thresholds for both developing and aging worms.

References

Bahar, R., C. H. Hartmann, et al. (2006). Increased cell-to-cell variation in gene expression in ageing mouse heart. *Nature* **441**(7096): 1011-1014.

Brem, R. B., G. Yvert, et al. (2002). Genetic Dissection of Transcriptional Regulation in Budding Yeast. *Science* **296**(5568): 752-755.

Capra, E. J., S. M. Skrovanek, et al. (2008). Comparative Developmental Expression Profiling of Two *C. elegans* Isolates. *PLoS ONE* **3**(12): e4055.

Collins, J. J., C. Huang, et al. The measurement and analysis of age-related changes in Caenorhabditis elegans (January 24, 2008), WormBook, ed. The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.137.1, http://www.wormbook.org

de Magalhaes, J. P., J. Curado, et al. (2009). Meta-analysis of age-related gene expression profiles identifies common signatures of aging. *Bioinformatics* **25**(7): 875-881.

Dean, M. and R. Allikmets (1995). Evolution of ATP-binding cassette transporter genes. *Current Opinion in Genetics & Development* **5**(6): 779-785.

Dillin, A., D. K. Crawford, et al. (2002). Timing Requirements for Insulin/IGF-1 Signaling in C. elegans. *Science* **298**(5594): 830-834.

Doroszuk, A., L. B. Snoek, et al. (2009). A genome-wide library of CB4856/N2 introgression lines of Caenorhabditis elegans. *Nucl. Acids Res.* **37**(16): e110-.

Durinck, S., Y. Moreau, et al. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* **21**(16): 3439-3440.

Emmons, S. W., M. R. Klass, et al. (1979). Analysis of the constancy of DNA sequences during development and evolution of the nematode Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America* **76**(3): 1333-1337.

Farber, C. R., J. E. Aten, et al. (2009). Genetic dissection of a major mouse obesity QTL (Carfhg2): integration of gene expression and causality modeling. *Physiol. Genomics* **37**(3): 294-302.

Fu, J. and R. C. Jansen (2006). Optimal Design and Analysis of Genetic Studies on Gene Expression. *Genetics* **172**(3): 1993-1999.

Geigl, J. B., S. Langer, et al. (2004). Analysis of Gene Expression Patterns and Chromosomal Changes Associated with Aging. *Cancer Research* **64**(23): 8550-8557.

Gerrits, A., Y. Li, et al. (2009). Expression Quantitative Trait Loci Are Highly Sensitive to Cellular Differentiation State. *PLoS Genet* **5**(10): e1000692.

Gutteling, E. W., J. A. G. Riksen, et al. (2007). Mapping phenotypic plasticity and genotype-environment interactions affecting life-history traits in Caenorhabditis elegans. *Heredity* **98**(1): 28-37.

Halaschek-Wiener, J., J. S. Khattra, et al. (2005). Analysis of long-lived C. elegans daf-2 mutants using serial analysis of gene expression. *Genome Research* **15**(5): 603-615.

Hill, A. A., C. P. Hunter, et al. (2000). Genomic Analysis of Gene Expression in C. elegans. *Science* **290**(5492): 809-812.

Jansen, R. C. and J.-P. Nap (2001). Genetical genomics: the added value from segregation. *Trends in Genetics* **17**(7): 388-391.

Jiang, M., J. Ryu, et al. (2001). Genome-wide analysis of developmental and sex-regulated gene expression profiles in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America* **98**(1): 218-223.

Kammenga, J. E., A. Doroszuk, et al. (2007). A Caenorhabditis elegans wild type defies the temperature-size rule owing to a single nucleotide polymorphism in tra-3. *Plos Genetics* **3**(3): 358-366.

Keurentjes, J. J. B., J. Y. Fu, et al. (2007). Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. *Proceedings of the National Academy of Sciences of the United States of America* **104**(5): 1708-1713.

Kurz, C. L., M. Shapira, et al. (2007). Caenorhabditis elegans pgp-5 is involved in resistance to bacterial infection and heavy metal and its regulation requires TIR-1 and a p38 map kinase cascade. *Biochemical and Biophysical Research Communications* **363**(2): 438-443.

Li, Y., O. A. A. Alvarez, et al. (2006). Mapping determinants of gene expression plasticity by genetical genomics in C-elegans. *Plos Genetics* **2**(12): 2155-2161.

Lu, T., Y. Pan, et al. (2004). Gene regulation and DNA damage in the ageing human brain. *Nature* **429**(6994): 883-891.

Lund, J., P. Tedesco, et al. (2002). Transcriptional Profile of Aging in C. elegans. *Current Biology* **12**(18): 1566-1573.

Mahajan-Miklos, S., M.-W. Tan, et al. (1999). Molecular Mechanisms of Bacterial Virulence Elucidated Using a Pseudomonas aeruginosa Caenorhabditis elegans Pathogenesis Model. **96**(1): 47-56.

Maydan, J. S., S. Flibotte, et al. (2007). Efficient high-resolution deletion discovery in Caenorhabditis elegans by array comparative genomic hybridization. *Genome Research* **17**(3): 337-347.

McCarroll, S. A., C. T. Murphy, et al. (2004). Comparing genomic expression patterns across species identifies shared transcriptional profile in aging. *Nat Genet* **36**: 197 - 204.

Palopoli, M. F., M. V. Rockman, et al. (2008). Molecular basis of the copulatory plug polymorphism in Caenorhabditis elegans. *Nature* **454**(7207): 1019-1022.

Petretto, E., J. Mangion, et al. (2006). Heritability and Tissue Specificity of Expression Quantitative Trait Loci. *PLoS Genet* **2**(10): e172.

Ramoni, M. F., P. Sebastiani, et al. (2002). Cluster analysis of gene expression dynamics. *Proceedings of the National Academy of Sciences of the United States of America* **99**(14): 9121-9126.

Rockman, M. V. and L. Kruglyak (2006). Genetics of global gene expression. *Nat Rev Genet* **7**(11): 862-872.

Schadt, E. E., C. Molony, et al. (2008). Mapping the Genetic Architecture of Gene Expression in Human Liver. *PLoS Biol* **6**(5): e107.

Schadt, E. E., S. A. Monks, et al. (2003). Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422**(6929): 297-302.

Seidel, H. S., M. V. Rockman, et al. (2008). Widespread Genetic Incompatibility in C. Elegans Maintained by Balancing Selection. *Science* **319**(5863): 589-594.

Smith, E. N. and L. Kruglyak (2008). Gene-Environment Interaction in Yeast Gene Expression. *PLoS Biol* **6**(4): e83.

Smyth, G. K. (2005). Limma: Linear models for microarray data. *Bioinformatics and Computational Biology Solution Using R and Bioconductor*. New York, Springer**:** 397-420.

Smyth, G. K. and T. Speed (2003). Normalization of cDNA microarray data. *Methods* **31**(4): 265-273.

Somel, M., P. Khaitovich, et al. (2006). Gene expression becomes heterogeneous with age. *Current Biology* **16**(10): R359-R360.

Southworth, L. K., A. B. Owen, et al. (2009). Aging Mice Show a Decreasing Correlation of Gene Expression within Genetic Modules. *PLoS Genet* **5**(12): e1000776.

Wu, D., J. R. Cypser, et al. (2009). Multiple mild heat-shocks decrease the Gompertz component of mortality in Caenorhabditis elegans. *Experimental Gerontology* **44**(9): 607-612.

Zhan, M., H. Yamaza, et al. (2007). Temporal and spatial transcriptional profiles of aging in Drosophila melanogaster. *Genome Research* **17**(8): 1236-1243.

Zhao, Z., L. L. Fang, et al. (2004). ATP-binding cassette protein E is involved in gene transcription and translation in Caenorhabditis elegans. *Biochemical and Biophysical Research Communications* **323**(1): 104-111.

Chapter: 5

Heritability and Expression Quantitative Trait Loci in aging *C. elegans*

Ana Viñuela, L. Basten Snoek, Joost A.G. Riksen, Jan E. Kammenga

Manuscript in preparation

Abstract

Gene expression variation is a heritable trait in segregating populations. This allows for genetic mapping of genomic regions associated with variation in gene expression, called expression quantitative trait loci (eQTLs). Mapping of eQTLs is a powerful tool to characterize the loci underlying complex disease traits. However, heritability of any phenotypic trait is not necessarily constant with age. The variances determining heritability become a function of age since different genetic and environmental effects can influence the traits measured. We recently reported that age influences the distribution of eQTLs and that the temporal dynamics of regulatory loci are heritable and can become more prominent at older ages in the nematode *Caenorhabditis elegans* (Viñuela *et al.*, 2010). Now, we investigate how heritability of gene expression changes with age and how those changes may affect the detection of eQTL in aging *C. elegans.* We found that the average broad sense heritability of gene expression did not change with age. But, the average heritability in developing and aging worms was determined by different genes. Also, we identified a stronger imbalance between highly heritable genes and eQTLs in aging than in developing worms, indicating that heritability is a poorer predictor of eQTL detection at older ages. We propose that gene expression regulation becomes more polygenic in aging worms. As a result, heritability values of gene expression change with age, and our ability to detect eQTLs, decreased. We believe that our study has important implications for the analysis of complex diseases and the problem of the missing heritability in complex traits.

Introduction

Variation in gene expression, or transcript abundance, is an important contribution to phenotypic diversity observed within and between populations. It is closely associated with many complex traits, including a great number of common human diseases, and provides the basis for evolutionary change (Wray *et al.*, 2003). Many studies have shown that individual variation in gene expression is heritable in segregating populations, i.e. it is genetically determined (Brem *et al.*, 2005). But evidence is emerging that heritability is a poor predictor of eQTL detection, although detailed studies are sparse. A low heritability means that most of the phenotypic variation among individuals in a population may be due to environmental factors. A high heritability means that most of the observed variation can be attributed to variation in the individual genotypes. This allows for genetic mapping of genomic regions associated with variation in gene expression, called expression quantitative trait loci (eQTLs) (Flint *et al.*, 2005). Mapping of eQTLs is a powerful tool to characterize the loci underlying complex disease traits. It can be used as the link between genetic markers of disease which are identified in genome-wide association studies and the transcript abundance of one or more genes. For instance, eQTL were found to be associated with the risk of lung cancer (Li *et al.*, 2010). The strategy has been successfully applied to several other complex diseases such as celiac disease (Hunt *et al.*, 2008) and asthma (Moffatt *et al.*, 2007). Moreover, eQTLs were detected affecting levels of gene expression in *cis* or in *trans*, thus providing means to identify potential underlying genetic pathways.

However, the relationship between heritability and the mapping of eQTL is not well understood. Petretto *et al.* (2006) applied genome-wide expression profiling and eQTL mapping to investigate the regulation of gene expression in different tissues within rat recombinant inbred lines (Petretto *et al.*, 2006). It appeared that the proportion of heritable expression traits was similar across tissues; yet, heritability itself was not a good predictor for detection of an eQTL. The efficiency to map eQTL depended to a large extent on the P-value thresholds at which eQTLs were mapped. Additionally, their results illustrate that heritability is an unreliable parameter for predicting putative *trans*-eQTLs.

We recently reported that age influences the distribution of eQTLs and that the temporal dynamics of regulatory loci are heritable and can become more prominent at older ages in

the nematode worm *C. elegans* (Viñuela *et al.*, 2010). Here, we investigate for the first time how heritability changes with age and to which extent this affected the detection of eQTL in *C. elegans*. Understanding how heritability changes with age is not only important for eQTL mapping. Age-specific heritability determines also how populations respond to natural selection and has only rarely been studied, so far only for life-history traits in cattle (Albuquerque and Meyer, 2005), fruitflies (Snoke and Promislow, 2003) and in natural populations (Charmantier *et al.*, 2006).

We used microarrays to measure genome-wide gene expression in two of the most genetically divergent wildtypes Bristol N2 and Hawaii CB4856 at three different ages: young worms, t1, age was 40 h; for reproductive worms, t2, age was 96 h; and for old worms, t3, age was 214 h. We investigated how heritability of gene expression changes with age and how these correlated with mapped eQTL (Viñuela *et al.*, 2010).

Results

Age and genotype effects on gene expression in N2 and CB4856.

First, we studied the effect of the parental genotypes on gene expression levels in the three age groups separately, juveniles, reproductive and old worms. We identified 11.3 % (2151), 9.8 % (1868) and 14.1 % (2679) genes significantly affected by the genotype for each age group, respectively (Supplementary Table ST1). Yet, the average lifespan differed between both strains, 16 d. for N2 and 13 d. for CB4856 (Viñuela *et al.*, 2010). Age at maturity on the other hand, did not differ (Gutteling *et al.*, 2007). In other words, physiological age differed between wild types at the time of RNA harvest. To correct for this discrepancy, we treated age differences with a two-time-model for developing worms (juveniles and reproductive worms) and for aging worms (reproductive and old worms) (Viñuela *et al.*, 2010) (Supplementary Table ST2). Three factors were considered: relative age, genotype (marker) and their interaction. In this way we found 50.1% (11653) and 40.7% (9458) of all transcripts to have a significant age effect in developing and aging worms, respectively. We also found transcript effects linked to genotype (15.3% (2899) and 10.5% (1999) for

developing and aging worms, respectively) and genotype-by-age effect (20.5% (3888) and 15.2% (2886)).

A Gene Ontology (GO) enrichment analysis of the regulated genes (Supplementary Table ST3) showed genotypic differences in regulation of calcium transport, and structure molecules as collagen or extracellular matrix in developing worms. For aging worms, we found genotypic differences in regulation of response to oxidative stress, and glycoprotein catabolic process. The genotype-by-age interaction affects signaling pathways mediated by cell surface receptors, metabolic processes for nitrogen, glutamine, glyoxylate, isocitrate or malate in developing worms. And in aging worms, regulated processes included cell division, and lipid storage.

Complexity of genetic variation in gene expression

We investigated the overlap between differentially expressed genes for N2 and CB4856 and the genes for which an eQTL was mapped (Viñuela *et al.*, 2010). We identified three possible relevant gene categories (Figure 1): A) Genes differentially expressed between the parental lines that also had an eQTL. B) Genes differentially expressed between the parental lines without an eQTL. C) Genes not differentially expressed between the parental lines but with an eQTL.

Category A genes (Figure 1A) included those with a different level of expression between the parental lines and statistical evidence for genomic linkage in the recombinant genotypes. This was the expected behavior of polymorphic traits that segregate in the RILs. It involved strong regulatory elements influencing gene expression which activity was detectable after the recombination events. We found 496 and 266 genes in this category for genotypic effect in developing and aging worms, respectively. Genes with a genotype-by-age effect were 342 and 288 genes, for developing and aging worms (Supplementary Table 1 and Supplementary Figures SF3 to SF6). A Gene Ontology (GO) enrichment analysis (Supplementary Figure SF1 and Supplementary Table ST4) showed ATPase activity, calcium transport, and protein kinase actitivy as strongly regulated processes in developing worm with genotype effect. Genotype-by-age regulated genes in developing worms involved ATPase activity, but also lipid storage and lipid glycosilation, oxidoreductase activity and regulation of locomotion. In aging worms we found enrichment for response to

oxidative stress, metal ion binding and hydrolase for genotypic effect. Genes with a genotype-by-age effect were related to oxidative stress responses, but also to metabolism of glutamine and glycoproteins, and ATP-binding.

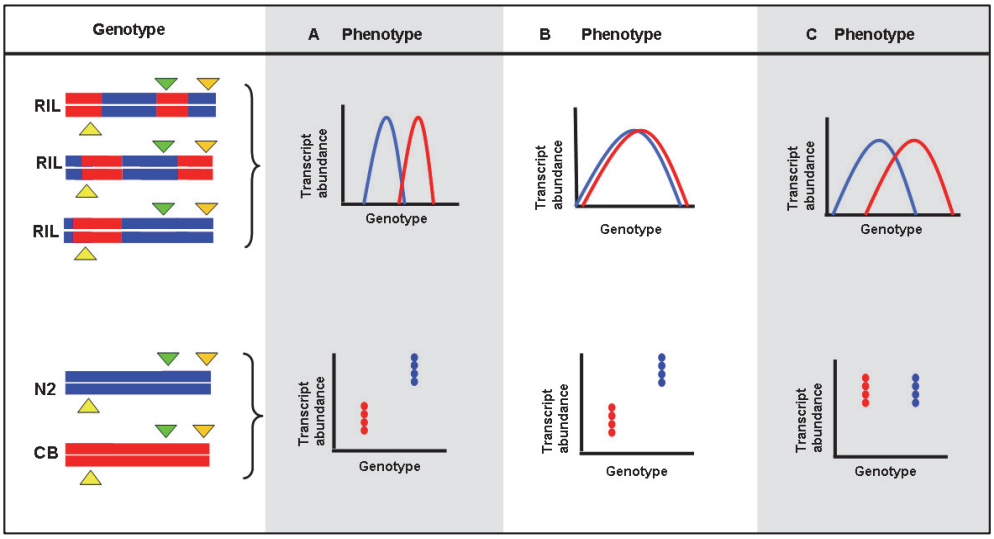


Figure 1.- Interpretation of gene expression differences in segregating populations. Three different features could be identified when comparing mapping data from RILs and expression level differences between parental strains. **A.-** Genes with a statistical evidence for genomic linkage (eQTL) and different levels of transcript abundance between the parental lines. The regulatory elements (green and orange triangles) may have different effects (red or blue) over the phenotype (gene expression). Reshuffling the different regulatory elements in the RILs induced phenotypes in the population within the range of the parental strains (from red to blue). **B.-** Genes with no statistical evidence for genomic linkage (eQTL) and different level of expression between the parental lines. The regulatory elements may have similar and/or opposite small effects. Their recombination induces intermediate phenotypes in the RILs. None of the single effects of any of the multiple regulators was large enough to be indentified. Different mapping strategies or complementary experimental data may be able to identify multiple regulatory elements affecting the expression levels. **C.-** Genes with statistical evidence for genomic linkage (eQTL) and with similar levels of transcript abundance between the parental strains. The genomic recombination may had induce extreme phenotypes when compare to the parental lines. The expression level of some of those genes may show signs of transgressive segregation.

In category B (Figure 1B), we identified 2535 genes for developing, and 1803 genes for aging worms that were different between the parental lines but with no statistical evidence for genomic linkage associated to genotypic effect (Table 1). For genotype-by-age interaction 3650 (developing) and 2741 (aging) genes were identified. The differences

between the parental were mainly genetic, but we failed to detect the corresponding eQTLs or $_{gxa}$eQTL. Failure in linkage detection in heritable traits may be explained by many loci with small effects, none of which is strong enough to be detected. Thus, those genes were likely to have a polygenic control of their transcript abundance.

Table 1.- Number of overlapping genes between parental and segregants analysis

	Developing		Aging	
	genotype	**age*genotype**	**genotype**	**age*genotype**
Genes with a significant eQTLs	1401	935	867	844
Differentially expressed the parents	2899	3888	1999	2886
Detected eQTL & differentially expressed in the parents	496	342	266	288
NO detected eQTL & differentially expressed in the parents	2535	3650	1803	2741
Detected eQTL & NO differentially expressed in the parents	1041	691	673	688

Table 1.- Number of genes overlapping between differentially expressed genes in the parental strains and genes for which an eQTL was mapped. First two rows show the number of genes with an eQTL in the RILs and differentially expressed in the parental strains with genotype and age-by-genotype effects. The next threes rows shows the number of genes within the three categories considered: Category A) Genes with at least an eQTL and differentially expressed between the parents. Category B) Genes without an eQTL but differentially expressed in the parents; and Category C) Genes with an eQTL but no differentially expressed in the parents. Thresholds for eQTL detection –log10 P-value > 3, and –log10 P-value > 2 for parental analysis.

Category C (Figure 1C), contained genes with at least an eQTL but no differences in expression between the parental strains. In total we found 1041 (developing) and 673 (aging) genes with an eQTL; and 691 (developing) and 688 (aging) genes with an $_{axg}$eQTL in this group (Table 1). Contrary to genes in category B, we identified a genomic linkage that was not observed in the parental strains. This may be the result of new epistatic interactions in the segregants or complementary additive effects of the new allelic combinations (Rieseberg *et al.*, 1999; Brachi *et al.*, 2010). Both possibilities involved multiple loci regulation, and suggested transgression of gene expression due to higher or lower transcript abundance in the segregants than in either parent.

Patterns of gene expression showed signs of transgressive segregation

Transgressive segregation in the RILs implies that the allelic combination produced higher or lower expression values than in either parent. To detect transgression, we compared expression values of parents and segregant strains in juvenile, reproductive and old worms. 8205 (43.4 %), 9109 (48.2 %) and 8484 (44.9 %) genes which expression levels transgressed were identified. Moreover, we investigated whether changes in gene expression over time could also transgress with respect to the parental lines. Genes with expression levels outside or inside the parental ranges in two time-points were considered. In this way, we identified 1032 genes (5.4 %) with significant transgression of expression changes in developing worms, and 1122 genes (5.9 %) in aging worms. The results indicated that transgressive segregation of gene expression was widespread and similar in number of genes at different ages in *C. elegans*. Moreover, patterns of expression transgressed in the segregants, suggesting multiple regulators controlling those patterns.

Heritability of gene expression changes with age

We investigated how heritability (H^2) of gene expression changes with age and how this correlated with mapped eQTL. The heritability is the fraction of variation in gene expression that can be attributed to genotypic differences in the segregants. Therefore, highly heritable genes may be more likely to have a detectable eQTL and highly heritable genes without eQTL may suggest complex and non detectable regulation in gene expression.

The heritability of each transcript was estimated as in Keurentjes, *et al*. (2007) from the pooled within line variance of the parents. As expected, we observed that the relative number of genes for which an eQTL can be found increases with a higher H^2, for both the developing worms. However, for old and for aging worms (reproductive to old worms) the relative number of eQTLs hardly increased with an increasing H^2 (Figure 2 and Supplementary Figure SF6). Furthermore, we identified highly heritable patterns of gene expression for developing worms (9924 genes with $H^2 > 0.69$) and for aging worms (7338 genes with $H^2 > 0.77$) (See methods and materials for threshold determination). A GO term analysis of those genes (Supplementary Tables ST5 and ST6) indicated that in developing

worms unique highly heritable functions were calcium-transporting ATPase activity, carboxylestarase actitivy, and dendrite development among others. Some of the highly heritable functions identified only in aging worms were gluconeogenesis, calcium channel activity, cholesterol binding or protein kinase activities.

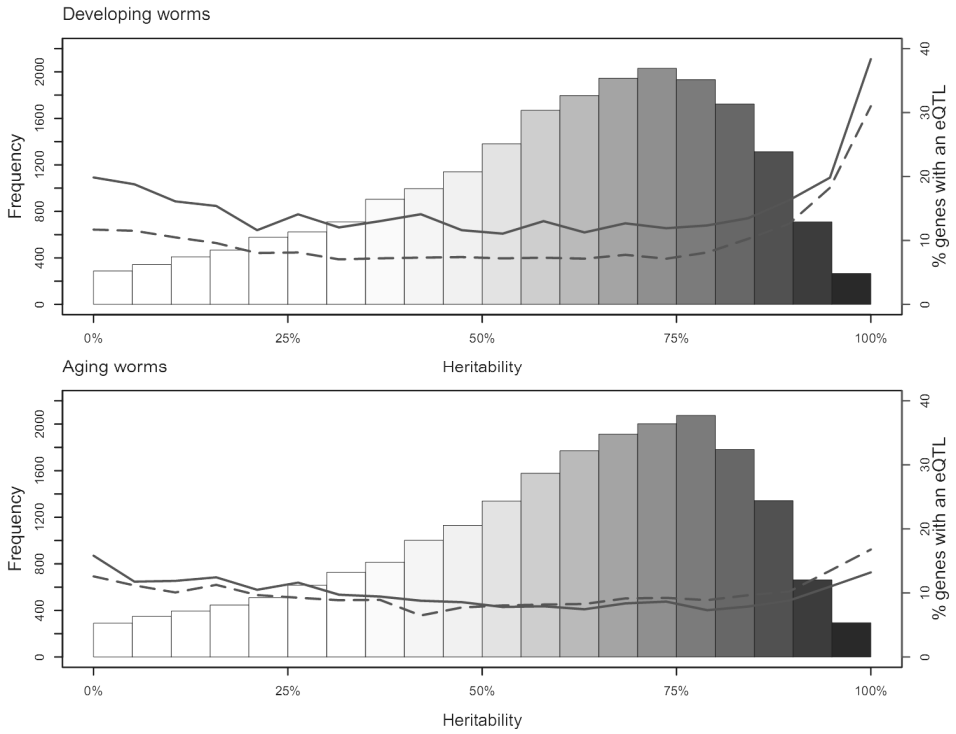


Figure 2.- Heritability and eQTL. Heritability of gene expression in developing and aging worms was plotted. *y-axis* (left side) shows the frequency of transcripts within a range of heritability values. *x-axis* percentage of heritability per gene. The lines indicate (*y-axis*, *right side*) the number of genes with an eQTL (solid line) and number of genes with an $_{axg}$eQTL (dashed line) within the range of heritability.

Figure 3 shows the overlap between highly heritable genes ($H^2 > 0.69$ and $H^2 > 0.77$), number of eQTLs and transgressive genes in developing and aging worms. As expected from the comparison between segregants linkage and differentially expressed genes in the parental lines (Table 1 and Figure 3), there was a large group of highly heritable genes without detectable eQTLs, at both age groups (79% and 78.8%, respectively). The analysis also showed that high heritability was not a good predictor for genomic linkage, since

33.7% (620) and 47.6% (632) genes with eQTLs in developing and aging worms did not fall into the high heritability groups of genes (see also Figure 2). Moreover, it suggested

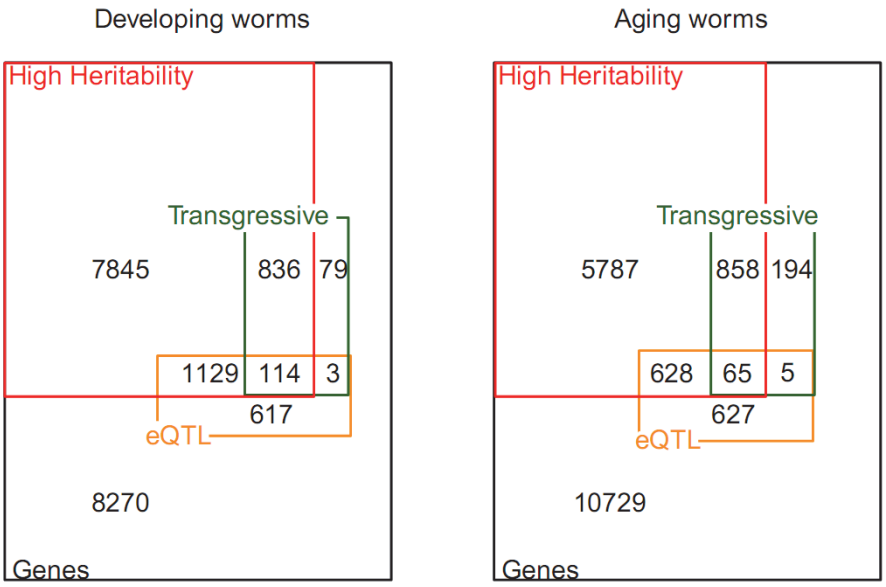


Figure 3.- Overlapping features analyzed in developing and aging worms. From almost nineteen thousand analyzed genes (black square), we identified highly heritable genes (red square), genes with at least one detectable genomic linkage (eQTL, orange square) and genes which expression values transgressed from the parental values (Transgressive, green square).

that age influenced the relationship between heritability and the ability to detect genomic linkage. This age influence was also observed in the comparison of transgressive genes and heritable genes. In developing worms 7.2% (82) genes which expression values transgressed with respect to the parents did not have high heritability values, as compared to 17.7% (199) genes in aging worms. Still, genes with signs of transgressive segregation tended to have higher heritability values (Supplementary Figure SF7). Altogether, we conclude that filtering highly heritable genes for eQTLs analysis will result in a decrease in linkage detection close to 50% in aging worms. However, the approximately 80% of genes with highly heritable values without genomic linkage at both age groups suggested lack of eQTL detection.

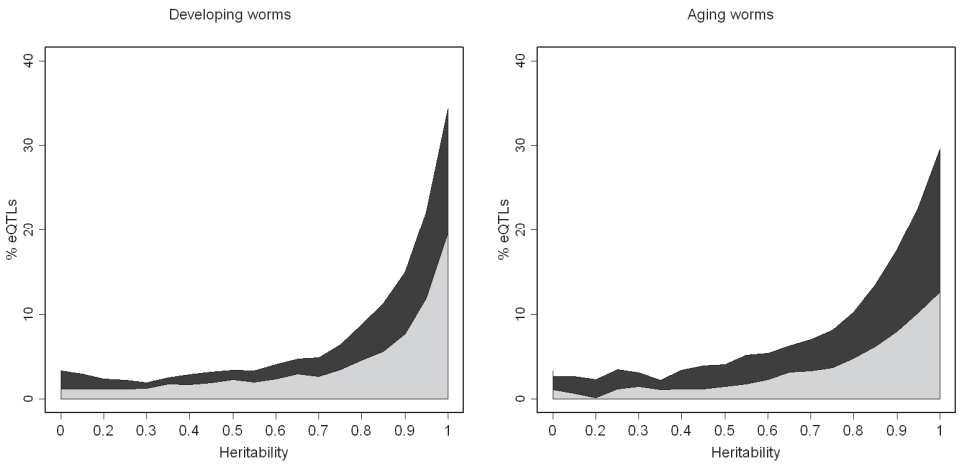


Figure 4.- eQTLs *vs.* heritability. Percentage of eQTLs (*x-axis*) detected in transcripts ordered by heritability values (*x-axis*). Grey area indicates the percentage of genomic linkage (markers) detected considering the effect of only one marker. Red area shows the increment in number of genomic linkage (markers) detected when two or more markers effect were allowed in the model. The total percentage of genomic linkage (markers) detected at each age group with one or more markers linkages are the sum of both colored areas (black and grey) and marked in the *x-axis*.

The number of polymorphic regulators increased with age

Failure to detect eQTLs may be explained by multiple regulatory elements which effect can not be detected using the current analysis. Moreover, the differences in number of genes with eQTLs, signs of transgression and high heritability between developing and aging worms, also suggested that the power of detection of genetic linkage was affected by the differences in gene expression regulation induced by age. To test this hypothesis we applied a forward and backward marker selection approach in the linear model previously used. The procedure identified per gene the linear model that better explained the variation in expression considering more than one marker effects and paired interactions between markers. This strategy allowed identifying genes that were likely to have more than one polymorphic regulatory element affecting their expression values at different ages. A summary of the results is shown in Figure 4, where the number of transcripts was ordered by heritability values (*x-axis*) and the total percentage of detected eQTL per transcript groups was plotted (*y-axis*). In developing worms, we found that the number of transcripts

with an eQTL due to the effect of one marker (grey area) increased when two or more markers were fitted into the model (black area). This increment was bigger in transcripts with higher heritability values as the red area is larger at the right side of the graph. On the other hand, the increase on eQTLs detected with two or more markers in the model was larger in aging worms (bigger black area). This suggests again that heritable regulation of gene expression patterns becomes more complex in aging worms due to involvement of multiple loci.

Discussion

Heritability provides a measurement of the phenotypic variation that can be attributed to genotypic variation and is an indicator of the relative importance of genes and environment in traits variation (Visscher *et al.*, 2008). The power to detect gene variants that affect gene expression (eQTL) depends to a large extent on heritability. The average heritability of gene expression in developing worms was 0.64. This is comparable to average heritability values found in RIL populations of yeast (Brem *et al.*, 2005; Zou *et al.*, 2009) and Arabidopsis (Keurentjes *et al.*, 2007). In developing worms we found that the fraction of genes with an eQTL increased with heritability. This is according to theory and was also reported for Arabidopsis seedlings. Yet, Petretto *et al.* (2006) showed that heritability alone was not a reliable predictor of whether an eQTL can be detected. They observed that heritability used as a filter reduces the ability to discover eQTLs with small effects.

Heritability of any phenotypic trait might not be constant with age. Traits measured during an individual's lifetime can have different genetic and environmental effects influencing it, so that the variances determining heritability become a function of age. We found an average heritability of 0.67 for aging worms which was not significantly different from heritability in developing worms. But we found that the heritability in developing and aging worms was determined by different genes having different GO terms. We estimated heritability considering changes on gene expression and we accounted for variance in gene expression over two age points. This method leads to an overestimation of the genetic variance. However, our approach provides a valid source to compare the relative genetic

and environmental influence on patterns of gene expression at different ages. Recently, we showed that the number of detected heritable gene expression changes declined with age in worms (Viñuela *et al.*, 2010). Linkage analysis indicated that the decline in eQTL number in older worms was stronger in *cis*-acting linkages than in *trans*-acting. In line with those results, the data presented here indicated that gene expression regulation became more polygenic in aging worms, suggesting that multiple intermediates involved in distant regulation became relatively more influential with age. Likewise, changes in gene regulatory mechanisms may have an effect on our ability to detect eQTLs. In that sense, it is known that the efficiency to map eQTLs depends to a large extent on the P-value thresholds at which eQTLs are mapped. This is especially relevant in *trans*-eQTLs, because they exhibit smaller genetic effects (Petretto *et al.*, 2006). Their smaller effect has been related to the number of protein intermediates (Smith and Kruglyak, 2008). The genetic variance effect may be diluted in a larger number of intermediates, and therefore it would have been more difficult to map. Moreover, *trans*-regulation is more likely to have intermediate regulators with opposing effects on transcripts levels. Such opposite effects can cancel each other, limiting our ability to detect genetic linkage at different older ages (Brem *et al.*, 2005; Smith and Kruglyak, 2008). It is possible; therefore, that failure to detect eQTL in old worms was linked to a change in the proportion of *cis* and *trans* regulation, with a relative increase in the latest that are more difficult linkages to detect. Furthermore, previous studies have shown that *cis*-acting linkages are more consistent across environmental conditions (Li *et al.*, 2006; Petretto *et al.*, 2006; Smith and Kruglyak, 2008). Accordingly, the relative influence of the environment on gene expression might be increased in old worms, affecting our ability to detect genetic linkage.

The comparison of genomic linkage between segregants and parental strains supports a model for polygenic gene expression regulation in aging worms. We identified a similar number of transgressed gene expression values at both age groups. Transgressive segregation is attributed to epistatic interactions between alleles or to opposites additive effect of segregating alleles (Rieseberg *et al.*, 1999). Therefore, the high heritability values in transgressive genes that we observed were expected. However, the number of transgressed genes for which we were able to identify genomic linkage was lower in aging worms, indicating a more complex polygenic regulation at these ages. Moreover, we were

able to identify three categories of genes: (A) genes with genomic linkage in parents and segregants, (B) genes with genomic linkage in the parents, but not in the segregants, and (C) genes with genomic linkage in the segregants but not in the parents. The overall number of genes in those categories decreased with age for all three gene categories (Table 1). This suggested an equal decrease in power detection for expression genomic linkages in parents and segregants. It also indicates that our ability to detect eQTLs was affected by changes in the gene expression regulation with age.

Organisms are genetically programmed to survive. Longevity is determined by processes that maintain somatic tissues functioning, and it is limited by the energy consumed in the process. The aging process depends largely on the accumulation of molecular damage (Kirkwood, 1977; Kirkwood, 2002). Longevity programs therefore, include damage repair, toxicant excretion and metabolism to minimize age accumulation damage and cover energy requirements (Kuningas *et al.*, 2008). Detoxification and toxicant excretion genes have been identified among molecular mechanism involved in longevity of *C. elegans* (McElwee *et al.*, 2004; Gems and McElwee, 2005). Previously, we identified detoxification genes as regulated in aging worms by our linkage analysis (Viñuela *et al.*, 2010). Here, we also identified in the group of genes with gene expression genomic linkage in parental and segregant strains (Category A) enrichment for oxidative stress response and some metabolic pathways in aging worms. Moreover, in the heritability analysis we observed the expression of genes involved in gluconeogenesis and other metabolic processes as highly heritable only in aging worms compared to developing worms. This suggested genetic determination in longevity associated processes. Together, our results support a model of longevity where expression levels of genes involved in damage repair and energy distribution are strongly regulated at old ages. However, heritability and genomic linkages are genetic parameters influenced by population structure and environmental factors. Our analysis was limited by the variation expressed in the RIL population at certain ages (juveniles, (late) reproductive and old) and in laboratory (stable) conditions. Therefore, the genetic variation necessary to detect the genomic linkage in longevity genes was revealed in aging worms.

In conclusion, heritability of gene expression changes with age. The genome wide broad sense heritability was similar for both age groups; but at the single gene level, the

expression variance that we can attribute to genetic variance changes with age. A reason for that may be changes in gene expression regulation. A scenario where regulation became more polygenic in aging worms explains the decrease in eQTLs. Likewise, it explains the imbalance between highly heritable genes and eQTLs in aging worms; as well as the smaller genetic linkages identified in transgressive genes.

Methods & Materials

C. elegans culturing

C. elegans nematodes strains N2 and CB4856 were cultured on standard nematode growth medium (NGM) with *E. coli* OP50 as food source and a constant temperature of 24ºC. Populations were started with non-mated hermaphrodites and screened regularly to remove any occurring males. Reproductive nematodes from both wildtypes were bleached (0.5 M NaOH, 1% hypochlorite) in order to collect age synchronized eggs (Emmons *et al.*, 1979) which were inoculated (t0) in 9cm petri dishes. After 40 hours (t1) nematodes in late L4 stage from 6 dishes were collected as one sample, frozen in liquid nitrogen and stored at -80ºC until RNA extraction. The remaining 18 dishes were kept in culture until hour 41 when the nematodes were transferred to fresh NGM dishes (with *E. coli* OP50) treated with 0.05-0.01 mg/ml of FUDR (fluorodeoxyuridine) in order to avoid hatching of the eggs. After 30 hours the nematodes with FUDR were transferred to fresh dishes (without FUDR) to prevent starvation and remove the FUDR. After 23 hours, 96 hours of total culture time (t2), nematodes from 6 dishes were collected and frozen in liquid nitrogen prior to RNA extraction. The remaining 12 dishes were kept at constant temperature until 214 hours of culture (t3), when they were harvested and frozen in liquid nitrogen. All the dishes were visually inspected before harvest. Any population with infection, more than one generation (reproduction) or starving nematodes (lacking bacteria) were discarded.

Microarray experiments

RNA from nematodes was extracted following the Trizol method, followed by the RNeasy Micro kit (Qiagen, Valencia, CA, USA) to clean up the samples. Labeled cDNA was produced with an Array 900 HS kit from Genisphere and Superscript II from Invitrogen. The Nucleospin kit (Bioké, Leiden, NL) was used to clean the cDNA samples in order to reduce unspecific binding to the arrays. The two colors 60-mers arrays were obtained from Washington University.

N2 *versus* CB4856 samples were hybridized to each array, with 6 replicates of t1, t2 and t3 in a dye swap design. The microarrays were hybridized following the Genisphere Array 900 HS protocol. The differential hybridization due to SNP differences between N2 and CB4856 is low in these 60-mers arrays (Li *et al.*, 2006). All microarray data have been deposited in Gene Expression Omnibus (GEO) with the common accession number GSE22887.

Microarray normalization

A Perking & Elmer scanner was used to extract raw intensities. The R software was use for preprocessing and normalization (www.r-project.org) using the Limma package (Smyth, 2005). The Loess method (Smyth and Speed, 2003) was used for normalization within arrays and normalization between arrays was done using the quantile method (Yang, 2003), both methods are included in the Limma package. The expression data from N2 and CB4856 were normalized together with the segregant expression data (GEO accession number GSE17071).

Outliers were removed prior normalization. From RILs data, outliers were considered values lower or higher than two times standard deviation of the mean, per spot per stage. Outliers from N2 and CB4856 arrays were identified using a linear model. Expression values of juveniles and reproductive worms in one model, and reproductive and old worms in other were treated together in linear models. The models fit expression values according to the genotype (N2 or CB4856) in two time points (t1-t2, and t2-t3) and remove values outside the 0.995 confidence interval, one spot at the time, recursively. No more than 6 values were allowed to be removed. Values from reproductive worms were considered outliers if they were removed from any of the two linear models.

Statistical analysis of the parental lines

Differential expression between genotypes (N2 and CB4856) and age was determined as in Viñuela & Snoek *et al.* (2010). In short, using the log2 single channel normalized intensities and a linear model (gene expression = marker (effect) + error), we calculated how well the genotype could explain the variation in measured expression levels for each gene for each of the time points separately. To quantify the heritable differences in gene expression that are age dependent we extended our linear model by analyzing two age-groups at once and including the relative age of N2 and CB4856 at the time of harvest as an explanatory factor. Relative ages were calculated as time of harvest (1.7 d., 3.91 d, and 8.91 d.) divided by average life span of N2 (16.35 d.) and CB4856 (13.07 d.). The model (gene

expression = marker (effect) + relative age + interaction (marker*relative age) + error) was used on the age-groups t1 and t2 (developing worms), and the age-groups t2 and t3 (aging worms). The significance and effect of each genotype, relative age and for the interaction between the two was obtained for all genes on the array.

We used a permutation approach to determine the thresholds for the differentially expressed genes (as in (Li *et al.*, 2006) and (Viñuela *et al.*, 2010)). The expression intensities per spot were randomly distributed over the replicates (and time points). Each spot was only picked once, from a total of 23232 permutations. In this way we obtained a threshold for each of the explanatory factors. For the single time-points we used a genome-wide threshold of –log10 P-value = 2, which resembles a false discovery rate (FDR) of 0.0129, 0.0118 and 0.0136 for each of the three time-points, respectively. For the combined models (t1 to t2 and t2 to t3) we used a genome-wide threshold of –log10 P-value of 2, which resembles an FDR of 0.0111, 0.0108 and 0.0113 for genotype, age and the interaction age-genotype for the developing worms and FDR of 0.0108, 0.0103 and 0.0105 for genotype, age and age-genotype for the aging worms.

Gene Ontology analysis

Gene Ontology (GO) data and functional domain data were extracted from Wormbase release WB210. GO terms with less than 2 genes were discarded. Over-represented groups of GO terms and domains were identified using a hypergeometric test (P-value < 0.01). In this way we analyzed 2109 unique GO terms form 18,312 annotated genes, respectively.

Comparative analysis between parents and segregant lines

We normalized the parental strains and segregant data sets together (see Microarray normalization). eQTLs were re-mapped using the same linear model as in Viñuela & Snoek, *et al.* and in the parental strains analysis (Viñuela *et al.*, 2010). Thereby, we first obtained genome-wide eQTL profiles for all genes for the three age-groups. Secondly, eQTL profiles for all the genes in developing and aging worms were obtained, with relative age, genotype and the age-genotype interaction as explanatory factors. We compared the number of genes with at least one significant eQTL (see below for significant levels) with the genes significantly different between the parents. Than, we compared and quantified the number of regulated and non-regulated genes between the RILs and the parents. For determining the thresholds we used a permutation approach as described above. Therefore, we calculated P-values from permuted data for eQTL mapping and for genotypic effect in the parents at the three time points and in the two age-group combined models. The

probabilities to find genes that were significantly different between the parental genotypes and also have an eQTL by chance were determined. This approach allowed us to relax the threshold for the linkage mapping from –log10 P-value 3.8 in the original paper (Viñuela *et al.*, 2010). For simplicity, we decided to use a level of significance of –log10 P-value = 3 for eQTL mapping and a –log10 P-value = 2 for the parents analysis for all time points and models. Those thresholds resembled the following joined-FDR for the three time-points: 0.0129, 0.0118 and 0.0136, respectively. The joined-FDR in the combined models for genotype, age and the interaction in developing worms were 0.0111, 0.0108 and 0.0113. In aging worms the FDR were 0.0108, 0.0103 and 0.0105 respectively.

Multiple regulatory elements analysis

To investigate if genes without an identified eQTL in the single marker model could have one or more eQTL we use a forward & backward marker selection approach in the linear model described in the previous section. We only did this for the models including relative age effects. To make this procedure possible we selected four markers per chromosome, obtaining a total of 24 markers. All these 24 markers, including paired interactions, relative age and the interaction between marker and relative age were in the starting model. The Bayesian Information Criterion (BIC) was used as a selection method to identify the best model explaining the origin of the variation per gene expression. We allowed 2000 steps and no more than 6 explanatory variables to identify the best model. From each model the number of markers explaining the variation in gene expression was determined and used in further analysis.

Transgressive segregation

Transgression to higher or lower expression, or to both extremes were defined for those genes with at least n=6 transcripts out of the 2 times standard deviation of the mean expression of the higher or the lower parent. The FDR and the limit for the number of transcripts (n=6) were computed as (Keurentjes *et al.*, 2007). In the three time point analysis, a threshold of 6 transcripts resembles to a FDR of 0.049, 0.046 and 0.045. Results were averaged over 100 permutations.

To identify transgressive alleles over time we first counted those genes with at least 6 transcripts (n = 6) out of the 2 times standard deviation of the mean of the higher or the lower parent in both t1 and t2 for developing worms, and in both t2 and t3 for aging worms. The FDR for this analysis were: 0.0099 and 0.0028. Results were averaged over 100 permutations.

Heritability

We calculated the heritability of each transcript as $H^2 = (V_{RIL} - V_P) / V_{RIL}$ (Keurentjes *et al.*, 2007), where V_{RIL} was the variance among the segregants and V_P was the pooled within line variance of the parents. We used the same approach to estimate the heritability of gene expression over time; that is, the patterns of gene expression changes. For each transcript we calculated the variance among the segregants as the variance due to genotypic effect over time (developing and aging worms); and for the parents variance over time. In this way the effect of the interaction between relative age and genotype on the variance was excluded. In the single time-point models 35.23%, 30.57% and 25.46% transcripts had a negative heritability. The two age-group models had 4.54% and 4.47% transcripts with negative heritability in developing and aging worms, respectively.

To identify highly heritable genes we used a permutation test, as in the previous analysis. Transcript values were randomized prior to calculating permutated heritable values per transcript for developing and for aging worms. From the permutated values we identified the higher heritability value that allow less than 1% false positive (FDR = 0.01), $H^2 > 0.69$ for developing worms, $H^2 > 0.77$ for aging worms. Results were averaged from 100 permutations per gene.

Supplementary files

All the supplementary tables and figures are available from the author on http://www.anavinuela.com/thesis/thesis.html

Supplementary Table ST1: List of differentially expressed genes between N2 and CB4856 in juvenile (t1), reproductive (t2) and old worms (t3). Genes with a –log10 P-value > 2 in a linear model (gene expression = marker (effect) + error) were considered significant. Array ID = identification number for microarray spot. CDS name = coding sequence name. Gene name = public gene name. P-value = -log10 P-value for the significant test.

Supplementary Table ST2: List of differentially expressed genes between N2 and CB4856 in developing (t1 vs. t2) and aging worms (t2 vs. t3). Genes with a –log10 P-value > 2 in a linear model (gene expression = marker(effect) + relative age + interaction (marker*relative age) + error) were considered significant. *Array ID* = identification number for microarray spot. *CDS name* = coding sequence name. *Gene name* = public gene name. *P-value* = -log10 P-value for the significant test. Worksheets = *Developing (marker)* or genes differentially regulated by genotype in developing worms; *Developing (marker-*

age) for genes differentially regulated by genotype*age interaction in developing worms; *Aging (marker)* for genes differentially regulated by genotype in aging worms; *Aging (marker-age)* for genes differentially regulated by genotype*age interaction in aging worms.

Supplementary Table ST3: Significantly enriched Gene Ontology (GO) terms and functional domains. *GO terms worksheets* contain full GO enriched analysis for differentially expressed genes between parental strains in developing and aging worms, genotype effect and genotype*age interacting effect. *GO terms* are the GO id numbers. Total genes are the total number of genes in *C. elegans* assigned to each GO term. *Number of genes* are the number of regulated genes by genotype or genotype*age in each GO term. *P-values* were calculated using hypergeometric test, values lower that 0.01 and more than two genes in a category were considered significant. The hypergeometric test compared a group of 2779 GO terms, with 18,312 annotated genes. Gene annotation with GO terms were extracted from Wormbase release WB210.

Supplementary Table ST4: Table with GO enrichment analysis of overlapping genes between differentially expressed genes in the parental strains with at least an eQTL in developing and aging worms.

Supplementary Table ST5: Table with GO enrichment analysis of highly heritable genes in developing and aging worms.

Supplementary Table ST6: Table summary with Go terms enrichment analysis summary for highly heritable genes in developing and aging worms. First table shows GO terms significantly enriched in both groups of genes with a P-value < 0.01. Second table shows GO terms significantly enriched (P-value < 0.01) in developing worms genes but not in aging worms. Third table showed GO terms significantly enriched (P-value < 0.01) in aging worms genes but not in developing worms

Supplementary Figure SF1: Highly significant GO terms overrepresented in genes differentially expressed between parental strains and with at least a significant eQTL. Only GO terms with at least one P-value < 0.01 (-log10 P-value = 2) are shown. Full list of GO terms at 0.05 significance can be found in Supplementary Table ST3.

Supplementary Figure SF2: Number of overlapping genes between differentially expressed genes in the parental strains and with at least an eQTL at different thresholds in developing worms with genotype effect. *Bottom bar graphs* show number of genes with at

least an eQTL at different thresholds (from –log10 P-value =1 to –log10 P-value = 7). *Left bar graphs* show number of differentially expressed genes between parental strains at different thresholds. The number of overlapping genes is showed in the *centered graph*. The dark square indicates the number of overlapping genes at the joined-FDR (–log10 P-value =3 for eQTL analysis, –log10 P-value = 2 for parental strains analysis). A hypergeometric test was used to determine significance of the overlapping genes. The color legends for the significance are showed at the *Right graphs.*

Supplementary Figure SF3: Number of overlapping genes between differentially expressed genes in the parental strains and with at least an eQTL at different thresholds in developing worms with genotype*age interacting effect. See Supplementary Figure SF1 legend for details.

Supplementary Figure SF4: Number of overlapping genes between differentially expressed genes in the parental strains and with at least an eQTL at different thresholds in old worms with genotype effect. See Supplementary Figure SF1 legend for details.

Supplementary Figure SF5: Number of overlapping genes between differentially expressed genes in the parental strains and with at least an eQTL at different thresholds in old worms with genotype*age interacting effect. See Supplementary Figure SF1 legend for details.

Supplementary Figure SF6: Heritability and eQTL. Heritability of gene expression at juveniles (t1), reproductive (t2) and old (t3) worms was plotted. *y-axis* (left side) shows frequency of transcripts within a range of heritability. *x-axis* percentage of heritability per gene. The red line indicates (*y-axis*, *right side*) number of genes with an eQTL within the range of heritability.

Supplementary Figure SF7: Transgressive segregation, heritability and eQTLs. Upper graphs shows values for developing worms, lower graphs for aging worms.Left graphs (upper and lower) show the number of segregants which expression values transgress from the higher or lower parent value (*y-axis*) *versus* their heritability values (*x-axis*). Middle and right graphs plots the number of genes with eQTLs (*trans*-acting middle, *cis*-acting right) detected, their heritability values and their transgressive segregants. We detected more transgressive segregation in genes with *trans*-acting regulation and in developing worms than in aging worms and *cis*-acting eQTLs. Red line indicates the threshold for significant transgressive segregation.

References

Albuquerque, L. G. and K. Meyer (2005). Estimates of covariance functions for growth of Nelore cattle applying a parametric correlation structure to model within-animal correlations. **93**(3): 213-222.

Brachi, B., N. Faure, et al. (2010). Linkage and Association Mapping of *Arabidopsis thaliana* Flowering Time in Nature. *PLoS Genet* **6**(5): e1000940.

Brem, R. B., J. D. Storey, et al. (2005). Genetic interactions between polymorphisms that affect gene expression in yeast. *Nature* **436**(7051): 701-703.

Charmantier, A., C. Perrins, et al. (2006). Age-dependent genetic variance in a life-history trait in the mute swan. *Proceedings of the Royal Society B: Biological Sciences* **273**(1583): 225-232.

Emmons, S. W., M. R. Klass, et al. (1979). Analysis of the constancy of DNA sequences during development and evolution of the nematode Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America* **76**(3): 1333-1337.

Flint, J., W. Valdar, et al. (2005). Strategies for mapping and cloning quantitative trait genes in rodents. *Nat Rev Genet* **6**(4): 271-286.

Gems, D. and J. J. McElwee (2005). Broad spectrum detoxification: the major longevity assurance process regulated by insulin/IGF-1 signaling? *Mechanisms of Ageing and Development* **126**(3): 381-387.

Gutteling, E. W., J. A. G. Riksen, et al. (2007). Mapping phenotypic plasticity and genotype-environment interactions affecting life-history traits in Caenorhabditis elegans. *Heredity* **98**(1): 28-37.

Hunt, K. A., A. Zhernakova, et al. (2008). Newly identified genetic risk variants for celiac disease related to the immune response. *Nat Genet* **40**(4): 395-402.

Keurentjes, J. J. B., J. Y. Fu, et al. (2007). Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. *Proceedings of the National Academy of Sciences of the United States of America* **104**(5): 1708-1713.

Kirkwood, T. B. L. (1977). Evolution of ageing. *Nature* **270**(5635): 301-304.

Kirkwood, T. B. L. (2002). Evolution of ageing. *Mechanisms of Ageing and Development* **123**(7): 737-745.

Kuningas, M., S. P. Mooijaart, et al. (2008). Genes encoding longevity: from model organisms to humans. *Aging Cell* **7**(2): 270-280.

Li, Y., O. A. A. Alvarez, et al. (2006). Mapping determinants of gene expression plasticity by genetical genomics in C-elegans. *Plos Genetics* **2**(12): 2155-2161.

Li, Y., C.-C. Sheu, et al. (2010). Genetic variants and risk of lung cancer in never smokers: a genome-wide association study. *The Lancet Oncology* **11**(4): 321-330.

McElwee, J. J., E. Schuster, et al. (2004). Shared Transcriptional Signature in Caenorhabditis elegans Dauer Larvae and Long-lived daf-2 Mutants Implicates Detoxification System in Longevity Assurance. *J. Biol. Chem.* **279**(43): 44533-44543.

Moffatt, M. F., M. Kabesch, et al. (2007). Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* **448**(7152): 470-473.

Petretto, E., J. Mangion, et al. (2006). Heritability and Tissue Specificity of Expression Quantitative Trait Loci. *PLoS Genet* **2**(10): e172.

Rieseberg, L. H., M. A. Archer, et al. (1999). Transgressive segregation, adaptation and speciation. *Heredity* **83**(4): 363-372.

Smith, E. N. and L. Kruglyak (2008). Gene-Environment Interaction in Yeast Gene Expression. *PLoS Biol* **6**(4): e83.

Smyth, G. K. (2005). Limma: Linear models for microarray data. *Bioinformatics and Computational Biology Solution Using R and Bioconductor*. New York, Springer**:** 397-420.

Smyth, G. K. and T. Speed (2003). Normalization of cDNA microarray data. *Methods* **31**(4): 265-273.

Snoke, M. S. and D. E. L. Promislow (2003). Quantitative genetic tests of recent senescence theory: age-specific mortality and male fertility in Drosophila melanogaster. *Heredity* **91**(6): 546-556.

Viñuela, A., L. B. Snoek, et al. (2010). Genome-wide gene expression regulation as a function of genotype and age in C. elegans. *Genome Research* **20**(7): 929-937.

Visscher, P. M., W. G. Hill, et al. (2008). Heritability in the genomics era - concepts and misconceptions. *Nat Rev Genet* **9**(4): 255-266.

Wray, G. A., M. W. Hahn, et al. (2003). The Evolution of Transcriptional Regulation in Eukaryotes. *Molecular Biology and Evolution* **20**(9): 1377-1419.

Yang, Y. H., and Thorne, N. P. (2003). Normalization for two-color cDNA microarray data. *D. R. Goldstein (ed.), Science and Statistics: A Festschrift for Terry Speed* **40**: 403-418.

Zou, Y., Z. Su, et al. (2009). Uncovering genetic regulatory network divergence between duplicate genes using yeast eQTL landscape. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **312B**(7): 722-733.

Chapter: 6

General discussion: Gene expression changes and robustness theory

Ana Viñuela

The main objective of my thesis was to study the genetic architecture of gene expression in *C. elegans* by means of perturbation. By exposing the worm to pesticides, different temperatures and their combinations, we perturbed *C. elegans*. An additional perturbation consisted of ageing experiments in recombinant inbred lines, where different alleles are sources of perturbations. We followed the reasoning that changes in gene expression may lead to differential phenotypes revealing elements of the genetic architecture of gene expression. However, we found that multiple and diverse changes in gene expression, which were induced by perturbations, resulted in relatively small phenotypic changes. In other words, we observed that the worm appeared to be a robust biological system in its final output, despite induced changes at the gene expression level. This robustness seems to be a property of many biological systems. It allows them to maintain their functions despite perturbations and to produce relatively invariant phenotypes (Kitano, 2004). Phenotypic robustness has been observed at many levels of biological organization and phenotypes including development, protein folding or physiological homeostasis, but the actual mechanism behind it is not well understood (de Visser *et al.*, 2002; Masel and Siegal, 2009). Because perturbation experiments are the main experimental approach to study robustness, our results are relevant to discuss the implications of robustness theory in the genetic architecture of gene expression in *C. elegans.*

Phenotypic robustness may be observed in different forms depending on the type of perturbations and the phenotype itself. Since the main parameters that define a phenotype are genotype and environment, it is common to distinguish between environmental and genetic robustness. De Visser *et al.* (2002) describe environmental robustness as any kind of biological mechanism that reduces phenotypic sensitivity to non-heritable perturbations. Genetic robustness, on the other hand, refers to the biological elements that confer insensitivity to heritable variation. As a result, mechanisms behind robustness allow the accumulation of cryptic genetic variation (de Visser *et al.*, 2002). This variation is considered hidden since it is not observable in the form of phenotypic changes. However, cryptic variation may be revealed by means of a perturbation, environmental or genetic, strong enough to break down the robust system. As a consequence, new and possibly advantageous phenotypic variation may become apparent. For this reason, cryptic variation

is considered the raw material for evolution (Masel and Trotter, 2010). Yet, the specific molecular mechanisms behind the phenotypic robustness and the accumulation of hidden variation are not well understood.

Definitions of environmental robustness usually distinguish between external and internal perturbations (de Visser *et al.*, 2002). Examples of robustness to external environmental perturbations like temperature and toxicants are investigated in Chapter 2 and 3. There, we first analyzed the transcriptional responses to the pesticides chlorpyrifos (CPF) and diazinon (DZN). The main conclusion was that relatively similar environmental stressors can induce the expression of different genes. Still, the different genes produce similar toxic responses and survival rate (Chapter 2, Figure 4). If we consider a lack of toxicant environment as the norm, the exposure to pesticides regulated additional genes and pathways to maintain the robust survival phenotype. But it is unfeasible for an organism to have an additional response, pathway or group of genes for each environmental factor to which it becomes exposed during its lifetime. As an alternative, redundant genes or alternative metabolic pathways allow to maintain the phenotype despite the environmental perturbations. Moreover, the different environmental variables may differentially affect the system. Chapter 3 described the effects of multiple environmental factors on genome wide gene expression. There, environmental perturbations like pesticides and temperature were combined. We observed a differential influence in gene expression by the studied factors, suggesting a different reaction of the system in response to perturbations. The transcriptional reactions were buffered by the system that maintains the robust phenotype despite the external interacting environmental insults.

Phenotypic robustness requires behavioral and molecular mechanisms to maintain a stable phenotype. Assuming a simplified situation in which an organism cannot change its behavior and avoid the source of external environmental perturbation; the molecular mechanism becomes essential to buffer responses and maintain the phenotype. The biological switches involved in environmental and genetic robustness have been called *capacitors* for their ability to reveal cryptic variation and capacitate adaptation to perturbations (de Visser *et al.*, 2002; Masel and Siegal, 2009). It has been suggested that

the ability to connect to a large number of genes and to function on mRNA production are common characteristics of phenotypic capacitors (Levy and Siegal, 2008; Masel and Siegal, 2009). Those genes may act as genetic buffers due to their capability to interact with many genes in many different pathways (Lehner *et al.*, 2006). In this regard, we describe also in Chapter 3 the effect of multiple environmental factors on nuclear receptors (NRs). We were interested in their transcriptional response because NRs seem to play an important role in the diversifying responses to environmental perturbations. Indeed, the large number of genes coding for NRs in *C. elegans* compared to other species has been suggested as a sophisticated mechanism to ensure growth and reproduction in its natural environment (Sluder *et al.*, 1999). Therefore, their regulatory activity of responses to environmental insults suggested that NRs may act to buffer environmental perturbations in *C. elegans*. Moreover, functional redundancy has been proposed as a molecular mechanism for buffering. Redundancy results in robustness against mutations, environmental changes or internal noise in situations where part of the system compensates totally or partially the loss of another (Wagner, 2005; Wang and Zhang, 2009). The relatively high number of NRs in the *C. elegans* genome supports the idea that some NRs may contribute to environmental robustness, acting as molecular switches of transcriptional responses with redundant functions.

Internal environmental perturbation can not be avoided. If protein concentrations are relevant for regulating age specific processes; stochasticity of gene expression and fluctuations in gene product concentration can be a source of perturbation (Masel and Siegal, 2009). Developmental phenotypes are probably the most well studied processes and clear examples of robustness. In *C. elegans* for instance, vulva development and the final cell fate output pattern is highly stable. The worm uses cellular and pathway redundancy to maintain the phenotypic stability despite differences in environmental conditions. Likewise, mutational and environmental perturbations have revealed genetic variation in gene pathways involved in development, and have suggested underlying cryptic variation and robustness (Braendle *et al.*, 2007). But, during development the organisms age as well. The molecular mechanisms that have established robustness in developmental phenotypes may be broken down later in life revealing senescent phenotypes. Age, in that case, may be

considered an additional environmental capacitor that reveals cryptic variation changing the internal environment. Because aging is not a regulated process as compared to survival (Lithgow, 2006), phenotypes linked to aging may be considered an outcome of the breakdown of robustness. Consequently, cryptic variation with neutral effects in young robust organisms became non-neutral in old organisms as they are revealed by age. In conclusion, robustness to internal environmental perturbations in young organisms and during development has been well studied. However, the relation between robustness and aging is not evident. It seems clear though, that survival depends on how robust the biological system is to internal and external perturbations. And likewise, that age is an internal source of perturbation which affects robustness and determines life span.

Our results on heritable gene expression regulation with age revealed different levels of robustness as a consequence of genetic perturbations during aging. An organism is alive as long as the survival phenotype is robust against perturbations. In that sense, variation in life span among the RILs (Chapter 4, Figure 3) suggests different sensitivities to perturbations. Since all segregants were maintained in optimal growing conditions, short-living RILs imply low levels of robustness and high sensitivity to perturbations. On the other hand, long-living RILs manifested robustness to perturbation resulting in a longer lifespan. At the gene transcription level, differences in robustness may be explained as differences in gene regulatory network structure (Gjuvsland *et al.*, 2007). Network properties like feedbacks, connectivity, modularity or redundancy have been associated with robustness in the biological system (Albert *et al.*, 2000; Kitano, 2004; Venkatesh *et al.*, 2004; Ma *et al.*, 2006). Moreover, robustness can also be shown in response to heritable genetic perturbations. These include not only mutations, but also recombinations and patterns of epistatic interactions produced by new allelic combinations (de Visser *et al.*, 2002; Schlichting, 2008; Masel and Siegal, 2009). In that sense, we observed changes in the gene regulatory network of *C. elegans* with age in our eQTL study. Such a difference has been also reported in mice, where changes in coexpression network connectivity have been associated to aging (Southworth *et al.*, 2009). Thus, network structure and robustness differences might explain life span variation in the RILs. We attributed those differences in robustness to less strong regulatory elements active in aged organism. Initially, the number

of detected heritable patterns of gene expression was reported to decrease with age (Chapter 4). We suggested that gene expression regulation might as well decrease as a consequence of aging. However, eQTL detection is quite limited to detection of strong regulatory elements and prone to false positives (Carlborg *et al.*; Breitling *et al.*, 2008). For this reason, we considered it likely a decrease in the power of eQTL detection as a result of changes in the gene regulatory network structure in old worms. We investigated this possibility in Chapter 5 where differentially expressed genes between the parental strains indicated the existence of undetected regulation differences (Chapter 5, Table 1). Moreover, the imbalance between highly heritable genes and eQTLs in old worms suggested a decrease in linkage detection (Chapter 5, Figure 2). And finally, the alternative mapping strategy promotes a hypothesis in which multiple regulatory elements with small and additive effect regulate gene expression in old worms; as opposed to the more common strong single regulations active in young worms (Chapter 5, Figure 4). Altogether, our results support the idea that gene expression regulation changes in response to internal environmental perturbations induced by age. Likewise, it is suggested that different allelic combinations introduced genetic perturbations that modified the robustness capacity among the RILs. Therefore, the capacity to maintain a robust phenotype despite internal and external perturbations was different and leads to different life span values.

Longevity might be a phenotypic consequence of robustness. For this reason, longevity models may provide an explanation of the molecular mechanism underlying robustness differences among the RILs. The so-called *longevity* genes are associated to damage repair and energy distribution functions. Life span is limited by the available energy invested into repairing the molecular damage accumulated with age (Kirkwood, 2002). In Chapter 5, we detected gene expression variation among parental and segregants in longevity genes. This phenotypic variation was observed in old worms because cryptic genetic variation was revealed at that age. Whether cryptic variation is a consequence of or a cause for evolutionary adaptation is not clear, and this has been a matter of intense discussion over the past years (Gibson and Wagner, 2000; de Visser *et al.*, 2002; Braendle *et al.*, 2007; Manolio *et al.*, 2009; Masel and Trotter, 2010). Likewise, the molecular mechanisms underlying robustness are not well understood and hotly debated. For some authors,

robustness to environmental and genetic perturbations requires different mechanisms (Masel and Siegal, 2009), while others have indentified coupling functions in genes involved in genetic, environmental and stochastic robustness (Lehner, 2010). In both cases, robustness provides the potential to survive severe environmental and genetic perturbations in the form of cryptic genetic variation. Within this perspective, *longevity* genes emerge as relevant molecular mechanism underlying robustness. And it suggests that the evolutionary forces acting on robustness might be as well decisive on the evolution of aging.

Finally, it is important to stress that our study was limited to the gene transcription level. The genotype-phenotype map involves many other levels that make it difficult to connect gene expression to a complex phenotype such as life span. Indeed, it has already been shown for plants that many genetic effects detected in eQTLs may not be linked to physiological or life history phenotypes. In other words, the variation in gene expression does not always propagate to phenotypic variation (Fu *et al.*, 2009). This is a consequence of phenotypic buffering; and therefore, of robustness.

Conclusions and final remarks

C. elegans has developed multiple mechanisms to survive under the influence of highly variable environments. The genotype-phenotype map has proved to be robust to environmental changes in order to maintain development, reproduction and other physiological functions. In this thesis, I showed how the transcriptional system is highly sensitive to external and internal perturbations. Still, the variation observed in gene transcripts not always translates to phenotypic variation, probably due to buffering mechanisms. In some cases however, the buffering mechanisms may fail and the robustness breaks. Then, phenotypic variation is observed and genetic cryptic variation is revealed. The molecular mechanisms underlying robustness are not well understood; but it seems likely that robustness provides the potential to survive severe environmental and genetic perturbations in the form of cryptic genetic variation. Such variation may be reveal under perturbation as new molecular and physiological phenotypes. Future work might focus on other levels of biological organization to better explain robustness, since gene transcription is only one step within the long genotype-phenotype path.

References

Albert, R., H. Jeong, et al. (2000). Error and attack tolerance of complex networks. *Nature* **406**(6794): 378-382.

Braendle, C., J. Milloz, et al. (2007). Mechanisms and Evolution of Environmental Responses in Caenorhabditis elegans. *Current Topics in Developmental Biology*, Academic Press. **Volume 80:** 171-207.

Breitling, R., Y. Li, et al. (2008). Genetical Genomics: Spotlight on QTL Hotspots. *PLoS Genet* **4**(10): e1000232.

Carlborg, Ã., D. J. De Koning, et al. Methodological aspects of the genetic dissection of gene expression. *Bioinformatics* **21**(10): 2383-2393.

de Visser, J. A. G. M., J. Hermisson, et al. (2002). Perspective: Evolution and Detection of Genetic Robustmess. *Evolution* **57**(9): 1959-1972.

Fu, J., J. J. B. Keurentjes, et al. (2009). System-wide molecular evidence for phenotypic buffering in Arabidopsis. *Nat Genet* **41**(2): 166-167.

Gibson, G. and G. Wagner (2000). Canalization in evolutionary genetics: a stabilizing theory? *BioEssays* **22**(4): 372-380.

Gjuvsland, A., E. Plahte, et al. (2007). Threshold-dominated regulation hides genetic variation in gene expression networks. *BMC Systems Biology* **1**(1): 57.

Kirkwood, T. B. L. (2002). Evolution of ageing. *Mechanisms of Ageing and Development* **123**(7): 737-745.

Kitano, H. (2004). Biological robustness. *Nat Rev Genet* **5**(11): 826-837.

Lehner, B. (2010). Genes Confer Similar Robustness to Environmental, Stochastic, and Genetic Perturbations in Yeast. *PLoS ONE* **5**(2): e9035.

Lehner, B., C. Crombie, et al. (2006). Systematic mapping of genetic interactions in Caenorhabditis elegans identifies common modifiers of diverse signaling pathways. *Nat Genet* **38**(8): 896-903.

Levy, S. F. and M. L. Siegal (2008). Network Hubs Buffer Environmental Variation in *Saccharomyces cerevisiae*. *PLoS Biol* **6**(11): e264.

Lithgow, G. J. (2006). Why aging isn't regulated: A lamentation on the use of language in aging literature. *Experimental Gerontology* **41**(10): 890-893.

Ma, W., L. Lai, et al. (2006). Robustness and modular design of the Drosophila segment polarity network. *Mol Syst Biol* **2**.

Manolio, T. A., F. S. Collins, et al. (2009). Finding the missing heritability of complex diseases. *Nature* **461**(7265): 747-753.

Masel, J. and M. L. Siegal (2009). Robustness: mechanisms and consequences. *Trends in Genetics* **25**(9): 395-403.

Masel, J. and M. V. Trotter (2010). Robustness and Evolvability. *Trends in Genetics* **26**(9): 406-414.

Schlichting, C. D. (2008). Hidden Reaction Norms, Cryptic Genetic Variation, and Evolvability. *Annals of the New York Academy of Sciences* **1133**(1): 187-203.

Sluder, A. E., S. W. Mathews, et al. (1999). The Nuclear Receptor Superfamily Has Undergone Extensive Proliferation and Diversification in Nematodes. *Genome Research* **9**(2): 103-120.

Southworth, L. K., A. B. Owen, et al. (2009). Aging Mice Show a Decreasing Correlation of Gene Expression within Genetic Modules. *PLoS Genet* **5**(12): e1000776.

Venkatesh, K. V., S. Bhartiya, et al. (2004). Multiple feedback loops are key to a robust dynamic performance of tryptophan regulation in Escherichia coli. *FEBS Letters* **563**(1-3): 234-240.

Wagner, A. (2005). Robustness, evolvability, and neutrality. *FEBS Letters* **579**(8): 1772-1778.

Wang, Z. and J. Zhang (2009). Abundant Indispensable Redundancies in Cellular Metabolic Networks. *Genome Biology and Evolution* **1**: 23-33.

Summary

Most organisms are exposed to a continuously changing environment throughout their life. For instance the ambient temperature is usually not constant and many species are exposed to a diverse range of anthropogenic stressors like toxic compounds. Moreover, individuals are prone to genetic changes due to mutation and allelic recombinations. All these factors might affect particular phenotypes, while others remain unchanged. This thesis provides insight into how phenotypic traits are affected by external stress factors and allelic recombinations in the nematode *Caenorhabditis elegans* (Nematoda; Rhabditidae). Because phenotypes and their variation may be explained by variation in gene expression, this thesis explored the architecture of gene expression and some of the elements that contribute to gene expression.

Chapters 2 and 3 focus on environmental stressors with i) a specific target (two organophosphorus pesticides) and ii) a non-target mode of action (temperature) to study their influence on gene expression. A single genotype, the canonical wild type strain Bristol (N2) was used to study the effect of interacting pesticides by exposing nematodes to a toxicant mixture at two different temperatures Analysis revealed common transcriptional responses related to detoxification, stress, innate immunity, and transport of lipids to all treatments. It was found that for both pesticides these similar processes were regulated by different gene transcripts in single and combined treatments. These results also showed that the effect of a mix of low doses of pesticides is not a summed effect of the single components. Moreover, increased temperature elevates the toxic consequences to the pesticides exposures. This toxicity gain is attributed to an elevated uptake and accumulation of the toxicants in the organisms. These results support the idea that the observed higher toxicity of pesticides with temperature might be a consequence of gene-environment interactions affecting detoxification genes. Together, the first part of this thesis illustrates the intense crosstalk between gene pathways in response to interacting environmental stressors in *C. elegans*.

The second part of the thesis elaborates on the influence of different genotypes as multiple perturbations on gene expression. How the genotype-phenotype relationship progresses with age was investigated using a quantitative genetics approach (genetical genomics). We

performed a genetic mapping strategy of gene transcription variation (expression-QTL, eQTL) to explore the dynamics of regulatory loci affecting genome-wide gene expression at three different ages. We used a recombinant inbred line (RIL) population generated from a cross between the *C. elegans* strain N2 and the wild type CB4856 in Chapter 4. Also, we investigated the influence of age to reveal a genotype-by-age effect ($_{gxa}$eQTL) on gene expression. The total number of detected eQTL decreased with age whereas the variation in expression increased. In developing worms, the number of genes with increased expression variation (1282) was similar to the ones with decreased expression variation (1328). In aging worms the number of genes with increased variation (1772) was nearly 5 times higher than the number of genes with a decreased expression variation (373). Furthermore, the number of *cis*-acting eQTL in juveniles decreased by almost 50% in old worms whereas the number of *trans*-acting loci decreased by ~27%, indicating that *cis*-regulation becomes relatively less frequent than *trans*-regulation in aging worms. Our findings demonstrate that eQTL patterns are strongly affected by age and suggest that gene network integrity declines with age. To better understand the changes in the gene network with age, gene expression profiles of N2 and CB4856 were generated for Chapter 5. We explored gene expression heritability and transgression as genetic parameters for the analysis of gene expression divergence in different genotypes. The average broad sense heritability was similar in developing and aging worms; but the gene expression variance that can be attributed to genetic variance in each gene changes with age. It can be proposed that regulation became more polygenic in aging worms. These changes explain the decrease in detected eQTLs. Likewise, it explains the imbalance between highly heritable genes and eQTLs in aging worms.

Chapter 6 discusses the main conclusion of this thesis in the context of the robustness theory. Robustness in biological systems provides the potential to survive severe environmental and genetic perturbations in the form of cryptic genetic variation. The variation we observed in gene transcripts due to external and internal perturbations not always translated to physiological phenotypic variation. In some cases however, the mechanisms underlying phenotypic robustness failed and phenotypic variation was observed. Such genetic cryptic variation was revealed as new molecular and physiological phenotypes.

Samenvatting

De meeste organismen worden gedurende hun hele leven blootgesteld aan milieuomstandigheden die continu veranderen. De omgevingstemperatuur is bv. vaak niet constant en vele soorten worden blootgesteld aan diverse anthropogene stressfactoren zoals toxische stoffen. Tevens zijn individuen onderhevig aan genetische veranderingen zoals mutaties en nieuwe allelische recombinaties in het genoom. Dit alles kan tot gevolg hebben dat bepaalde fenotypische kenmerken veranderen terwijl anderen ongewijzigd blijven. Dit proefschrift geeft inzicht hoe fenotypische eigenschappen veranderen onder invloed van externe stressfactoren en allelische recombinaties in de nematode *Caenorhabditis elegans* (Nematoda; Rhabditidae). Omdat fenotypes voor een groot deel bepaald worden door variatie in gen expressie is de invloed op de architectuur van gen expressie bestudeerd.

In hoofdstuk 2 en 3 wordt de invloed van twee verschillende externe stressfactoren bestudeerd op de genoomwijde gen expressie. De eerste factor heeft effect op een specifieke target-receptor (twee verschillende organofosfor pesticiden) en de tweede factor heeft een aspecifiek effect (temperatuur). Bestudeerd is de invloed van de pesticiden en de temperatuur op gen expressie in één enkel genotype, genaamd Bristol N2. Gevonden werd dat blootstelling aan beide pesticiden leidde tot vergelijkbare gen expressie patronen. Deze waren gerelateerd aan detoxificatie, stress response, "innate" immuniteit en lipiden transport. Voor beide pesticiden werden deze processen evenwel gereguleerd door verschillende genen. De resultaten laten ook zien dat het effect van een zgn. lage-concentratie mix van beide stoffen niet gelijk is aan de som van de enkele stoffen. Ook bleek dat een hogere temperatuur gepaard ging met een toename van de toxiciteit, wat vermoedelijk het gevolg was van een toename van opname en accumulatie van de toxicanten. Deze resultaten suggereren dat de hogere toxiciteit bij een toename van de temperatuur een gevolg is van gen-omgevings interacties van genen betrokken bij detoxificatie. Dit eerste deel van het proefschrift laat de intensieve interactie zien tussen gen "pathways" en milieuomstandigheden in *C. elegans*

In het tweede deel van dit proefschrift komt de invloed van verschillende genotypes op de gen expressie aan de orde. Hierbij wordt ingegaan op de genotype-fenotype relatie als

functie van de leeftijd met behulp van een kwantitatieve genetische benadering (genetical genomics). Genetische kartering van gen expressie (detectie van expressie-QTL) werd gebruikt voor het bestuderen van de dynamiek van gen expressie regulatie in jonge en oude wormen. Hiervoor is gebruik gemaakt van een recombinante inteelt populatie gebaseerd op een kruising tussen de *C. elegans* isolaten Bristol N2 en CB4856 (Hoofdstuk 4). Het totale aantal gedetecteerde eQTL nam af met de leeftijd terwijl de variatie in gen expressie toenam. Het aantal genen met toegenomen expressie variatie in jonge wormen (1282) was ongeveer gelijk aan het aantal met afgenomen expressie variatie niveaus (1328). In oudere wormen was het aantal genen met toegenomen variatie (1772) bijna 5 keer hoger dan het aantal met afgenomen variatie (373). Het aantal *cis-located* eQTL in jonge wormen nam met 50% af in oude wormen terwijl het aantal *trans-located* eQTL afnam met 27%. Dit laat zien dat *cis* regulatie relatief minder vaak voorkomt in oudere wormen. De resultaten suggereren dat de integriteit van gen netwerken sterk afneemt met de leeftijd. Om meer inzicht te krijgen in de netwerk veranderingen, zijn de gen expressie profielen van de wilde isolaten Bristol N2 en CB4856 bestudeerd (Hoofdstuk 5). Hierin wordt de erfelijkheid en de transgressie als genetische parameters voor de analyse van gen expressie verschillen tussen de isolaten bestudeerd. De gemiddelde "broad sense heritability" was gelijk in jongen en oude wormen maar de variatie in gen expressie die kan worden toegeschreven aan genetische variatie in elk gen verandert met de leeftijd. De resultaten suggereren dat de regulatie meer polygeen wordt in oudere wormen. Dit verklaart de afname van het aantal gedetecteerde eQTL en de onevenredige verdeling tussen genen met hoge overerfbaarheid en het aantal eQTL in oudere wormen.

Hoofstuk 6 bediscussieert de resultaten van dit proefschrift in het licht van de zgn. "robustness theory". Robuustheid in biologische systemen is de basis voor het vermogen om te overleven onder verschillende milieuomstandigheden en genetische veranderingen in de vorm van cryptische variatie speelt hierin een belangrijke rol. De variatie in gen expressie door interne en externe verstoringen leidden niet in elk geval tot fenotypische veranderingen. In sommige gevallen faalde de robuustheid en werd fenotypische verandering waargenomen. Deze cryptische variatie werd zowel op moleculair als fysiologisch niveau waargenomen.

Resumen

Casi todas las formas de vida están continuamente expuestas a cambios ambientales. Por ejemplo, la temperatura ambiental no suele ser constante, o se exponen a un amplio rango de sustancias tóxicas, o sufren cambios genéticos debido a mutaciones o recombinaciones alélicas. Todos estos factores pueden afectar a ciertos fenotipos, mientras que otros permanecen invariables. Esta tesis doctoral proporciona una visión de cómo los fenotipos se pueden ver afectadas por factores de estrés externos e internos en el nematodo *Caenorhabditis elegans* (Nematoda; Rhabditidae). Para ello, se centra en la exploración de la arquitectura de la expresión genética y algunos de los elementos que contribuyen a ella, debido a que en muchas ocasiones los fenotipos y sus variaciones se pueden explicar en base a la variación en expresión genética.

Los capítulos 2 y 3 se centran en aquellas perturbaciones que: a) afectan a dianas moleculares concretas (pesticidas órganofosfatos) y b) afectan a dianas no específicas (temperatura), para estudiar su influencia sobre la expresión genética. Para ello, la línea estándar de laboratorio N2 se usó en la generación de perfiles de transcripción del genoma completo de *C. elegans* expuesto a diferentes pesticidas y temperaturas. Los resultados describen respuestas comunes a distintos pesticidas y relacionadas con detoxificación, estrés, inmunidad innata y transporte de lípidos, pero regulados por genes diferentes. Igualmente se estudiaron las diferencias y similitudes entre las respuestas a los tratamientos individuales de varios pesticidas y la mezcla de ellos. Los resultados muestran que el efecto de una mezcla de pesticidas no es la suma exacta de los efectos de los tratamientos individuales a nivel de transcripción génica. Además, muestran como una elevada temperatura puede aumentar las consecuencias tóxicas de la exposición a tóxicos. Los resultados apoyan la idea de que la ganancia en toxicidad con la temperatura se relaciona con interacciones gen-ambiente durante la expresión de genes de detoxificación.

Los capítulos 4 y 5 analizan la influencia de las diferencias genotípicas como múltiples perturbaciones de la expresión genética. Usando técnicas de genética cuantitativa (genómica genética) y una estrategia de mapeo de la variación de la expresión genética (expresión-QTL, eQTL) exploramos la dinámica de su regulación a tres diferentes edades. En el capítulo 4 se usó una población de líneas recombinantes (RIL) generadas a partir de

los cruces entre N2 y la línea CB4856. Los resultados revelaron la existencia de locus cuya influencia sobre expresión genética es dependiente del genotipo y de la edad ($_{gxa}$eQTL). El número total de eQTLs detectados disminuyó con la edad, mientras que la variación en expresión genética aumentó. Además, la cantidad de eQTLs con actividad en *cis* (local) en nemátodos juveniles disminuyó casi un 50% en los envejecidos, mientras que el número de locus con actividad reguladora en *trans* (distante) disminuyó en un ~27% indicando, que la regulación en *cis* resulta relativamente menos frecuente en nemátodos envejecidos. Nuestros resultados demuestran que los patrones heredables de expresión genética (eQTL) están fuertemente afectados por la edad y sugieren, que la integridad de la red de conexiones genéticas disminuye con la edad. Para entender mejor estos cambios, generamos para el capitulo 5 perfiles de expresión genética de las líneas parentales de la población recombinante (N2 y CB4856). Comparando dichos perfiles, exploramos la heredabilidad y la transgresión de la expresión genética así como el efecto de la edad en ellos. La media de heredabilidad de la expresión genética en sentido amplio para todo el genoma fue similar en nemátodos en desarrollo y en envejecimiento, pero a nivel de genes la variación de la expresión genética que puede ser atribuida a variación genética cambió con la edad. Es por ello que se considera que la regulación de la expresión genética se modifica, volviéndose más poligénica y por tanto compleja en nemátodos envejecidos, lo que explicaría la reducción en el número de eQTLs detectados. Simultáneamente, estos cambios explican el desequilibrio entre el número de genes con valores de expresión altamente heredable y el número de eQTLs en nemátodos envejecidos, así como el menor número de eQTLs atribuidos a genes cuya expresión transgrede la media de los padres.

El capitulo 6, discute las principales conclusiones de esta tesis en el contexto de la teoría de la robustez. La robustez en los sistemas biológicos proporciona el potencial para sobrevivir a condiciones ambientales severas y a perturbaciones genéticas. La variación observada en la expresión genética no siempre es traducida a variación fenotípica, porque se acumula en forma de variación genética críptica. En algunos casos, los mecanismos de robustez fenotípica fallan y se puede observar la variación fenotípica oculta. Esa variación críptica se muestra como nuevos fenotipos moleculares y fisiológicos que pueden ser, en ciertos ambientes, ventajosos.

Acknowledgments

Life in a foreign country is a job in itself. You have to learn the language (upps!), to cycle (under the rain!!), to eat pepernoten (but only during two months a year!!), or to drink beers in special glasses (which is the hardest task, obviously). But at the same time, It is a very rewarding experience, because you have the chance to meet great people and learn from them. I want to thank all of you who made me feel at home in Wageningen during the last four years.

But, first of all I want to thank Jan K., because you gave me opportunity to become a scientist. Not only that, you showed me the kind of scientist I aspire to become one day. You are always enthusiastic, passionate and eternally optimistic. And that of course makes me mad at times. But, you are a great person and that makes you a great supervisor. Thank you for everything you taught me, about science and about myself. Joost, my *pouring* boy. You help me so much, and not only pouring my dishes. To say "thank you" would not be enough. So, let's relax like worms and have some beers now that the work is done. Basten, you are a great and incredible patient teacher who manages to teach me that science is just a toy for grow-ups, with a lot of pieces and joy in it. To the three of you, thank you for helping me do a good job (and on time!), always with a smile on your faces.

Jaap, it took me three months in Nema to figure out that you are THE professor. That already says a lot about you. You built here something unique and it will always be in my heart. Agsniezka, you helped me a lot at the very beginning, more than you think. Your advices will always travel with me. Lisette, bedankt voor uw geduld. Een kleuterschool was nog nooit zo goed georganiseerd. Roel, your enthusiasm and your *retardedness* helped me remember why I like science so much. It was really fun to have you around for work and even better to have you around for beers. Rita, It was very nice to share the office with you during my last year. Good luck with your own book, you are almost there.

To the downstairs team, a.k.a. the Gang: Sven (The Cheshire Cat, always with you mischievous smile), Paul (the best *annoying* teacher ever!), Hans, Jet, Kasia. You made my work in the lab much easier than it was, but please take singing lessons. Hanny, it was a pleasure and an honor to share the office with you for two years. I miss our little chats in the morning and the peace and calm that you bring to the office. Mike Herman, you showed me other science, other music and your great country. I just need to know your trick to receive packages from Sinteklaas….

To LMA, just one message: Put a nematode in your life!! Lotte and Ruud, you know it. Arjen, thank you so much for teasing me. You know that too much courtesy could have being very unhealthy for me. Thank you also for the Friday beers at the Vlam, I would not have survived without them. Jan vdV., thank you for being you.

Of course many people make Nema a special place, but not all fit in here. To all of you here is my gratitude for your help, the coffee breaks, the nice environment and the great *labuitjes*. Special thanks to Rikus, because nothing works in Nema without machines or beers.

To my Dutch family: Lotte, Ruud and Annemarie. You have been with me all the way along in this crazy trip called PhD. I may be the first, but I am certainly not the last. For all the bad moments that you helped me get through, for all the good moments that you enjoyed with me, thank you so much. You made Wageningen *mijn thuis*. Tineke, you own a warm place in my heart, good luck with everything. Kamila and Tomek, you are always in a positive mood and willing to enjoy a good party. I wish you the best and of course good luck with Nina Alice. Lenny, it was great to share space and time with you. Good luck on your next step, I am sure you will do great.

Alicia, Asier, Amaia, Jonatan y Alicia, gracias chicos por estar ahí, aconsejarme de vez en cuando y tomaros unas cañitas conmigo en cualquier parte del mundo. Fletcher, Nacho, Pablo, Antonio, Jose, Rafa, Mª Carmen, Nesli y David. Estáis aquí por que me seguís llamando amiga a pesar de la distancia. Por qué una vez os dije que me iba y no volvía, y al día siguiente me llamasteis para ir de cañas. Estáis aquí por que somos todos unos perros y nadie más puede entenderlo. Primos, primas, tios y tias, en Sevilla, Madrid o Sanlucar. Es que al final sois muchos y no me cavéis todos, pero gracias por vuestro apoyo. Abu, un beso muy grande.

Tato, Nena y Fer, estar lejos de vosotros es más difícil de lo que os pueda parecer. Sin vuestro apoyo nunca habría dado el primer paso de irme fuera y mi doctorado nunca habría sido posible. Este libro está dedicado a vosotros.

Wageningen, 4 de Marzo de 2011.

Publications

Articles:

Viñuela, A.*; Snoek, L.B.*; Riksen, J.A.G.; Kammenga, J.E. (2010) Genome-wide gene expression regulation as a function of genotype and age in *C. elegans*. **Genome Research (20): 929-937.** * equal contribution.

Viñuela, A.; Snoek, L.B..; Riksen, J.A.G.; Kammenga, J.E. (2010) Genome-Wide Gene Expression Analysis in Response to Organophosphorus Pesticide Chlorpyrifos and Diazinon in *C. elegans*. **PLoS One, 5(8): e12145. doi:10.1371/journal.pone.0012145**

Conference proceedings and meetings:

Viñuela, A.; Snoek, L.B.; Riksen, J.A.G.; Kammenga, J.E.. Heritable expression regulation in aging *C. elegans*. Contributed talk in: *GENESYS workshop at the 11th International Conference on Systems Biology*, 10-16 October 2010. Edinburgh, UK.

Viñuela, A.; Snoek, L.B.; Riksen, J.A.G.; Kammenga, J.E.. Heritable expression regulation in aging *C. elegans*. Contributed talk in: *Evolutionary Biology of Caernorhabditis and other Nematodes*, 5-8 June, 2010. Wellcome Trust Conference Centre. Hinxton, Cambridge. UK.

Viñuela, A.. Natural variation in the dynamics of global gene transcription regulation during aging in *C. elegans*. Invited talk in: *ECOGEN and MCDB Seminar*, 2010-01-28. Kansas University (USA)

Viñuela, A.; Snoek, L.B.; Riksen, J.A.G.; Kammenga, J.E. (2010) Transcriptional consequences of recombination in aging *C. elegans* Ana Viñuela, L. Basten Snoek, Joost A. G. Riksen, Jan E. Kammenga. Poster presentation in*: Evolutionary Biology of Caernorhabditis and other Nematodes*, 5-8 June, 2010. Wellcome Trust Conference Centre. Hinxton, Cambridge, UK.

Viñuela, A.; Snoek, L.B.; Riksen, J.A.G.; Kammenga, J.E. (2010) Natural variation of the dynamics of global gene transcription regulation in aging worms. Poster presentation in *Keystone Symposia, New Insights into Healthspan and Diseases of Aging: From Molecular to Functional Senescence*, Tahoe City, California, 2010-01-31/ 2010-02-05.

Viñuela, A.; Snoek, L.B.; Riksen, J.A.G.; Kammenga, J.E.. Natural variation of the dynamics of global gene transcription regulation in aging worms. Contributed talk in: *16th Benelux Congress of Zoology on Zoology: shaped by evolution*, Wageningen, The Netherlands, 28-30 October 2009. - KNDV, NIOO, 16th Benelux Congress of Zoology, Wageningen, the Netherlands, 2009-10-28/ 2009-10-30

Viñuela, A.; Snoek, L.B.; Riksen, J.A.G.; Kammenga, J.E.. Molecular mechanisms of toxicity at the gene transcription level. Progress presentation in: *5th Ecogenomics annual meeting,* 2009-04-16. Amsterdam, the Netherlands.

Snoek, L.B.; **Viñuela, A.**; Gutteling, E.W.; Riksen, J.A.G.; Kammenga, J.E. (2009) Natural variation reveals the autophagy gene unc-51 as a key connection between sex and death in *C. elegans*. Poster presentation in: *17th International C. elegans meeting*, University of California, Los Angeles, June 24-28, 2009. - Los Angeles: 17 th International C. Elegans Meeting, 2009-6-24/ 2009-06-28

Viñuela, A.; Snoek, L.B.; Riksen, J.A.G.; Kammenga, J.E. (2009) Natural variation of the dynamics of global gene transcription: regulation in aging worms. Poster presentation in*: 17th International C. elegans meeting*, University of California, Los Angeles, June 24-28, 2009. - Los Angeles: 17th International C. Elegans Meeting, 2009-06-24/ 2009-06-28

Viñuela, A.; Snoek, L.B.; Riksen, J.A.G.; Kammenga, J.E. (2009) The transcriptional mechanisms of organophosphorus pesticide mixtures in *C. elegans.* Poster presentation in*: 17th International C. elegans Meeting*, Los Angeles, California, 24-28 June 2009. - Los Angeles: 17th International *C. elegans* Meeting, 2009-06-24/ 2009-06-28

Viñuela, A.; Snoek, L.B.; Riksen, J.A.G.; Kammenga, J.E. (2009) Natural variation of the dynamics of global gene transcription: regulation in aging worms. Poster presentation in*: GENSYS satellite meeting. European Conference on Complex Systems' 09*, Warwick, UK, 21-25 September 2009. - Warwick: European Conference on Complex Systems' 09, Warwick, UK, 21-25 September 2009, 2009-09-21/ 2009-09-25

Viñuela, A.; Snoek, L.B.; Riksen, J.A.G.; Kammenga, J.E. (2009) Natural variation of the dynamics of global gene transcription regulation in aging worms. Poster presentation in*: 1st symposium on Systems Genetics: from man to microbe, from genotype to phenotype*, Groningen, The Netherlands, 1-2 October 2009.

Viñuela, A.; Snoek, L.B.; Riksen, J.A.G.; Kammenga, J.E.. Towards global gene transcription profiling in natural populations of *C. elegans.* Progress presentation in: *4th Ecogenomics annual meeting*, 2008-04-17. Amsterdam, the Netherlands.

Viñuela, A.; Riksen, J.A.G.; Kammenga, J.E.. Molecular mechanisms of toxicity at the gene transcription level. Contributed talk in: *Workshop of Genetics of Thermal adaptation in ectotherms*, 2008-10-05. Wageningen, the Netherlands.

Viñuela, A.; Riksen, J.A.G.; Kammenga, J.E.. Mapping determinants of gene expression plasticity to toxicants in wild isolates of *C.elegans*. Progress presentation in: *3rd Annual Meeting of the Ecogenomics*, 2007-04-19. Amsterdam, the Netherlands.

Viñuela, A.; Riksen, J.A.G.; Kammenga, J.E.. Microarray analysis of Nickel and Chlorpyrifos exposed nematodes. Progress presentation in: *2nd NoMiracle Progress Workshop*, 2007-09-06. Salzburg, Austria.

Grants and awards:

GENESYS travel award and best PhD student's presentation. *GENESYS workshop at the 11th International Conference on Systems Biology*. 10-16 October 2010, Edinburgh, UK.

Keystone Symposia travel scholarship: *Keystone meeting: New Insights into Heath-span and Diseases of Aging: From Molecular to Functional Senescence*, Tahoe City, California, Jan 31-Feb 5, 2010. National Institutes of Health, National Institute of Aging Grant.

Curriculum vitae

Ana Viñuela Rodríguez was born on the 25th of February 1979 in Madrid, Spain, and shortly after she moved to Córdoba. In 1997 she started her university studies in Biology in the Faculty of Sciences. After finishing her Diplomatura (BSc) in 2001, she moved to Madrid to specialize in Genetics at the Universidad Complutense de Madrid. As part of her Licenciatura (MSc) in Biology, she spent an academic year as Erasmus student at the Friedrich-Schiller-Universität of Jena Germany. During that period, she worked as student assistant in a protein-protein interaction study of a porcine teschovirus (*PTV-1*) at the Institut für Virologie und Antivirale Therapie. She returned to Madrid in 2004 to finish her studies in genetics and to follow a Master in Bioinformatics and Computational Biology. In September 2005 and as part of her studies in bioinformatics, she moved to Denmark for four months for an internship in comparative analysis of bacterial genomes at the Center for Biological Sequence Analysis in the Denmark Technical University. She returned to Spain to graduate in 2006. In February 2007 she was appointed as PhD student (AIO) in the Laboratory of Nematology where she studied gene expression in the nematode *Caenorhabditis elegans*.

PE&RC PhD Education Certificate

With the educational activities listed below the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Review of literature (5.6 ECTS)

- Mapping gene expression plasticity in *C. elegans* (2007)

Writing of project proposal (7 ECTS)

- Gene-environment interactions in *Caenorhabditis elegans* (2009)

Post-graduate courses (4.9 ECTS)

- Statistical analysis of microarray data with R; statistics.com (2007)
- Mixture toxicity within a DEB context, experimental design and data analysis; NoMiracle project (2007)
- Quantitative genetics; Summer Institute in Statistical Genetics (2009)
- QTL Mapping; Summer Institute in Statistical Genetics (2009)

Deficiency, refresh, brush-up courses (6 ECTS)

- Modern statistics for the life science (2008)

Competence strengthening / skills courses (4.5 ECTS)

- PhD Competence assessment (2007)
- Information literacy for PhD + EndNote introduction (2007)
- Academic writing II (2009)
- Scientific writing (2009)

PE&RC Annual meetings, seminars and the PE&RC weekend (1.8 ECTS)

- PE&RC Weekend (2007)
- PE&RC Days (2007, 2009 and 2010)

Discussion groups / local seminars / other scientific meetings (7.1 ECTS)

- Experimental evolutionary discussion group (2007/2010)
- Ecogenomics (2008/2009)

International symposia, workshops and conferences (18.4 ECTS)

- 2nd NoMiracle Progress Workshop; Salzburg, Austria (2007)
- Ecogenomic consortium project progress meeting (2007, 2008 and 2009)
- 17th International *C. elegans* Meeting; Los Angeles, California, USA (2009)
- European Conference on Complex Systems' 09; GENSYS satellite meeting, Warwick, UK (2009)
- Keystone Symposia, New Insights into Healthspan and Diseases of Aging: From Molecular to Functional Senescence; Tahoe Lake, Reno, USA (2010)
- Evolutionary Biology of Caenorhabditis and Other Nematodes; Cambridge, UK (2010)
- GENESYS Workshop at the International Congress of System Biology; Edinburgh, UK (2010)

The research described in this thesis was financially supported by the Ecogenomics Consortium of the Netherlands Genomics Initiative (NGI) and the EU FP6 framework programme NoMiracle.

Cover: Ana Viñuela Rodríguez

Cover pictures, from the top to the bottom: (red) hermaphrodite *C. elegans* nematode; (white) hermaphrodite and male *C. elegans* and eggs; (blue) male *C. elegans*; picture of a two color microarray from hermaphrodites *C. elegans* exposed to pesticides; expression profiles and eQTLs profiles of *pgp-6* (left) and *daf-16* (right) and heatmap of Gene Ontology significant terms from genes significantly affected by pesticides on the background. Back cover: prompt of the command window in R.