Mechanisms of buffering of phenotypic variation in *A. thaliana*: Features of HSP90 and characterization of AGO1 as a new regulator of robustness

Tzitziki Janik Lemus Vergara

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington
2015

Reading Committee:

Christine Queitsch, Chair

Philip Green

Takato Imaizumi

Program Authorized to Offer Degree: Department of Genome Sciences

University of Washington

Abstract

Mechanisms of buffering of phenotypic variation in *A. thaliana*: Features of HSP90 and characterization of AGO1 as a new regulator of robustness

Tzitziki Janik Lemus Vergara

Chair of the Supervisory Committee:
Associate Professor, Christine Queitsch
Department of Genome Sciences

Biological systems have molecular mechanisms that allow them to maintain a constant phenotype despite environmental and genetic perturbations. These mechanisms can take the form of master regulator of robustness, gene products that are epistatic to environmental and genetic perturbations. When these regulators are perturbed, the phenotypic variation of the biological systems increases. The best example of a master regulator of robustness is HSP90. HSP90 interacts with proteins involved in different biological pathways, and thus is a central node of many genetic networks. HSP90 provides robustness to environmental and genetic perturbations, and it has been hypothesized that influences the evolution of its clients. Here, I provide evidence that HSP90 impacts the evolutionary trajectories of clients belonging to pairs of duplicate genes or gene families. MicroRNAs (miRNAs) have been shown to provide robustness to environmental and genetic perturbations of specific phenotypic traits. The central node of the miRNA pathway in *A. thaliana* is AGO1. In my thesis, I show that AGO1 is another master regulator of robustness, since AGO1 can buffer both environmental and genetic perturbations for a variety of traits. I also show that AGO1-dependent variation is background specific. Two of the AGO1-revealed phenotypes, lesions in cotyledons and lack of correlation between days to flowering and leaf number were analyzed to further detail.

Table of Contents

Chapter 1 7

Introduction 7

1.1 Masters regulators of robustness 8
1.2 Network architecture of robustness 10
1.3 The role of small RNAs in robustness 11
1.4 miRNAs in plants: functions and biogenesis 13
1.5 AGO1 and other *Arabidopsis* Argonaute proteins 15
1.6 miRNAs and stress response in plants 16
1.7 Flowering time regulation in *A. thaliana* 18
1.8 Research goals 20
1.9 Organization of the chapters and name conventions 20

Chapter 2 21

The protein Chaperone HSP90 can facilitate the divergence of Gene duplicates . 21

2.1 Introduction 23
2.2 Materials and Methods 24
2.3 Results 26
2.3.1 BES1 is an HSP90 client 26
2.3.2 HSP90 facilitates the divergence of gene duplicates 29
2.3.3 HSP90 client BES1 shows hallmarks of sub-and neofunctionalization 32
2.4 Discussion 37
2.5 Supplementary Materials and Methods 40

Chapter 3 52

3.1 Introduction 53
3.2 Methods 55
3.3 Results 58
3.3.1 Hsp90 client status contributes to dN/dS 58
3.3.2 Hsp90-associated effects on kinase evolutionary rate increase with increasing divergence 63
3.3.3 Hsp90 client status tends to be gained and maintained 65
3.4 Discussion 68
3.5 Supplemental information 73

Chapter 4 82

The mechanistic underpinnings of an *ago1*-mediated, environmentally-dependent, and stochastic phenotype 82

4.1 Introduction 83
4.2 Materials and Methods 86
4.1 Results 90
4.1.1 Argonaute 1 mutant seedlings develop lesions in full spectrum light conditions 90
4.1.2 ago1-27 cotyledons overexpress canonical SA, JA, and JA/ET signaling markers in full spectrum light 95
4.1.3 Exogenous JA and ET treatments, but not SA treatment, suffice to induce lesions in wild-type 99
4.1.4 Disruption of JA perception dramatically reduces lesions in ago1-27 mutant seedlings 101

4.1.5 ARGONAUTE1 and HSP90 interact genetically in producing the lesion phenotype 103
4.2 DISCUSSION **105**
4.3 Supplementary Information **108**

Chapter 5 ..... **115**

AGO1 Buffers genetic variation in the plant *Arabidopsis thaliana* **115**
5.1 Introduction **115**
5.2 Materials and methods **118**
5.3 Results **120**
5.3.1 ago1 mutant plants show increased phenotypic variation 120
5.3.2 AGO1 perturbation can reveal and conceal genetic variation 122
5.3.3 ago1 perturbation uncouples two closely related flowering time traits in a background-specific manner 126
5.3.4 Identifying the molecular mechanism(s) that could uncouple rosette leaf number and days to flowering. 127
5.3.5 Genotyping of the MIR156F locus. 130
5.3.6 MIR156F polymorphisms alter the predicted secondary structure of the miR156f precursor. 131
5.3.7 Genes in the ageing pathway have increased mRNA levels in Ler ago1-27 relative to Col-0 ago1-27. 131
5.3.8 FLC as a third epistatic factor contributing to rosette leaf number. 132
5.4 Discussion **134**
5.4.1 AGO1 buffers developmental noise and genetic variation 134
5.4.2 ago1-27 in Ler background uncouples days to flowering from rosette leaf number. 136
5.4.3 The involvement of ageing pathways in decreasing leaf number 136
5.4.4 Polymorphisms in MIR156F may disrupt MIR156F processing, decreasing the miR156 pool. 137
5.4.5 FLC and MIR156F contribute to rosette leaf number in a background-specific manner. 138
5.5 Supplementary information **139**

Chapter 6 ..... **152**

Discussion and future studies **152**
6.1 Implications of AGO1 buffering of environmental perturbations **152**
6.2 Implications of AGO1 buffering of genetic perturbations **153**
6.3 Extreme example of AGO1 revealed perturbations **154**
6.4 Implications of AGO1 in the evolution of genomes: a hypothesis **155**
6.5 AGO1 and HSP90 act in different pathways to confer robustness **156**

References **158**

VITA **177**

ACKNOWLEDGEMENTS

This thesis is the product of not only my efforts, but also of several people that help me along this exciting, professional and almost spiritual journey. I want to thank:

My advisor, Christine Queitsch, for all her support, empathy and encouragement during these years.

Members of the Queitsch lab, past and present, for helping me, question me, give me insight, and making science even more fun. Thanks to Sanna (A.M.O.) Sullivan, Jenn Lachowiec, Keisha Carlson, Alex Mason, Max Press, Mike Dorrity, Kerry Bub, Soledad Undurraga, Janne Lempe, Cris Alexandre, Pauline Rival, Elizabeth Morton, Karla Shultz, and James Urton. Thanks to our honorary member Josh Cuperus, for guiding me through the world of miRNA biogenesis.

Members of my committee, Robert Waterson, Celeste Berg, Phil Green, and Takato Imaizumi.

Genome Sciences graduate students, particularly to: Kat Claw, for welcomed me, invited me to socialize, and introduced me to SACNAS; Sharon Greenblum, for her friendship and support at the school, life, and soccer field; Aaron Miller, and Sam Lancaster, for making me laugh, and gave me life advices.

To my soccer team, for welcome me to play the beautiful game. To my Seattle friends, for their friendship, encouragement, and making sure I got my dose of fun.

To my family, and especially my mom for all her support, advice, encouragement, and unconditional love. To Daniel, for his love, support, and making me laugh every day.

Chapter 1

INTRODUCTION

The ability of biological systems to maintain a phenotype by concealing environmental and genetic variations is called phenotypic robustness, and the phenotype that is kept constant is deemed robust (Whitacre 2012; Queitsch et al. 2012; Lempe et al. 2013). Phenotypic robustness arises from a combination of features of the underlying biological networks, such as connectivity, feedback loops, redundancy, and from other non-genic mechanisms, such as epigenetic modifications (Waddington 1953; Lempe et al. 2013). When the mechanisms that confer robustness are challenged by environmental or genetic perturbations, the overall robustness of the system decreases, and the organisms display larger phenotypic variation than in control conditions (Queitsch et al. 2002; Lempe et al. 2013).

Robustness, as many other traits, can be measured quantitatively, although it is not trivial. Traditional measurements of robustness consisted in assessing the symmetry of organisms, with the most symmetrical individuals being the most robust (Clarke 1998). Another measurement system involves testing the accuracy by which a phenotype is formed among isogenic individuals. Examples of measurement systems for phenotypic robustness include number of bristles (Posadas and Carthew 2014) and wing shape in fruit flies (Waddington 1953), hypocotyl length in plants (Queitsch et al. 2002), and metabolite production in yeast (Breunig et al. 2014). One observation from these studies is that the robustness of one phenotype does not predict the robustness of another (Clarke 1998; Masel and Siegal 2009); which gives support to the idea that robustness tends to be modular (Lempe et al. 2013).

Loci that affect robustness can be mapped after perturbing the biological system (Jarosz and Lindquist 2010; Takahashi 2013). Likewise, it is possible to identify the underlying cryptic genetic variation that affects phenotypes when robustness is lost (Sangster et al. 2008a). Plants are an excellent system to study phenotypic robustness given their

continuous development, and their sessile life-style that forces them to adapt to environmental changes. Besides specific loci that can influence robustness, there are other robustness mechanisms including the fine-tuning of expression during development by regulatory RNA, features of genetic networks architecture, and robustness by duplicated genes.

1.1 MASTERS REGULATORS OF ROBUSTNESS

Gene products that affect multiple traits when perturbed are called 'masters regulators of robustness' (Rutherford and Lindquist 1998; Lempe et al. 2013). They are also referred to in the literature as capacitors, and network hubs (Rutherford and Lindquist 1998; Queitsch et al. 2002; Levy and Siegal 2008). One of the best-characterized master regulators of phenotypic robustness is the protein chaperone Heat Shock Protein 90 (Hsp90), which buffers phenotype against both environmental and genetic perturbations (Rutherford and Lindquist 1998; Queitsch et al. 2002; Yeyati et al. 2007; Sangster et al. 2008a; Specchia et al. 2010; Jarosz and Lindquist 2010; Casanueva et al. 2012).

Hsp90 is a highly conserved protein chaperone that helps its interacting proteins (clients) acquire an active conformation (Taipale et al. 2010). Hsp90 clients have a wide range of biological functions, including kinases, hormone receptors, transcription factors, ligases, mitochondrial proteins, and resistance proteins, among others (Taipale et al. 2010; Hartson and Matts 2012). A common feature of Hsp90 clients is their tendency to be long and disordered (Taipale et al. 2012; Karagoz and Rudiger 2015). Hsp90 also interacts with another set of proteins called HSP90 co-chaperones via specific protein domains. Hsp90 co-chaperones affect Hsp90 activity and client specificity (Mayer 2010). The numerous interactions of Hsp90 make it a 'hub' for many genetic networks, , which supports the observation that Hsp90 buffers many developmental phenotypes (Lempe et al. 2013).

Inhibition of Hsp90 reveals phenotypic variation in flies, plants, fish, worms and yeast (Rutherford and Lindquist 1998; Queitsch et al. 2002; Yeyati et al. 2007; Sangster et al.

2008a; Specchia et al. 2010; Jarosz and Lindquist 2010; Casanueva et al. 2012). This phenotypic variation can consist of new phenotypes, increased frequencies of spontaneous developmental abnormalities, and shifts in mean and variance of a particular quantitative trait. Within genetically divergent individuals of the same species, Hsp90-revealed phenotypes depend on the genetic background of the organism (Queitsch et al. 2002; Sangster et al. 2008b; Sangster et al. 2008a). In isogenic siblings, inhibition of Hsp90 increases developmental noise, and can modify the penetrance of partial loss of function mutations (Queitsch et al. 2002; Sollars et al. 2003; Casanueva et al. 2012).

The ability of Hsp90 to buffer genetic variation has led to the *capacitor hypothesis*, which is summarized below. The hypothesis states that interactions with the chaperone Hsp90 confer robustness to mutations, allowing clients to explore a greater sequence space without compromise their functions (Rutherford and Lindquist 1998). Then, upon Hsp90 reduction, cryptic genetic variations would be revealed and subject to selection. Then, favorable Hsp90 client changes can then be selected for. In this sense, Hsp90 could influence protein evolutionary processes.

Under the capacitor hypothesis, other proteins can become Hsp90 clients and then, after gaining new mutations, can become stable and loose Hsp90 client status. In this scenario, new Hsp90 clients would be highly dependent on Hsp90 function and availability. In support of the capacitor hypothesis, Citri et al. (Citri et al. 2006) showed that Hsp90 client status is dynamic, and that a protein can become an Hsp90 client by gaining one mutation. In addition, Taipale and collaborators (Taipale et al. 2012) showed that in the mammalian kinases superfamily, strong Hsp90 clients have more nonsynonymous mutations than non-clients. In prokaryotes, clients of GroEL/ES, the Hsp90 prokaryote counterpart, show different degrees of dependency on the chaperone for stability and folding (Tokuriki and Tawfik 2009). In this work I present further evidence in support of the capacitor hypothesis and the role of Hsp90 in protein evolution.

Besides Hsp90, other master regulators of robustness include ELF4, a plant circadian regulator that can also buffer phenotypic variation (Doyle et al. 2002). In yeast, more than

300 genes, when deleted, increase morphological variation (Levy and Siegal 2008). Despite advances in the discovery of phenotypic capacitors in single-celled organisms, there are still very few known in multicellular organisms. In addition, it is uncertain if these mechanisms share molecular features, and how their perturbation would impact phenotypic traits and disease susceptibility. Broadening the knowledge on master regulators, would allow addressing among other questions, whether phenotypic capacitors share molecular features, and to establish a set of molecular markers of robustness, that can be used in the analysis of complex traits and diseases.

1.2 NETWORK ARCHITECTURE OF ROBUSTNESS

In the quest to find more loci that affected robustness of traits in plants, research groups have use QTL mapping. Hall et al. used two population of *A. thaliana* Recombinant Inbred Lines (RILs) and measured five traits. They found 22 QTLs; most of the QTLs that affected the variance of a particular trait (measured as the average of each individuals Levene's statistic) also affected the mean of that trait. In 2008, Sangster and colleagues used three RILs population to find QTLs that affected hypocotyl (embryonic stem) and root length, in control HSP90-reduced conditions. In control conditions, several QTLs for variance did not coincide with mean QTLs. As mentioned earlier, hypocotyl length is highly susceptible to microenvironmental changes, which includes developmental noise in the genetic network during development. Among recombinant inbred lines, network interactions that were present in the parental lines might be lost, therefore accounting for a decrease in the correlation of variance and mean QTL. However, in HSP90 reduced conditions, mean and variance QTL co-occur most of the time. A similar phenomenon has been observed in yeast, where temperature stress and Hsp90 reduction increased the relationship between genotype and phenotype (Jarosz and Lindquist 2010). The results from these studies provide evidence that (i) newly released variants are poorly buffered, and (ii) that perturbation of trait robustness usually is coupled with a decrease in function. In addition, these studies suggest that most *A. thaliana* QTLs discovered were trait-specific, and that the gene network underlying phenotypic traits tends to be modular.

A study by Fu et al. (Fu et al. 2009) took another approach and mapped loci that affected many traits. They used 162 *A. thaliana* RILs derived from very divergent parents (strains Cvi and L*er*), and measured 139 phenotypic and 40, 580 molecular traits. They assayed approximately 50,000 SNPs, and found six 'QTLs hotspots', loci located in 5cM windows, that could influence up to 7,158 metabolite traits. These QTLs hotspots are located on chromosomes 1, 2, 3, and 5, and none of them map to *HSP90*. The fact that one of the six QTL hotspots influence 77% of 116 phenotypic traits with QTLs, suggests genetic networks are robust to a large number of small perturbations, unless the mutations are localized in one of the network hubs.

Gene duplicates also influence the robustness of genetic networks (Whitacre 2012); this is particularly true in plants, where most of the genes involved in certain biological pathways such as brassinosteroids and microRNAs, are all in gene families (Kim and Wang 2010; Bologna and Voinnet 2014). Individual gene mutants in these pathways, such as *beh2* and certain *dcl* mutant genes do not appear to have a phenotype, but a phenotype can be observed when mutant alleles are combined (Bologna and Voinnet 2014).

1.3 THE ROLE OF SMALL RNAS IN ROBUSTNESS

Phenotypic variation among isogenic siblings can be due to changes in their microenvironment, stochastic fluctuations and gene expression noise during development. One of the mechanisms that biological systems use to buffer gene expression noise are microRNAs (miRNAs) (Rubio-Somoza and Weigel 2012; Posadas and Carthew 2014). MiRNAs regulate the gene expression of their targets by adjusting the mean expression and reducing the variance of their target mRNA levels (Hornstein and Shomron 2006; Wu et al. 2009a; Ebert and Sharp 2012). MiRNAs were hypothesized to play a role in robustness by existing in network motifs that facilitate the production of a constant output despite variation in their input levels (Hornstein and Shomron 2006). One such motif is the incoherent feed-forward loop, in which an initial transcription factor promotes the expression of both a gene product and a miRNA that targets the gene product (Ebert and Sharp 2012; Posadas and Carthew 2014).

Several studies have showed that miRNAs can buffer phenotypic variation in plants and flies (Sieber et al. 2007; Li et al. 2009b; Hilgers et al. 2010; Cassidy et al. 2013). In plants, Sieber and collaborators showed that miR164 buffers phenotypic variation by decreasing and restricting the gene expression of its target genes CUP-SHAPED COTYLEON1 (CUC1), and CUC2 (Sieber et al. 2007). In the absence of miR164, CUC expression randomly expands from its normal tissues to inflorescence meristems and flower primordia, tissues where miR164 is normally expressed. Thus miR164 helps maintain organ boundaries. In *Drosophila*, *miR-263a/b* protects sense organ cells from apoptosis, and absence of *miR-263a/b* causes loss of sensory bristles at random (Hilgers et al. 2010). Recently, Schmiedel et al. (Schmiedel et al. 2015) showed that miRNAs reduce protein expression noise for lowly expressed genes, and that genes targeted by multiple miRNAs have the least protein noise.

Other plant-specific small RNAs buffer phenotypic variation, such as trans-acting siRNAs (tasiRNAs). TasiRNAs are involved in leaf morphogenesis, and in the juvenile-to-adult developmental transition (Rubio-Somoza and Weigel 2012) TasiRNA biogenesis involves the production of transcripts from the TAS3 locus, that are then targeted by miR390 and ARGONAUTE7 (AGO7). This triggers the production of tasiRNAs that then repress the expression of AUXIN RESPONSE FACTOR 3 (ARF3) and ARF4 (Pulido and Laufs 2010). TAS3 transcripts originate in the lower side of the leaf and travel to the upper (adaxial) part of the leaf, establishing a gradient of TAS3 transcripts. Since *AGO7* is only expressed in the adaxial part of the leaf, tasiRNAs restrict *ARF3* and *ARF4* expression only on this side of the leaf, establishing tissue boundaries (Chitwood et al. 2009). Mutations in the genes involved in the biogenesis of tasiRNAs lead to defects in leaf polarity (Pulido and Laufs 2010), and an earlier transition to the adult phase.

1.4 miRNAs IN PLANTS: FUNCTIONS AND BIOGENESIS

Plant miRNAs participate in a wide variety of biological pathways, they regulate developmental transitions, leaf morphology, floral development, root and meristem architecture, and participate in stress and hormone responses (Poethig 2013; Leung and Sharp 2010; Sunkar et al. 2012; Spanudakis and Jackson 2014; Peláez and Sanchez 2013). To add to the complexity, miRNA range in size from 21-24 nucleotides (nt), and their biogenesis and silencing machinery involves the function of a multitude of paralogous genes (Chen 2009; Bologna and Voinnet 2014).

In *A. thaliana*, the biogenesis of 21-nt miRNAs starts with the transcription of MIRNA genes by RNA Polymerase II (Pol II). Then, the transcripts are capped and polyadenylated, generating pri-miRNAs with self-complementarity, resulting in imperfect hairpin structures commonly composed of an upper terminal loop, an upper stem, the miRNA miRNA/miRNA* duplex, a lower stem, and flanking single-stranded basal segments known as a hairpin precursor or pre-miRNA (**Figure 1.1**) (Zhu et al. 2013). This structure is recognized by DICER LIKE1 (DCL1) protein, which process the pre-miRNA and generates a mature miRNA (Bologna et al. 2013). In plants, MIRNA transcripts range in size from 49 to 900 nt, and are processed by two major mechanisms: base-to-loop, and loop-to-base (Bologna et al. 2013; Bologna and Voinnet 2014).

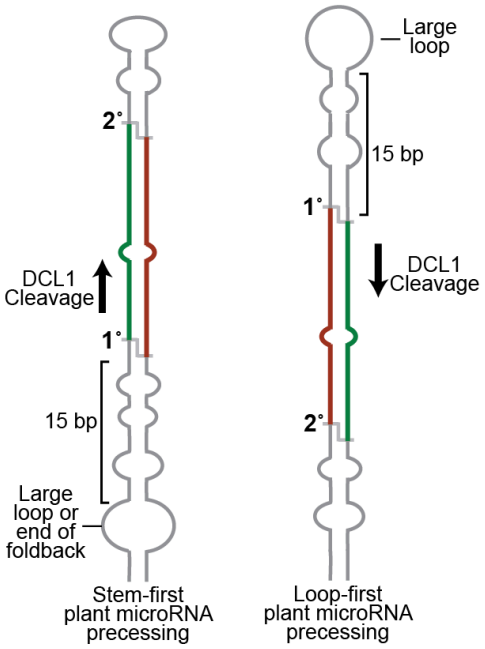


Figure 1.1 Representative structure of a pri-miRNA. Red marks the position of the miRNA* sequence, green marks the position of the miRNA.

In the base-to-loop mechanisms, pre-miRNA processing starts from the lower stem and proceeds to the upper terminal loop. DCL1 and auxiliary proteins recognize an internal loop, followed by a stem of approximately 15-nt. Then, DCL1 makes the first cleavage at the end of the 15-nt stem, and a second cleavage at a distance of approximately 21-nt (Cuperus et al. 2010; Todesco et al. 2012; Cuperus et al. 2011). In the loop-to-base mechanism, an upper stem segment guides the processing of pre-MIRNAs, and two consecutive cuts release the miRNA/miRNA*. *MIR156* and *MIR160* miRNA families are processed by this mechanism (Bologna et al. 2013). In chapter five I discuss the implications of polymorphisms in the processing of miR156f.

Methylated miRNA/miRNA* duplexes are then transported to the cytoplasm and commonly loaded into ARGONAUTE1 (AGO1), one of 10 AGO proteins in *A. thaliana* to form the RNA induced Silencing complex (RISC) (Axtell 2013). Most AGO1-associated

miRNAs have a uracil at their 5'end (Mi et al. 2008). The RISC complex then represses the expression mRNAs in a sequence specific manner (Axtell 2013), primarily by cleaving the target mRNA or inhibiting its translation (Brodersen et al. 2008). The mechanism that determines which mode of repression the RISC complex used was unknown, until recently. Reis and collaborators showed that the interaction of DCL1 with the DOBLE-STRANDED RNA BINDING1 (DRB1) is necessary for the regulation by slicing, whereas DCL1 interaction with DRB2 is required for regulation by translational inhibition (Reis et al. 2015).

1.5 AGO1 AND OTHER *ARABIDOPSIS* ARGONAUTE PROTEINS

Eukaryotic Argonaute proteins have four domains: a variable N-terminal domain, and highly conserved PAZ, MID, and PIWI domains. The MID domain recognizes the small RNAs 5' end, while the PAZ domain binds to the 3' end. The PIWI domain exhibits an endonuclease activity (Hock and Meister 2008; Czech and Hannon 2011) (**Figure 1.2**).

Figure 1.2 Representative structure of *A. thaliana* AGO1. The different protein domains are represented in colored boxes.

There are 10 *AGO* genes in *Arabidopsis*, and they can be divided in three different clades (Vaucheret 2008; Hock and Meister 2008). Argonautes in the *AGO1*, *AGO10* and *AGO5* clade associate with 21-nt miRNAs. AGO1 is the major protein in the miRNA pathway, and associates with 80% of the 21-nt miRNAs (Mi et al. 2008; Czech and Hannon 2011; Wang et al. 2011). Besides miRNAs, AGO1 participates in the silencing of viral genes by loading 21 to 22 nt viral RNA (viRNA) (Bologna and Voinnet 2014). A second clade includes *AGO2*, *AGO3* and *AGO7*. AGO2 also binds to 21-nt miRNA, but is mostly associated with siRNAs. AGO2 function is also needed for the resistance to several viruses. AGO7 is involved in the tasiRNA pathway and interacts almost exclusively with miR390.

The function of AGO3 is currently unknown. *AGO4*, *AGO6*, *AGO8* and *AGO9* comprise another clade. *AGO4*, *-6*, and *-9* interact with 24-nt siRNAs that have an adenosine in their 5' end (Mi et al. 2008; Wang et al. 2011; Zheng et al. 2007). *AGO4* is required for the RNA-dependent DNA methylation (Wu et al. 2010).

In yeast, flies, humans, and Tetrahymena, Argonaute function requires the activity of HSP90 (Johnston et al. 2010; Iwasaki et al. 2010; Gangaraju et al. 2011; Miyoshi et al. 2010; Wang et al. 2013; Woehrer et al. 2015). In plants, Iki et al. used tobacco cell extracts to show that HSP90 interacted physically with AGO1, and that this interaction was required for the loading of both siRNA and miRNA (Iki et al. 2010). Further studies are needed in order to test the requirements of HSP90 for the activity of other plant Argonaute proteins.

Since AGO1 is an essential protein, the study of *AGO1* genetic interactions usually involves the usage of partial mutant *ago1* alleles with various degrees of phenotype severity. Common hypomorphic *ago1* alleles ordered by the strength of their phenotype are: *ago1-46*, *ago1-27*, and *ago1-25*. The three mutants have a point mutation in AGO1 PIWI domain. Mutants *ago1-46* and *ago1-27* have a mutation that affect the same 993 residue, but while both *ago1-46* and *ago1-27* maintain a portion of their slicing activity, *ago1-27* allele is impaired for siRNA regulation (Morel et al. 2002). The three mutants are smaller, have greater serration on their leaves, and flower later than wild type (Morel et al. 2002; Smith et al. 2009).

1.6 miRNAs AND STRESS RESPONSE IN PLANTS

In chapter four I describe the effects of AGO1 perturbation in response to abiotic stresses such as high light and hormone treatment. Thus, I provide a brief background of Arabidopsis stress responses.

Given their sessile life style, plants have to contend with an ever-changing environment. Several studies have reported changes in miRNA levels in heat, cold, drought, osmotic, high salt, hypoxia and oxidative stresses, abscisic acid treatment, and ultra-violet

light exposure (Sunkar 2010). Some miRNAs link the stress response to changes in development, for example, under cold stress miR156 levels increase, and this has been associated with a delay in flowering time, and an increase in rosette leaf number (Kim et al. 2012). Plant miRNA are also involved in the response to biotic stresses, such as pathogens and viruses. Most notably miR393, which is upregulated in response to bacterial flg22 (Zhang et al. 2011). Other miRNAs associated with pathogen response include miR160 and miR167 (Peláez and Sanchez 2013).

Upon pathogen contact, plants usually elicit a hypersensitive response, which consists of widespread cell death beginning at the point of infection, which can be visualized as necrotic spots in the leaf (Heath 2000). The hypersensitive response is an attempt to restrict the growth of the pathogen, and is genetically controlled by the resistance genes (R genes). When *Arabidopsis* senses pathogens, it triggers the expression of some R genes, which lead to an increase of reactive oxygen species (ROS) and eventually to cell death (Heath 2000). The plant hormones Jasmonates, ethylene, and salicylic acid are important for the accuracy of the hypersensitive response (Kazan 2015).

Jasmonates (JA) are small molecules involved in a variety of abiotic and biotic stress responses such as drought, salt, cold, and pathogens (Kazan 2015; Sunkar et al. 2012). JA modulate ROS production, and interact positively with ethylene to mediate the response to fungal pathogens, but JA and ethylene also interact negatively to mediate the response to numerous abiotic stresses (reviewed in Kazan, 2015). Ethylene (ET) can act positively and negatively in response to pathogens, and can globally repress salicylic acid-responsive genes. Salicylic acid (SA) biogenesis is a precursor of ROS, and SA detection can trigger hypersensitive response. Although JA, ET and SA have been traditionally linked to plant defense, recent studies have also shown their role in plant development (Kazan 2015).

1.7 FLOWERING TIME REGULATION IN *A. THALIANA*

In chapter five I describe the uncoupling of two flowering time traits, and test which of several pathways influencing flowering time are involved in the observed phenotype. I provide a brief description of how flowering time is regulated in *Arabidopsis*.

Reproduction is one of the main objectives of biological systems, and in the case of the plant *Arabidopsis thaliana*, reproduction can be achieved only once in its lifetime. Therefore, the transition from vegetative to reproductive development is highly regulated by several pathways that sense when internal and environmental conditions are optimal. These pathways converge in floral integrator genes such as *SUPPRESOR OF OVEREXPRESSION OF CO 1* (*SOC1*), and *FLOWERING LOCUS T* (*FT*) (Posé et al. 2012; Pajoro et al. 2014), activating the meristem identity genes *APETALA 1* (*AP1*), *FRUITFUL* (*FUL*), and *LEAFY* (*LFY*). The meristem identity genes transform the shoot apical meristem, from producing leaves, to an inflorescence meristem, which develops into reproductive organs (Andres and Coupland 2012).

In *Arabidopsis*, flowering is promoted by the exposure to light during long days. This exposure translates, by a series of exquisitely regulated pathways, into the activation of *FT* (Song et al. 2013; Song et al. 2014). In the photoperiod pathway, CONSTANS (CO) promotes the transcription of *FT. CO* transcription and protein levels are also tightly regulated to assure that CO can only activate *FT* under long days (Song et al. 2013; Song et al. 2014). Once FT is produced in the leaves, it travels through the phloem to the shoot apical meristem. In the meristem, FT activates the expression of SOC1, AP1, and *SQUAMOSA PROMOTER BINDING LIKE 3* (*SPL3*) (Andres and Coupland 2012; Pajoro et al. 2014). SPL3 then activates the expression of the meristem identity genes (Yamaguchi et al. 2009) (**Figure 1.3**).

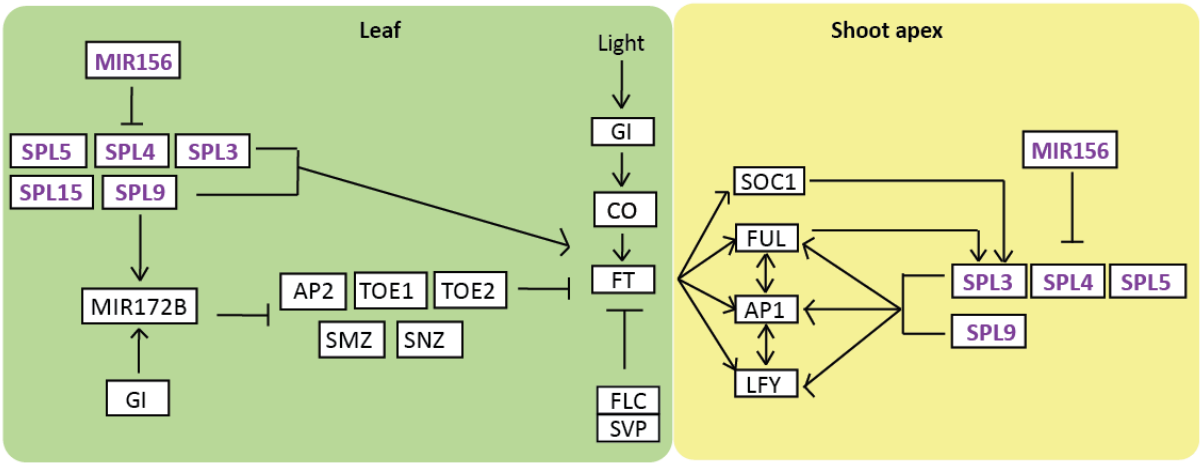


Figure 1.3 Simplified version of flowering time regulation of *A. thaliana*. Genes in purple indicate genes involved in the ageing pathway.

FLOWERING LOCUS C (*FLC*) also regulates the levels of *FT* mRNA. *FLC* is a major repressor of flowering, and is responsible from many of the difference in flowering time among *A. thaliana* strain (Michaels et al. 2003; Werner et al. 2005; Gan et al. 2011). When plants are exposed to long periods of cold temperature (vernalization) *FLC* expression is epigenetically repressed, and *FT* levels increase (Michaels et al. 2003; Helliwell et al. 2015). The ageing flowering pathway also alters the levels of *FT*, albeit indirectly. The ageing pathway involves the action of conserved miR156 and miR172. In young seedlings, levels of miR156 are high whereas levels of miR172 are low (Aukerman and Sakai 2003; Wu and Poethig 2006; Wu et al. 2009b). As the plant gets older, miR156 levels decrease and the expression of its target genes, including *SPL3*, and *SPL9*, increase. SPL9 activates the expression of *MIR172B* (Wu et al. 2009b). MiR172b then represses the expression of *APETALA2* (*AP2*)-like genes, which in turn repress the expression of *FT* (Aukerman and Sakai 2003; Chen 2004; Mathieu et al. 2009).

In addition to *SPL3* and *9*, miR156 also represses the expression of *SPL2*, *SPL4*, *SPL5*, *SPL6*, *SPL10*, *SPL11*, *SPL13*, and *SPL15* (Wu et al. 2009b; Poethig 2013). SPL genes involved in the ageing pathway are *SPL4*, *SPL3*, *SPL5*, *SPL9* and *SPL15* (Wang 2014; Spanudakis and Jackson 2014). In the leaves, SPL9 and SPL15 function is important to

confer the leaves adult characteristics (Wang et al. 2008; Spanudakis and Jackson 2014). In the meristem, SPL9 promotes the expression of *SOC1* and *MIR172b*, and also of *FUL*, *LFY* and *AP1*. *SPL3*, *SPL4*, and *SPL5* promote the expression of *FT*, *LFY* and *AP1* (Yamaguchi et al. 2009; Poethig 2013). In addition to participate in the ageing pathway, miR156 and miR172 have shown to be involved in the response to flowering time during stress temperature conditions (Lee et al. 2010; Stief et al. 2014).

1.8 RESEARCH GOALS

The present dissertation examines the extent of the buffering abilities of two essential, highly connected genes: the protein chaperone HSP90 and AGO1. In chapters two and three I aim to provide evidence that dispute or support the HSP90 capacitor hypothesis, which states that HSP90 can influence the evolution of its interacting proteins (clients). In chapters four and five I focus on AGO1, and test if AGO1, like HSP90, is a phenotypic capacitor by (i) testing if AGO1 buffers phenotypic variation in isogenic seedlings in *A. thaliana,* (ii) whether AGO1 can buffer genetic variation present in divergent *Arabidopsis* strains, and (iii) if AGO1-dependent loci overlap with HSP90-dependent loci. In the process of testing AGO1 buffering, I observed strain-specific phenotypic variation under AGO1 perturbation. I then examine molecular mechanisms responsible for these phenotypes, revealing several genetic components responsible for this phenotypic variation.

1.9 ORGANIZATION OF THE CHAPTERS AND NAME CONVENTIONS

In the present dissertation, all the chapters except chapter one and six have an introduction followed by a materials and methods section, results, discussion, and supplementary data. Figures and tables that are not listed as supplementary information (S) are embedded in the results.

Regarding gene nomenclature, in *Arabidopsis thaliana*, gene names are written in capital letters and italics (*i.e. AGO1*), whereas protein names are written in capital letters (*i.e.* AGO1). Mutant alleles, as in other organisms, are written in lower case and italics (Meinke and Koornneef 1997).

Chapter 2

THE PROTEIN CHAPERONE HSP90 CAN FACILITATE THE DIVERGENCE OF GENE DUPLICATES[1]

This chapter is included in the present thesis because I collaborated in the analysis that led to its publication. My contribution to this chapter was the acquisition and analysis of yeast HSP90 clients. The results of the analysis allowed us to generalize the statement that HSP90 can facilitate the divergence of gene duplicates in other organisms, and not just in *Arabidopsis thaliana*. I also contributed figures 2.2b, S2.4a and b, and data for Table S2.2. The first person "we", and the "our" pronoun indicates the work of J. Lachowiec. My work is stated with the "I" person, and the "my" pronoun.

Abstract

The heat shock protein 90 (HSP90) acts as a chaperone by ensuring proper maturation and folding of its client proteins. The HSP90 capacitor hypothesis holds that interactions with HSP90 allow proteins to accumulate mutations while maintaining function. Following this logic, HSP90 clients would be predicted to show relaxed selection compared with non-clients. In this study, we identify a new HSP90 client in the plant steroid hormone pathway: the transcription factor BES1. Its closest paralog, BZR1, is not an HSP90 client. This difference in HSP90 client status in two highly similar proteins enabled a direct test of the capacitor hypothesis. We find that *BES1* shows relaxed selection compared to *BZR1,* hallmarks of neo- and subfunctionalization, and dynamic HSP90 client status across independent evolutionary paths. These results suggested that HSP90's influence on gene evolution might be detectable if we compare gene duplicates, because duplicates share most

[1] This chapter was published in the journal *Genetics* as "The protein chaperone HSP90 can facilitate the divergence of gene duplicates", by J. Lachowiec, T Lemus, JH Tomas, PJM Murphy J Nemhauser and C Queitsch.

other properties influencing evolutionary rate that might otherwise conceal the chaperone's effect. I tested this hypothesis using systematically identified HSP90 clients in yeast, and observed a significant trend of HSP90 clients evolving faster than their non-client paralogs. This trend was not detected when yeast clients and non-clients were compared without considering paralog status. Our/my data provide evidence that HSP90 influences selection on genes encoding its clients and facilitates divergence between gene duplicates.

2.1 INTRODUCTION

The phenotypic capacitor HSP90 is thought to influence evolutionary processes through its ability to both conceal and release genetic variation (Rutherford and Lindquist 1998; Queitsch et al. 2002; Yeyati et al. 2007; Jarosz and Lindquist 2010). Perturbation of this conserved and essential chaperone reveals cryptic genetic and epigenetic variation in flies, plants, fish, and yeast (Rutherford and Lindquist 1998; Queitsch et al. 2002; Yeyati et al. 2007; Jarosz and Lindquist 2010; Sollars et al. 2003). In worms, HSP90 affects the penetrance of partial loss of function mutations (Burga et al. 2011). As expected under the capacitor hypothesis, worms with naturally lower HSP90 levels show significantly higher mutation penetrance (Casanueva et al. 2012). HSP90-dependent variation can be revealed by moderate environmental stress alone, providing a plausible release mechanism for this concealed variation in nature (Jarosz and Lindquist 2010; Rutherford and Lindquist 1998; Queitsch et al. 2002). Previous work from the lab showed that HSP90-dependent variation is common in natural plant populations, implicating the chaperone as an important player in shaping phenotype and evolutionary trajectories (Sangster et al. 2008a). Together, these findings have prompted a longstanding debate about the importance of HSP90 in evolutionary processes and the magnitude of its effect (Bergman and Siegal 2003; Meiklejohn and Hartl 2002; Rando and Verstrepen 2007).

It is well-established that HSP90 recognizes metastable proteins and facilitates their folding and stability (Taipale et al. 2010). Recent studies demonstrate that protein stability is a major constraint on protein evolution (Bloom et al. 2006; Pena et al. 2010). Stable proteins tend to evolve faster as they can explore greater sequence space without losing function (Bloom et al. 2006). In prokaryotes, overexpression of the chaperonin GroEL/ES allows the evolution of a far greater number of highly active enzyme variants by compensating for their reduced stability (Peña et al. 2010). We hypothesized that in eukaryotes HSP90 may facilitate gene divergence by similarly relaxing constraints on protein stability. If so, HSP90 clients should show greater tolerance to mutations and evolve faster than non-clients facing similar evolutionary pressures. To address HSP90's role in gene evolution in this context, we focused on recent gene duplicates, which initially face similar selection pressures and encode proteins of similar stability. We used the brassinosteroid (BR) pathway in *Arabidopsis*

thaliana as an experimental model because every step in BR signaling is encoded by gene families (Kim and Wang 2010) and previous studies suggested that the BR pathway may require HSP90, although no HSP90 client had been identified (Sangster and Queitsch 2005; Sangster et al. 2007b).

We demonstrate here that only one of two paralogous transcription factors in the BR pathway is an HSP90 client and that its encoding gene shows relaxed purifying selection compared to its non-client paralog. Gene duplicates diverge through sub- and neofunctionalization. Consistent with subfunctionalization, only the HSP90 client is temperature sensitive; consistent with neofunctionalization, the gene encoding the HSP90 client contains a novel exon and non-synonymous polymorphisms in divergent *A. thaliana* strains. HSP90-facilitated divergence of gene duplicates is widespread, because in the yeast *Saccharomyces cerevisiae*, genes encoding HSP90 clients tended to evolve faster than those encoding their non-client paralogs. Together, our/my data provide strong evidence for HSP90-faciliated evolution in extant genomes and hence strong support for the capacitor hypothesis.

2.2 MATERIALS AND METHODS

Plant growth conditions and treatments. J. Lachowiec grew the plants. Columbia-0 (Col-0) was used as wild type (WT). *bes1-D*, *bzr1-D*, *DWF4*-ox, *bin2-1*, *BRI1*-ox, and *bes1-2* (WiscDsLox246D02) were in the Columbia-0 (Col-0) background. The *bes1-D* mutant is a recapitulation line using a transgene to constitutively express the mutant form of *bes1-D* in a Col-0 background. Seedlings were grown for seven days on media with DMSO (mock) or geldanamycin, brassinolide, and brassinazole, dissolved in DMSO. Statistical significance of response of hypocotyl length of seedlings from 2-4 replicates of 10-60 seedlings was determined using standard least square linear regression.

Biochemistry. J. Lachowiec performed the western blots. For western blot, seven-day-old seedlings, grown in red LED light, were ground in liquid nitrogen. Buffer (0.15M Tris pH

6.8) was added, and extracted protein was quantified using Bradford's assay. Proteins were resolved using SDS-PAGE, transferred to nitrocellulose and probed with anti-BES1 antibody. For co-immunoprecipitation, ground rosette tissue was used. Extracted protein was incubated with Protein L Agarose, which was pre-incubated with anti-HSP90 3G3 antibody. Beads were pelleted and washed in buffer. Anti-BES1 antibody was used to detect BES1 in the input and pellet.

Phylogenetic tree and dN/dS analysis. J. Lachowiec did the phylogenetic analysis. Sequences for *BZR/BEH* family members in available sequenced plants were acquired from http://phytozome.net v5.0 from a BLAST search for gene families with similarity to *BES1*. MUSCLE 3.7 was used for amino acid alignment of the identified sequences, and Gblocks was used to remove regions with poor conservation. The remaining 86 sequences were re-aligned and neighbor-joining was used to create a distance tree. For the *BZR/BEH* tree, the outgroup was identified as a *BZR/BEH* family member that was closely related, but an outgroup to all *A. thaliana BZR/BEH* family members. Sequences were aligned in MUSCLE 3.7 and PhyML was used for maximum likelihood tree (Guindon and Gascuel 2003). For dN/dS analysis, codeml from PAMLv4.4b was run using models 0, 1, and 2 (Bielawski and Yang 2003).

Yeast data analyses. I analyzed the yeast data. Published HSP90 interactors were used (Zhao et al. 2005). The branch length of HSP90 interactors in three-member and two-member families was obtained from Ensembl Compara (release 61).

2.3 RESULTS

2.3.1 BES1 is an HSP90 client

Inhibition of HSP90 yields a wide variety of morphological phenotypes in *A. thaliana* plants (Queitsch et al. 2002; Whitesell et al. 1994; Sangster et al. 2007b; McLellan et al. 2007). Among these phenotypes, we previously noted severely dwarfed plants, which closely resembled known BR mutants. To directly test whether the BR pathway requires HSP90 function, we grew seedlings in the presence of exogenous BR (brassinolide, the most biologically active BR) with and without the highly specific HSP90 inhibitor geldanamycin (GdA) (Queitsch et al. 2002). Inhibition of HSP90 function significantly interfered with response to BRs (Figure S1A, R^2=0.68, p<0.0001, linear regression model, standard least square fit). Consistent with this finding, GdA also reduced seedling response to brassinazole, an inhibitor of BR biosynthesis (Asami et al. 2000), (**Figure S2.1B**, red light, R^2=0.76, p<0.0001; **Figure S2.1C**, dark R^2=0.81, p<0.0001).

We next addressed what step in the BR signaling pathway was most responsive to a loss of HSP90 function. The best-characterized HSP90 clients are the mammalian steroid hormone receptors and kinases (Whitesell et al. 1994; Picard et al. 1990; Taipale et al. 2010). The most common clients are transcription factors (Taipale et al. 2010). In *A. thaliana*, only a few clients are known, none of which function in the BR pathway (Hubert et al. 2003a; Hubert et al. 2003b; Ishiguro S 2002; Iki et al. 2010). As HSP90 clients do not share a common sequence or structural motif, client status is typically determined by a combination of genetic and biochemical analyses. Here, we took advantage of several well-characterized mutants in the BR pathway to test their response to HSP90 inhibition. We focused on the most likely clients: the steroid hormone receptor kinase BR INSENSITIVE1 (BRI1) and the transcription factors BES1 and BZR1 (Wang et al. 2001). To distinguish between HSP90 effects on BR signaling versus BR synthesis, we included a mutant in DWARF4 (DWF4), an enzyme that catalyzes a rate-limiting step of BR biosynthesis. Well-characterized gain-of-function mutants were used to bypass the extensive redundancy in the BR pathway. As had been shown previously, each mutant significantly increased hypocotyl length (He et al. 2005) (**Figure 2.1A, B**). Upon inhibition of HSP90 with 0.5 μM GdA, *BRI1-ox, DWF4-ox,* and

bzr1-D seedlings responded like wild-type seedlings (*BRI1-ox,* p=0.22; *DWF4-ox,* p=0.99, *bzr1-D*, p=0.6273), whereas *bes1-D* seedlings showed significant hypersensitivity to HSP90 inhibition (p<0.0001, **Table S2.1**, **Figure 2.1A-C**). *bes1-D* hypersensitivity to GdA suggests that BES1 may be an HSP90 client, while the wild-type response of *bzr1-D* mutants to GdA argues against an HSP90 client status for BZR1. Notably, the dominant *bes1-D* and *bzr1-D* mutants carry the identical amino acid change from proline to leucine (Tang et al. 2011; Wang et al. 2002; Yin et al. 2002; Albrecht et al. 2008; Guo et al. 2009; Li et al. 2009a). The *bes1-D* mutant used here is expressed from a constitutive 35S CaMV promoter, providing an alternative explanation for the increased GdA sensitivity of *bes1-D*. Further increases in GdA levels yielded a significant response in *BRI1-ox*, *DWF4-ox* (*BRI1-ox,* p<0.0001; *DWF4-ox,* p=0.0003), but not in *bzr1-D* seedlings (p=0.5918, **Table S2.1**), suggesting that there may be additional HSP90 targets upstream or alongside the transcription factors tested here (**Figure 2.1 A, B**). Similar results were obtained in the dark (**Figure S2A, B**). Together these data suggest that BES1, but not BZR1, is an HSP90 client.

Another criterion for HSP90 client status is proof of physical interaction (Taipale et al. 2010). Using a co-immunoprecipitation assay with an HSP90-specific antibody, we found that BES1 physically interacts with HSP90 in plants (**Figure 2.1D**). In the absence of BRs, BES1 is negatively regulated by BIN2 and several related kinases (Vert and Chory 2006). Cellular perception of BRs triggers inhibition of BIN2 and activates BES1. Hypophosphorylated, active BES1 can be detected as a fast mobility band on western blots (Yin et al. 2002) (**Figure 2.1E**). In this form, BES1 interacts with other transcription factors, binds DNA, and promotes plant growth (Cooper et al. 2005). In our studies, BES1 can interact with HSP90 independent of phosphorylation, as we were able to pull-down faster and slower mobility BES1 bands (**Figure 2.1D**).

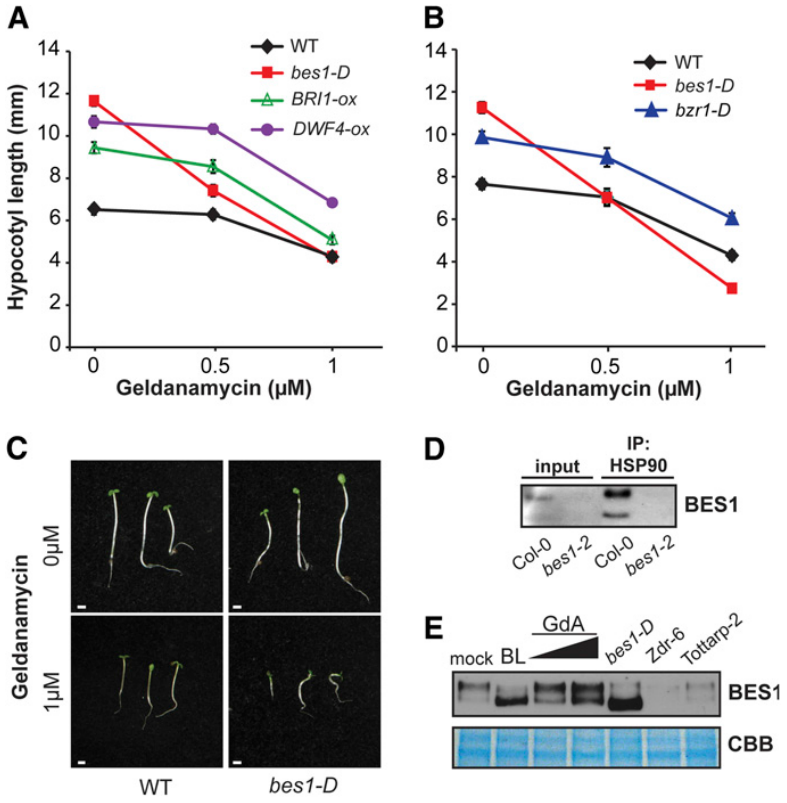


Figure 2.1 BES1 is an HSP90 client.

(A) Seedlings with increased BR signaling through overexpression of DWF4 (*DWF4-ox*) or overexpression of BRI1 (*BRI1-ox*) showed similar sensitivity to HSP90-inhibition by GdA compared to WT. In contrast, seedlings with constitutive activation of BES1 (*bes1-D*) were significantly more sensitive than WT in red light. Standard error is shown. **(B)** In contrast to the dramatic GdA hypersensitivity of *bes1-D* mutants, *bzr1-D* mutants respond like WT in red light. **(C)** Representative WT and *bes1-D* seedling phenotypes in red light. **(D)** BES1 interacts physically with HSP90. BES1 was immunoprecipitated (IP) with an HSP90 antibody in WT, but not in loss-of-function *bes1-2* mutants, confirming that this antibody is specific to BES1. Input and IP are shown for both. **(E)** GdA treatment caused a shift of BES1 mobility. Unlike a similar shift caused by brassinolide (BL) treatment, the GdA-induced mobility shift was associated with decreased hypocotyl length. Zdr-6 and Tottarp-2 show

reduced levels of BES1; detected protein is presumably due to the presence of other splice forms. Coomassie Brilliant Blue (CBB) is shown as a loading control.

While BRs and GdA have opposite effects on plant growth, and likely on BES1 function, inhibition of HSP90 produced a similar shift in BES1 mobility as BR treatment (**Figure 2.1E**). Like other well-established clients, such as the *Drosophila* Argonaute Piwi and the human transcription factor HSF, BES1 was not degraded upon HSP90 inhibition (**Figure 2.1E**) (Gangaraju et al. 2011; Zou et al. 1998). It appears that while the hypophosphorylated form of BES1 accumulated upon HSP90 inhibition, it is non-functional (**Figure 2.1A-D**). Our data suggest that BES1 may require HSP90 for its activation, perhaps by facilitating BES1 dimerization, promoting nuclear translocation, or interfering with phosphorylation by BIN2. If GdA indeed reduced the functional pool of BES1, we predicted that any increase in BIN2 function would sensitize plants to inhibition of HSP90. *bin2-1* mutants are semi-dominant hypermorphs with increased levels of phosphorylated BES1 leading to strongly reduced BES1 activity and repressed BR signaling (Kim and Wang 2010). As predicted, loss of HSP90 activity sensitized seedlings to a gain of BIN2 activity, as evidenced by a dramatically increased proportion of severely dwarfed seedlings in a segregating population of *bin2-1* mutant seedlings (**Figure S2.2C**). These results also suggest that BIN2 does not require HSP90 activity for its function; such a scenario would lead to suppression not enhancement of the *bin2-1* phenotype.

2.3.2 HSP90 facilitates the divergence of gene duplicates

If HSP90 indeed allows its clients to explore a wider range of sequence space, the *BES1* gene would be predicted to show evidence of relaxed selection compared to the *BZR1*gene. We created a phylogenetic tree of the *A. thaliana BZR*/*BEH* gene family using a gene from *Aquilegia coerulea* as outgroup (Figure 2A, S3). *BES1* and *BZR1* are the most recently diverged paralogs among the six *A. thaliana BZR*/*BEH* family members, with 88% amino acid identity (Wang et al. 2002). We determined the ratio of the rate of non-synonymous substitutions to the rate of synonymous substitutions (dN/dS) for *BES1* and *BZR1*, under a model allowing all branches of the *A. thaliana BZR*/*BEH* tree to evolve at

different rates. Consistent with our prediction, the gene encoding the HSP90 client BES1 (dN/dS=0.09) shows relaxed purifying selection compared to the gene encoding the non-client BZR1 (dN/dS=0.04) (**Figure 2.2A**). As *BES1* and *BZR1* diverged recently and only differ in a small number of amino acids, their difference in dN/dS was not significant. However, we took advantage of the entire *BEH* gene family to test whether *BES1* shows a different evolutionary rate compared to other family members. Using a maximum likelihood approach (Bielawski and Yang 2003), we showed that *BES1* indeed exhibits a significantly different evolutionary rate, if all other branches are assumed to evolve at the same rate (2δ=5.232, df=1, p=0.0222). In contrast, no significant difference was found for the evolutionary rate of *BZR1* under the same assumptions (2δ=3.244, df=1, p=0.0717).

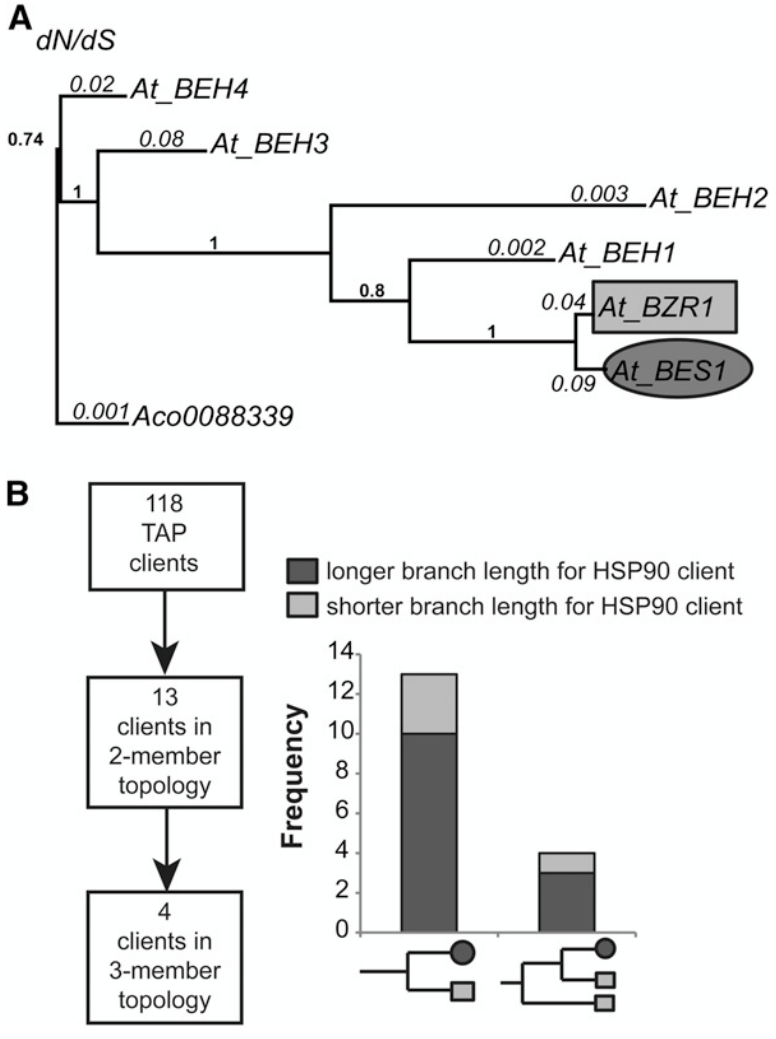


Figure 2.2 HSP90 clients show relaxed selection compared to their paralogs.
(A) The branch leading to *BES1* has a larger dN/dS ratio than the branch leading to *BZR1*. dN/dS ratios are in italics. Branch lengths represent the difference in the number of amino acid substitutions among family members from an unrooted tree. The fraction of 100 bootstraps supporting each branch are shown in bold. **(B)** *S. cerevisiae* HSP90 clients tend to evolve faster than their paralogs in two-member and three-member families. Clients are dark-grey circles; non-clients are light-grey squares. Star denotes significance. Significance was determined by using a one-sample Wilcoxon test, which tests the deviation from the expected ratio of client/non-client branch length of 1, n=13, 95% confidence interval 1.02-1.64, p=0.002.

To address whether HSP90 clients generally evolve faster, I analyzed a data set of systematically identified *Saccharomyces cerevisiae* HSP90 clients (Zhao et al. 2005). Proteins that physically interact with HSP90 by tandem affinity purification-tagged (TAP) mass spectrometry (Zhao et al. 2005) were considered as likely HSP90 clients. Likely HSP90 co-chaperones, identified by the TPR domain (Wegele et al. 2004), were removed from the analysis. Unlike clients, co-chaperones interact with HSP90 through the TPR domain and modulate HSP90 activity. As expected, the evolutionary rates of HSP90 clients did not differ significantly from all other yeast genes (**Figure S2.4A**). Factors such as differences in selection pressure, protein stability, or codon bias, among others, likely obscure any impact on evolutionary rate by HSP90. To compare genes well-matched for these factors, I identified gene duplicates in which one paralog encoded a likely HSP90 client. Consistent with our results for *BES1*, genes encoding yeast HSP90 clients showed significantly longer branch length than their respective closest paralog (**Figure 2.2B**, table S2, n=13, 95% confidence interval 1.02-1.64, p=0.002, one-sample Wilcoxon test, testing the deviation from the expected ratio of client/non-client branch length of 1). Likely non-clients, identified by synthetic genetic interaction with an *hsp90* mutation (SGI), did not show this trend (**Figure S2.4B**, n=27, p=0.97). I then tested a more stringent situation. In cases in which yeast HSP90 interaction status was the derived state (i.e., not present in a common ancestor), three out of four HSP90 clients showed longer branch length than their respective closest paralog (**Figure**

2.2B, three-member families). Next, we addressed whether these differences in evolutionary rate were due to expression differences between clients and their respective non-client paralogs. Genes that evolve faster tend to be expressed at a lower level (Drummond et al. 2005). In contrast, many clients are significantly higher expressed than their non-client paralogs across nearly 200 environmental conditions (Swarbreck et al. 2008) (**Table S2.2**, **Figure S2.5**). We observed no correlation of expression levels and branch lengths between clients and their respective paralogs (**Figure S2.4C**, **Figure S2.5**, **Table S2.2**). Taken together, our/my data suggest that HSP90 can facilitate the divergence of gene duplicates in yeast and plants.

2.3.3 HSP90 client BES1 shows hallmarks of sub-and neofunctionalization

Evolutionary theory holds that after gene duplication, one copy dies off quickly or changes function (Conant and Wolfe 2008). A surviving gene copy can retain part of the ancestral gene function, such as expression in fewer tissues or under certain environmental conditions (subfunctionalization) and/or acquire a novel beneficial function (neofunctionalization) (Conant and Wolfe 2008). An obvious subfunctionalization path for an HSP90 client is loss of function under environmental conditions that challenge HSP90 chaperone activity, such as increased temperature. HSP90 clients are typically less stable than other proteins and hence lose function at increased temperature despite induced HSP90 expression (Taipale et al. 2010). We grew seedlings at 27ºC, a temperature known to challenge HSP90 function but not induce heat stress in *A. thaliana* (Queitsch et al. 2002). The temperature response of *BRI1-ox*, *DWF4-ox*, and *bzr1-D* mutants closely resembled the response of wild-type seedlings **(Figure 2.3D)**. In contrast, not one of more than 100 *bes1-D* seeds in multiple independent experiments germinated at 27ºC (**Figure 2.3D, E**). This germination phenotype was completely suppressed at standard growth conditions (22ºC) **(Figure 2.3D**, **E**, **Figure S2.2D**). The loss of BES1 function at moderately elevated temperature, likely a direct result of challenged HSP90 function, is strong support for subfunctionalization.

If genes encoding HSP90 clients evolve faster than their non-client paralogs, neofunctionalization may be facilitated. Although there appears to be extensive redundancy between the HSP90 client BES1 and the non-client BZR1 (Cooper et al. 2005), we found evidence of BES1 neofunctionalization. First, the major *BES1* splice variant At1g19350.3 encodes a novel exon not found in *BZR1* or any of the other *BEH/BZR* family genes (**Figure 2.3A**). This exon shares significant homology with intergenic regions on chromosomes 3 and 5 (**Figure 2.3B**). This exon together with adjacent intron sequence has additional matches to the 5'UTR *BZR1* sequence and to another intergenic region (**Figure 2.3C**). Gene chimeras are a hallmark of neofunctionalization (Hahn 2009), and these findings suggest that distant genomic regions may have contributed to the novel *BES1* exon.

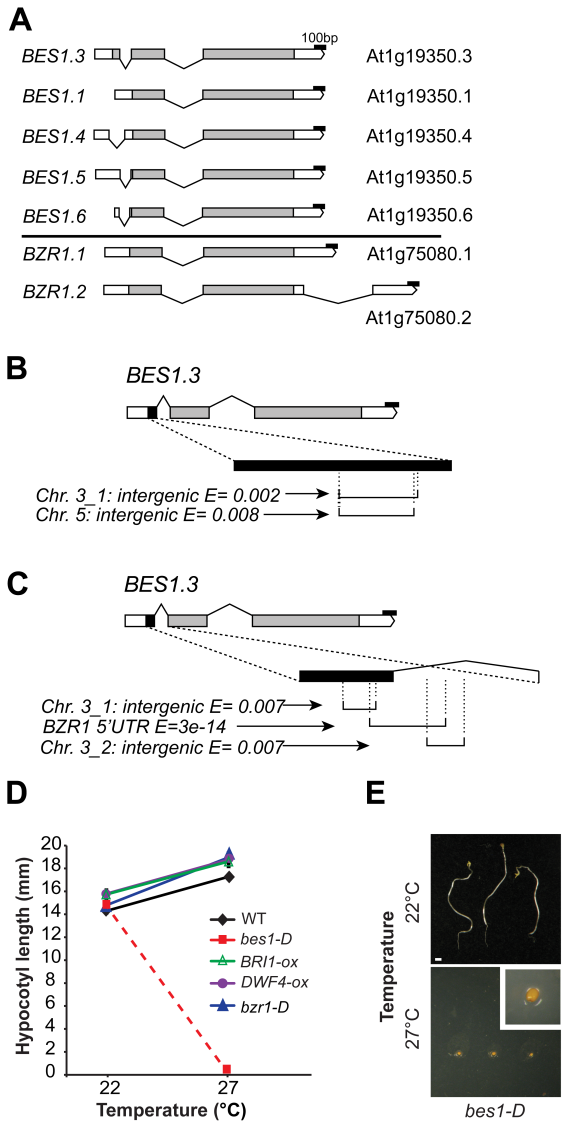


Figure 2.3. BES1 shows evidence of neo- and subfunctionalization.

(A) The major *BES1* splice variant, *BES1.3,* encodes a novel exon. Grey boxes are exons, black lines are introns, and white boxes are UTRs. **(B)** A BLASTn search for regions with homology to the novel *BES1.3* exon identifies intergenic loci. **(C)** A BLASTn search for regions with homology to the novel exon and first intron of *BES1.3* identifies *BZR1* 5' UTR sequence in addition to intergenic loci. **(D)** At 27°C, the *bes1-D* mutant failed to germinate, while all other mutants showed a wild-type response. **(E)** Representative *bes1-D* seedlings and seeds at 22°C and 27°C, respectively. Enlarged image of *bes1-D* (inset) shows a seed that failed to germinate at 27°C.

Second, we found that *BES1* polymorphisms across divergent *A. thaliana* accessions were significantly associated with the phenotypic variation these strains showed in response to HSP90 inhibition. Wild *A. thaliana* accessions harbor considerable genetic variation, yet due to *A. thaliana's* inbreeding life-style, individual accessions are nearly isogenic. Sensitivity to HSP90 inhibition varied dramatically among accessions (**Figure 2.4A**). Some accessions grouped with the hypersensitive *bes1-D* mutant (**Figure 2.4A**, red bar), whereas others responded very little. Accessions that grouped with the hypersensitive *bes1-D* mutant contained three intronic polymorphisms; but this association was not significant (**Figure 2.4A**, **B**, light-blue). In contrast, we found significant associations between decreased sensitivity to HSP90 inhibition and two other *BES1* polymorphisms (**Figure 2.4B**, **Figure S2.6A**). The first polymorphism is a frameshift mutation in the strains Zdr-6 and Tottarp-2 that results in an early stop codon (**Figure 2.4A**, black bars, **Figure 2.1E**). Consistent with a severely truncated BES1 protein, these strains responded little to HSP90 inhibition. The second polymorphism was found in four different strains (**Figure 2.4A**, green bars). This synonymous polymorphism in the 5' end of *BES1* alters the preferred codon for alanine to a rarely used codon (Wright et al. 2004). Changes in codon usage can alter translation efficiency and protein folding kinetics, potentially making stabilization by HSP90 superfluous. A similar change in codon usage has been observed for evolutions of viral proteins under conditions of reduced HSP90 in mammalian cells (Vaughan et al. 2010). Thus, in strong support of the capacitor hypothesis, HSP90 inhibition revealed phenotypic differences among divergent *A. thaliana* strains that are associated with *BES1* polymorphisms. Moreover, HSP90 client status appears to be highly dynamic with BES1 losing HSP90 dependence in some strains.

Another non-synonymous polymorphism in *BES1* did not correlate with response to HSP90 inhibition (**Figure 2.4B**). Consistent with HSP90 facilitating *BES1* evolution, this polymorphism may represent a step towards a novel phenotype through acquisition of a second mutation, with which it interacts epistatically (**Figure 2.4B**) (Salverda et al. 2011). Neutral non-synonymous mutations, such as mutations that increase protein stability without affecting function, can facilitate ascent to new fitness optimum (Peña et al. 2010). In contrast,

all twenty *BZR1* polymorphisms were synonymous (**Figure 2.4B**), consistent with *BZR1*'s lower evolutionary rate compared to *BES1*. None of these were associated with sensitivity to HSP90 inhibition, supporting BZR1's non-client status (**Figure 2.4B**, **Figure S2.6B**). Eight *BZR1* polymorphisms were unique to the strain Uod-7, which carries a large intronic insertion in addition to other intronic polymorphisms (**Figure 2.4B**). These polymorphisms may lead to mis-regulation of *BZR1* (Le Hir et al. 2003), thereby increasing the need for BES1. This scenario is consistent with Uod-7's hypersensitivity to GdA (**Figure 2.4A**). Taken together; our data suggest that HSP90 facilitates divergence of gene duplicates by promoting sub-functionalization through temperature sensitivity of its clients and neofunctionalization through their increased tolerance of mutations.

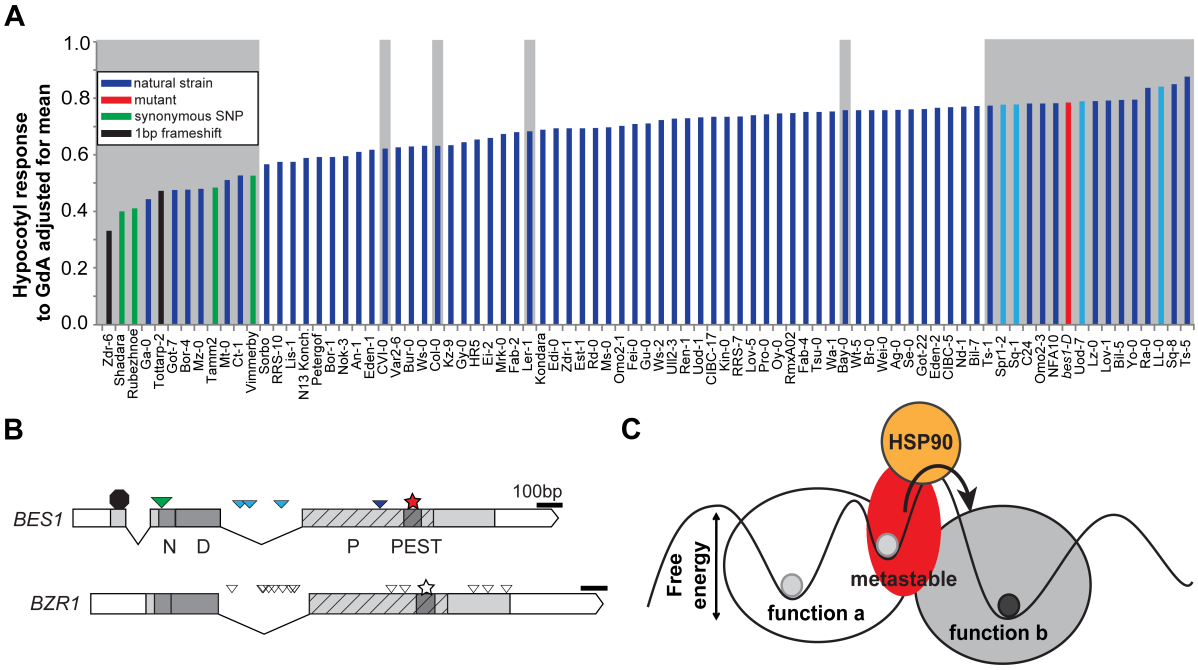


Figure 2.4. BES1 HSP90 client status is dynamic.

(A) Hypocotyl length GdA sensitivity of divergent *A. thaliana* strains and *bes1-D* (red). *BES1* and *BZR1* were sequenced in strains highlighted in grey. Strains with frameshift mutations are in black; strains with *BES1*-synonymous polymorphisms are green. Both polymorphisms are significantly associated with GdA sensitivity. Strains with *BES1* intronic polymorphisms are light blue. **(B)** *BES1.3* and *BZR1.1* polymorphisms. Grey boxes are

exons, black lines are introns, and white boxes are UTRs. The octagon marks the frameshift mutation, the green triangle marks the synonymous SNP, the blue triangle marks the nonsynonymous polymorphism, white triangles mark non-coding or synonymous polymorphisms, light blue triangles mark intronic polymorphisms in *BES1*, and stars mark the *bzr1-D* and *bes1-D* dominant mutations. The domains are N-nuclear localization signal, D-DNA binding domain, P-phosphorylation domain, PEST-PP2A interaction domain. **(C)** A protein that exists in a free energy minimum (function a) acquires mutations that render it metastable and thus recognized by HSP90. As an HSP90 client, the protein can visit a greater sequence space, increasing the chance of reaching another free energy minimum associated with a novel function (function b) and loss of client status.

2.4 DISCUSSION

With BES1, we have identified a novel HSP90 client in a crucial plant growth pathway (**Figure 2.1**). By necessity, plant growth must be finely tuned to the environment. The temperature sensitivity of BES1 tightly links BR pathway function to the ambient environment. We and others showed previously that HSP90 plays an important role in defenses against herbivores and microbial pathogens, (Sangster et al. 2007b; Hubert et al. 2003b; Hubert et al. 2003a; Sangster and Queitsch 2005), as well as the timing of flowering (Sangster et al. 2007b). Our findings add an important layer of complexity to known hormone-environment interactions (Nemhauser 2008; Robert-Seilaniantz et al. 2011a), especially in light of HSP90's known role in resource allocation among defense and growth pathways. Our data also add to the emerging evidence of functional divergence between BES1 and BZR1—the critical downstream targets of BR signaling. For example, recent genome-wide chromatin immunoprecipitation experiments show only partial overlap of gene targets in BES1 and BZR1 (Yu et al. 2011; Sun et al. 2010). This partial overlap highlighted key unresolved questions about BR transcriptional responses. Specifically, what distinguishes BES1- or BZR1-specific targets from targets regulated by both proteins, and what determines whether BES1 and BZR1 act as repressors or activators? As BES1, but not BZR1, is an HSP90 client, we speculate that interaction with the chaperone may facilitate

association of BES1 with specific partner proteins, resulting in BES1-specific functions (Cooper et al. 2005).

In addition to identifying a novel plant HSP90 client, the present study provides support for the capacitor hypothesis by garnering evidence of HSP90-facilitated evolution in extant genomes. When I compared paralogs, I found a significant trend that genes encoding diverse HSP90 clients evolved faster than non-clients (**Figure 2.2**). In contrast, when I compared diverse yeast HSP90 clients and non-clients as aggregate groups, I could not detect any significant difference in evolution rate (**Figure S2.4B**). This result is consistent with a similar analysis comparing prokaryotic genes encoding diverse clients of the bacterial chaperonin GroEL/ES and non-clients (Williams and Fares 2010). On the contrary, within the superfamily of mammalian kinases, strong HSP90 clients carry more non-synonymous mutations than non-clients (Taipale et al. 2012). These different results are not surprising because many other factors or gene properties influence evolutionary rate, most of which will differ for a diverse set of genes. Combined, these factors can outweigh and conceal effects of HSP90 and GroEL/ES client status on evolutionary rate. By focusing on gene duplicates of diverse HSP90 clients, which share many properties influencing evolutionary rate, I found support for HSP90's previously hypothesized effect on gene evolution. As overexpression of GroEL/ES also increases the evolutionary rates of its clients (Peña et al. 2010), our study supports an ancient and conserved role for protein chaperones in gene evolution.

Our/my findings also help resolve an apparent paradox about the fate of gene duplicates. Gene duplicates are functionally redundant immediately after duplication, rendering one copy superfluous or even harmful (Papp et al. 2003; Lynch and Conery 2000). Evolutionary theory predicts non-functionalization—one gene copy is silenced and subsequently lost—as the fate of most duplicated genes. Recent studies have challenged this view by revealing that gene duplicates with partially redundant function are maintained much longer than expected (Conant and Wolfe 2008; Lynch and Conery 2000; Maere et al. 2005; Dean et al. 2008; DeLuna et al. 2008). Moreover, in yeast, plants, insects, and humans, genes that can be maintained in duplicate show strong functional bias, with transcription factors and kinases significantly over-represented (Conant and Wolfe 2008; Wapinski et al. 2007; Guan

et al. 2007; Maere et al. 2005; Aury et al. 2006). Both observations raise questions about the molecular mechanism(s) that aid the initial preservation, continued maintenance, and eventual divergence of a specific subset of gene duplicates. Our/my data are consistent with a model in which acquisition of HSP90 client status is a molecular mechanism aiding each of these steps. Acquisition of HSP90 client status can occur through a single mutational step (Taipale et al. 2010; Citri et al. 2006). A new HSP90 client will be subject to immediate environmental subfunctionalization. As we observed for BES1, HSP90 clients are exquisitely sensitive to environmental conditions affecting protein folding (Nathan et al. 1997). Immediate and efficient subfunctionalization will counteract deleterious dosage effects and foster long-term maintenance of gene duplicates, and hence provide opportunity for their functional divergence. In fact, *bes1-D* mutants were more sensitive to a temperature increase than to the HSP90 inhibition, which is consistent with the fact that temperature change is a more complex perturbation compared to the specific inhibition of a single chaperone. In addition to providing evidence for HSP90-faciliated subfunctionalization, we show that HSP90 client status correlates with hallmarks of neofunctionalization and increased evolutionary rates of the genes encoding them. HSP90 recognizes metastable signal transduction proteins, most of which are transcription factors and kinases (Taipale et al. 2010). We speculate that HSP90's specificity for these substrates contributes to the observed functional bias among gene duplicates. Additional support for HSP90's role in gene duplicate divergence comes from our observation that HSP90 client status is dynamic in the BZR/BEH family and in yeast gene families, with frequent gains and losses (**Figure 2.4A, B**). Dynamic client status is also observed in the mammalian kinome (Taipale et al. 2012). In fact, even across wild *A. thaliana* accessions BES1 HSP90 client status itself appears to be dynamic.

The presence of HSP90 clients and non-clients in gene families with partially redundant function has important implications for another aspect of HSP90-mediated capacitance. Inhibition of HSP90 increases phenotypic variation even in the absence of genetic variation (Queitsch et al. 2002; Sangster et al. 2008a). This increase of phenotypic variation in isogenic lines has been attributed to an increased frequency of stochastic events in development or a greater sensitivity to microenvironments. As recently shown, significant phenotypic variation arises in isogenic worms due to the loss of functional redundancy

between paralogs (Burga et al. 2011; Casanueva et al. 2012). Under conditions that challenge HSP90 function, HSP90-dependent paralogs will become inactive, thereby reducing functional redundancy and increasing phenotypic variation. We suggest a general role for HSP90 in maintaining a reservoir of phenotypic variation through facilitating conditional functionality of gene duplicates.

2.5 Supplementary Materials and Methods

Plant growth conditions and treatments

J. Lachowiec grew the plants. Columbia-0 (Col-0) was used as wild type. *bzr1-D*, *DWF4*-ox, *bin2-1*, *BRI1*-ox, and *bes1-2* (WiscDsLox246D02) were in the Columbia-0 background. The recessive, loss-of-function mutant *bes1-2* was isolated in this study, using a T-DNA insertion line (WiscDsLox246D02). The *bes1-D* mutant is a recapitulation line using a transgene to constitutively express the mutant form of *bes1-D* in Col-0. Seeds were stratified for three nights at 4°C and grown vertically in the dark or red LED light for seven days on 0.5x LS 0.08% bactoagar growth media at 22°C. Geldanamycin (Sigma #G3381), brassinolide (Chemiclones #101), and brassinazole (Chemiclones #117) were dissolved in DMSO. Drugs and mock (DMSO only) treatments were added to growth media. For growth at 27°C, seedlings were placed in 22°C for 24 hours and then transferred to 27°C. For response to GdA in accessions, 15 seeds from 96 accessions were grown in the dark for seven days on medium containing DMSO or 1μM GdA. Accessions in which less than seven seeds germinated were removed from further analysis. Difference between hypocotyl length in mock and treatment was taken and adjusted by the mock mean per genotype.

Statistical analyses

J. Lachowiec performed the statistical analyses. Analyses of data were performed with JMP7 (SAS Institute). To measure the effect of drug treatments and interactions on hypocotyl length, standard least square linear regression was used. Each drug, drug by drug interaction (all fixed effects), and replicate (random effect) was modeled (2-4 replicates for each experiment, 10-60 seedlings each). When comparing the effect of genotype on response to drug treatment, standard least square linear regression was used. Genotype, drug, genotype

by drug interaction (all fixed effects), and replicate (random effect) was modeled (2-4 replicates of each experiment, 10-60 seedlings each). To associate *BES1* and *BZR1* SNPs with response to GdA, the GdA response was normalized using the mock treatment mean for each strain. Each SNP in the 30 strains was tested for association with GdA response using ANOVA (One-way).

Western blot

J. Lachowiec did the western blots. Seedlings were grown for seven days in red LED light and then ground in liquid nitrogen. Buffer (0.15M Tris pH 6.8) was then added, and extracted protein was quantified using Bradford's assay (Pierce #1856210). Approximately 15ug of protein was loaded per lane. Gel Code Blue Stain Reagent (Pierce #24590) was used to determine loading. Proteins were transferred to nitrocellulose and probed with anti-BES1 antibody (gift of Yanhai Yin).

Co-immunoprecipitation

J. Lachowiec did the co-immunopresipitations. Col-0 (WT) and *bes1-2* rosette leaves grown in long day conditions were harvested, frozen in liquid nitrogen, ground, and resuspended in HEM buffer (10mM HEPES, 1mM EDTA, 20mM sodium molybdate, 1mM PMSF (Fluka 78830) and protease inhibitor (Roche 11-836-153-001)). The solution was ultracentrifuged for 100,000 x g for 30 minutes at 4°C. The resulting supernatant was incubated for three hours with Protein L Agarose (Pierce 20510), which was pre-incubated with anti-HSP90 3G3 antibody (Enzo Lifescences ALX-804-079-R400). Beads were pelleted and washed with HEM buffer. Anti-BES1 antibody (gift of Yanhai Yin) was used to detect BES1 in the pellet. The newly isolated *bes-2* mutant (recessive, loss-of-function) was used to confirm that this antibody is specific to BES1.

Phylogenetic tree and dN/dS analysis

J. Lachowiec performed the phylogenetic analyses. Sequences from genomes that were not publically available were removed manually. MUSCLE 3.7 was used for amino acid alignment of the identified sequences, and Gblocks was used to remove regions with poor conservation. Sequences with low similarity in the high conservation region identified with

Gblocks were removed manually. The remaining 86 sequences were re-aligned and neighbor-joining was used to create a distance tree. For the *BZR/BEH* tree, the outgroup was identified as a *BZR/BEH* family member that was closely related, but an outgroup to all *A. thaliana BZR/BEH* family members. MUSCLE 3.7 was used for amino acid alignment of outgroup and *A. thaliana* family members using default parameters and manually examined for errors. PhyML was used for maximum likelihood tree (Guindon and Gascuel 2003). For dN/dS analysis, PAL2NALv13 was used to convert amino acid alignment to codon alignment. The codeml program from PAMLv4.4b was run with gaps removed using models 0, 1, and 2 (Bielawski and Yang 2003). The branch leading to *BES1* or *BZR1* was allowed to vary in model 2. The dN/dS values for the *BZR/BEH* family came from model 1.

Yeast data analyses

I performed the yeast data analyses. For this study, published HSP90 interactors were used (Zhao et al. 2005). TPR-domain containing proteins were curated from available literature and sequence information and excluded from further analysis. I used the Ensembl Compara (release 61) database to obtain the branch length of HSP90 interactors in three-member and two-member families. For the three-member families, I filtered out genes in which the inner or outer-paralog were also HSP90 interactors. For the two-member families I filtered out those in which both genes were HSP90 interactors. To determine whether there was a significant difference in branch length between clients, non-clients, and their paralogs, I calculated the proportion of total branch length for each client or non-client and their paralogs. A one-sided Wilcoxon test was used to determine significance of detected differences in proportional branch length between client and paralog or non-client and paralog.

Table S2.1 Significance of response to HSP90 inhibition

	0.5uM GdA		1.0uM GdA	
genotype	R^2	p-value	R^2	p-value
bes1-D	0.73	<0.0001	0.89	<0.0001
bzr1-D	0.26	0.63	0.63	0.59
BRI1-ox	0.48	0.22	0.74	<0.0001
DWF4-ox	0.75	0.99	0.8	0.0003

Seedlings with increased BR signaling through overexpression of DWF4 (*DWF4-ox*) or overexpression of BRI1 (*BRI1-ox*) showed WT-like sensitivity to HSP90-inhibition at 0.5µM GdA. In contrast, seedlings with constitutive activation of BES1 (*bes1-D*) were significantly more sensitive than WT at 0.5µM GdA in red light. At 1.0 µM GdA, *BRI1-ox, DWF4-ox*, but not *bzr1-D* showed increased sensitivity compared to WT.

Table S2.2 Yeast HSP90 clients tend to evolve faster than their respective non-client paralogs independent of expression levels.

client			non-client paralog			comparison		
gene	expression mean	branch length	gene	expression mean	branch length	client minus paralog expression mean		client branch/non-client branch
YBR172C	-0.3258	1.0072	YPL105C	-0.1654	0.7407	-0.1603	*	1.3598
YCL024W	-0.3181	0.3717	YDR507C	-0.4492	0.2050	0.1311		1.8136
YDL025C	0.3978	0.5842	YOR267C	0.2715	0.4900	0.1263		1.1923
YDL199C	0.5594	1.7737	YFL040W	0.0277	1.7386	0.5317	*	1.0202
YDR001C	0.7171	0.2090	YBR001C	0.5592	0.1304	0.1579	***	1.6024
YFL011W	-0.1650	0.2859	YMR011W	-0.4621	0.1549	0.2971	***	1.8461
YFR024C-A	0.0450	0.3320	YHR016C	0.3144	0.2701	-0.2694		1.2291
YGL077C	-0.3462	0.4449	YNR056C	0.0736	1.0123	-0.4198		0.4395
YHR080C	0.3019	0.5174	YDR326C	-0.3168	0.2316	0.6187	***	2.2340
YKR003W	0.0862	0.2907	YHR001W	0.1561	0.2927	-0.0700	***	0.9931
YMR192W	-0.0287	1.6764	YPL249C	0.0550	0.8823	-0.0837		1.9000
YNR011C	-0.1769	1.0256	YKL078W	-1.1110	1.2136	0.9340	***	0.8451
YNR031C	0.0252	0.5748	YCR073C	-0.2379	0.5540	0.2630	***	1.0375

Mean expression was calculated from (Gasch *et al*. 2000), using all their reported conditions. Differences in mean expression were calculated using a Wilcoxon paired sign rank test. Significance values are abbreviated as * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

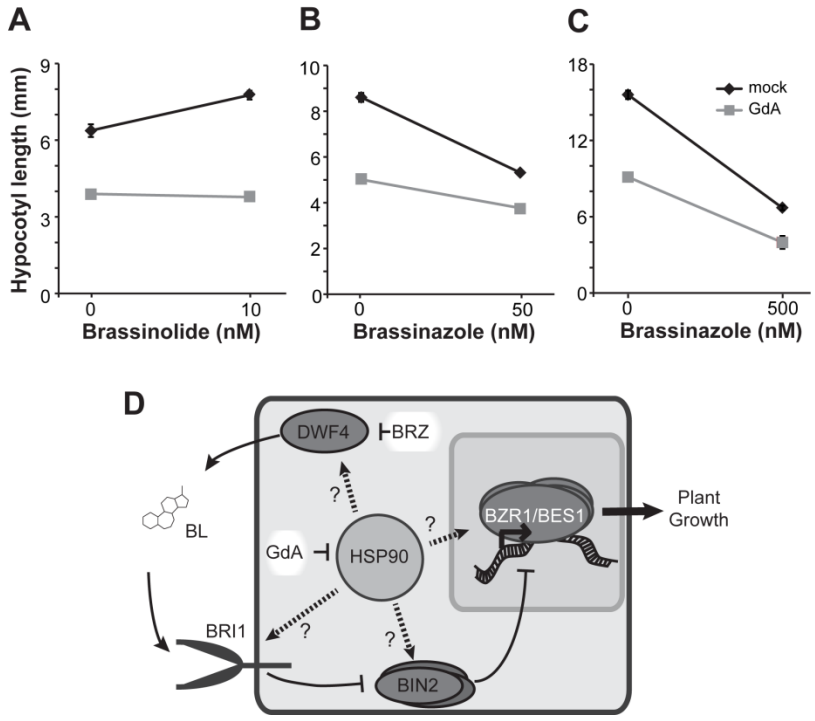


Figure S2.1 HSP90 is required for BR signaling.

(A) GdA reduces seedling response to brassinolide (BL), the most biologically active BR in red light. Red light was used because GdA decays rapidly in white light. Standard error is shown for all values. **(B)** Treatment with brassinazole (BRZ), a BR biosynthetic inhibitor, greatly reduces the effect of GdA treatment on seedling growth in red light. **(C)** Treatment with BRZ greatly reduces the effect of GdA treatment on seedling growth in the dark. **(D)** The BR signaling pathway. BRs (brassinolide-BL) are synthesized through a number of enzymatic reactions, including the rate-limiting enzyme DWF4, the target of BRZ. When BR levels are low, BIN2 and related kinases inhibit the activity of a family of transcription factors, including BES1 and BZR1. As BRs accumulate, they are detected by the plasma-membrane localized receptor BRI1. Activated BRI1 triggers a series of phosphorylation and dephosphorylation events that ultimately inhibit the activity of BIN2 and its paralogs. Hypophosphorylated BES1 and BZR1 then can bind DNA and trigger changes in target gene transcription. Each of the indicated proteins is a potential HSP90 client.

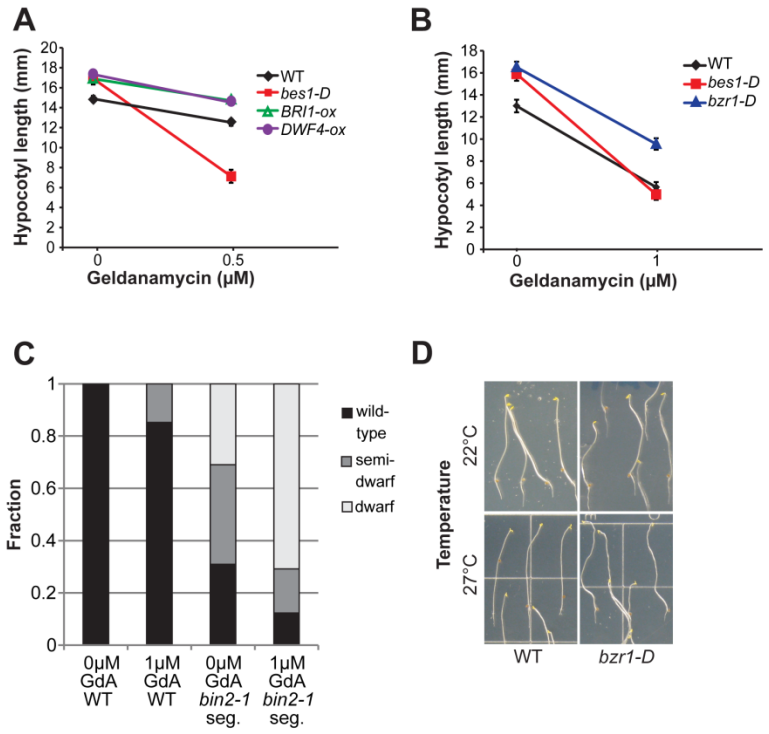


Figure S2.2 BES1 is an HSP90 client.

(A) Seedlings with increased BR signaling through overexpression of DWF4 (*DWF4-ox*) or overexpression of BRI1 (*BRI1-ox*) showed similar sensitivity to HSP90-inhibition by GdA compared to WT in the dark. In contrast, seedlings with constitutive activation of BES1 (*bes1-D*) were significantly more sensitive than WT in the dark. Standard error is shown. **(B)** In contrast to the dramatic GdA hypersensitivity of *bes1-D* mutants, *bzr1-D* mutants respond like WT in the dark. **(C)** BIN2 is not a likely HSP90 client. The *bin2-1* mutation results in a dwarfed phenotype compared to WT. The phenotypes of a segregating population of *bin2-1* were enhanced rather than alleviated upon inhibition of HSP90. Plant growth phenotypes were categorized as wild-type, semi-dwarf, or dwarf and assessed for 60 seedlings per genotype and condition. **(D)** Hyper-sensitivity to increased temperature is not observed in WT and *bzr1-D* seedlings grown at 22°C and 27°C.

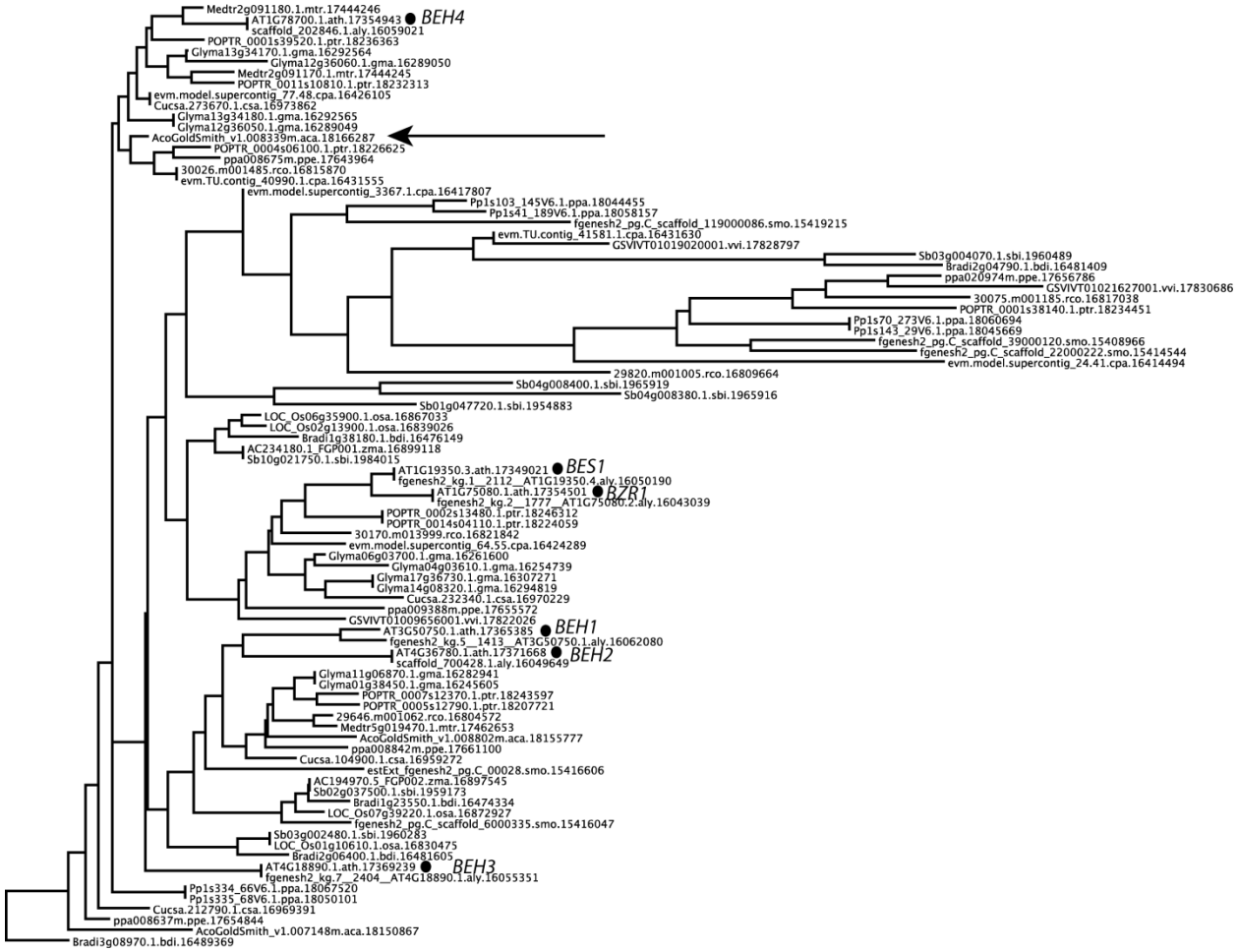


Figure S2.3 BZR/BEH family genes are found in many plant species. A neighbor-joining tree, using percent amino acid identity among genes identified as potential BZR/BEH family members among 16 plant species, shows that a gene from *Aquilegia coerulea* (arrow) is the most closely related outgroup to all BZR/BEH family members in *A. thaliana* (indicated by "•").

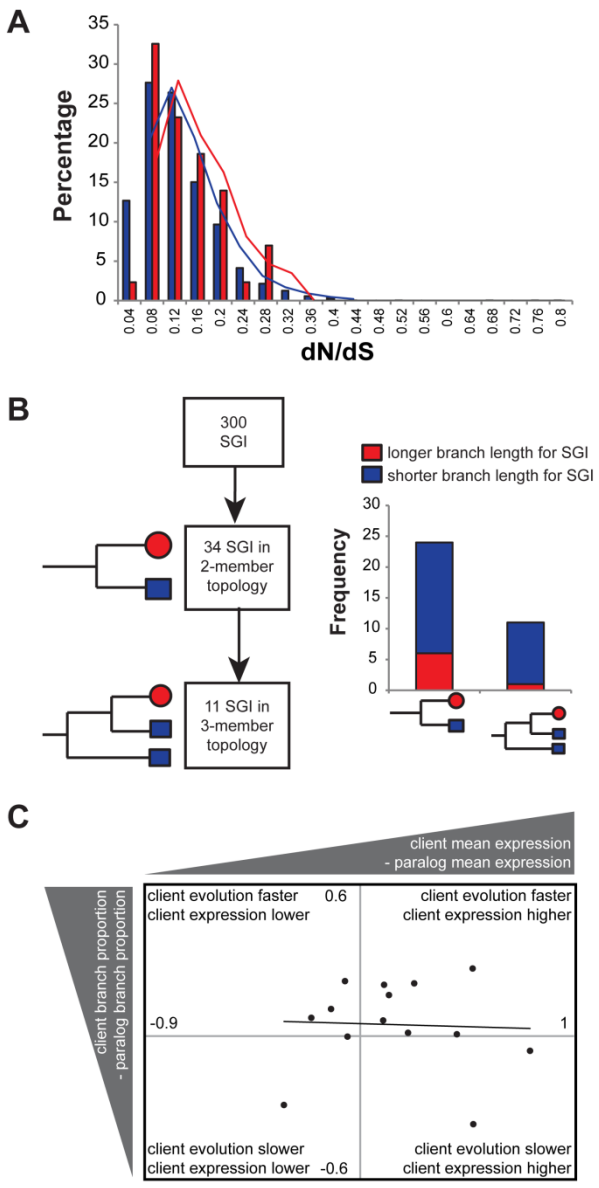
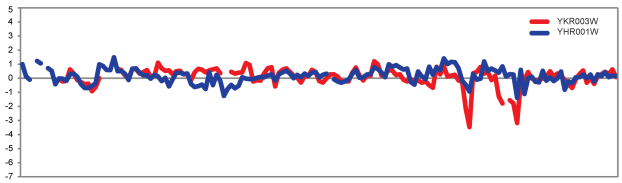
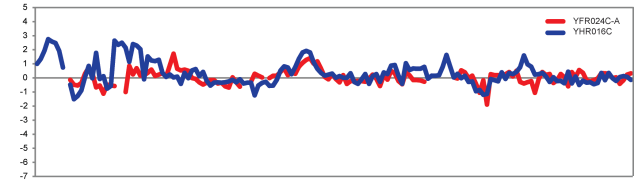
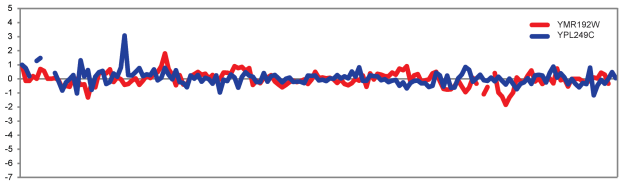


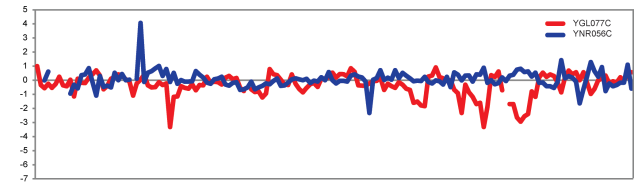
Figure S2.4 ***S. cerevisiae*** **HSP90 genetic interactors do not show elevated rates of evolution. (A)** *S. cerevisiae* HSP90 TAP-identified clients do not show a significant difference from all other yeast genes in their dN/dS ratios. **(B)** *S. cerevisiae* HSP90 synthetic genetic interactors (SGI) do not show longer branch lengths than their paralogs in two-member and three-member families. In fact, they tend to show significantly shorter branch lengths (p=0.03, Wilcoxon, one-sided). SGI are red circles; SGI paralogs are blue squares. **(C)** Increased evolutionary rate of yeast HSP90 clients is not explained by lower expression.

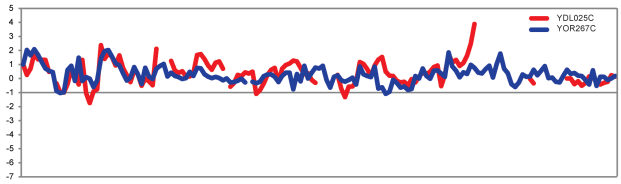
X-axis shows client mean expression minus non-client paralog mean expression across many environmental conditions (Gasch et al. 2000). Y-axis shows client evolutionary rate minus non-client paralog evolutionary rate (using proportional branch length for both). Labels in quadrants indicate relationship of client evolutionary rate and expression level. Numbers indicate scale. Solid trend line shows lack of correlation between evolutionary rate and expression levels ($p=0.882$, $R^2=0.0019$).

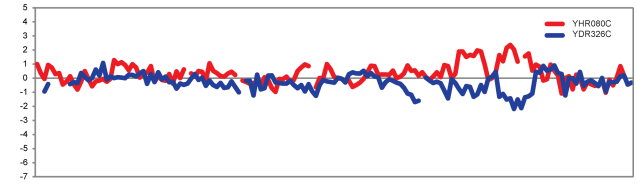

YKR003W
YHR001W

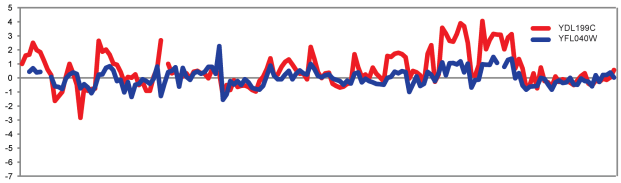

YFR024C-A
YHR016C

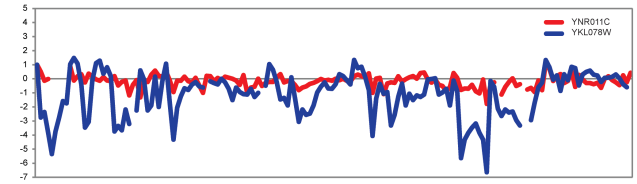

YMR192W
YPL249C

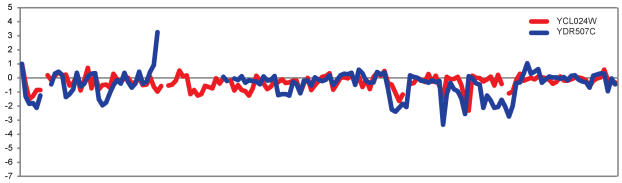

YGL077C
YNR056C

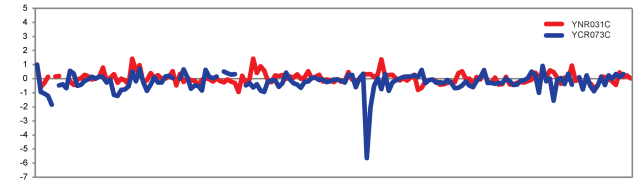

YDL025C
YOR267C

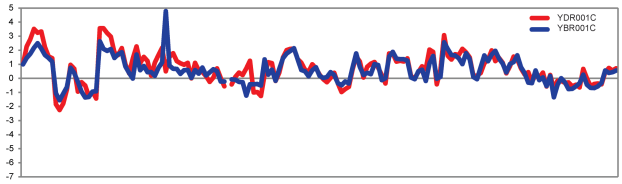

YHR080C
YDR326C

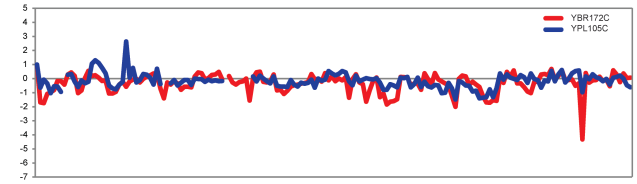

YDL199C
YFL040W


YNR011C
YKL078W


YCL024W
YDR507C


YNR031C
YCR073C


YDR001C
YBR001C


YBR172C
YPL105C

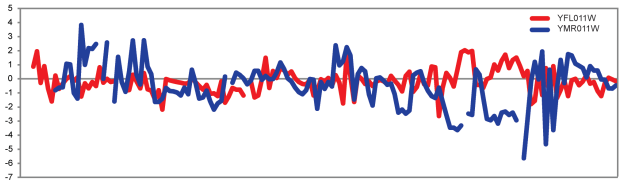

YFL011W
YMR011W

Figure S2.5 Expression profiles of yeast HSP90 clients and their respective non-client paralogs across 173 environmental conditions (Gasch et al. 2000). Client expression is in red; non-client expression is in blue. All 14 pairs are shown. Statistics for observed differences appear in **Table S2.2**.

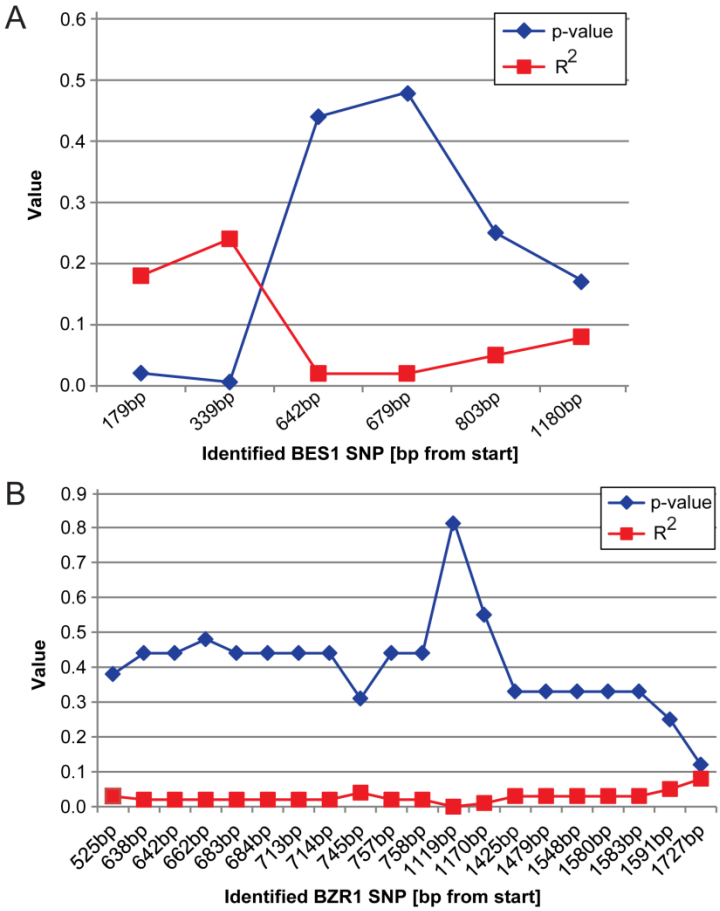


Figure S2.6 Significant associations were observed between SNPs and hypocotyl response to HSP90 inhibition in *BES1* (A) but not *BZR1* (B) among the 30 sequenced accessions. ANOVA was used to assess significance.

Chapter 3

THE PROTEIN CHAPERONE HSP90 AND THE EVOLUTION OF THE HUMAN KINOME[2]

This chapter was included in the thesis because I performed analyses that contributed to the publication of the present article. Specifically, I tested whether HSP90 client kinases had greater nucleotide diversity and damaging mutations than non-client kinases. The results of these analyses led J. Lachoweic to ask if the HSP90's effect on kinase evolutionary rate increased with time of divergence. I also contributed to data for table S3.1. The first person "we", and the "our" pronoun indicates the work of J. Lachowiec. My work is stated with the "I" person, and the "my" pronoun.

Abstract

Heat-shock protein 90 (Hsp90) promotes the maturation and stability of its client proteins, including many kinases. In doing so, Hsp90 may allow its clients to accumulate mutations as previously proposed by the capacitor hypothesis. If true, Hsp90 clients should show increased evolutionary rate compared to non-clients; however, other factors, such as gene expression and protein connectivity, may confound or obscure the chaperone's putative contribution. Here, we compared the evolutionary rates of many Hsp90 clients and non-clients in the human protein kinase superfamily. We show that Hsp90 client status promotes evolutionary rate independently of, but in a similar magnitude to, gene expression and protein connectivity. Hsp90's effect on kinase evolutionary rate was detected across mammals and increased with time of divergence. Hsp90 clients also showed increased nucleotide diversity and harbored more damaging variation than non-client kinases across humans. These results are consistent with the central argument of the capacitor hypothesis

[2] This chapter was published in Molecular Biology and Evolution as "Hsp90 promotes kinase evolution" by J Lachowiec, T Lemus, E Borenstein, C Queitsch.

that interaction with the chaperone allows its clients to harbor genetic variation. Hsp90 client status is thought to be highly dynamic with as few as one amino acid change rendering a protein dependent on the chaperone. Contrary to this expectation, we found that across protein kinase phylogeny Hsp90 client status tends to be gained, maintained, and shared among closely related kinases. We also infer that the ancestral protein kinase was not an Hsp90 client. Taken together, our results suggest that Hsp90 played an important role in shaping the kinase superfamily.

3.1 INTRODUCTION

The conserved heat shock protein Hsp90 facilitates the proper folding and stability of its substrates (clients) (Taipale et al. 2010), many of which are kinases with important roles in growth and development. Hsp90 perturbation increases the penetrance of expressed genetic variants and reveals cryptic genetic variation in genetically divergent populations of plant, fly, yeast, and fish (Queitsch et al. 2002; Jarosz and Lindquist 2010; Rutherford and Lindquist 1998; Yeyati et al. 2007). In worms, naturally varying Hsp90 levels predict mutation penetrance with lower Hsp90 levels resulting in greater penetrance (Burga et al. 2011; Casanueva et al. 2012). These observations with traditional model organisms prompted the controversial hypothesis that Hsp90 plays an important evolutionary role, allowing genetic variation to remain phenotypically silent and releasing it in environments that perturb Hsp90 function (Rutherford and Lindquist 1998). Consistent with this hypothesis, perturbing Hsp90 function in surface-dwelling *Astyanax mexicanus* fish results in eye phenotypes that are reminiscent of the natural adaptation of eye loss in the cave-dwelling fish of the same species, presumably due to release of Hsp90-dependent standing variation (Rohner et al. 2013). Hsp90-dependent standing variation occurs frequently in natural strains of plants, flies, and yeast and often affects complex traits (Sangster et al. 2008a; Jarosz and Lindquist 2010; Rutherford and Lindquist 1998), consistent with a significant role of Hsp90 in evolution, especially in the evolution of genes encoding its client proteins.

The evolutionary rate of protein coding genes is commonly measured as the ratio of non-synonymous changes to synonymous changes, dN/dS. Using this measure, we recently reported that genes encoding Hsp90 clients tend to evolve faster than genes encoding their

non-client paralogs (Lachowiec et al. 2013). This trend was not observed without considering paralog status, presumably because many other factors, some of which are shared among paralogs, influence evolutionary rate. One drawback of this study was the small number of available client and non-client paralog pairs. Recently Taipale et al. (2012) systematically annotated Hsp90 clients in the human kinome, using a high-throughput assay to assess kinase and Hsp90 interactions. Of the 314 tested kinases, 98 are classified as strong Hsp90 clients and 95 as weak clients. Notably, the dN of strong Hsp90 clients is greater than the dN of non-clients, suggesting that interaction with Hsp90 may indeed allow for increased accumulation of non-synonymous genetic variation (Taipale et al. 2012).

There is precedent for chaperone-facilitated evolution of client proteins. In prokaryotes, GroEL/ES clients show different degrees of dependency on the chaperone for stability and folding. Genes that encode proteins with greater GroEL/ES dependency show greater dN/dS after correction for confounding factors, using appropriate statistical models (Fares et al. 2002; Bogumil and Dagan 2010, 2012; Williams and Fares 2010; Bogumil et al. 2012). Similarly, overexpression of GroEL/ES promotes tolerance of nonsynonymous mutations (Tokuriki and Tawfik 2009). Factors that may confound putative chaperone effects on evolutionary rate include absolute expression levels (Pál et al. 2001; Wall et al. 2005) and a protein's connectivity in protein–protein interaction (PPI) networks (Fraser et al. 2002). Highly expressed genes tend to evolve slower than genes with lower expression levels as do genes encoding proteins that interact with many other proteins, presumably due to the selective pressure to maintain all functional interactions. Neither of the studies discussed above (Lachowiec et al. 2013; Taipale et al. 2012) controlled for these factors when examining the role of Hsp90 on protein evolution. Another factor contributing to dN/dS is protein stability; genes encoding stable proteins tend to evolve faster (Bloom et al. 2006). Many Hsp90 clients are inherently unstable and are rapidly degraded in Hsp90-limited conditions (Taipale et al. 2010). Interaction with the chaperone Hsp90 promotes client protein stability (Taipale et al. 2012), suggesting a mechanistic basis for the chaperone's putative effect on dN/dS.

Here J. Lachowiec and I set out to dissect and compare the contributions of Hsp90 client status, gene expression levels, and protein interaction degrees to the evolutionary rate of kinases. We evaluate the role of Hsp90 in kinome evolution within and across lineages. Hsp90 client status is thought to be dynamic throughout gene family evolution with as few as one amino acid change resulting in a switch from non-client to client (Citri et al. 2006; Lachowiec et al. 2013; Taipale et al. 2012). We assess this unexplored dynamic by quantifying the transition rates between client and non-client states throughout kinase evolution and show that client status switching is infrequent. We also infer the ancestral state of client status among kinases. Taken together, J. Lachowiec and my results support the controversial hypothesis that Hsp90 plays an important role in the evolutionary processes shaping large gene families.

3.2 METHODS

Data sources and estimating contributions to dN/dS

J. Lachowiec and I obtained Hsp90 interaction scores, client category, and number of connections in the protein-protein interaction (PPI) network were obtained from (Taipale et al. 2012). J. Lachowiec obtained the gene expression data and estimated HSP90 contributions to the dN/dS. For gene expression levels in human, RNA-seq counts normalized for read depth (RPKM, Reads per kilo-base per million) across eleven normal human tissues from (Castle et al. 2010) were downloaded (http://medicalgenomics.org/rna_seq_atlas/ [September, 2012]). The average expression was taken per gene (across all splice forms) and the maximum and mean expression across all tissues was calculated. Only the 210 kinases that overlapped between the Hsp90 kinome analysis (Taipale et al. 2012) and the eukaryotic kinase tree (Manning et al. 2002) and that had values for expression (Castle et al. 2010), dN/dS, and PPI connectivity were used in the downstream analyses. Since the phylogenetic relatedness of the kinases may influence these contributions, we also considered each variable in a phylogenetic context using phylogenetically independent contrasts (PIC). Phylogenetic contrasts for each variable were estimated in R using *pic* in the package ape (Paradis et al. 2004) with the kinome tree from (Manning et al. 2002) (http://kinase.com/human/kinome/groups/ePK.ph). J. Lachowiec

conducted linear regression analyses with the intercept set to 0 (Garland et al., 1992) since the order of subtraction to calculate the PIC was arbitrary and calculated the association between each pair of variables. Partial correlations among the variables were calculated with R using *pcor* in the package ppcor (http://cran.r-project.org/web/packages/ppcor/index.html).

dN/dS comparisons across species

J. Lachowiec obtained the list of ortholog genes and their sequence. For each kinase, orthologs for the human kinases and their respective dN/dS values were identified using Ensembl release 70, (http://www.ensembl.org [March, 2013]) in the following species: *Pan troglodytes*, *Pongo abelii*, *Nomascus leucogenys*, *Macaca mulatta*, *Callithrix jacchus*, *Mus musculus*, *Rattus norvegicus*, *Ictidomys tridecmlineatus*, *Oryctolagus caniculis*, *Canis familiaris*, *Bos taurus, Sus scrofa*, and *Ailuropoda melanoleuca*. If multiple orthologs were identified, the kinase was removed from the analysis to avoid double counting it. Time to common ancestor was obtained from compiled studies curated at TimeTree.org (Hedges et al. 2006).

Hsp90 client divergence timing among humans

I acquired human exome sequences from the Exome Variant Server (Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (http://evs.gs.washington.edu/EVS/ [November, 2012]) using an in-house Perl script.
In addition to the filters from the ESP Server, I incorporated the filters used in (Tennessen et al. 2012) to account for possible copy number variation: all SNVs had to have a quality score above 20, allele balancing above 65%, read depth between 10 and 1000x, I also excluded genotype qualities values of zero. To calculate nucleotide diversity, I used

$$\pi_{SNV} = 2f(1-f)\frac{n}{n-1}$$

where f is the frequency of the major allele and n is the number of haploid genomes (Hernandez et al. 2011). For each gene π_{SNV} was summed for each nucleotide position and normalized for gene length. For gene length I used the length of the longest CDS reported by Ensembl release 69. I extracted GERP scores for each SNV from ESP, calculated the average GERP score for each kinase gene, and normalized the GERP counts by the length of the gene (CDS sequence). PolyPhen scores were also obtained from ESP.

Examining client status dynamics and estimating transition rates between client and non-client states.

J. Lachowiec performed the analyses of this section. She estimated the phylogenetic signal of client status across the kinome, Pagel's λ, using the function *fitDiscrete* in the package geiger (http://cran.r-project.org/web/packages/geiger/index.html). She modeled a kinase tree with all branches leading to tips of equal length (λ = 0) eliminating the phylogenetic signal. J. Lachowiec then compared this model to a model with the true kinase phylogeny (i.e. an optimized λ) using likelihood-ratio tests, where the likelihood ratio approximates a χ^2 distribution with one degree of freedom (Harmon et al. 2008; Motani and Schmitz 2011). She used BayesTraits v2 (Pagel 1994) (http://www.evolution.rdg.ac.uk/BayesTraits.html) to estimate transition rates along the kinase tree using a MultiState model of evolution. She used the human kinase tree and the client status of each kinase, coding both weak and strong clients (according to Taipale *et al.*, 2012) as 'A' and non-clients as 'B'. She then used maximum likelihood with the parameter rate set to 2, representing the two transition rates 1) client to non-client and 2) non-client to client. She tested various hypotheses about transition rate parameters by restricting transition rates: 1) rate of gain of client status to 0, 2) rate of loss of client status to 0, and 3) equal gain and loss rates. She compared the log(likelihood) of the restricted models to one another and to the log(likelihood) of the unrestricted models. For these comparisons she used likelihood-ratio tests, where the likelihood ratio follows a χ^2 distribution with degrees of freedom equal to the number of restricted parameters. She repeated the analysis using Markov Chain Monte Carlo (MCMC) implemented in BayesTraits v2. She had appropriate levels of acceptance with the option rateDev set to 2. She let the chain run for 100,000 iterations with a uniform prior distribution between 0 and 100. She examined the transition rates after a 20,000 iteration burn-in period.

To determine the client status of the root kinase, she used MCMC implemented in BayesTraits v2. She used the kinase tree and client status coding described above. She fossilized the root as either A or B for the whole tree. She found appropriate levels of acceptance with the option rateDev set to 1. She ran the chain for 10 million iterations, and compared the two different client states of the root by calculating the Bayes Factor based on

the harmonic means (twice the difference between the two harmonic means of the model likelihood) (Pagel et al. 2004).

Examining non-kinase Hsp90 interactors

J. Lachowiec and I analyzed the non-kinase interactors. J. Lachowiec removed all genes that were found in more than one functional category: E3 ligases, transcription factor, or kinase. The client status as defined in Taipale *et al.* (2012) was used for each gene, with E3 ligases and transcription factors (TFs) categorized as "not significant interactor", removed from further analyses. She used pairwise dN/dS values with mouse as an outgroup to examine the dN/dS among the Hsp90 interactors and non-interactors. Neither TFs nor E3 ligases are monophyletic. She subdivided the TFs into phylogenetically related families (Vaquerizas et al. 2009) and compared the dN/dS between Hsp90 interactors and non-interactors within each family and in aggregate. Subdividing E3 ligases into phylogenetic groups based on sequence similarities has not been previously conducted, so she subdivided E3 ligases based on domain presence (Kelch, WD40 from (Taipale et al. 2012) or RING, U box, HECT, F box, SOCS box, BTB, DDB1-like, ZnF A20 from (Albrecht et al. 2008)) and compared the dN/dS between Hsp90 interactors and non-interactors within each group. Because no differences in dN/dS were found between Hsp90 interactors and non-interactors within groups defined by individual domains, she also clustered E3 ligases based on presence or absence of many domains simultaneously. I identified the E3 domains using PFAM version 26. Clustering was completed using hierarchical clustering in R with the function *heatmap*.

3.3 RESULTS

3.3.1 Hsp90 client status contributes to dN/dS

Taipale *et al*. (Taipale et al. 2012) previously reported that strong Hsp90 kinase clients acquire more non-synonymous mutations than non-client kinases, using pairwise dN values from human and mouse. We first extended these analyses by examining evolutionary rate (pairwise dN/dS between human and mouse) for non-, weak, and strong clients based on

strength of kinase interaction with Hsp90 as defined by (Taipale et al. 2012). Evolutionary rate analysis (dN/dS) is the gold standard for assessing the types of selection potentially acting on proteins. The dN/dS for strong Hsp90 clients was significantly greater than the dN/dS for non-clients (**Figure 3.1a**, p = 0.0004, Wilcoxon rank sum test); no significant dN/dS difference was observed between weak clients and non-clients. Both of these findings were consistent with the prior findings observed with dN alone. Second, we found that dN/dS values of non-clients are significantly different from the combined dN/dS values of strong and weak clients (p =0.01221, Wilcoxon rank sum test), which was not previously addressed (Taipale et al. 2012). None of the kinases showed dN/dS > 1, which would indicate positive selection. Rather, the Hsp90 client kinases showed relaxed purifying selection compared to non-clients, consistent with previous observations in plants and yeast (Lachowiec et al. 2013).

The observed greater dN/dS of Hsp90 clients may be an indirect consequence of other factors that strongly influence evolutionary rate, specifically gene expression levels (Pál et al. 2001; Wall et al. 2005) or protein interaction degree (Fraser et al. 2002). If genes encoding Hsp90 clients are expressed at lower levels (Taipale et al. 2012) or if client proteins are less connected in protein interaction networks, the correlation between Hsp90 client status and dN/dS can be explained without invoking Hsp90 as a contributor to evolutionary rate. We examined whether gene expression level or PPI connectivity explains the observed contribution of Hsp90 client status to kinase dN/dS, by conducting a linear regression analysis. To this end, we calculated the mean and maximum gene expression of client and non-client kinases across RNA-seq experiments for 11 different human primary tissue samples (Castle et al. 2010) and used PPI values from Taipale *et al*. (2012). To account for the phylogenetic relatedness among these kinases, we used the kinase tree to calculate phylogenetically independent contrasts (PIC) (Felsenstein 1985) to correct for non-independence of all variables under study: kinase dN/dS, Hsp90 client status, gene expression levels, and PPI values. For this analysis, Hsp90 client status for each kinase was measured by its quantitative interaction score with the chaperone (Hsp90 interaction score, HIS) (Taipale et al. 2012). We considered whether expression breadth, measured as the

number of tissues in which a kinase is expressed, was correlated with dN/dS. In this data set, however, 187 of the tested 210 kinases were expressed in all tissues, precluding analysis.

Table 3.1 dN, dS, and dN/dS Values for Human kinases

	dN	95% CI for dN	dS	95% CI for dS	dN/dS	95% CI for dN/dS
Nonclients	**0.043**	**(0.0018, 0.1497)**	**0.605**	**(0.2704, 1.0417)**	**0.069**	**(0.0043, 0.2239)**
All clients	**0.055**	**(0.0025, 0.1811)**	**0.619**	**(0.3106, 1.2571)**	**0.088**	**(0.0053, 0.3148)**
Weak clients	**0.047**	**(0.0011, 0.1780)**	**0.629**	**(0.2823, 1.2331)**	**0.073**	**(0.0036, 0.2475)**
Strong clients	**0.063**	**(0.0050, 0.1813)**	**0.609**	**(0.3486, 1.1787)**	**0.104**	**(0.0091, 0.3176)**

Table 3.2 Hsp90 interaction score is positively associated with dN/dS.

Model	Regression coefficient		R^2
PIC[1] dN/dS ~ PIC Hsp90 interaction score	0.007462	***	0.06447
PIC Hsp90 interaction score ~PIC expression maximum[2]	-0.003507		-0.004182
PIC dN/dS ~ PIC expression maximum	-0.0006343	*	0.02056
PIC Hsp90 interaction score ~ PIC PPI	-0.35586	****	0.08779
PIC dN/dS ~ PIC PPI	-0.008626	***	0.06256

[1]phylogenetic independent contrasts
[2]maximum expression across 11 tissues
p-values:*0.05, **0.01, ***0.001, ****0.0001

We used linear regression to understand the relationships among the PIC for each pair of variables. The PIC of kinase evolutionary rates and HSP90 Interaction Score (HIS) were positively associated (p = 0.0001, **Table 3.2**), consistent with our phylogenetically naïve, categorical analysis (**Figure 3.1a**). As expected the PIC of kinase evolutionary rate and gene expression levels were negatively associated (p = 0.02, **Table 3.2**) (Pál et al. 2001); the PIC of kinase evolutionary rate and PPI connectivity were also negatively associated (Fraser et al.

2002) (p = 0.0001, **Table 3.2**). Notably, we did not observe a significant association between HIS and expression contrasts (p = 0.7, **Table 3.2**), suggesting that there are no systematic differences in gene expression levels that may drive the observed differences in evolutionary rate between Hsp90 client and non-client kinases. However, we found a negative correlation between PPI connectivity and HIS contrasts (**Table 3.2**), raising the possibility that the greater evolutionary rate of Hsp90 kinase clients derives from fewer protein interactions.

To further disentangle the contributions of Hsp90 interaction score, gene expression levels, and PPI connectivity to kinase evolutionary rate, we also calculated partial correlations among the contrasts for the four variables. We found that HIS was positively correlated with kinase evolutionary rate when controlling for both gene expression levels and PPI connectivity (r = 0.19, p = 0.003, Pearson correlation, **Table 3.3**). The relative contribution of HIS and gene expression levels to kinase evolutionary rate was comparable when controlling for the respective other variables; the relative contribution of PPI connectivity was marginal (**Table 3.3**). To quantify the combined explanatory power of Hsp90 interaction score, gene expression levels, and PPI connectivity to kinase dN/dS, we tested the extent to which the PIC of all three variables combined could explain dN/dS PIC. This model explained only a small component of the dN/dS PIC (R^2= 0.11).

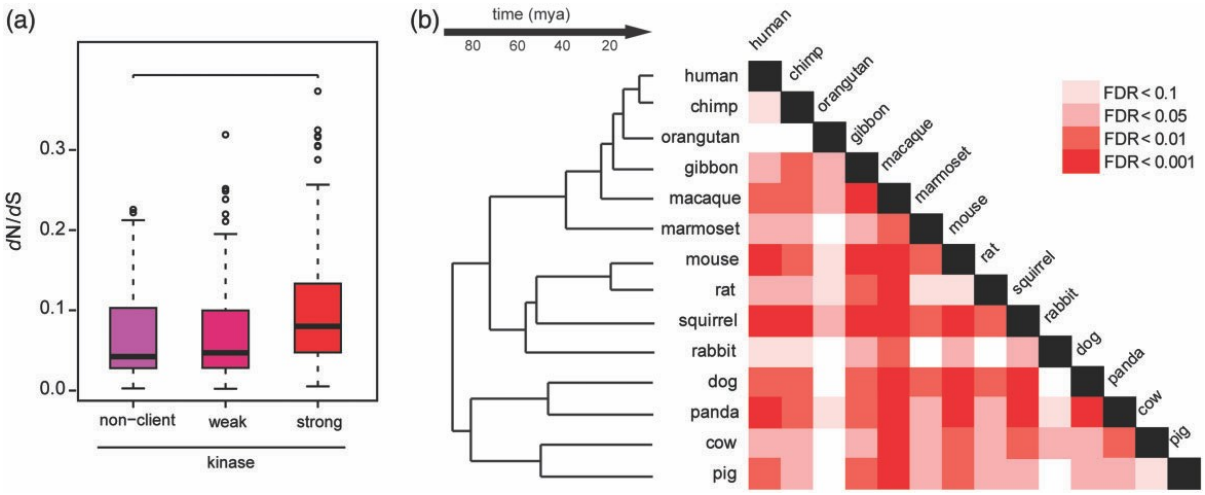


Figure 3.1. Hsp90 client and nonclient kinases differ significantly in evolutionary rate. (a) The difference for human-mouse pairwise dN/dS for client and nonclients kinases is driven by strong clients (p = 0.0004211, Wilcoxon rank-sum test). (b) Significant differences between strong and nonclient kinase dN/dS were observed across mammals in a core set of kinases found in all mammalian species examined (Wilcoxon rank-sum test). Shades of red indicate false discovery rate (FDR).

In summary, we found that 1) Hsp90 client status is positively associated with kinase dN/dS, 2) this association appears to be independent of gene expression levels or PPI connectivity, and 3) the strength of association is comparable for Hsp90 client status and gene expression levels. Notably, none of the tested factors alone or in combination explained a large proportion of the observed variation in kinase evolutionary rate.

Table 3.3. Hsp90 is correlated with dN/dS when controlling for PPI and expression[a]

PIC HIS	PIC HIS	PIC dN/dS	PIC PPI[b]
PIC dN/dS	0.197**		
PIC PPI	-0.105	-0.0958	
PIC max expr[c]	-0.139*	-0.212**	0.027063

[a]Partial correlation are shown between two variables, controlling for the other two, Pearson correlation.

[b]PPI, number of PPIs.

[c]max expr, maximum expression across 11 tissues.

P-values: *0.05, **0.01.

3.3.2 Hsp90-associated effects on kinase evolutionary rate increase with increasing divergence

Thus far, we have analyzed kinase dN/dS values that were calculated from human and mouse, species which diverged approximately 75Ma (Chinwalla 2002). As mice are known to be a long-branching clade (Wu and Li 1985; Rat-Genome-Sequencing-Project-Consortium 2004), we evaluated whether Hsp90's effect on kinase dN/dS values was lineage- or time-dependent. To do so, we determined the core set of kinases conserved across all tested mammals and obtained pairwise dN/dS values for the one-to-one orthologs of strong human Hsp90 kinase clients and nonclients. In total, 70 kinases were conserved, comprised of 28 nonclients, 20 weak clients, and 22 strong clients. Consistent with the mouse-human comparisons, strong Hsp90 clients evolved faster than non-clients across these species, including several monophyletic lineages such as Primates, Glires, and Laurasitheria (**Figure 3.1b**).

Across all comparisons, the average ratio of strong client dN/dS to nonclient dN/dS was 1.63±0.25 (standard deviation), indicating that Hsp90 clients undergo relaxed selection compared with nonclients (**Figure S3.1a**). Our results suggest that Hsp90 client kinases tend to accumulate nonsynonymous variation compared with nonclient kinases regardless of species compared and time of divergence.

Several previous studies have demonstrated that Hsp90-dependent standing variation is common in yeast, plant, fly, and fish populations (Yeyati et al. 2007; Queitsch et al. 2002; Sangster et al. 2008a; Rutherford and Lindquist 1998; Jarosz and Lindquist 2010). Using genetic or pharmaceutical perturbation of Hsp90, these studies found significantly increased trait heritability upon Hsp90 inhibition, presumably due to the many identified Hsp90-responsive loci. For the vast majority of these Hsp90-dependent loci the identity of the underlying polymorphism (*i.e.* the gene or regulatory region affected) remains unknown and hence evolutionary rate analysis of these loci has yet to be conducted. However, a detailed genetic study found that an Hsp90 client showed increased accumulation of non-synonymous variation compared to its non-client paralog within *A. thaliana* across divergent strains (Lachowiec et al, 2013).

I therefore hypothesized that Hsp90's effect on kinase evolutionary rate should be detectable among humans. To test this hypothesis, I took advantage of the thousands of sequenced human genomes and examined nucleotide diversity of kinase clients and non-clients as a measure of accumulated genetic variation. Specifically, I assessed nucleotide diversity for kinase clients and non-clients in approximately 6500 individuals (Exome Variant Server), accounting for relatedness and using HIS contrasts as measure of Hsp90 client status.

Indeed, the PIC of kinase HIS and nucleotide diversity were positively correlated ($p = 0.04965$, $R^2 = 0.025$, **Table S3.1**), suggesting that Hsp90 client kinases harbor greater genetic variation than non-client kinases in humans. This positive correlation suggests that purifying selection is reduced among Hsp90 clients compared with nonclients, resulting in higher levels of slightly deleterious mutations. To determine the type of genetic variation that Hsp90 buffered, we examined nucleotide diversity in nonsynonymous and synonymous sites separately. The correlation between the PIC of client score and PIC of nucleotide diversity was stronger when only considering nonsynonymous sites ($r = 0.23$, $p = 0.003$). In contrast, there was no significant correlation when only considering synonymous sites ($r = 0.03$, $p = 0.69$). This result is consistent with the notion that it is Hsp90's function in protein folding that enables greater dN/dS of its clients rather than the rise of new mutations, epigenetic phenomena, or other, yet unidentified mechanism. To further explore this aspect, I examined whether Hsp90 clients tolerated potentially more deleterious variation across humans, using both genomic evolutionary rate profiling (GERP) scores (Cooper et al. 2005) and Polymorphism Phenotyping (PolyPhen) scores (Adzhubei et al. 2010) as measures of the predicted phenotypic effect of individual single nucleotide variants (SNVs). GERP ascertains the degree of past purifying selection on a site through examination of rejected nucleotide substitutions across homologous sequences (Cooper et al. 2005), whereas PolyPhen categorizes non-synonymous SNVs by likelihood of damaging protein structure and function (Adzhubei et al. 2010). Although there was no significant correlation between GERP and HIS contrasts ($r = 0.02$, $P = 0.7593$, Pearson correlation, **Table S3.1**), PolyPhen and HIS contrasts were positively correlated ($r = 0.21$, $p = 0.00987$, Pearson correlation, **Table S3.1**).

To summarize, Hsp90 kinase clients appeared to harbor more genetic variation and more damaging mutations than non-clients across many human individuals.

3.3.3 Hsp90 client status tends to be gained and maintained

Our observation that Hsp90's effect on kinase evolutionary rate increases with divergence time suggests that client status is stable over long time periods. This interpretation contrasts with experimental findings that mutating a single amino acid residue can suffice to dramatically alter a protein's dependence on Hsp90 (Citri et al. 2006). Similarly, paralogs or otherwise related proteins that differ only by a few amino acids can differ in Hsp90 client status (Citri et al. 2006; Lachowiec et al. 2013), suggesting that Hsp90 client status can be highly dynamic. To resolve this apparent contradiction, we analyzed the distribution of Hsp90 client status among gene duplicates in the kinase superfamily. Gene duplicates tended to share Hsp90 client status (**Figure 3.2a**, χ^2-test, $p = 1.837e^{-05}$) indicating that client status has phylogenetic signal.

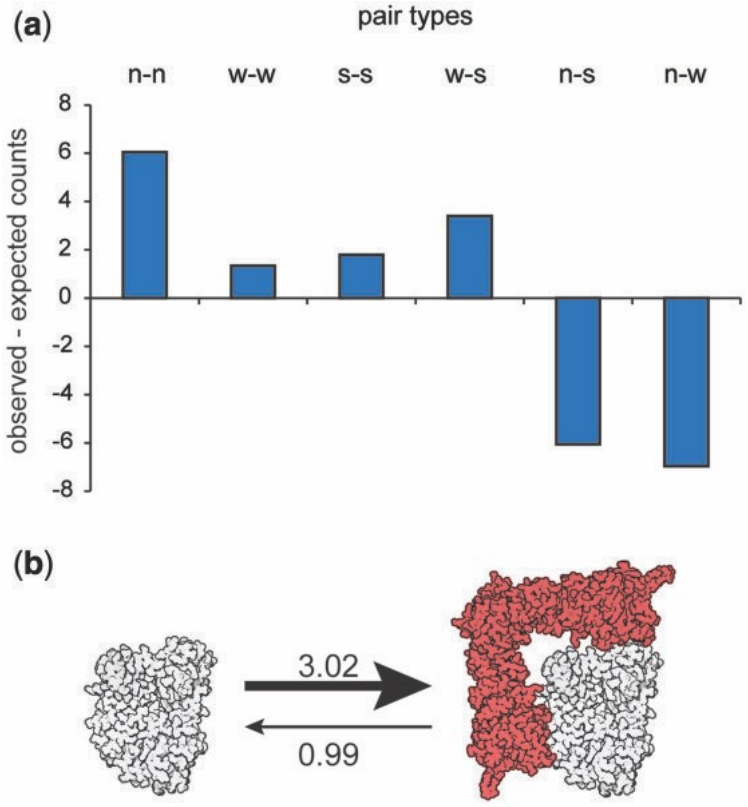


Figure 3.2. Hsp90 client status dynamics across the kinome. (a) Using the three client classes defined by Taipale et al. (2012), we categorized pairs of duplicated genes into six classes. We find that gene duplicates share client status (χ^2-test, P = $1.837e^{-05}$) more often than expected by chance. The number of expected pairs was calculated based on a random distribution of clients across the tree. n-n stands for pair of non-Hsp90 clients; w-w, weak Hsp90 clients; s-s, strong clients; w-s, pair of weak and strong clients; n-s, pair of a strong and a non-client; n-w, pair of a weak and a non-client. (b) Kinases (white) are three times more likely to gain Hsp90 client status (red) than lose it based on ML and MCMC estimates of state transition rates using Bayes Traits.

This analysis only considered genes with one paralog, eliminating one-third of the available kinome data. Therefore, we explicitly tested whether there was phylogenetic signal in client status patterns across the entire kinase phylogeny using Pagel's λ (Pagel 1999).

Pagel's λ is a statistic that tests whether phylogeny correctly predicts the patterns of covariance among species on a given trait. Here, we estimated this statistic to test whether kinase phylogeny predicts the patterns of covariance among kinases with regard to Hsp90 client status. Assuming that kinases were either Hsp90 clients or non-clients, we compared a model in which λ was computed using the known kinase phylogeny (Manning et al. 2002) (λ = 0.74) to a model assuming no phylogenetic signal (λ = 0, star phylogeny). The former model including phylogenetic signal was a better fit for the observed pattern of client status across the kinase phylogeny than the model without phylogenetic signal ($2\delta = 8.94$, df = 1, p = 0.00279). We conclude that client status is not randomly distributed, and that more closely related kinases tend to share client status across the whole kinase superfamily.

Having established that client status tends to be shared among related kinases, we next explored the evolutionary paths by which client status is acquired and changes through time.

To do so, we estimated the transition rates between client and non-client states along the kinase phylogeny, using maximum likelihood (ML) implemented in the software BayesTraits (Pagel et al. 2004; Pagel 1999). We found that the rate of the transition from non-client to client along the kinase phylogeny was 3.02, compared to the reverse rate of 0.99. Restricting these rates to equal values significantly worsened the model fit ($2\delta = 8.066$, df = 1, p=0.0045), hence we conclude that kinases are more likely to become Hsp90 clients than to lose client status (**Figure 3.2b**). To examine whether these ML-derived rates were robust, we also used the Markov Chain Monte Carlo (MCMC) analysis implemented in BayesTraits to estimate the rates of gain and loss of client status. The rates of transition for non-client to client (3.11) and client to non-client (0.98) indicated that rate estimates were robust (**Figure 3.2b**). The tendency of non-client kinases to become clients and of client kinases to remain clients agrees with previous suggestions that kinase dependence on Hsp90 may be 'addictive' and may contribute to the greatly increased sensitivity of cancer cells to Hsp90 inhibitors (Workman et al. 2007).

If protein kinases indeed become 'addicted' to being Hsp90 clients, one may expect the ancestral kinase to be a non-client. To infer the client status of the ancestral kinase, we compared a tree with a non-client kinase root to a tree with a client kinase root using MCMC. We found strong support for the tree with a non-client root (Bayes factor ~9.1). Both Hsp90 and kinases were present in the eukaryotic common ancestor (Bogumil et al. 2014; Manning and Hunter 2009); however, our results would suggest that they likely did not interact. Although, Hsp90 client predictions and confirmed clients in the prokaryote *E. coli* include kinases (Press et al. 2013), these prokaryotic protein kinases belong to a different kinase family and differ in structure (Manning and Hunter 2009). Taken together, our results are consistent with a scenario in which early kinases evolved independently of Hsp90 with subsequent multiple independent gains of Hsp90 client status, increasing Hsp90's effect on kinase evolution and adding to the overall chaperone dependence of kinases for their function.

3.4 DISCUSSION

Although previous studies showed that Hsp90-dependent standing variation is common in natural populations (Rutherford and Lindquist 1998; Queitsch et al. 2002; Salathia et al. 2007; Sangster et al. 2008a; Rohner et al. 2013; Jarosz and Lindquist 2010), the chaperone's impact on protein evolution, especially in comparison to established factors such as gene expression, has remained unknown. Here, we present evidence for a direct role of Hsp90 in kinase evolution and compare its impact to gene expression levels and PPI connectivity.

We find that interaction with Hsp90 is associated with higher dN/dS for kinases and that Hsp90 client status appears to contribute to kinase evolutionary rate independently of gene expression and PPI connectivity. Each of these factors contributed to a similar degree to kinase evolutionary rate; combined they explained only about 10% of kinase dN/dS variation. Adding additional variables may decrease the amount of unexplained variation, but many variables are known to covary. For example, Bloom and Adami (2003) have argued that PPI connectivity is confounded by protein abundance, with highly abundant proteins

engaging in a larger number of interactions (Bloom and Adami 2003). Protein abundance arises as a combination of gene expression levels, and rates of translation and protein degradation. Others have argued that the dominant role of gene expression in driving evolutionary rate is at least in part due to selection for translational robustness, which encompasses selection for increased translational accuracy by optimizing codon usage, and selection to increase the number of proteins that fold properly despite mistranslation (Drummond et al. 2005). Previously, optimized codon usage, as measured by codon adaptation index, was observed for stochastic clients of GroEL/ES (Warnecke and Hurst 2010). However, we did not observe an association between CAI and Hsp90 client score ($P = 0.34$, $R^2 = 3.6e^{-03}$), possibly reflecting that Hsp90 does not act as a cotranslational chaperone (Taipale et al. 2010) like GroEL/ES (Ying et al. 2006). Further, large-scale studies of translation efficiency and its correlation with CAI in human have yielded contradictory results, with some studies failing to find significant correlations (Chamary et al. 2006; Waldman et al. 2010). We also note that the kinase Hsp90 clients analyzed here are generally less abundant in steady-state protein levels (Taipale et al. 2012), presumably due to their enhanced structural liability. Nevertheless, by using partial contrasts, we find evidence for a contribution of Hsp90 to kinase dN/dS independent of expression level and PPI degree.

As suggested by prior studies of Hsp90-dependent variation within diverse populations of plants, yeast, fly and fish (Rutherford and Lindquist 1998; Yeyati et al. 2007; Queitsch et al. 2002; Sangster et al. 2008a; Sangster et al. 2008b; Jarosz and Lindquist 2010), we found support for Hsp90-dependent genetic variation across humans. Both nucleotide diversity and PolyPhen scores were significantly correlated with kinase HISs; yet effect sizes were modest. No significant correlations were observed for GERP scores. The GERP score is a position-specific estimate of evolutionary constraint using maximum likelihood evolutionary rate estimation (Cooper et al. 2005). High evolutionary constraint is often interpreted as functional relevance. As we do not expect Hsp90-dependent variation to reside in highly constrained sites, it is not surprising that GERP scores for Hsp90 kinase clients and non-clients did not differ significantly.

In agreement with the modest signal of increased nucleotide diversity within human, we observed an increased evolutionary rate for Hsp90 client kinases across divergent mammalian lineages, consistent with Hsp90 allowing for greater accumulation of non-synonymous changes in the genes encoding its clients. In fact, Hsp90's effect on kinase evolutionary rate increases when considering species that are more distantly diverged. The greater accumulation of non-synonymous changes may allow client kinases to explore a wider sequence space and potentially acquire novel functions at a faster rate.

Taipale et al. (2012) also determined HISs for a large number of transcription factors (TFs) and E3 ligases. In contrast to our findings for client and non-client kinases, TF and E3 ligase HIS were not significantly associated with their evolutionary rates (**Supplementary Text**, **Figure S3.2**). Unlike the kinases analyzed, the transcription factors are not monophyletic (Vaquerizas et al. 2009) and hence may differ more in other factors influencing evolutionary rate, such as protein structure and stability. As for the E3 ligases, others have suggested that Hsp90 works in concert with E3 ligases to promote proteasome-dependent degradation rather than chaperoning these enzymes (Murata et al. 2001; McClellan et al. 2005; Morishima et al. 2008; Ehrlich et al. 2009; Taipale et al. 2010) . Hsp90 works in concert with many other proteins that enable and modify its function and generate client specificity, such as diverse Hsp70s, co-chaperones, and immunophilins (Taipale et al. 2010). Although these "collaborating" proteins physically interact with Hsp90, they do so in a sequence-or domain-specific manner, and Hsp90 does not facilitate their folding (Murata et al. 2001; McClellan et al. 2005; Morishima et al. 2008; Ehrlich et al. 2009; Taipale et al. 2010); hence they are unlikely to experience relaxed selection due to this interaction. Following this line of reasoning, we previously excluded known Hsp90 co-chaperones and Hsp70s from evolutionary rate analyses (Lachowiec et al. 2013).

Previously, we found that Hsp90 clients in plants and yeast showed significantly greater evolutionary rates than their non-client paralogs (Lachowiec et al. 2013). The kinome data set only contained five pairs of strong kinase clients and non-clients. Although, in four of these pairs the strong kinase client showed greater dN/dS than its non-client paralog, this difference was not significant, presumably due to small sample size (n=5, 95% confidence

interval 0.6-12.6, p = 0.125, one-sample Wilcoxon test, testing the deviation from the expected ratio of client/non-client dN/dS of 1). We did, however, detect a significant trend for duplicate genes such that paralog pairs that contained at least one strong Hsp90 client showed significantly greater divergence than pairs that did not (**Figure S3.3**). We speculate that interaction with Hsp90 allows clients to tolerate slightly deleterious, but non-lethal mutations without losing function, thereby facilitating the emergence of novel functions over time (Lachowiec et al. 2013).

Beyond paralog pairs, we further explored the relationship of Hsp90 client status and kinase evolutionary rate among kinase families. The effect of Hsp90 client status on dN/dS was consistent across all kinase groups (**Figure S3.4a**) with greater dN/dS observed for Hsp90 client kinases compared to non-client kinases. The TK and TKL families were enriched for strong Hsp90 clients (**Figure S3.4b**), and possibly in part due to this enrichment, both families showed the highest evolutionary rates among the tested kinase families (**Figure S3.4c**). TK and TKL kinase families are evolutionarily young and have been implicated in the rise of multicellularity (Lim and Pawson 2010). In contrast, of all families, the CAMK family was most depleted for strong Hsp90 clients. Unlike in all other kinase families, the few strong CAMK clients did not evolve faster than the weak clients (**Figure S3.5**). As almost a third of the CAMK family genes are pseudogenes (not included in this analysis), contrasting with only about one-fifth of kinase pseudogenes overall (Manning et al. 2002), we speculate that the "missing" strong CAMK clients have become pseudogenes.

These family-specific observations and our finding that kinases tend to acquire and maintain Hsp90 client status prompt our speculation that Hsp90 may play a complex role in the birth and death of kinases. The most common outcome after gene duplication is pseudogenization of one copy (Nei and Roychoudhury 1973). Acquiring Hsp90 client status likely leads to instant sub-functionalization due to the temperature sensitivity of clients and may facilitate gene copy maintenance (Lachowiec et al. 2013). The observed greater accumulation of non-synonymous variation in genes encoding Hsp90 client may also facilitate neo-functionalization. At the same time, the greater accumulation of more harmful variation may predispose genes encoding clients to the pseudogene fate. In other words,

acquiring Hsp90 client status may be akin to a delayed death sentence on a long evolutionary time scale.

In fact, the suggested "Hsp90 addiction" of kinases (Workman et al. 2007), especially in cancer cells with their mutated oncogenic kinases, is reminiscent of a recent argument by Fernandez and Lynch (Fernandez and Lynch 2011). These authors attributed the increasing complexity of protein interaction networks from bacteria to human to compensation for the decreased stability of proteins over evolutionary time. They posited that through drift proteins would be exposed to destabilizing mutations and hence become susceptible to aggregation and malfunction. Interaction in homo-and hetero protein complexes will then compensate for the stability deficits of proteins in higher organisms (Fernandez and Lynch 2011). Of course, chaperones also prevent aggregation and stabilize proteins (Taipale et al. 2010). The small evolutionary snapshot of the kinome with its tendency to acquire and maintain Hsp90 client status fits well within this framework of thought.

3.5 SUPPLEMENTAL INFORMATION

Table S3.1. Hsp90 client kinases harbor more and more harmful genetic variation across humans.

	p-value	Rho (Pearson's)
PIC[1] nucleotide diversity ~ PIC HIS[2]	0.04965	0.157
PIC PolyPhen probable ~ PIC HIS	0.0099	0.206
PIC GERP ≥ 5 ~ PIC HIS	0.7593	0.02473

[1]Phylogenetic independent contrasts

[2]HIS- Hsp90 interaction score

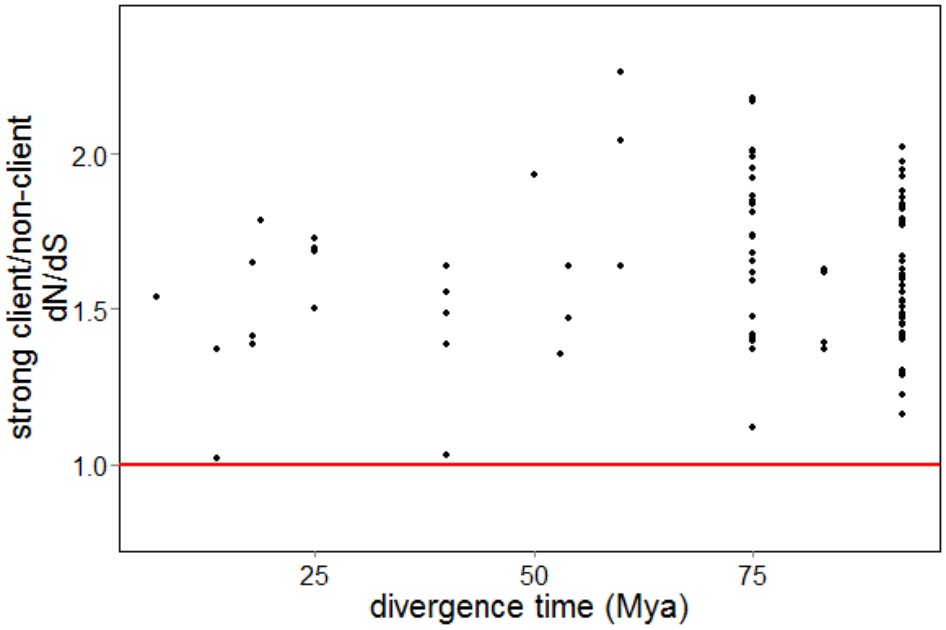


Figure S3.1. The average strong client kinase dN/dS was always greater than the non-client kinase dN/dS across mammals, regardless of species used for pairwise dN/dS comparison (ratios were always greater than 1, red line).

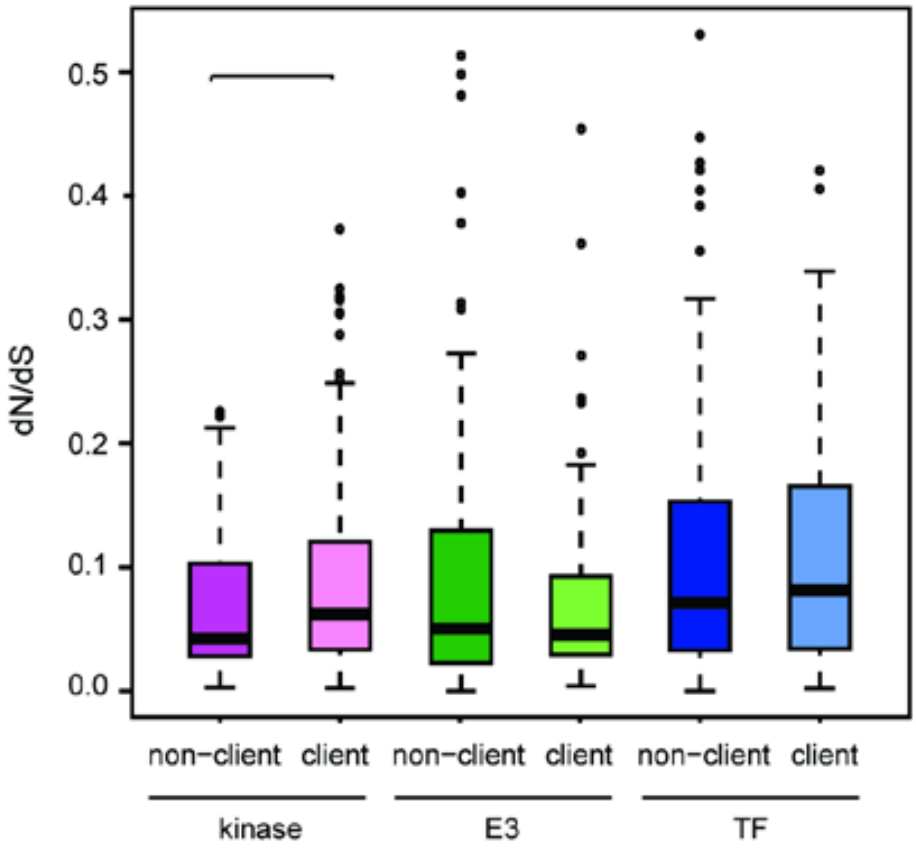


Figure S3.2. E3 and TF clients do not evolve faster than their respective non-clients. In contrast to kinases ($p = 0.01221$, Wilcoxon rank-sum test), transcription factor and E3 ligase clients do not show significantly greater dN/dS than transcription factors and E3 ligase non-clients.

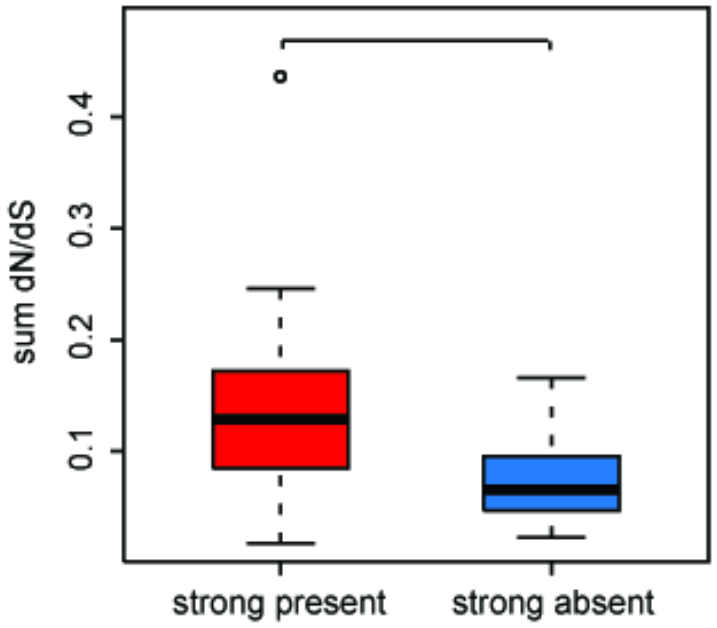


Figure S3.3. Hsp90 client status is associated with divergence of gene duplicates. Divergence was measured as the sum dN/dS for each gene duplicate pair. Pairs of gene duplicates that encode at least one strong Hsp90 client (n=19) diverge faster than pairs without an Hsp90 client (n=27) ($p = 0.001183$, Wilcoxon rank-sum test).

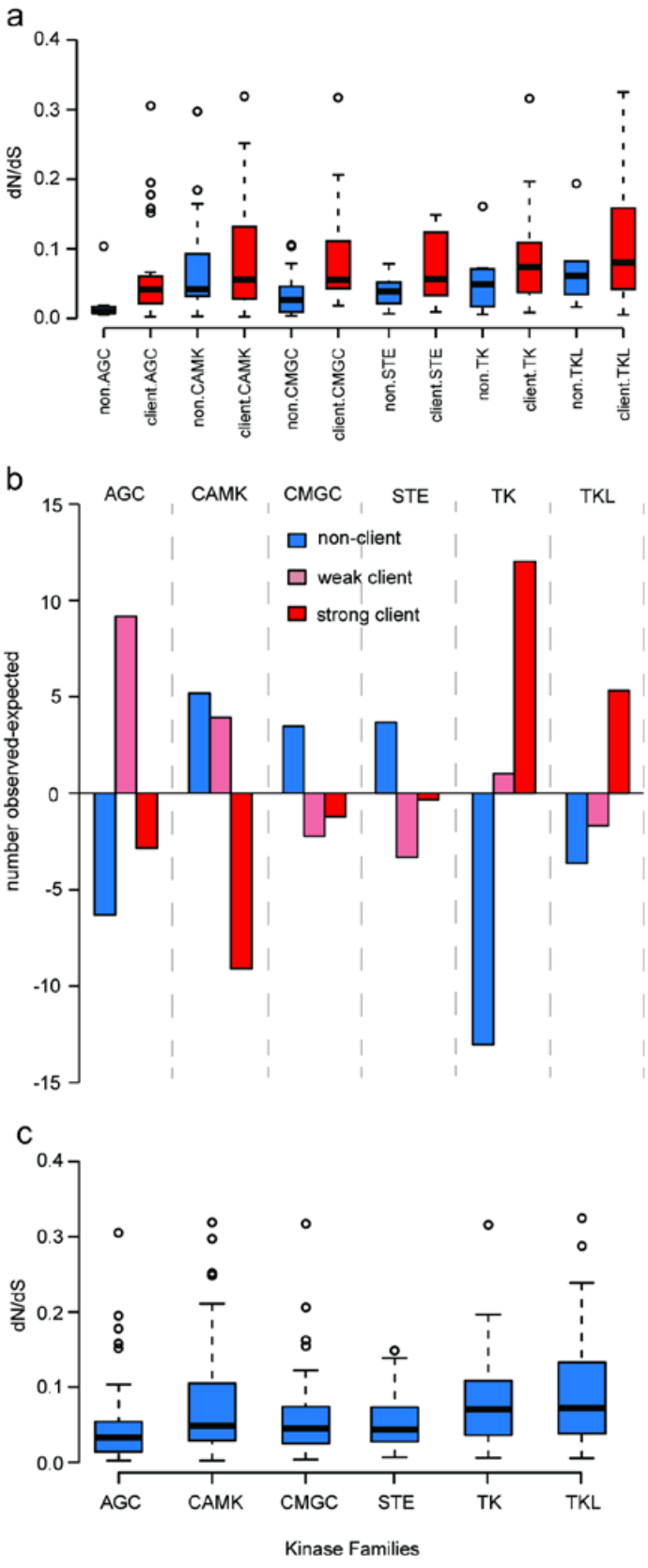

a
dN/dS
0.4
0.3
0.2
0.1
0.0
non.AGC
client.AGC
non.CAMK
client.CAMK
non.CMGC
client.CMGC
non.STE
client.STE
non.TK
client.TK
non.TKL
client.TKL
b
AGC
CAMK
CMGC
STE
TK
TKL
non-client
weak client
strong client
number observed-expected
15
10
5
0
-5
-10
-15
c
dN/dS
0.4
0.3
0.2
0.1
0.0
AGC
CAMK
CMGC
STE
TK
TKL
Kinase Families

Figure S3.4. Hsp90's effect on dN/dS is observed in all kinase groups.

a) In each kinase group, the Hsp90 clients tended to show greater dN/dS than non-clients. b) The TK and TKL groups were enriched for strong Hsp90 clients; the CAMK group was most depleted for strong Hsp90 clients. The expected number of non-clients, weak clients, and strong clients was calculated by assuming an equal distribution of all client states across all kinase groups. c) The TK and TKL families show the largest median values of dN/dS, albeit their dN/dS did not significantly differ from the other kinase groups. AGC stands for proteinase A, G and C families; CAMK, Calmodulin/Calcium regulated kinases; CMGC, group of kinases that comprise the families CDK, MAPK, GSK3, and CLK; STE, yeast homologs of STE7, STE11, and STE20; TK, tyrosine kinase; TKL, tyrosine kinase-like.

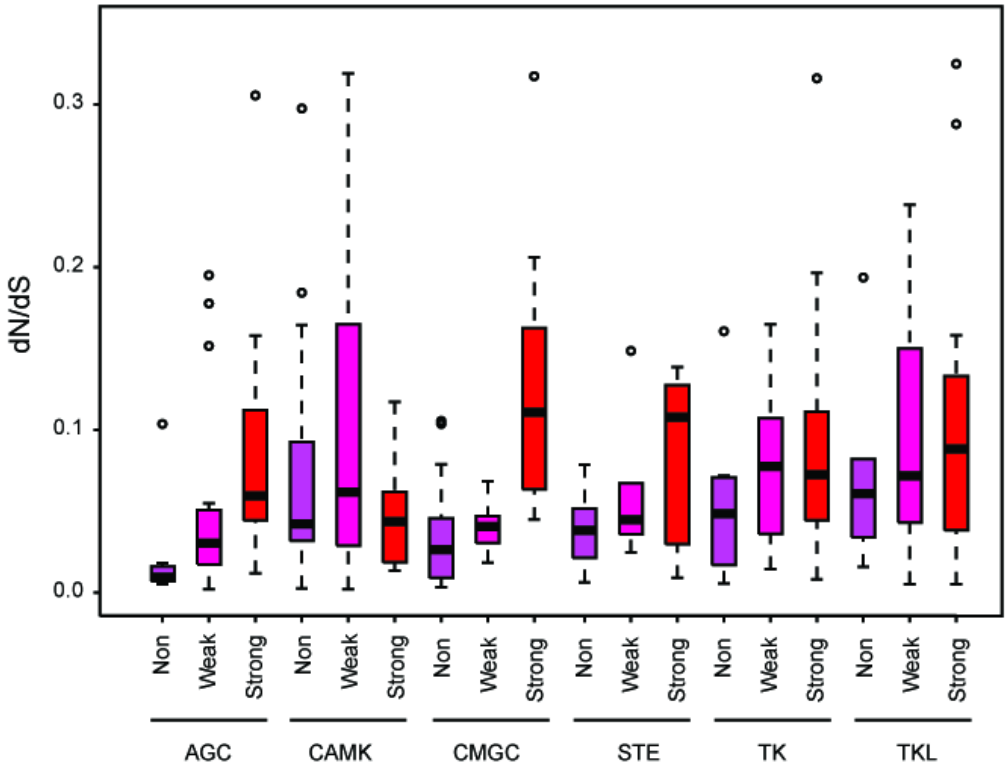


Figure S5. Family-specific differences for the association of Hsp90 client status and dN/dS. In general, strong kinase clients (red) tended to have greater dN/dS than weak kinase clients (pink), and weak kinase clients (purple) tended to have greater dN/dS than kinase non-clients. The TK and CAMK groups did not follow this pattern, with weak kinase clients tending to have greater dN/dS, presumably due the broad distribution of dN/dS for weak kinases in these groups. AGC stands for proteinase A, G and C families; CAMK, Calmodulin/Calcium regulated kinases; CMGC, group of kinases that comprise the families CDK, MAPK, GSK3, and CLK; STE, yeast homologs of STE7, STE11, and STE20; TK, tyrosine kinase; TKL, tyrosine kinase-like.

Supplementary Text

Interaction with Hsp90 is associated with faster evolutionary rates in clients, but not co-chaperones or otherwise collaborating proteins.

We failed to detect an effect of Hsp90 client status on the evolutionary rate of transcription factors and E3 ligases (Figure S2). In contrast to the tested kinases, which evolved from a common ancestor, transcription factors and E3 ligases are not monophyletic (Li et al. 2008b; Vaquerizas et al. 2009) and hence likely differ greater in other features influencing evolutionary rate. Hence, we subdivided the tested transcription factors (Vaquerizas et al. 2009) and E3 ligases (Li et al. 2008b) into phylogenetically related groups for further tests.

For transcription factors, this subdivision approach was hampered by the extremely low number of transcription factors that interact with Hsp90 (58 out of 843 tested) (Taipale et al. 2012). Comparing evolutionary rates of clients and non-clients in phylogenetically related groups was not feasible (average group size 18.4, average number of clients 1.8, and non-clients 16.7). As a side note, the interaction of TF with HSP90 was tested using a small number of cochaperones. It is therefore possible that many more HSP90-TF interactions could be detected if the assay is repeated with different HSP90 cochaperones.

For E3 ligases, we used previously described subdivisions and tested for association of Hsp90 client status and dN/dS. Specifically, we divided the E3 ligases into RING finger domain-containing (RNF) and non-RNF proteins (Li et al. 2008b). We found no significant effect of Hsp90 client status on RNF and non-RNF E3 ligases. Similarly, testing only E3 ligases with a Kelch fold (Taipale et al. 2012) yielded no evidence that Hsp90 interactors evolved faster than non-interactors. E3 ligases also contain small domains with roles in recognition of ubiquitinated substrates. We attempted to cluster the tested E3s by these domains as identified in Pfam, using several different thresholds for domain identity. Again, we found no significant differences in dN/dS associated with Hsp90 interaction in these groups.

This persistent failure in light of the clear and consistent signal in kinases prompted us to further explore the literature on Hsp90 and E3 ligase interaction. In contrast to TFs and kinases, which require Hsp90's assistance to reach their mature fold, E3 ligases may interact physically with Hsp90 for another reason. Indeed, some E3 ligases have been shown to collaborate with Hsp90 in the degradation of other proteins, rather than being chaperoned by Hsp90 (Murata et al. 2001; Giannini and Bijlmakers 2004; Morishima et al. 2008). If E3 ligases generally are Hsp90-collaborating proteins, we would not expect an Hsp90-associated effect on E3 ligase dN/dS, akin to prior observations with other Hsp90-associated proteins such as Hsp70 and various Hsp90 co-chaperones, all of which are rather highly conserved across eukaryotes.

Chapter 4

THE MECHANISTIC UNDERPINNINGS OF AN *AGO1*-MEDIATED, ENVIRONMENTALLY-DEPENDENT, AND STOCHASTIC PHENOTYPE[3]

This chapter was included in the thesis because I discovered the phenomenon and did the initial characterization analyses. This includes recording the proportion of wild type, *ago1*, and *hsp90* mutant plants that harbored lesions, drafting the initial hypothesis, and formulating the experiments to test whether the lesions could be the result of a hypersensitive response, or to the loss of tissue specification. I also contributed to the data and pictures for figures 4.1, 4.2, and 4.3. G. Alex Mason performed most of the experiments to test the initial hypothesis, and then took the project and found the mechanism through which the lesions arise in *ago1-27* plants. The first person "we", and the "our" pronoun indicates the work of G. Alex Mason. My work is stated with the "I" person, and the "my" pronoun.

Abstract

ARGONAUTE 1 (AGO1) is an essential regulator of organismal development through its role in microRNA processing. In plants, continuous morphogenesis and exposure to environmental insults have invested AGO1 with additional roles in response to environmental stimuli. I observed that hypomorphic *ago1* mutant *Arabidopsis thaliana* seedlings occasionally develop lesions on cotyledons, but that these lesions become much more common upon exposure to unfiltered light. The histology of these lesions showed that they are necrotic growths resulting from the hypersensitive immune response, indicating that this pathway is dysregulated in *ago1* mutants in an environmentally-responsive fashion. We found that immune response genes associated with the salicylic acid, ethylene, and jasmonate pathways are all up-regulated in *ago1* seedlings, and that treatment with jasmonate or ethylene are sufficient to induce lesions in the place of unfiltered light. However, salicylic acid treatment actually repressed lesion formation. We were able to mostly rescue the lesion

[3] Most of this chapter is part of a manuscript that is going to be re-submitted to *Plant Physiology*. The list of authors is G A Mason, T Lemus, and C Queitsch.

formation phenotype by crossing *ago1* with a mutant of *coi1*, a component of the jasmonate signaling pathway, indicating that jasmonate is the causative agent. Lastly, we investigated the genetic interaction between AGO1 and HSP90, which has been shown to regulate the jasmonate pathway. Unexpectedly, lesions responded synergistically, indicating that *ago1* lesion formation is not only a function of HSP90 dysregulation, or vice versa, but that the responses are related. We expect that lesions form because UV radiation acts as an elicitor of the jasmonate response, which is normally repressed by AGO1-dependent mechanisms. AGO1's precise regulatory role in repressing jasmonate signaling will need to be elucidated in further studies. Altogether, our results describe a novel role for AGO1 as an integrator of environmental stress signals controlling the hypersensitive response.

4.1 INTRODUCTION

As an essential component of the microRNA (miRNA) pathway, ARGONAUTE 1 (AGO1) is a key regulator of development and environmental response in *Arabidopsis thaliana*. As a member of the clade of 10 AGOs in Arabidopsis (Vaucheret 2008), AGO1 is an ancient, highly conserved, and essential protein. In plants, AGO1 is the main effector for miRNA activity. AGO1 uses 21-25nt long miRNAs that are transcribed from specific miRNA genes to negatively regulate gene expression.

MiRNAs and AGO1 are critical for buffering abiotic environmental stresses. Expression levels of miRNAs respond to heat shock, changes in salinity, cold, drought, oxidative stress, microbial defense, abscisic acid treatment, ultra-violet light exposure, and hypoxia (Sunkar et al., 2012). In fact, many single miRNA families are responsive to multiple environmental stresses, indicating that one miRNA family can mediate the proper responses to many environmental perturbations (Sunkar et al., 2012). MiRNAs allow for plants to orchestrate the proper response to environmental cues, such as defense.

The hypersensitive response (HR) in plants is the rapid accumulation of dead cells in response to pathogen invasion. The cell death associated with the HR is regulated genetically and is thought to be a form of programmed cell death (Heath 2000; Lam et al. 2001). HR

occurs as a plant's means of restricting pathogen growth, and perception of a pathogen triggers the expression of R-genes, which allow for the plant to signal for changes in ion fluxes, which are then followed by production of reactive oxygen species (Heath, 2000). Generation of an oxidative burst is one of the first signals used by plants when they encounter a pathogen (Orozco-Cárdenas et al. 2001). The breakdown of cellular components eventually leads to localized cell death. The phytohormones salicylic acid, jasmonate, and ethylene are of particular importance in HR and are critical for defense against pathogens and/or herbivores. These hormone pathways interact at several points to coordinate the HR.

Salicylic acid (SA) plays a central role in resistance against biotrophic pathogens, such as *Pseudomonas syringae* and turnip crinkle virus (Koornneef and Pieterse 2008). SA biosynthesis is a pre-cursor to production of ROS, and the perception of SA can lead to hypersensitive-response-mediated localized cell death. Specific to AGO1, SA is required for response to biotic environmental perturbations, and there is evidence linking miRNAs to disease resistance pathways. MiR393 is involved in the response to bacterial flg22; mir393 then represses several auxin related response genes (Zhang et al. 2011). This repression allows the plant to upregulate the necessary responses for resistance to the microbial threat.

Jasmonates (JAs) are small signaling molecules that regulate a wide breadth of processes, from development (e.g. fertility) to response to wounding by herbivores (Balbi and Devoto 2008; Gfeller et al. 2010; Koo and Howe 2009; Memelink 2009; Turner et al. 2002; Wasternack 2007). JA is known to modulate ROS production (Overmyer et al. 2003), specifically hydrogen peroxide (Orozco-Cárdenas et al. 2001). JA and SA are often considered to behave antagonistically, however, expression studies indicate that crosstalk between the two hormones may be more complex (Robert-Seilaniantz et al. 2011b).

Ethylene (ET) is best known as a positive regulator of ripening and senescence (Dahl and Baldwin 2007). In addition to these roles, ET can act both positively and negatively on plant immunity. In *Arabidopsis*, ET potentiates some branches of SA-responsive gene expression (Lawton et al. 1994; De Vos et al. 2006). However, ET also globally represses SA-responsive genes through the action of the transcription factors ETHYLENE

INSENSITIVE 3 and ETHYLENE INSENSITIVE-3 LIKE 1 (Chen et al. 2009) ET also interacts with JA to repress or potentiate different defense responses. When acting together, JA and ET synergistically promote defense against necrotrophic fungal pathogens through the expression of the ERF-regulated transcription factor branch of the JA pathway. However, ET also represses JA-herbivory responses (Ballaré 2011) by antagonizing the transcription factor MYC-regulated branch of the JA pathway (Anderson et al. 2004; Lorenzo et al. 2004; Pré et al. 2008).

Here we provide evidence that AGO1 is required for environmental buffering and integration of environmental signals in canonical defense pathways. I show that *ago1* mutant plants develop lesions on their embryonic leaves (cotyledons) when grown under full (white) light conditions on sucrose plates. We demonstrate lesions are a form of localized cell death and form due to an up-regulated hypersensitive defense response. We show that gene expression markers of SA, JA, and JA/ET signaling pathways are significantly up-regulated in white light grown *ago1* cotyledons, and that this response requires the JA-dependent pathway.

In the plant *Arabidopsis thaliana*, *HSP90* also serves as regulator of robustness to environmental perturbations. HSP90 is known to be required for buffering temperature changes, resistance to herbivory, and potentiating immune responses (Liu et al. 2004c; Lu et al. 2003; Sangster et al. 2007a). HSP90 and AGO1 directly interact as shown by co-purification studies of both proteins (Landthaler et al. 2008; Liu et al. 2004a; Maniataki and Mourelatos 2005) and additional studies (Iki et al. 2010) indicate that HSP90 is specifically required for the loading of small RNAs into RNA-induced silencing complexes, suggesting that AGO1 is an HSP90 client. We therefore predicted that HSP90 reduction might influence how AGO1 is used to interpret a plant's surroundings. In this paper, we show that AGO1 and HSP90 may act together non-additively in defense signaling and response to the environment. Further, I also show that decreased *HSP90* levels increases the penetrance of the *ago1* mutation.

Our results highlight that AGO1, a protein that can buffer cryptic genetic variation in *Arabidopsis thaliana* (see next chapter), is critical for integrating environmental cues properly, much like HSP90. We describe how *ago1* is sensitized to its surroundings and we also provide evidence that translational inhibition is critical for responding to environmental cues, as evidenced by the hypomorphic *ago1-27* mutant.

4.2 MATERIALS AND METHODS

Plant Growth Conditions:

I grew *ago1-27, ago1-25,* HSP90 A1 RNAi, HSP90 C1 RNAi genotypes for the initial characterization of *ago1* mutants, when I counted the number of lesions. Then, G. Alex Mason grew subsequent seed batches for her experiments.

Columbia-0 (Col-0) was used as wild-type reference. *ago1-27, ago1-25,* HSP90 A1 RNAi, HSP90 C1 RNAi, *coi1-1* and *ein2-1* were in the Col-0 background (table S1). For experiments, seeds were sterilized with ethanol and plated onto 1× Murashige and Skoog (MS) basal salt medium supplemented with 1× MS vitamins, 1% (wt/vol) sucrose, 0.05% MES (wt/vol), and 0.24% (wt/vol) phytagel. After stratification in the dark at 4 °C for 5 d, plates were transferred to an incubator (Conviron) that was set to long day (LD) (16L: 8D at 22 °C: 20 °C), with light supplied at 100 μmol·m−2·s−1 by cool-white fluorescent bulbs; for control plates (filtered light), a long-pass yellow filter that blocks 454 nm light was placed in front of the bulbs.

Genotyping and gene expression analyses:

RNA Extractions and Quantitative Real-Time PCR. Alex G. Mason and I collected the tissue and RNA for the initial experiments in which we both tested whether the lesions were cells undergoing cell death or dedifferentiation. For all the other experiments, Alex Mason collected the tissue and extracted the RNA.

Total RNA was extracted from 30-mg frozen tissue using the SV Total RNA Isolation System (Promega). Subsequently, 2 µg of RNA were subjected to DNase treatment using Ambion Turbo DNA-free Kit (Applied Biosystems). RNA integrity and purity were checked with an Agilent Bioanalyzer using the RNA 6000 Nano Kit (Agilent Technologies). For cDNA synthesis, 200 ng of DNase-treated RNA was reverse-transcribed using the Transcriptor First Strand cDNA Synthesis Kit (Roche) and oligo dT primers. Transcript abundance was determined by real-time quantitative PCR using the LightCycler 480 system (Roche), with LightCycler 480 SYBR Green I Master (Roche) and the following PCR conditions: 5 min at 95 °C, followed by 35 cycles of 15 s at 95 °C, 20 s at 55 °C, and 20s at 72 °C. To ensure that PCR products were unique, a melting curve analysis was performed after the amplification. UBC21 expression (At5g25760) was used as a reference. All quantitative RTPCR primers were designed with Primer3Plus software (http://www.primer3plus.com/). Relative quantification was determined with the ΔΔCT Method (Fryer et al. 2011). Error was calculated using standard error of the mean across at least three biological replicates.

DNA Extractions and genotyping. I designed the CAPS markers to genotype *ago1-27* plants. Alex Mason extracted the DNA and did the genotyping.
Arabidopsis genomic DNA was extracted for genotyping analysis by PCR using the CTAB method (Weigel and Glazebrook 2002). The *ago1-27* and *coi1-1* alleles were genotyped by using the primers in Table S4. PCR conditions for *ago1-27* genotyping is as follows: 5' at 94 °C, followed by 35 cycles at 30 s at 94 °C, 30 s at 55 °C, 1 min at 72 °C. PCR product was then digested with Bsp1286I at 37 °C, which cuts wild-type sequence. PCR conditions for *coi1-1* genotyping is as follows: 5' at 94 °C, followed by 35 cycles at 30 s at 94 °C, 30 s at 55.7 °C, 1 min at 72 °C. PCR product was then digested with XcmI, which cuts wild-type sequence.

Hormone response assays. Methyl Jasmonate: Alex Mason performed these experiments. Seedlings were grown as previously described in **Plant Growth Conditions** under un-filtered light conditions. 5-day post germination seedlings in groups of 36 were sprayed with

1 ml of water, 1ml of 10 μM methyl jasmonate solution (Sigma Aldrich), or 1 ml of 50 μM methyl jasmonate solution. Seedlings were then scored for lesions at day 10.

Salicylic acid: Seedlings were grown as previously described in **Plant Growth Conditions** under un-filtered light conditions. 5-day post germination seedlings in groups of 36 were sprayed with 1 ml of 1xPBTx, 1ml of 10 uM salicylic acid in 1xPBTx, or 1 ml of 100 uM salicylic acid in 1xPBTx. Seedlings were then scored for lesions at day 10.

ACC (ethylene precursor): Seedlings were grown as previously described in **Plant Growth Conditions** under un-filtered light conditions. 5-day post germination seedlings in groups of 36 were sprayed with 1 ml of water, 1ml of 1 uM ACC, 1ml of 10 uM ACC or 1 ml of 20 uM salicylic acid in water. Seedlings were then scored for lesions at day 10.

Hypersensitive response assays: 1) Trypan Blue staining: I performed the initial staining, then, Alex Mason repeated the experiments. Fresh affected and unaffected cotyledons were collected and then treated with Trypan blue staining solution. Trypan blue staining solution was prepared as follows: add Trypan blue to lactophenol (10 ml lactic acid, 10 ml glycerol, 10 ml phenol, 10 ml water to a concentration 2.5 mg/ml. Then, add two volumes ethanol to Trypan blue-lactophenol solution. Staining solution was then added to fresh tissue and boiled for one minute. Tissue was then allowed to site for 5-15 minutes at room temperature. Stain solution was removed by washing with lactophenol and two volumes ethanol. Samples were then mounted on glass slides and photographed with Nikon whatever. Images were edited using Adobe Photoshop CS3.

2) DAB: Alex Mason performed these experiments. Fresh affected and unaffected cotyledons were collected and then treated with DAB (1mg/ml, pH 3.8) staining solution for 8 hours. Tissue was then de-stained with 100% ethanol. Samples were then mounted on glass slides and photographed with Nikon whatever. Images were edited using Adobe Photoshop CS3.

3) DAPI staining info: Alex Mason performed these experiments. Fresh affected and unaffected cotyledons were collected and then fixed in 3% (v/v) glutaraledhyde overnight. Tissue was then stained with DAPI solution (0.1% v/v in 5% DMSO, 1% (v/v) Tween-20). Tissue was then washed with PBST at least twice and mounted with 50% (v/v)

glycerol/PBST on glass slides. Samples were then mounted on glass slides and photographed. Images were edited using Adobe Photoshop CS3 and ImageJ.

Phenotyping assays: 1) Methyl jasmonate screening: Alex Mason did the phenotyping assays. 10-day old seedlings were sprayed with 50-100 uM methyl jasmonate solution in water. Seedlings were incubated for an additional three days and were then screened for response to jasmonate (i.e. anthocyanin accumulation).

2) ACC screening: Alex Mason did the screening. For experiments, seeds were sterilized with ethanol and plated onto 1× Murashige and Skoog (MS) basal salt medium supplemented with 1× MS vitamins, 1% (wt/vol) sucrose, 0.05% MES (wt/vol), and 0.24% (wt/vol) phytagel. Experimental plates were made to 10 uM ACC. After stratification in the dark at 4 °C for 5 d, plates were transferred to an incubator (Conviron) that was set to long day (LD) (16L:8D at 22 °C:20 °C), with light supplied at 100 μmol·m−2·s−1 by cool-white fluorescent bulbs; for control plates (filtered light) for three hours and then wrapped in aluminum foil. Plates were incubated for 7 days and then scored for triple response.

Statistical analysis: Alex Mason did the statistical analysis. All statistical analyses were performed using the software R. Comparisons of seedling lesion frequencies were performed with the χ^2-test. Seedling lesion formation was modeled with binomial regression as a response to genotype, hormone dose when relevant, replicate, individual, and the interaction between genotype and hormone.

4.1 RESULTS

4.1.1 Argonaute 1 mutant seedlings develop lesions in full spectrum light conditions.

In addition to previously characterized *ago1* phenotypes (Baumberger and Baulcombe 2005; Bohmert et al. 1998; Morel et al. 2002; Yang et al. 2006), such as hyponastic cotyledons and developmental delay (**Figure 4.1**), I observed that *ago1* mutant seedlings frequently produced white, often raised, lesions, which resided primarily on cotyledons (**Figure 4.1**). Wild-type seedlings (Col-0) showed these lesions at low frequency in both filtered light [yellow long-pass filters, blocking <454 nm, (Stasinopoulos and Hangarter 1990)] and full spectrum light (**Figure 4.1**). I grew an allelic series of *ago1* mutations in both light conditions and scored the frequency of affected seedlings (**Figure 4.1**). I scored affected seedlings rather than frequency of lesions because lesions often fused with each other, confounding frequency measures. For the weaker *ago1-27* and *ago1-46* mutant seedlings (Morel et al., 2002; Smith et al., 2009), I observed that the frequency of affected seedlings increased significantly in full spectrum light (**Figure 4.1**). In contrast, *ago1-25* mutant seedlings were significantly more affected than wild-type regardless of light conditions (**Figure 4.1**). Although the frequency of affected *ago1-25* mutant seedlings increased in response to full spectrum light, this increase was not significant compared to filtered light conditions. Alex Mason and I attribute this decreased environmental responsiveness to the fact that *ago1-25* is the most severe *ago1*-mutant tested. We selected the *ago1-27* mutant for our subsequent investigation of the molecular underpinnings of the lesion phenotype for two reasons. First, we were intrigued by its strong dependency on growth conditions (Generalized Linear Mixed Model, GLMM, GLMM, $p< 1x10^{-4}$). Second, as *ago1-27* is a weaker mutant allele than *ago1-25* (Morel et al., 2002), *ago1-27* mutant seedlings exhibited fewer and less severe unrelated developmental abnormalities.

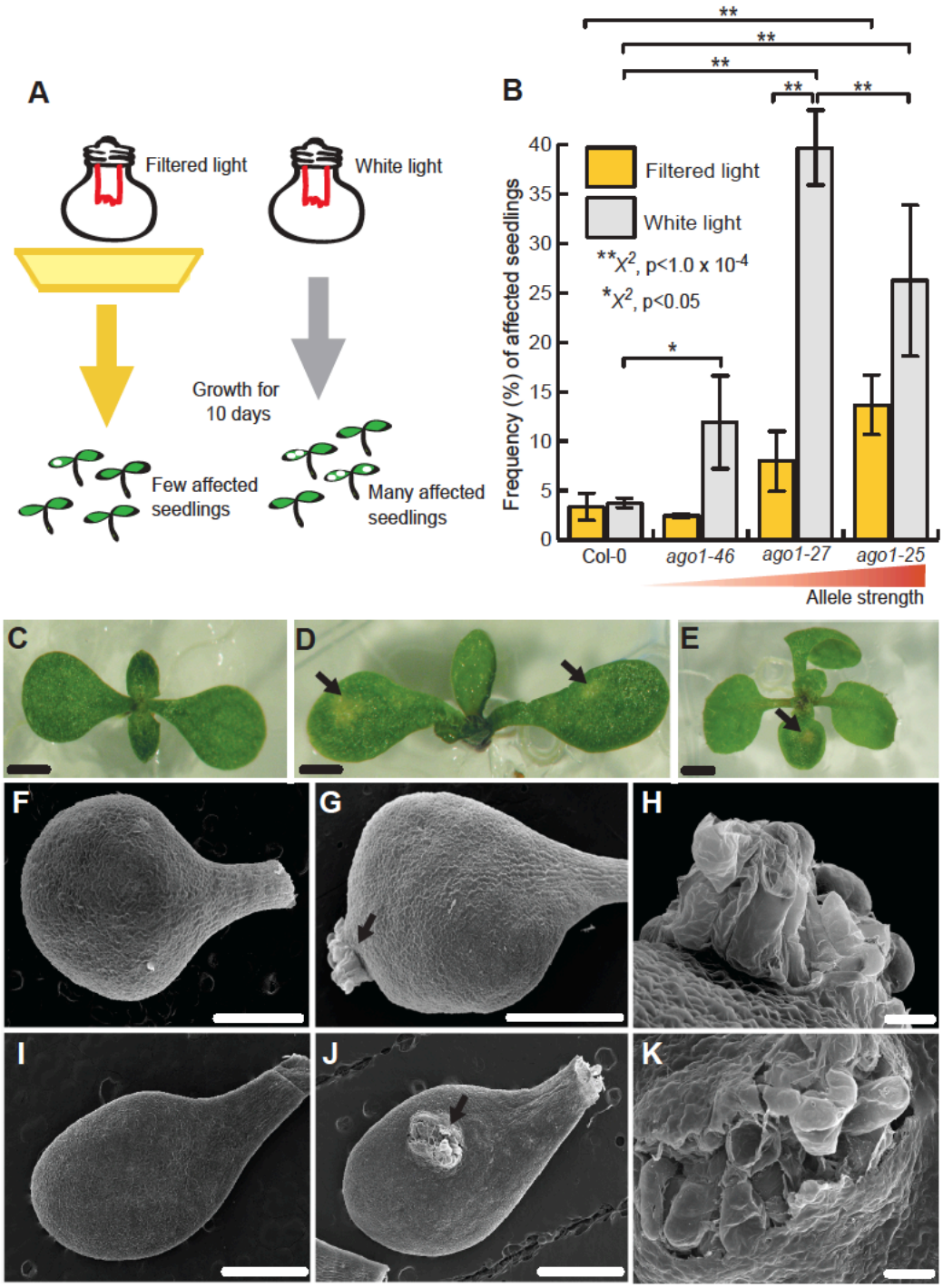


Figure 4.1 Hypomorphic argonaute 1 mutants produce seemingly stochastic lesions. (A) Experimental set-up: when grown in yellow, filtered light conditions for 10 days, few *ago1* seedlings with lesions on cotyledons were observed. When grown in full spectrum light, many more *ago1* seedlings developed lesions. (B) Under full spectrum light, *ago1-46*, *ago1-*

27, and *ago1-25* mutants showed significantly increased frequency of affected seedlings compared to wild-type. *ago1-27* mutants showed significantly more affected individuals compared to filtered light conditions; a similar, albeit non-significant tendency was observed for seedlings carrying either weaker (*ago1-46*) or stronger *ago1* allele (*ago1-25*). Error bars are standard error of mean across at least two replicates with n = 25-72 seedlings per replicate. Statistical significance was calculated using the χ2-test of significance (**p< 1x10-4, *p<0.05) (for all reported p-values, refer to **Table S4.4**). The difference in responsiveness between *ago1-25* and *ago1-27* was modeled using GLMM, testing for interaction between genotype and treatment with replicate and number of seedlings as random variables (p<1x10-4). Graded red triangle represents increasing strength of *ago1* alleles tested. (C-D) Lesions in wild-type and *ago1-27* mutant seedlings show distinct morphology. (C) Full spectrum light-grown, unaffected 10-day-old *ago1-27* seedling. (D) Full spectrum light-grown 10-day-old *ago1-27* seedling with lesions denoted by arrows. (E) Full spectrum light-grown, 10-day-old affected wild-type seedling with lesion denoted by arrow. (C-E) Scale bar: 1 mm. (F-K) Scanning electron micrographs of 10-day-old cotyledons grown under full spectrum light. (F) Unaffected wild-type cotyledon. (G) Wild-type cotyledon with lesion (denoted by arrow). (H) Magnified view of wild-type lesion from (G). (I) Unaffected *ago1-27* cotyledon. (J) *ago1-27* cotyledon with lesion tissue (denoted by arrow). (K) Magnified view of *ago1-27* lesion from (I). (F, G,I, and J) Scale bare: 1 mm. (H, K) Scale bare: 500 microns.

The raised, transparent morphology of the *ago1*-dependent lesions superficially resembled undifferentiated tissue, *i.e.* ectopic callus tissue (Boutilier et al. 2002). We tested this hypothesis by analyzing expression of genes implicated in callus formation (Sugimoto et al., 2010) and by testing the propensity of lesions to facilitate callus formation (Weigel and Glazebrook, 2002). However, genes implicated in callus formation and shoot apical meristem identity showed wild-type expression levels in *ago1-27* cotyledons (**Figure S4.1**). To test callus formation, Alex Mason and I dissected lesions from both *ago1-27* and wild-type and tested callus formation on hormone-containing media. At low hormone dose, *ago1-27* lesion tissue performed worse than wild-type or *ago1-27* unaffected tissue; at high hormone dose,

no difference was observed (**Figure S4.2**). We conclude that the observed lesions do not represent undifferentiated, callus-like tissue.

To explore the origin and identity of the lesion tissue, we employed scanning electron microscopy (SEM) (**Figure 4.1**). SEM revealed that both wild-type and *ago1-27* lesions 'erupt' from beneath the epidermis, suggesting that mesophyll cells are primarily affected (Figure 1). Cells within lesions appeared to be ruptured, suggesting that at least some lesions represented dead or dying tissue (van Doorn et al., 2011). To exclude the possibility that ruptured cells were an artifact of SEM sample preparation, Alex Mason and I used vital dye staining to validate our findings. Localized cell death is induced in plants as part of the hypersensitive response (HR) upon pathogen attacks. We hypothesized that the *ago1* lesions were a result of stochastically occurring, aberrant HR that was enhanced in *ago1* seedlings by exposure to full spectrum light. Localized cell death is examined by staining with vital dyes such as Trypan blue, which stains dead and dying tissue dark blue (Weigel and Glazebrook, 2002). We performed Trypan blue staining of unaffected (**Figure 4.2**) and affected cotyledon tissue (**Figure 4.2**). Cotyledons exhibiting lesions retained the dye, with lesions exhibiting dark blue staining indicative of ruptured cell walls (**Figure 4.2**). In fact, even unaffected *ago1-27* cotyledons showed deeper staining than wild-type cotyledons, suggesting increased susceptibility to cell wall rupture. To confirm these results, we performed 3,3'-diaminobenzidine (DAB) staining (**Figure 4.3**). DAB staining indicates the presence of reactive oxygen species, specifically hydrogen peroxide, which is a reliable marker of HR-mediated localized cell death (Weigel and Glazebrook, 2002). Indeed, we observed DAB staining coinciding with lesions (**Figure 4.3**). Based on both Trypan blue and DAB stainings, we conclude that *ago1-27* cotyledons produce lesions that resemble HR-mediated localized cell death.

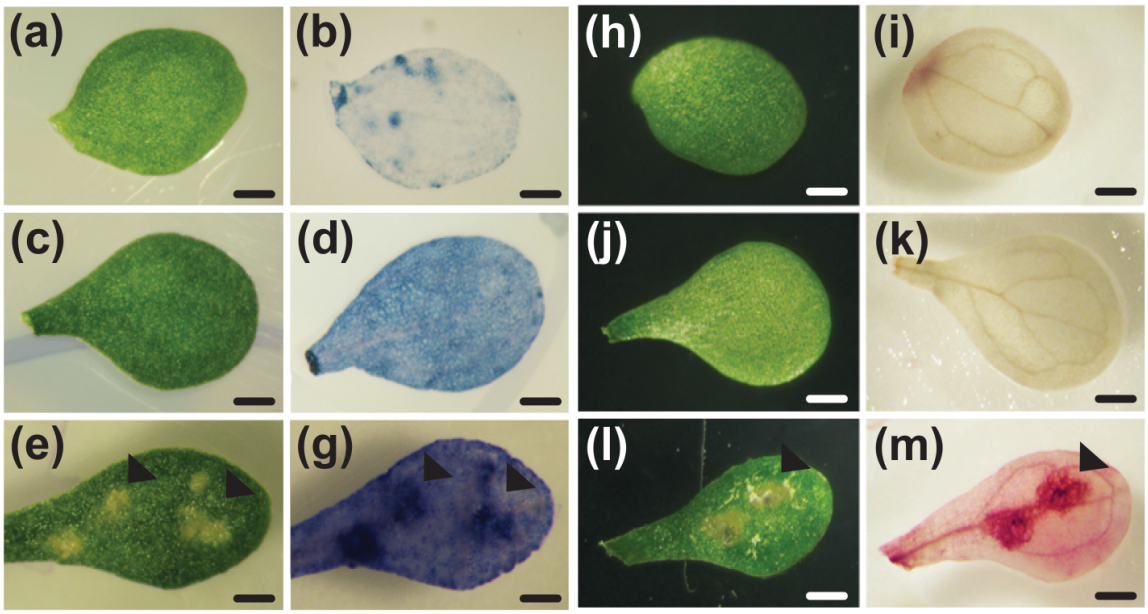


Figure 4.2 Vital staining of unaffected and affected tissue from wild-type and ago1-27 cotyledons. (A, G) Unaffected wild-type cotyledons prior to staining. (B) Unaffected wild-type cotyledon after Trypan blue staining showed little dye accumulation. (G) Unaffected wild-type cotyledon after DAB staining showed little dye polymerization.
(C, I) Unaffected *ago1-27* cotyledons prior to staining. (D) Unaffected *ago1-27* cotyledon after Trypan blue staining showed some dye accumulation. (L) Unaffected *ago1-27* cotyledon after DAB staining showed no dye polymerization. (E,K) Affected *ago1-27* cotyledons prior to staining. (F) Affected *ago1-27* cotyledon after Trypan blue staining showed strong accumulation of dye in lesions (arrows). (L) Affected *ago1-27* cotyledon after DAB staining showed strong polymerization of the dye in lesions (arrows). Scale bar: 1 mm.

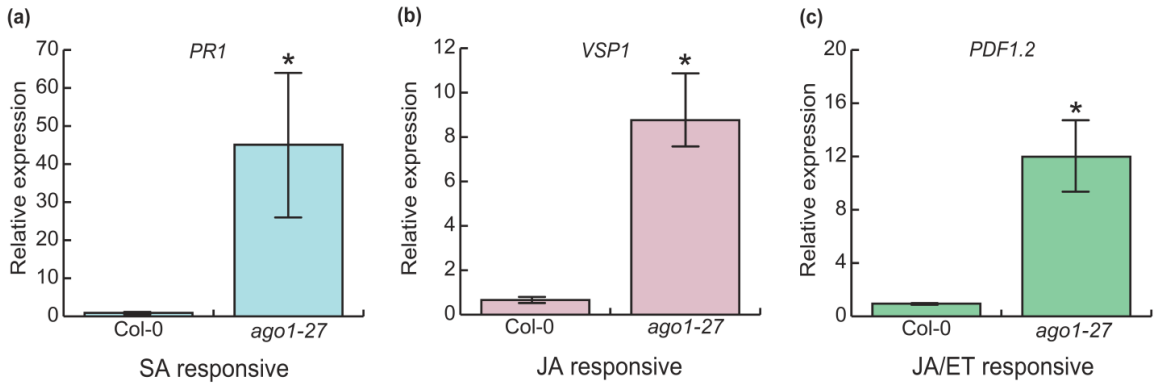


Figure 4.3 Marker genes of all three canonical defense hormone pathways were upregulated in *ago1-27* cotyledons. (A) PR1 (salicylic acid (SA) responsive); (B) VSP1 (jasmonate (JA) responsive); (C) PDF1.2 (jasmonate/ethylene (JA/ET) responsive) expression were dramatically upregulated in 10-day-old *ago1-27* cotyledons in full spectrum light. Results are from qRT-PCR; all assayed genes were normalized to UBQ10 expression. Error bars indicate standard error of the mean across three to four replicates of cotyledons harvested from 18 pooled seedlings.

4.1.2 ago1-27 cotyledons overexpress canonical SA, JA, and JA/ET signaling markers in full spectrum light

Phytohormones mediate all major aspects of a plant's response to environmental perturbations, including defense against biotic stresses. The three phytohormones salicylic acid (SA), ethylene (ET), and jasmonate (JA) are the canonical defense hormones. As previous results have linked light conditions (UV-B) with defense hormone response (A.-H.-Mackerness et al., 1999; Mazza et al., 1999), we set out to examine whether the SA, JA, or JA/ET pathways were responsible for the lesions on *ago1-27* cotyledons. We measured expression of defense markers for SA, JA, or JA/ET in 10-day-old cotyledons from wild-type and *ago1-27* mutant seedling grown in full spectrum light. Specifically, we assessed expression of *PATHOGENESIS-RELATED GENE 1(PR1)*, *VEGETATIVE STORAGE PROTEIN 1(VSP1)*, and *PLANT DEFENSIN 1.2 (PDF1.2)*, which are upregulated in

defense-triggered SA, JA, and JA/ET-signaling, respectively (Benedetti et al., 1995; Boter et al., 2004; Lorenzo et al., 2004; van Loon et al., 2006; Mur et al., 2006; Memelink, 2009). All three marker genes were greatly upregulated in *ago1-27* cotyledons compared to wild-type (**Figure 4.3**); hence, more detailed analysis was required to identify the involved pathway(s).

ago1 plants are susceptible to virus infection, implying a direct link between SA-mediated defense and AGO1 activity (Morel et al., 2002; (Ruiz-Ferrer and Voinnet, 2009). Because the SA marker gene *PR1* was most strongly upregulated in *ago1-27* cotyledons, we first wanted to examine the expression of additional genes in SA signaling. Although there are no known miRNA targets in the SA signaling pathway, altered light conditions change the expression of many miRNAs (Zhou et al. 2007). Based on previous studies, we selected three SA-signaling genes. The first*, NONEXPRESSOR OF PR GENES 1 (NPR1),* was selected because it contributes to activation of *PR*-genes after SA perception (Mukhtar et al. 2009). *NPR1* is also a key modulator in SA-JA crosstalk (Ballaré, 2011). The second, *SUPPRESSOR OF NPR1-1* (*SNC1),* was previously reported to be upregulated in *ago1-36* mutants (Yi and Richards, 2007). The third, *ENHANCED DISEASE SUSCEPTIBILITY 1* (*EDS1*), is a component of R gene-mediated disease resistance in *Arabidopsis thaliana* (Falk et al., 1999)*;* EDS1 is also required for production of SA in some conditions involving pathogen challenge (Rustérucci et al., 2001). All three genes were significantly upregulated in *ago1-27* cotyledons (**Figure 4.4**).

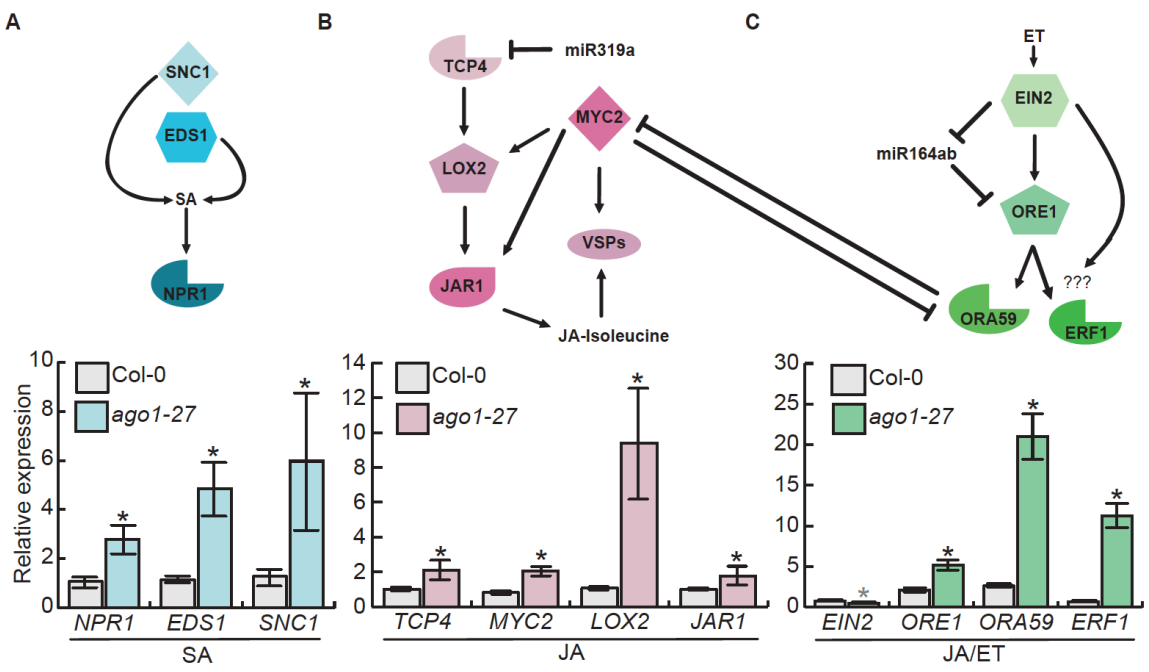


Figure 4.4 All three defense hormone pathways, including miRNA target genes, were upregulated in *ago1-27* cotyledons. (A) Top, pathway diagram indicates pathway connection of three tested SA genes, none of which are miRNA targets. Bottom, *NPR1*, *EDS1*, and *SNC1* were upregulated in *ago1-27* seedlings relative to wildtype. (B) Top, pathway diagram indicates pathway connection of four tested JA genes, the most upstream of which, *TCP4*, is a miRNA target. Bottom, the canonical jasmonate signaling and biosynthesis genes *TCP4*, *LOX2*, *MYC2*, *VSP2* and *JAR1* were strongly upregulated. (C) Top, pathway diagram indicates pathway connection of four tested JA/ET genes. The most upstream gene, *EIN2*, acts in an incoherent feed-forward loop with the mir164ab and its target *ORE1*. ET and JA pathways show cross-regulation via *ORA59* and *MYC2*. Bottom, (JA/ET) markers *ORE1*, *ORA59*, and *ERF1* were significantly upregulated in *ago1-27* cotyledons relative to wild-type; in contrast, *EIN2* was significantly downregulated in *ago1-27* cotyledons. All assayed genes are normalized to *UBQ10* expression. Error bars indicate standard error of the mean across three to four replicates of cotyledons harvested from 18 pooled seedlings. Black stars represent significant upregulation in *ago1-27* cotyledons; grey star (*EIN2*) represents significant downregulation in *ago1-27* cotyledons.

We next investigated additional marker genes in JA signaling. In *A. thaliana,* miR319a represses specific members of the *TEOSINTE BRANCHED/CYCLOIDEA/PCF (TCP)* family of transcription factors, which positively regulate JA biosynthesis (Danisman et al., 2012). TCP4 promotes JA biosynthesis by activating the expression of *LIPOXYGENASE2* (*LOX2*) (Schommer et al. 2008; Danisman et al. 2012). JASMONATE INSENSITIVE 1 (JIN1/MYC2), a master regulator of the response to herbivores, also activates expression of *LOX* genes (Boter et al., 2004; Lorenzo et al., 2004). TCP4-regulated LOX2 catalyzes an early step of jasmonate biosynthesis (Gfeller et al., 2010). Lastly, *JASMONATE RESISTANT 1* (*JAR1)* functions in the synthesis of biologically active jasmonate-isoleucine (Guranowski et al. 2007). All four tested genes -- *TCP4, LOX2, MYC2,* and *JAR1* -- were significantly upregulated in *ago1-27* cotyledons (**Figure 4**.**4**).

Lastly, we examined additional genes known to function in JA/ET signaling. We selected *ETHLYLENE INSENSITIVE 2 (EIN2)* because it is the key integrator of JA/ET stress responses (Alonso et al., 1999) and is required for *PDF1.2* expression (Alonso et al., 1999). We also examined the transcription factor *ORESARA 1*(*ORE1),* which is a direct target of miR164ab and which together with EIN2 regulates ethylene signaling via a trifurcate feed-forward loop (Kim et al., 2009). JA and ET also activate expression of AP2/ERF-domain transcription factors (Nakano et al., 2006) such as *OCTADECANOID-RESPONSIVE ARBIDOPSIS AP2/ERF* 59 (*ORA59*) and *ETHYLENE RESPONSE FACTOR 1* (*ERF1*) (Lorenzo et al., 2003; Pré et al., 2008). Overexpression of *ERF1* enhances resistance against *Botrytis cinerea;* mutation of *ERF1* increases susceptibility to this pathogen (Pré et al., 2008). Overexpression of *ORA59* yields similar results in fungal defense and suppresses MYC-dependent signaling in defense against herbivores, making plants more susceptible to herbivores (Verhage et al., 2011).

All but one gene were significantly upregulated in *ago1-27* cotyledons (**Figure 4.4**). The outlier, *EIN2,* was significantly downregulated at about half of wild-type levels. Given the strong upregulation of *PDF1.2*, this specific result is likely a consequence of the extensive feedback relationships. In summary, all three stress response pathways were upregulated, suggesting that *ago1* mutants no longer interpret environmental signals correctly

and may be overly sensitive to the subtle stress of full spectrum light (Stasinopoulos and Hangarter, 1990).

4.1.3 Exogenous JA and ET treatments, but not SA treatment, suffice to induce lesions in wild-type

The consistent upregulation of all three candidate pathways necessitated a different approach to identify the molecular origin(s) of the lesions in *ago1-27* cotyledons. We applied all three phytohormones separately and in combination and assayed the frequency of affected *ago1-27* and wild-type seedlings. We sprayed five-day-old seedlings grown in full-spectrum light with increasing concentrations of phytohormones (Genoud et al., 2002; Weigel and Glazebrook, 2002; Sangster et al., 2007) and scored affected seedlings at day 10. As SA signaling marker genes were upregulated, we predicted that exogenous SA treatment would increase the number of affected individuals for both *ago1-27* and wild-type. However, the frequency of affected *ago1-27* seedlings decreased; there was no significant response in wild-type (**Figure 4.5**). This result excludes SA as the cause of the observed lesions.

We next tested whether application of methyl jasmonate (MeJA) would increase the frequency of affected seedlings and whether we could detect differences in the response of *ago1-27* and wild-type. In wild-type, the frequency of affected individuals increased dramatically at 50 μM MeJA. Thus, JA treatment suffices to cause the observed lesions (**Figure 4.5**). In *ago1-27* seedlings, the number of affected individuals increased significantly at the low MeJA dosage (10 μM) but did not further increase at the high MeJA dosage (50 μM). This result suggests that *ago1-27* is hypersensitive to JA at low concentrations but its response plateaus at higher concentrations (**Figure 4.5**). We compared the responses of wild-type and *ago1-27* to MeJA using a generalized linear model with genotype and treatment as fixed effects and replicate as a random effect. Indeed, *ago1-27* mutants were significantly more responsive to MeJA at 10 μM conditions (GLMM, $p< 1x10^{-4}$), whereas at 50 μM wild-type was more responsive (GLMM, $p<1x10^{-4}$), indicating that *ago1-27* mutants have a lower threshold for JA pathway saturation.

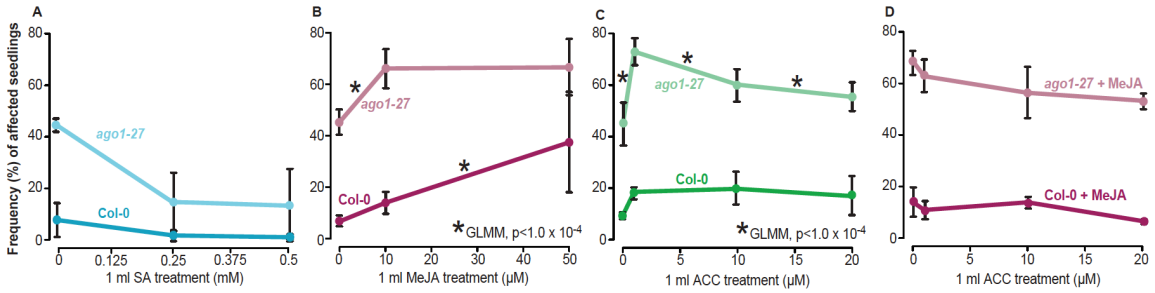


Figure 4.5 Methyl jasmonate (MeJA) treatment was sufficient to induce lesions. (A) Under high salicylic acid conditions, seedlings produced fewer affected individuals for both wild-type and *ago1-27* in full spectrum light. (B) At 10 μM MeJA, *ago1-27* was significantly more responsive compared to wild-type (GLMM, testing for interaction between genotype and treatment, p< 1x10-4) whereas at the 50 μM MeJA treatment, wild-type was more responsive (GLMM, p< 1x10-4). (C) At 1 μM ACC (ethylene pre-cursor molecule), *ago1-27* was more responsive than wild-type, (GLMM, p< 1x10-4), reaching a frequency of affected seedlings similar to *ago1-27* seedlings at 10 μM MeJA. At higher doses, ACC repressed the lesion phenotype in *ago1-27* (GLMM, p< 1x10-4) and had no further effect on wild-type. (D) Combining10 μM MeJA with varying ACC concentrations yielded no significant change, albeit frequencies of affected declined. n = 72-144 individuals per replicate, two to five replicates. The difference in responsiveness between genotypes was modeled using GLMM, testing for interaction between genotype and treatment with replicate and number of seedlings set as random variables (p< 1x10-4) (for all reported p-values, refer to **Table S4.4**)

Lastly, we tested whether ethylene treatment could increase the number of affected seedlings using the ethylene precursor 1-aminocyclopropane-1-carboxylic acid (ACC) (**Figure 4.5**). At the lowest dose of ACC (1 μM), the frequency of affected *ago1-27* seedlings increased sharply. Comparing *ago1-27* and wild-type response to ACC, we found that *ago1-27* mutants were significantly more responsive at 1 μM conditions (GLMM, $p< 1x10^{-4}$). However, with increasing ACC doses, the frequency of affected seedlings decreased subtly but significantly in *ago1-27* and remained unchanged in wild-type (GLMM, $p< 1x10^{-4}$). Thus, ET treatment does increase the number of affected *ago1-27* individuals at low doses

and decreases the frequency of affected *ago1-27* seedlings at high doses, but fails to fully mimic the mutant phenotype in wild-type.

As both JA and ET treatment produced an increase in affected seedlings, we wanted to test the effect of these hormones in combination. We performed these experiments with 10 μM MeJA, the highly responsive dose for *ago1-27* mutant seedlings, and at various concentrations of ACC. The presence of ACC did not significantly change the frequency of affected *ago1-27* and wild-type individuals observed at 10 μM MeJA (**Figure 4.5**). Thus, we conclude that JA, but not ET or SA, is likely the main driver of the environmentally-sensitive lesion phenotype in *ago1-27*.

4.1.4 Disruption of JA perception dramatically reduces lesions in ago1-27 mutant seedlings

Thus far, our expression results and exogenous hormone treatments point to JA as the key hormone underlying the lesion phenotype in *ago1-27* mutant seedlings. To test whether JA causes the lesion phenotype, we created a double mutant of *ago1*-27 and *coi1-1* to block JA perception. *CORONATINE INSENSITIVE 1* (*COI1*) encodes the main receptor of biologically active JA; the well-characterized *coi1-1* mutant carries an amino acid substitution in the F-box motif of COI1 (Feys et al. 1994; Yan et al. 2009). The *coi1-1* homozygous seedlings produced significantly fewer individuals with lesions than wild-type (**Figure 4.6**). The frequency of affected seedlings in the *ago1-27; coi1-1* double mutants was dramatically reduced compared to *ago1-27* single mutants, albeit higher than in wild-type (**Figure 4.6**). Disruption of JA perception and signaling appears to strongly suppress the environmentally-sensitive lesions in both *ago1-27* and wild-type.

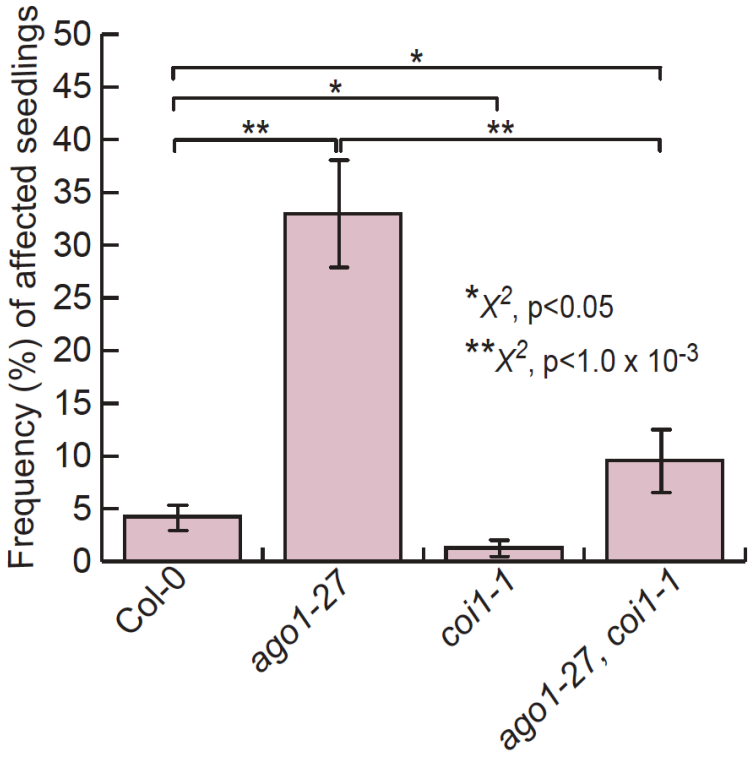


Figure 4.6 Disruption of JA perception suppresses lesion phenotype. *coi1-1* mutant seedlings produced significantly fewer affected seedlings than wild-type ($p<0.05$). *ago1-27*; *coi1-1* double mutant seedlings produced dramatically fewer affected seedlings than *ago1-27* single mutant seedlings ($p< 1x10-3$); however, frequency of affected seedlings was significantly higher than in wild-type ($p<0.05$). Statistical significance was calculated using the $\chi 2$-test of significance (**$p< 1x10-3$, *$p<0.05$) (for all reported p-values, refer to **Table S4.4**). n = 72-144 individuals per replicate, four replicates; error bars represent standard error of the mean across biological replicates.

To explore the possible interplay of JA and ET perception and signaling genetically, we created triple mutants of *ago1-27* with *coi1-1* and *ein2-1*. EIN2 is a membrane bound transport protein; *ein2-1* mutants carry an early stop mutation resulting in a truncated protein (Alonso et al., 1999). Consistent with our expression data for *EIN2* in *ago1-27* mutant seedlings, the *ein2-1* homozygous seedlings produced significantly more lesions than wild-type but were statistically similar to *ago1-27* mutant seedlings (**Figure S4.3**). Thus, loss of ET perception increases rather than decreases the lesion phenotype. Double mutants of *ago1-27* and *ein2-1* largely resembled *ago1-27* single mutants (**Figure S4.3**). The *coi1-1* mutation

was epistatic in both the double mutant with *ein2-1* and the triple mutant *ago1-27;coi1-1;ein2-1* (**Figure S4.3**). We conclude that the observed environmentally-sensitive, *ago1-27*-dependent lesions are largely driven by JA and that the effects of exogenous ET are likely due to the known cross-regulation of both pathways.

4.1.5 ARGONAUTE1 and HSP90 interact genetically in producing the lesion phenotype

Recent studies in both animals (Iwasaki et al., 2010; Johnston et al., 2010) and plants (Iki et al., 2010) demonstrate that the molecular chaperone HSP90 is required for different aspects of AGO function. HSP90-reduced plants show overexpression of genes involved in JA biosynthesis (Sangster et al., 2007) and are more resistant to herbivores than wild-type (Sangster et al., 2007). Therefore, we wanted to explore if the observed environmentally-and JA-dependent lesion phenotype was affected by perturbing HSP90. To do so, I created a double mutant of *ago1-27* with a previously published RNAi line, RNAi-A1 (Sangster et al., 2007). This line primarily affects levels of the environmentally-responsive, heat-inducible HSP90.1 paralog, which could be plausibly involved in the environmentally-dependent lesion phenotype. We refrained from using highly specific inhibitors of the chaperone, as these compounds are highly sensitive to full spectrum light (Queitsch et al., 2002). Indeed, Alex Mason and I observed a subtle but significant increase in affected seedlings in single HSP90-mutants (RNAi-A1) compared to wild-type seedlings (**Figure 4.7**). This result is consistent with the previously described role of HSP90 in JA signaling (Sangster et al., 2007). However, HSP90 perturbation produced far fewer affected seedlings than AGO1 perturbation, inconsistent with simple epistasis of both proteins. Assuming that the HSP90 RNAi-A1 lesion phenotype arises as the result of an AGO1-HSP90 interaction, there should be no further increase in the frequency of affected seedlings in the double mutant. If, however, HSP90 and AGO1 act independently, we should observe additive effects. These simple assumptions are complicated by the fact that both *ago1-27* and HSP90 RNAi-A1 are weak, partial loss-of-function mutants. Our result deviated subtly but significantly from additive expectations (GLMM, $p<1x10^{-4}$, **Figure 4.7**), suggesting a synergistic relationship between the chaperone and AGO1 in this phenotype. The non-additive higher incidence of affected seedlings in double mutants is consistent with previous observations that HSP90

perturbation increases the penetrance of genetic variants (Queitsch et al., 2012; Lempe et al., 2013). In summary, our results show that although AGO1 and HSP90 appear to be inter-dependent with regard to responding properly to environmental perturbations, neither's contribution is fully explained by the other. Furthermore, AGO1 is the principal integrator of environmental signaling for proper repression of lesion formation, with a much-reduced role for the environmentally responsive HSP90.1 paralog.

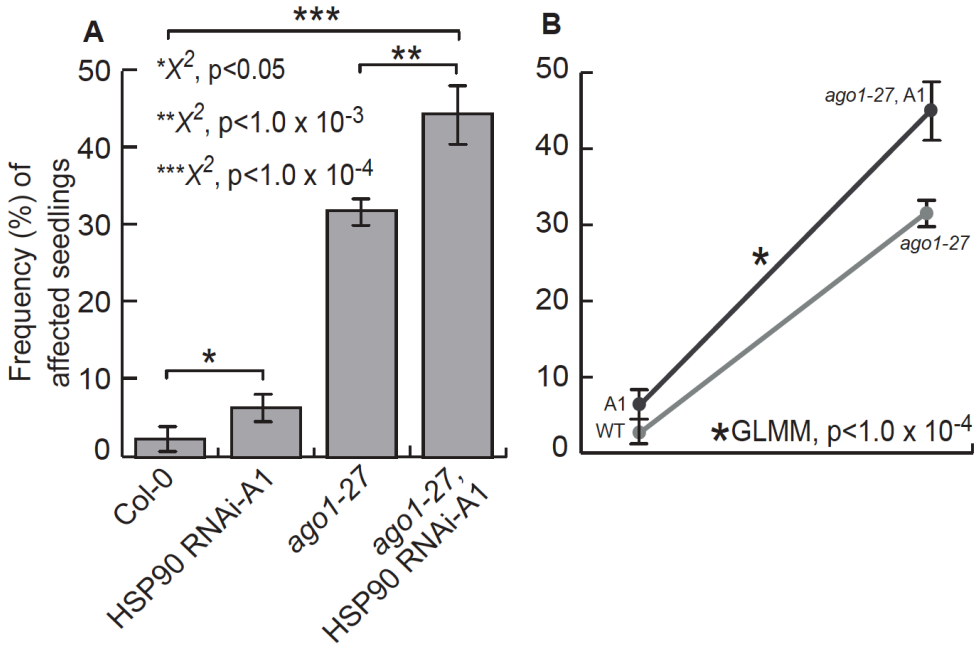


Figure 4.7 AGO1 and HSP90 genetically interact. (A) *HSP90 RNAi-A1* seedlings produced significantly more affected seedlings than wild-type ($p<0.05$). *ago1-27*; *HSP90 RNAi-A1* double mutant seedlings produced significantly more affected seedlings than *ago1-27* single mutant seedlings ($p< 1x10-4$). Error bars represent standard error of the mean across two replicates of n = 108 seedlings. (B) The *ago1-27* mutation showed higher penetrance in the presence of *HSP90 RNAi-A1* (A1) in full spectrum light (GLMM, $p< 1x10-4$). Statistical significance was calculated using the $\chi 2$-test of significance (***$p< 1x10-4$, ** $p< 1x10-3$, *$p<0.05$) (for all reported p-values, refer to **Table S4.4**). n = 72-144 individuals per replicate, two replicates. Error bars represent standard error of the mean across biological replicates.

4.2 DISCUSSION

Here, we present a previously undescribed phenotype in *ago1* mutant plants – apparently stochastic lesions in cotyledons – and uncover its molecular underpinnings. Several recent studies have revealed roles for specific miRNAs in development, hormone signaling, and defense. Our initial observation and subsequent investigation of the lesion phenotype identifies AGO1 as a key integrator of environmental signals and hormone signaling, providing robustness to micro-environmental challenges as was previously observed for the molecular chaperone HSP90. The spontaneous lesions that we observed in *ago1* seedlings, especially in response to full spectrum light, appear to be dead and dying tissue bearing the hallmarks of HR due to upregulated JA/ET signaling. In wild-type, functional AGO1 appears to play a key role in properly deploying HR only upon pathogen attack.

AGO1-dependent regulation of gene activity occurs through two principal mechanisms, mRNA cleavage (Voinnet, 2009) and translational repression (Iwakawa and Tomari, 2013). It was previously assumed that plant miRNAs primarily functioned through RNA cleavage due to their near perfect base pairing with their target sequences (Rhoades et al., 2002). More recently, however, several studies have provided evidence for translational repression (Aukerman and Sakai, 2003; Bari et al., 2006; Brodersen et al., 2008), notably with regard to the heat stress-induced downregulation of SPL transcription factors (Gandikota et al., 2007; Stief et al., 2014). Translational repression, which is reversible, rather than cleavage of mRNAs, has been proposed as a mechanism for facilitating recovery after environmental stress (Sunkar et al., 2007; Voinnet, 2008).

Notably, the hypomorphic *ago1-27* point mutation (Morel et al., 2002) is thought to primarily affect translational repression (Brodersen et al., 2008; Voinnet, 2009). In contrast, the *ago1-25* mutant (Morel et al., 2002) resides in close proximity to one of the three catalytic residues required for slicing. Complementation with *AGO1* transgenes containing wild-type catalytic residues (DDH, DDD) complement *ago1-25* mutant defects whereas

complementation with altered catalytic residues (ADH, DAH, DDA) do not, suggesting that *ago1-25* is primarily defective in slicing (Carbonell et al. 2012). *ago1-27* mutants were far more responsive to full spectrum light than *ago1-25* mutants, presumably due to the functional differences between both mutant alleles. Translational repression, which allows for faster recovery after stress than RNA cleavage that requires new transcription, appears to be of greater relevance for the response to full spectrum light. Our results are consistent with a recent report on the heat shock response in *ago1* mutants (Stief et al., 2014). *ago1-27* seedlings were more responsive to heat shock conditions compared to *ago1-25* seedlings (Stief et al., 2014).

Previous studies suggest a link between light conditions and JA signaling. Specifically, solar UV-B radiation renders plants more resistant to herbivory, primarily due to upregulation of the JA pathway (Ballare et al. 1996; Mazza et al. 1999; Demkura et al. 2010). Consistent with these prior findings, the *ago1*-dependent lesions were much less frequent in filtered light conditions which exclude UV-B radiation.

Similarly, HSP90 perturbation is known to increase JA signaling and yields greater resistance to herbivores (Sangster et al., 2007). As expected, mutants with reduced HSP90 levels show subtly but significantly greater frequency of affected seedlings. HSP90 plays important roles in plant defense, especially in defense against microbial pathogens (Hubert et al., 2003; Takahashi et al., 2003; Sangster and Queitsch, 2005; Zhang et al., 2015). HSP90 is also implicated in chaperoning AGO1 (Iki et al., 2010). Our double mutant analysis, however, suggest a synergistic, rather than an epistatic relationship. *ago1-27* single mutants show many more affected seedlings than HSP90 RNAi-A1 seedlings, and the number of affected seedlings increases further in double mutants beyond additive expectations. The interpretation of this double mutant result is complicated by the fact that both mutants represent weak partial mutants of essential genes. The increased penetrance of the *ago1-27* mutation upon HSP90 perturbation is consistent with the notion that HSP90 is critical for buffering deleterious mutations (Queitsch et al., 2012). Our results further suggest that fluctuations in environmental conditions are buffered by both AGO1 and HSP90.

In our gene expression studies, we observed that SA marker genes, just like JA and ET marker, were also highly upregulated. We conclude that *ago1* plants generally upregulate these hormone pathways, presumably in part due to their complicated cross-regulation. However, as our exogenous hormone treatments and genetic analyses demonstrate, JA is the main driver of the lesion phenotype. In summary, AGO1 and the miRNA pathway appear to minimize variation in phenotype among isogenic seedlings and suppress aberrant misregulation of plant defense pathways similar to previous observations for the phenotypic capacitor HSP90.

4.3 SUPPLEMENTARY INFORMATION

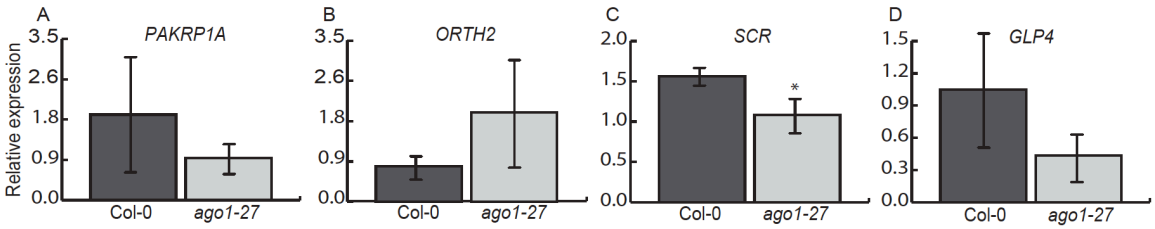


Figure S4.1. Plant stem cell marker genes are not upregulated in full spectrum light-grown *ago1-27* seedlings. (A, B, C, D) *ago1-27* cotyledons did not show significantly higher expression of the stem cell marker genes *PAKRP1A* (shoot apical meristem marker), *ORTH2* (shoot apical meristem marker), *SCR* (root apical meristem marker), and *GLP4* (root apical meristem marker). Expression of all assayed genes was normalized to *UBQ10* expression. Error bars indicate standard error of the mean across three to four replicates of cotyledons harvested from 18 pooled seedlings. Black star represents significant downregulation in *ago1-27* cotyledons.

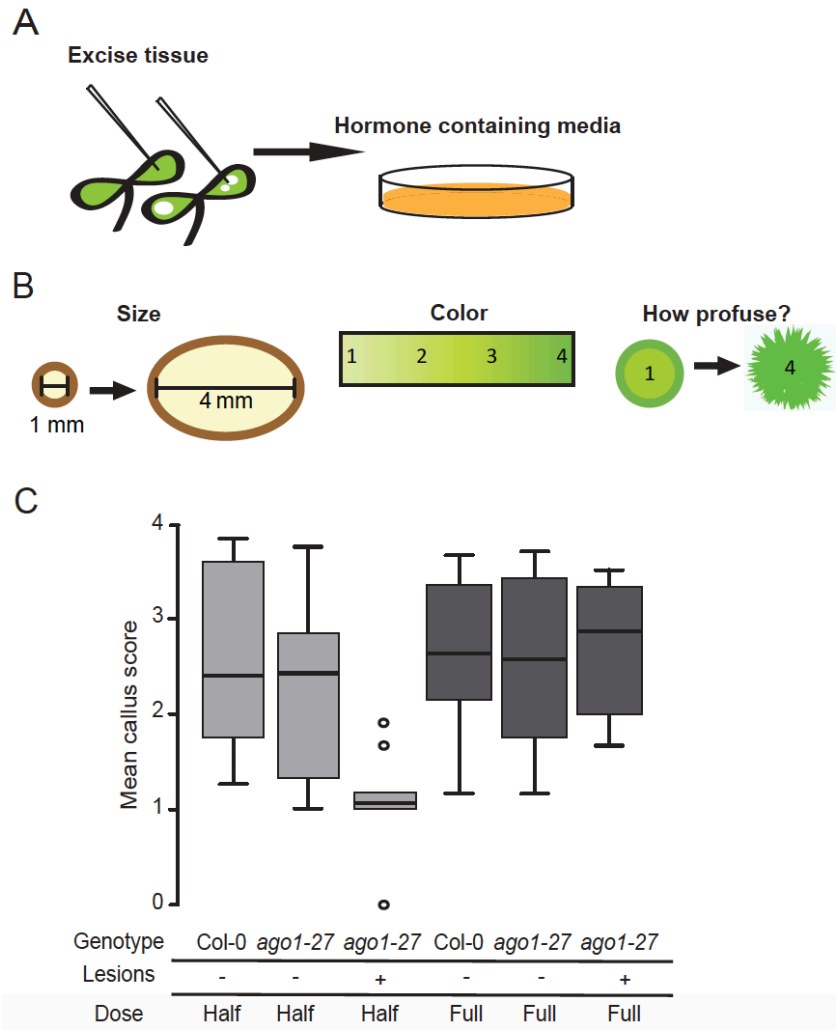


Figure S4.2. *ago1-27* lesion tissue did not show enhanced callus formation. (A) Schematic representation of experimental procedure. Affected and unaffected tissue was excised from 10-day-old *ago1-27* cotyledons and 10-day-old unaffected tissue was excised from wild-type cotyledons. Tissue was placed on callus-inducing media at the recommended hormone doses and half the recommended doses as described previously (Weigel and Glazebrook, 2002). Plates were scored five days later for callus induction. (B) Callus scoring index. Callus formation was scored based on three criteria: length at longest axis, color, and profuseness. These separate criteria were given a score of 1-4, with a score of 4 being the most callus-like of criteria. Scores were averaged to compute callus scores. (C) Mean callus score of tested genotypes and treatments. At the lower hormone dose, *ago1-27* lesion tissue does not form callus tissue at all.

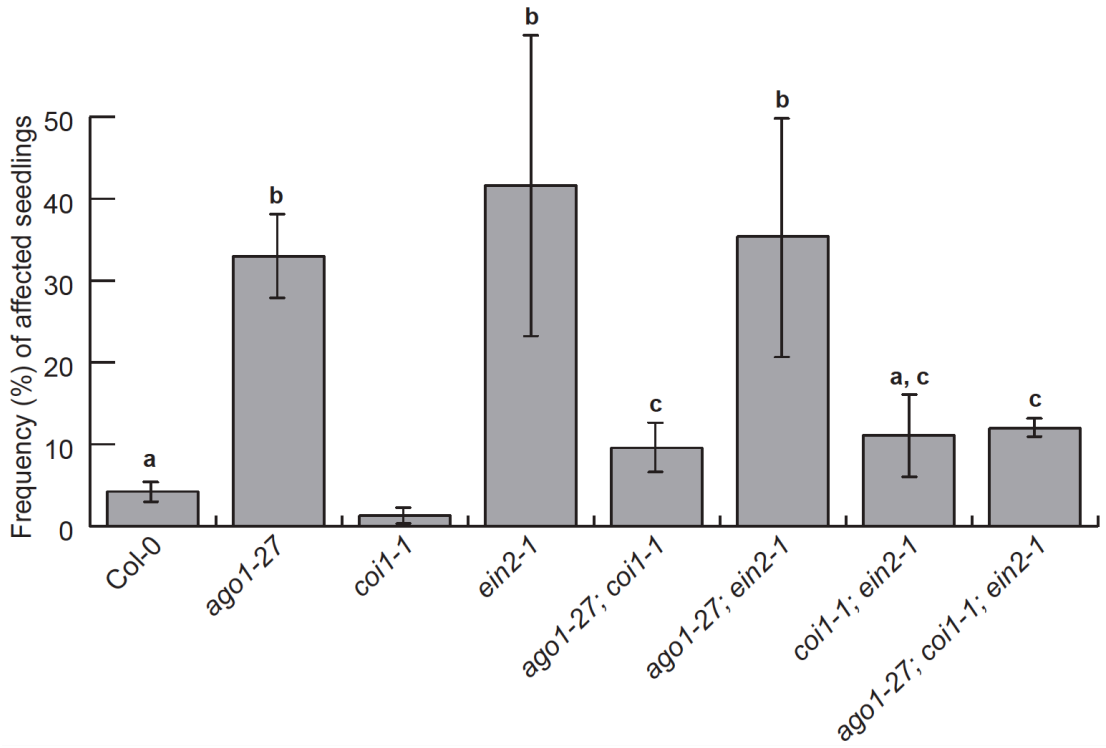


Figure S4.3. The disruption of JA perception in coi1-1 is epistatic to both ago1-27- and ein2-1-dependent lesion phenotypes. *ago1-27* and *ein2-1* seedlings produced similar frequencies of affected individuals. *ein2-1* single mutant and *ago1-27*; *ein2-1* double mutant seedlings produce similar frequencies of affected individuals. *coi1-1* suppressed the lesion phenotype in all three combinations: *coi1-1*;*ein2-1*, *ago1-27*; *coi1-1*, and *ago1-27*;*coi1-1*;*ein2-1*. Statistical significance was calculated using the χ^2-test of significance. Shared letters (a, b, c) denote statistically similar genotypes. Data for Col-0, *ago1-27*, *coi1-1*, and *ago1-27*; *coi1-1* also appears in **Figure 4.6.** For all reported p-values, refer to **Table S4.4**. n = 72-144 individuals per replicate, two-four replicates; error bars represent standard error of the mean across biological replicates.

Table S4.1: qRT-PCR primers

Gene	Accession number	Primer	Primer sequence
GLP4	AT1G18970	GLP4 F 1	ATGGCTTTCCATGCAAATCA
		GLP4 R 1	TCGAGATTCCTAAAGTGTTTAGACC
SCR	AT3G54220	SCR F 1	CTCAAACTTTCGAACCTCTCTATC
		SCR R 1	GGGAACAGTGGCCGTCA
ORTH2	AT1G57820	ORTH2 F 1	GTCTCACAAGGAGAAGCG
		ORTH2 R 1	GGCTCATTGTCACATCTAACGAA
PAKRP1A	AT4G14150	PAKRP1A F 1	GGATGTGTCATCATGCCC
		PAKRP1A R 1	TATTTCTAATCGTAGGAGAGACGC
PDF1.2	AT5G44420	PDF1.2 F 1	ATGGCTAAGTTTGCTTCCA
		PDF1.2 R 1	TTAACATGGGACGTAACAGATAC
PR1	AT2G14610	PR1 F 1	TCTTCCCTCGAAAGCTCAAG
		PR1 R 1	AAGGCCCACCAGAGTGTATG
VSP1	AT5G24780	VS1 F 1	ACCTCTTGGAACTCGGGATT
		VSP1 R 1	ACCACTTGCGTCAACTTCG
EDS1	AT3G48090	EDS1 F 1	TCTGTGGAAATGGCTGTGAG
		EDS1 R 1	CCAAAGGAGCTCCAAATGTC
NPR1	AT1G64280	NPR1 F 1	AGGCACTTGACTCGGATGAT
		NPR1 R 1	CCTCGGATTCCTATGGTTGA
SNC1	AT4G16890	SNC1 F 1	GCCGGATATGATCTTCGGAA
		SNC1 R 1	CGGCAAGCTCTTCAATCATGG
TCP4	AT3G15030	TCP4 F 1	CAACATCACCACCACACCTC
		TCP4 R 1	ACAGAAACCCTCCTCCGTTT
LOX2	AT3G45140	LOX2 F 1	TGAGGACTCATGCCTGTACG
		LOX2 R 1	ATTCCACCTCCGTTGACAAG
JAR1	AT2G46370	JAR1 F 1	TACGCGGGTGATCTACCTCT
		JAR1 R 1	AACCAACCGGTTTCTCCTCT
MYC2	AT1G32640	MYC2 F 2	GTGACGGATACGGAATGGTT
		MYC2 R 2	ATGCATCCCAAACACTCCTC
EIN2	AT5G03280	EIN2 F 1	AAAACCGGCTAAAGGCAAAT
		EIN2 R 1	TCAAAACCGAAGCCAAATTC
ORE1	AT5G39610	ORE1 F 1	CTGCTACTGCCATTGGTGAA
		ORE1 R 1	TCCAATAACCGGCTTCTGTC
ERF1	AT3G23240	ERF1 F 1	GATGGTTGTTCTCCGGTTGT
		ERF1 R 1	CCCAAAAGCTCCTCAAGGTA
ORA59	AT1G06160	ORA59 F 1	GGCTCTCGCTTATGATCAGG
		ORA59 R 1	GGACGGTTTCTCATGGAGTG

Table S4.2: Accession numbers of genes analyzed

Gene	Accession number
AGO1	AT1G48410
COI1	AT2G39940
EIN2	AT5G03280
GLP4	AT1G18970
SCR	AT3G54220
ORTH2	AT1G57820
PAKRP1A	AT4G14150
PDF1.2	AT5G44420
PR1	AT2G14610
VSP1	AT5G24780
EDS1	AT3G48090
NPR1	AT1G64280
SNC1	AT4G16890
TCP4	AT3G15030
LOX2	AT3G45140
JAR1	AT2G46370
MYC2	AT1G32640
EIN2	AT5G03280
ORE1	AT5G39610
ERF1	AT3G23240
ORA59	AT1G06160
HSP90.1	AT5G52640
UBC10	AT5G25760

Table S4.3: Genotyping primers

Gene	Accession number	Primer	Primer sequence	Restriction enzyme/allele cut
AGO1	AT1G48410	F	ACCACGTTCTTTGGGATGAG	Bsp1286I/wild-type
		R	TCTACCCATTCCACCTC	
COI1	AT2G39940	F	GGGGAGATAAGGGATATGAATGC	Xcm1/wild-type
		R	TTGTGGAAACCCCAAAACTC	
EIN2	AT5G03280	F	CGCCATCTTTGTTTCAACAATCAGATCC	BsrB1/wild-type
		R	CCAGAGGAAAGAGAGTTGGATGTAAAGTACTCTACCGCT	
HSP90 RNAi-A1	AT5G52640	Lw1	GATGCGGGATATGGCTGGACT	Produces 1194 bp band with Rc3 primer for wild-type
		Lm1	CGGCAGTTCATCAGGGCTAA	Produces 850 bp band with Rc3 primer for mutant
		Rc3	TATCACCTCGTTAGGGGCCA	

Table S4.4: Reported p-values

Figure	Comparison	Test	P-value
1	Col-0 Yellow vs. Col-0 White	Chi-square	0.8916
1	Col-0 Yellow vs. *ago1-27* Yellow	Chi-square	0.7677
1	Col-0 Yellow vs. *ago1-46* Yellow	Chi-square	**5.00E-06**
1	Col-0 White vs. *ago1-27* White	Chi-square	**4.00E-21**
1	Col-0 White vs. *ago1-46* White	Chi-square	0.0077
1	*ago1-27* Yellow vs. *ago1-27* White	Chi-square	**9.00E-26**
1	*ago1-46* Yellow vs. *ago1-46* White	Chi-square	2.12E-01
1	*ago1-25* White vs. Col-0 White	Chi-square	**1.00E-05**
1	*ago1-25* Yellow vs. *ago1-25* White	Chi-square	7.01E-01
1	*ago1-25* Yellow vs. Col-0 Yellow	Chi-square	**1.00E-06**
1	*ago1-27* White vs. *ago1-25* White	Chi-square	**1.00E-09**
1	*ago1-27* vs. *ago1-25*	GLMM	**< 1.0E-04 (towards *ago1-27*)**
5	Col-0 vs. *ago1-27*, 0 and 10 µM MeJA	GLMM	**< 1.0E-04 (towards *ago1-27*)**
5	Col-0 vs. *ago1-27*, 0 and 50 µM MeJA	GLMM	**< 1.0E-04 (towards Col-0)**
5	Col-0 vs. *ago1-27*, 10 and 50 µM MeJA	GLMM	**< 1.0E-04 (towards Col-0)**
5	Col-0 vs. *ago1-27*, all MeJA treatments	GLMM	**< 1.0E-04 (towards Col-0)**
5	Col-0 vs. *ago1-27*, 0 and 250 µM SA	GLMM	0.7077
5	Col-0 vs. *ago1-27*, 0 and 500 µM SA	GLMM	0.858
5	Col-0 vs. *ago1-27*, 250 and 500 µM SA	GLMM	0.5089
5	Col-0 vs. *ago1-27*, all SA treatments	GLMM	0.9123
5	Col-0 vs. *ago1-27*, 0 and 1 µM ACC	GLMM	**< 1.0E-04 (towards *ago1-27*)**
5	Col-0 vs. *ago1-27*, 0 and 10 µM ACC	GLMM	0.3521
5	Col-0 vs. *ago1-27*, 0 and 20 µM ACC	GLMM	**< 1.0E-04 (towards *ago1-27*)**
5	Col-0 vs. *ago1-27*, 1 and 10 µM ACC	GLMM	**< 1.0E-04 (towards *ago1-27*)**
5	Col-0 vs. *ago1-27*, 1 and 20 µM ACC	GLMM	**< 1.0E-04 (towards *ago1-27*)**
5	Col-0 vs. *ago1-27*, 10 and 20 µM ACC	GLMM	0.065
5	Col-0 vs. *ago1-27*, all ACC treatments	GLMM	0.093
5	Col-0 vs. *ago1-27*, 0 and 1 µM ACC, 10 µM MeJA	GLMM	0.2211
5	Col-0 vs. *ago1-27*, 0 and 10 µM ACC, 10 µM MeJA	GLMM	0.1482
5	Col-0 vs. *ago1-27*, 0 and 20 µM ACC, 10 µM MeJA	GLMM	0.3834
5	Col-0 vs. *ago1-27*, 1 and 10 µM ACC, 10 µM MeJA	GLMM	0.2118
5	Col-0 vs. *ago1-27*, 1 and 20 µM ACC, 10 µM MeJA	GLMM	0.8751
5	Col-0 vs. *ago1-27*, 10 and 20 µM ACC, 10 µM MeJA	GLMM	0.2002
5	Col-0 vs. *ago1-27*, all ACC treatments, 10 µM MeJA	GLMM	0.0732
6	Col-0 vs. *ago1-27*	Chi-square	**1.15E-55**
6	Col-0 vs. *ago1-27; coi1-1*	Chi-square	0.0036
6	Col-0 vs. *coi1-1*	Chi-square	0.022
6	ago1-27 vs. *ago1-27; coi1-1*	Chi-square	**2.69E-15**
6	ago1-27 vs. *ago1-27; ein2-1*	Chi-square	0.35
Supp. 4	Col-0 vs. *ago1-27; coi1-1; ein2-1*	Chi-square	0.0018
Supp. 4	Col-0 vs. *coi1-1; ein2-1*	Chi-square	0.76
Supp. 4	Col-0 vs. *ein2-1*	Chi-square	**1.30E-27**
Supp. 4	ago1-27 vs. *ago1-27; coi1-1; ein2-1*	Chi-square	**1.76E-07**
Supp. 4	ago1-27 vs. *coi1-1; ein2-1*	Chi-square	**3.86E-21**
Supp. 4	ago1-27 vs. *ein2-1*	Chi-square	0.031
7	Col-0 vs. RNAi-A1	Chi-square	0.0394
7	Col-0 vs. *ago1-27*	Chi-square	**4.67E-09**
7	Col-0 vs. *ago1-27*; RNAi-A1	Chi-square	**3.98E-25**
7	*ago1-27* vs. *ago1-27*; RNAi-A1	Chi-square	0.0076
7	*ago1-27* White vs. *ago1-27; RNAi-A1*	GLMM	**< 1.0E-04 (towards *ago1-27;* RNAi-A1)**

Chapter 5

AGO1 BUFFERS GENETIC VARIATION IN THE PLANT *ARABIDOPSIS THALIANA*

5.1 INTRODUCTION

Genetic networks involve feedback loops and regulatory mechanisms to produce developmental and physiologic robustness in spite of environmental or genetic perturbations (Rutherford and Lindquist 1998; Queitsch et al. 2002; Lempe et al. 2013). Network disruptions decrease robustness and produce increased phenotypic variation in either a particular trait or by affecting organism phenotype more broadly. One key component of network regulation is the protein chaperone Heat Shock Protein 90 (HSP90); its loss of function broadly disrupts the genetic network, thereby affecting genetic variation at many other loci and producing greatly increased phenotypic variation (Rutherford and Lindquist 1998; Queitsch et al. 2002; Yeyati et al. 2007; Sangster et al. 2008a; Jarosz and Lindquist 2010). The phenomenon that functional HSP90 keeps genetic variation silent and HSP90 perturbation releases this formerly hidden variation has been termed phenotypic capacitance (Masel and Siegal 2009) – a different term for epistasis. In contrast to traditional epistasis, which describes the non-reciprocal interaction of two loci, phenotypic capacitance is an epistasis phenomenon in which one locus, *e.g.* HSP90, interacts with many others.

MicroRNAs (miRNAs) also act as phenotypic capacitors by buffering stochastic (Hilgers et al. 2010), environmental (Li et al. 2009b), and genetic variation (Cassidy et al. 2013), raising the question as to their overlap or unique characteristics compared with HSP90. This question became particularly compelling with recent evidence that ARGONAUTE 1 (AGO1), a key enzyme in miRNA metabolism, is chaperoned by HSP90 (Iki et al. 2010). MicroRNAs are small 21-24 nucleotide RNAs that regulate the expression of target genes in a sequence-specific manner, by degrading their mRNA (Axtell 2013; Bologna and Voinnet 2014), or inhibiting their translation into proteins (Reis et al. 2015). In

plants, miRNA biogenesis starts with Polymerase II (Pol II) transcription of pri-miRNAs, which fold in a hairpin-structure consisting of an upper terminal loop and stem, followed by the miRNA, its complementary sequence (miRNA*) duplex, and a lower stem (Bologna and Voinnet 2014). The hairpin structure is recognized by DICER-LIKE 1 (DCL1) and cleaved twice to release a 21-24 nucleotide double stranded RNA with 3' overhangs (Bologna and Voinnet 2014). The miRNA/miRNA* duplex is loaded onto an Argonaute protein to form the RNA-induced silencing complex (RISC). RISC then targets the miRNA-complementary mRNAs. In *A. thaliana*, 21-nt miRNAs are loaded primarily onto AGO1 (Axtell 2013; Bologna and Voinnet 2014).

Several recent studies in various organisms suggest that AGO1 is chaperoned by HSP90, albeit with different, organism-specific functional roles. HSP90 physically interacts with Argonaute proteins in yeast (Wang et al. 2013), flies (Miyoshi et al. 2010; Iwasaki et al. 2010; Gangaraju et al. 2011), humans (Johnston et al. 2010), *Tetrahymena* (Woehrer et al. 2015), and in tobacco cell extracts (Iki et al. 2010). Here, I investigate to which extent perturbation of AGO1 (i) affects phenotypic variation in isogenic *A. thaliana* seedlings, (ii) releases genetic variation in divergent backgrounds, and (iii) and whether AGO1-dependent loci overlap with HSP90-dependent loci.

I found that AGO1 perturbation alone sufficed to significantly increase phenotypic variation in in several seedlings traits (**figure 5.1**). Next, I set out to test whether AGO1 perturbation releases cryptic genetic variation. To do so, crossed the *ago1-27* partial mutant with Col-0 lines with partial chromosome substitutions from the divergent *A. thaliana* strain Ler (STAIRS, STepped Aligned Inbred Recombinant Strains) (Koumproglou et al. 2002). I measured five quantitative traits and found instances of AGO1-dependent variation. Specifically, I observed that the presence of *ago1-27* combined with the presence of L*er* DNA on the top of chromosome V uncoupled days to flowering from rosette leaf number, the two highly correlated traits that are used interchangeably to assess the onset of flowering (Lempe et al. 2005; Salomé et al. 2011).

The transition from the vegetative to the reproductive stage in *A. thaliana* is irreversible and hence of utter importance for fitness. Therefore, this transition is highly regulated and relies on integrating a multitude of environmental and organismal signals such as light, day length, temperature, season (vernalization), plant hormones, and ageing pathways. These signals and their respective pathways converge in integrator genes such as *FLOWERING LOCUS T* (*FT*), and *SUPPRESOR OF OVEREXPRESSION OF CO 1* (*SOC1*) (Song et al. 2013; Pajoro et al. 2014), which in turn activate the floral meristem identity genes APETALA1 (AP1), FRUITFUL (FUL), and LEAFY (LFY). These inputs change the shoot meristem into an inflorescence meristem, initiating the plant reproductive stage and the development of flowers (Andres and Coupland 2012; Huijser and Schmid 2011; Poethig 2013).

A major player in flowering time phenotypes is *FLOWERING LOCUS C* (*FLC*), which prevents flowering by repressing the expression of *FT*. *FLC* expression is repressed when plants are exposed to cold temperatures for a long period of time (*i.e.,* vernalization or winter period) (Andres and Coupland 2012). Many natural strains that do not require vernalization to flower carry *FLC* mutations, as is the case in L*er*. The vernalization pathway intersects with the ageing pathway in the regulation of *FT*, *AP1*, *FUL* and *LFY* (Posé et al. 2012; Wang 2014). The principal regulators of the ageing pathway are two miRNAs, miR156 and miR172. MicroR156 represses the expression of several *SQUAMOSA PROMOTER BINDING LIKE* (SPL) transcription factors (miR156-SPL module), which in turn activate the expression of *FUL*, *AP1* and *LFY* (Yamaguchi et al. 2009). One of these, SPL9, promotes the expression of *miR172* (Wu et al. 2009b), which represses *APETALA2* (*AP2*), and *AP2*-like genes (*miR172-AP2* like module) (Aukerman and Sakai 2003), that in turn repress the expression of *FT* (Wu et al. 2009b).

Here, I show that the ageing pathway, and specifically, the miR156-SPL module, uncouples the closely related traits days to flower and rosette leaf number in *ago1-27* in a background-dependent fashion. Using several lines of complementary evidence, I demonstrate that the AGO1-dependent polymorphism(s) likely reside in miR156f and that these polymorphisms occur frequently in natural strains.

5.2 MATERIALS AND METHODS

Plant Materials. The following parental lines were used: Col-0, *ago1-27* in the Col-0 background (Morel et al. 2002), and STAIRS N9448, N9456, N9472, N9501(Koumproglou et al. 2002). *ago1-27* plants were crossed into the STAIR lines and F2s that had the wild type and *ago1-27* allele in both Col-0 and the STAIRS backgrounds were isolated. Selected F2s and their progeny were used to perform the described experiments.

Genotyping of F2 plants. I used PCR to genotype the F2s from each STAIRS – *ago1-27* cross. PCR conditions for *ago1-27* genotyping is as follows: 5' at 94 °C, followed by 35 cycles at 30 s at 94 °C, 30 s at 55 °C, 1 min at 72 °C. PCR product was then digested at 37°C with Bsp1286I, which cuts wild-type sequence.

Agar media. For the hypocotyl and root length assays, the plants were grown on MS media containing 0.0005% MES hydrate, 0.004% vitamin solution, 3% phytagar, and 1% sucrose.

Hypocotyl and root length assays. Seeds from different genotypes were laid on agar plates in a 10-seeds per plate arrangement. The plates were stacked in racks, wrapped in foil, and transferred to 4˚C for five days. Then they were unwrapped, and exposed to light for two hours. After that, the plates were wrapped in foil again, to prevent further light exposure and were transferred to a 23˚C tissue culture incubator for seven days. The plants were grown vertically. After that, the plates were taken out, and photographed. The photographs were used to measure the seedlings' hypocotyls and roots using the ImageJ software (http://rsbweb.nih.gov/ij/).

Early morphology traits analysis. Seeds from the different genotypes were plated on agar (36 seeds/per plate). The plates were wrapped in foil and transferred to 4˚C for five days. Then, they were unwrapped and transferred to long days (LD) at 23˚C tissue culture incubator for 10 days. The plants were grown horizontally. The plates were rotated every day to prevent light-growth biases. On the 10th day the seedlings were scored for their morphological traits.

Flowering time experiments. Seeds from different genotypes were embedded in 1mL of 0.1% agar, and then stratified for 5 days at 4˚C. Then, they were sowed on soil in 36-pot trays. Flowering time was measured by scoring both the number of rosette leaves and days to flowering. The last one was recorded when the primary inflorescence of the plant had reached a height of 1cm. Flowering time experiments were performed in long days (LD, 16 hours of light, 8 hours of dark), at 23˚C. For the cold and hot temperature experiments, plants were stratified as above, then grown at 23˚C for one day, and then transferred to 16˚C or 27˚C, respectively. All of them were grown in LD.

Rosette diameter measurements. The diameter of the rosette was measured on the day that the primary inflorescence of the plant reached a height of 1cm.

Vernalization treatment. Seeds were stratified for 5 days at 4˚C and then laid on soil. They were allowed to grow for 5 days at 23˚C in LD or short days (SD) conditions and then transferred to 4˚C for forty days, according to recommendations from Sung *et al.*, 2006 (Sung et al. 2006).

Gene expression analysis. To determine the expression levels via qPCR, total RNA was isolated from the aerial parts of 14-day old plants at ZT16 using the SV Total Isolation System (Promega). RNA quality was determine using a Nanodrop and only qualified samples (A260/A230 > 1.8 and A260/A280> 1.8) were used for subsequent qPCR experiments. To remove possible DNA contamination I treated the DNA with DNaseI (Ambion) for 60 minutes at 37˚C. I used the Transcriptor First Strand cDNA Synthesis Kit (Roche) for cDNA synthesis. The qPCR primers were designed using the Universal ProbeLibrary Assay Design Center tool (Roche), and Primer3 (Untergasser et al. 2012). Specific amplification was confirmed previous to the qPCR experiments. The qPCRs experiments were carried out in 96-well plates with a LightCycler480 (Roche) using SYBR green. The following program was used for the amplification: pre-denaturation for 5 min at 95˚C, followed by 35 cycles of denaturation for 15 s at 95˚C, annealing for 20 s at 55˚C, and elongation for 30 s at 72˚C. All

qPCR experiments were carried out on two biological replicates (independently harvested samples on different days), and three technical replicates per sample.

miRNA gene expression. Mature forms of miR156 and miR172b were measured by qPCR using hairpin oligonucleotides, as described (Varkonyi-Gasic et al. 2007) (**Table S5.1**).

Sequencing of miR156F, D and E in diverse *A. thaliana* strains. The genes *MIR156F*, *MIR156D* and *MIR156E* were amplified using the primers listed on **Table S5.1**. Then, the PCR products were sequenced by the Sanger method. The sequences were then aligned using T-coffee.

Northern blots. RNA was extracted from mostly aerial parts of 14-day old seedlings using Trizol according to the manufacturer's instructions. 10 ug of RNA was used for all samples. The samples were run in a denaturing 17% polyacrylamide gel, 7M urea, and then transferred to a nitrocellulose membrane as previously described (Cuperus et al. 2010). Blots were hybridized with oligo probes complementary to miR172b and exposed for one day.

5.3 RESULTS

5.3.1 ago1 mutant plants show increased phenotypic variation

I examined several morphological and quantitative phenotypes of hypomorphic *ago1* mutants, *ago1-46*, and *ago1-27,* the former being a less severe mutant than the latter. I noted that 10-day old isogenic seedlings of *ago1-46* and *ago1-27* showed increased phenotypic variation in morphological features such as lesions in cotyledons, rosette symmetry, and organ defects compared to isogenic wild type seedlings (**Figure 5.1a**). The more severe *ago1* mutant, *ago1-27*, showed subtly higher numbers of abnormal phenotypes compared to the less severe *ago1-46* mutant.

Next, I turned to examining hypocotyl length in the dark; a readily measured quantitative trait that shows increased variation in response to HSP90 perturbation. Similar to our earlier results with HSP90 perturbation, I found that *ago1-27* dark-grown seedlings showed a different mean ($p < 2.26e^{-16}$, Wilcoxon test) and significantly greater variance of hypocotyl length than wild type ($p = 0.0002391$, Levene's test) (**Figure 5.1b**). Two other *ago1* mutants, *ago1-46* and *ago1-25*, also showed greater variance in hypocotyl length compared to wild type seedlings (data not shown). I conclude that fully functional AGO1 maintains phenotypic robustness and buffers developmental noise among isogenic seedlings for both morphological and quantitative traits, as previously observed for HSP90.

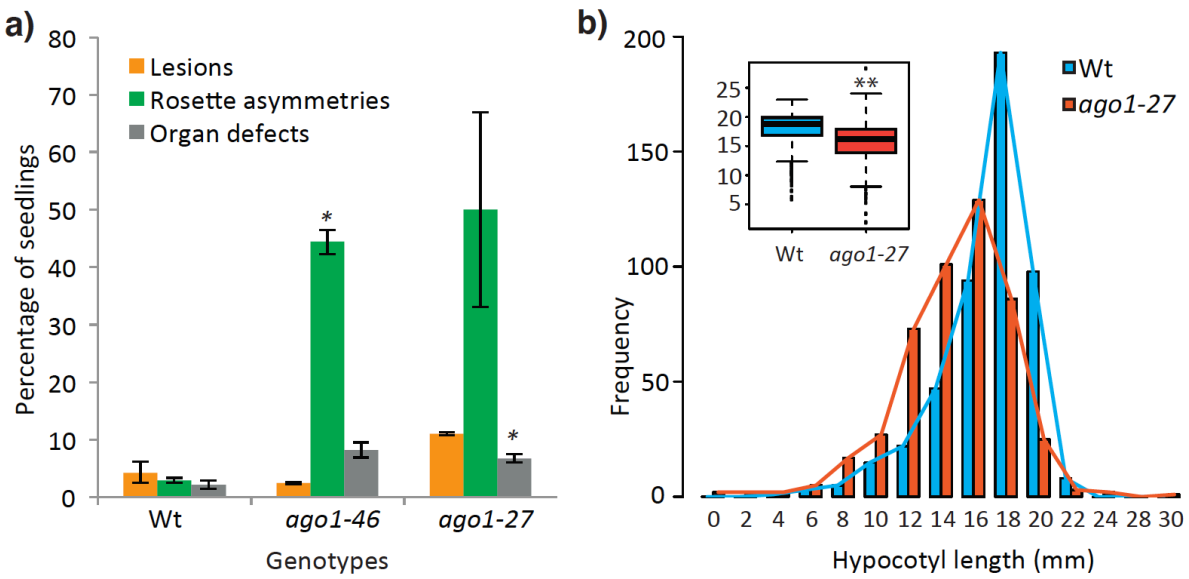


Figure 5.1. *ago1* plants show more variation in qualitative and quantitative traits. **(a)** Early morphology trait measures for wild type (Wt), *ago1-46*, and *ago1-27* seedlings. Ten-day old seedlings were scored for three different morphological traits. The data represent two biological replicates ($n = 144$ for *ago1* mutants, and $n = 216$ for Wt, * = $p < 0.05$, χ^2 - test). **(b)** Hypocotyl length for 7-day old, dark-grown seedlings. The variance for Wt and *ago1-27* distributions is different (Levene's test $p < 0.001$; $n = 475$ for *ago1-27*, $n = 486$ for Wt). **Inset:** boxplots of hypocotyl length distribution. Y-axis represents hypocotyl length (mm), ** = $p < 2.2e^{-16}$, Mann-Whitney Wilcoxon test.

5.3.2 AGO1 perturbation can reveal and conceal genetic variation

HSP90 buffers both developmental noise and genetic variation (Queitsch et al. 2002; Sangster et al. 2008b). Therefore, I next addressed if *AGO1* perturbation could reveal genetic variation and whether *ago1*-dependent loci overlap or differ from known HSP90-dependent loci. To do so, I crossed the *ago1-27* mutant (Col-0 background) to STAIRS lines (Koumproglou et al. 2002)(**Figure S5.1**) and analyzed the phenotypes of their segregating offspring. The rationale to cross the *ago1-27* mutant to the STAIRS line is that the resulting offspring should contain recently acquired genotypic differences that have not had time to evolve new network regulations to compensate for these differences, and therefore depend on capacitors like HSP90 to attain organismal robustness.

STAIRS lines exist for chromosomes 1, 3 and 5. Since *AGO1* is located on chromosome 1, I excluded the respective STAIRS lines from our analysis. For chromosomes 3 and 5, I selected two STAIRS lines each (chr3; N9448 and N9459, chr5; N9472 and N9501). There are three HSP90-dependent loci in the Ler region on the selected STAIRS lines that affect the selected phenotypes (Sangster et al. 2008a; Sangster et al. 2008b). I selected the following genotypes in the F_2 generation: wild type in the Col-0 background, *ago1-27* in the Col-0 background, wild type in the STAIRS background (hereafter referred to as L*er* background), and *ago1-27* in the STAIRS (L*er*) background, controlling for possible recombination events in the introgressed L*er* segment. For these four genotypes, I measured hypocotyl and root length, rosette diameter, and the closely correlated flowering time traits days to flowering and rosette leaf number. I selected these phenotypes because they are readily measurable and show evidence of Hsp90 buffered variation in previous studies, enabling the planned comparisons of HSP90-dependent and putative AGO1-dependent variation.

AGO1 reveals genetic variation affecting a quantitative trait X if one of these scenarios is met:

1) a) wild-type plants of Col-0 and L*er* backgrounds are similar in X, ($X_{\text{Col-0 wt}} - X_{L\textit{er}\text{ wt}} \approx 0$), and b) Col-0 *ago1-27* and L*er ago1-27* plants differ significantly in X ($X_{\text{Col-0 }\textit{ago1-27}} - X_{L\textit{er ago1-27}} \neq 0$).

2) a) wild type plants of Col-0 and L*er* differ significantly in X phenotypic, ($X_{\text{Col-0 wt}} - X_{L\textit{er}\text{ wt}} \neq 0$), yet b) Col-0 *ago1-27* and L*er ago1-27* plants show a significantly different relationship than observed for wild-types, *e.g.* different sign ($X_{\text{Col-0 wt}} - X_{\text{Col-0 }\textit{ago1-27}}$) ≠ ($X_{L\textit{er}\text{ wt}} - X_{L\textit{er ago1-27}}$).

Fundamentally, I assume that *AGO1* perturbation may alter the contributions of underlying genetic variation to a given quantitative trait. Therefore, revealing variation (*i.e.* increasing the contribution of a particular locus to a quantitative trait with a certain overall genetic contribution) should go hand in hand with concealing other variation (*i.e.* decreasing the contribution of other loci). Indeed, this phenomenon has been previously observed for HSP90 perturbation across many traits in *A. thaliana* recombinant inbred lines (Sangster et al. 2008a). If *AGO1* conceals genetic variation, I expect that for a quantitative trait X a) the wild type plants of both backgrounds show significantly different values in X ($X_{\text{Col-0 wt}} - X_{L\textit{er}\text{ wt}} \neq 0$), yet b) Col-0 *ago1-27* and L*er ago1-27* have similar X phenotypic values ($X_{\text{Col-0 }\textit{ago1-27}} - X_{L\textit{er ago1-27}} \approx 0$). Lastly, I may observe that the L*er* introgression acts epistatically and suppresses phenotypic differences between $X_{\text{Col-0 wt}}$ and $X_{\text{Col-0 }\textit{ago1-27}}$.

Indeed, I observe all three scenarios of epistasis (**Table 5.1**). Despite the mounting evidence that HSP90 facilitates some aspects of AGO1 function in many organisms, including plants, I did not observe an overlap of HSP90 and AGO1-dependent loci (Sangster et al. 2008a) (**Figure S5.2**). This apparent lack of HSP90-AGO1 epistasis agrees with the complexity we observed when analyzing double mutants for morphological traits and gene expression phenotypes. Although my findings could indicate that AGO1 and HSP90 simply buffer different variants, this interpretation is compromised by the fact that both HSP90 previous work (Sangster et al. 2008a) and my analyses used partial mutants. Hence, my results may be the consequence of differential sensitivity such that identified HSP90-dependent loci are those most strongly and directly affected by HSP90 perturbation whereas I may miss those loci that are indirectly affected because perturbation of the chaperone decreases AGO1 function.

Table 5.1. AGO1 buffering of genetic variation is common.

Trait	L*er* on Chr 3 1 - 9,742,000	L*er* on Chr 3 18,230,000 - 23,459,830	Ler on Chr 5 1 - 9,479,000	Ler on Chr 5 9,479,000 - 26,975,502
Rosette leaf number	epistatic	--	reveals	conceals
Days to flowering	conceals	epistatic	epistatic	--
Rosette diameter	reveals	conceals	--	--
Hypocotyl length	--	--	--	--
Root length	reveals	--	--	--

The portions of the Ler chromosome in each stair line are indicated. Red marks instances in which *ago1-27* reveals genetic variation; 'epistatic' means instances in which the Ler background was epistatic to the *ago1-27* mutation for the trait tested.

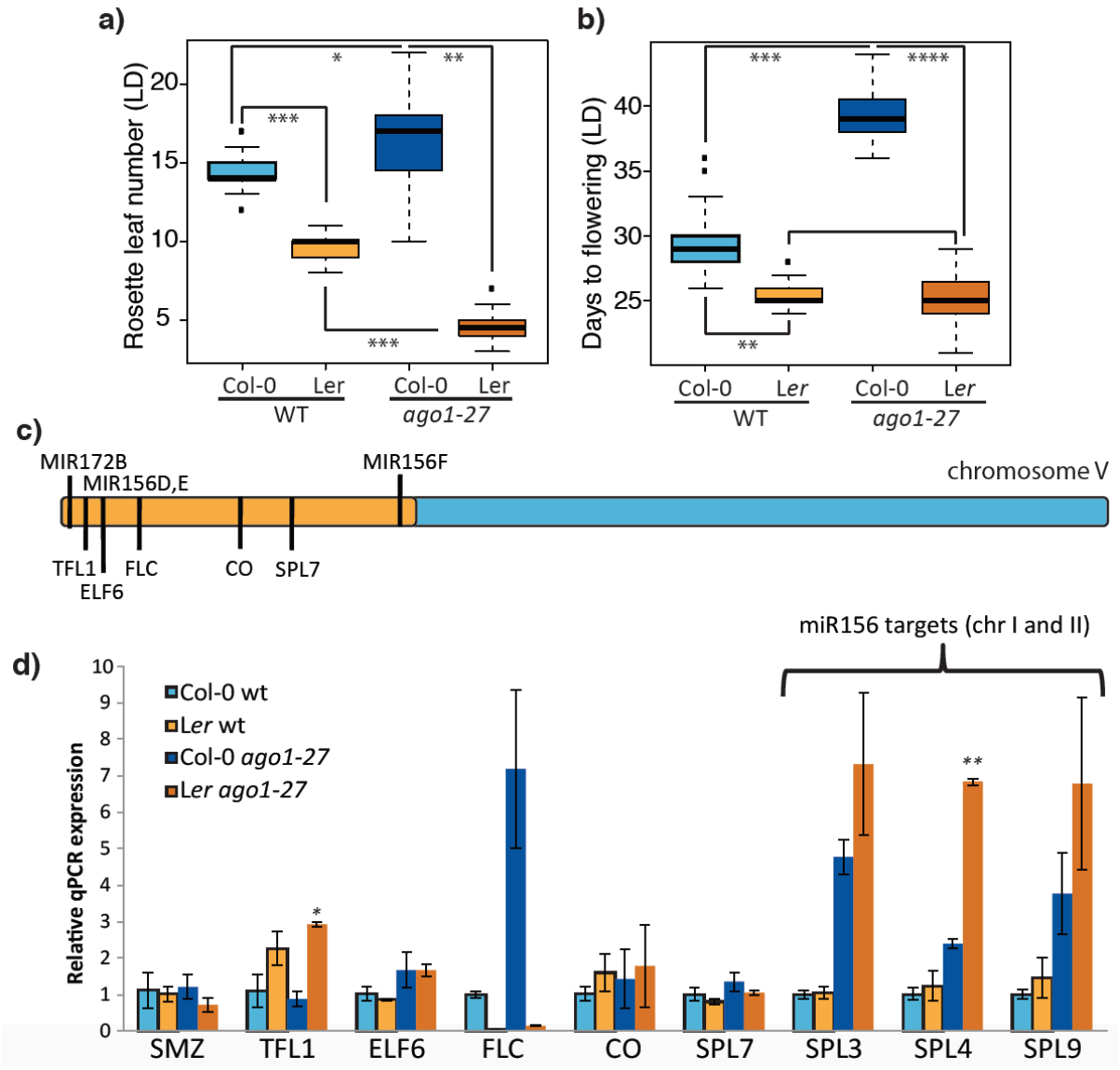


Figure 5.2. Looking for candidate genes to explain the decoupling of days to flowering and rosette leaf number. Plants for wild type Col-0, wild type L*er*, Col-0 *ago1-27* and Ler *ago1-27* were grown on soil in LD at 23˚C, n = 30-36. **(a)** Rosette leaf number. * = $p < 0.0001$, ** = $p < 0.5e^{-11}$, *** = $p < 0.3e^{-11}$, Mann-Whitney Wilcoxon test. **(b)** Days to flowering . ** = $p < 0.5e^{-10}$, *** = $p < 5.5e^{-12}$, **** = $p < 4.5e^{-12}$, Mann-Whitney Wilcoxon test. **(c)** Approximate location of candidate genes in chr 5 involved in the decoupling of days to flowering and rosette leaf number. **(d)** miR156-target SPL gene expression is upregulated in the Ler *ago1-27 r* relative to Col-0 *ago1-27*. 14-day old plant tissue was collected at ZT16 (Zeitgeber 16; 16 hours after the plants sensed the light). Mean expression data represent

two biological replicates, each with three technical replicates. Standard error is pictured. * = $p < 0.5$, ** = $p < 0.005$, T-test).

5.3.3 ago1 perturbation uncouples two closely related flowering time traits in a background-specific manner

One example of AGO1-dependent genetic variation particularly captured my attention because in this case AGO1 perturbation uncoupled the typically closely linked flowering time traits rosette leaf number and days to flowering. The close link between these traits makes intuitive sense, as the resources for flower and seed development are generated in rosette leaves. In *A. thaliana*, days to flowering and rosette leaf number are so closely linked that they are often used interchangeably to measure the onset of flowering (Lempe et al. 2005; Salomé et al. 2011). There are only a few reported cases in which such uncoupling has been observed, involving certain early flowering mutants (Pouteau et al. 2004) and in response to a particular environmental perturbation (Takahashi and Morikawa 2014); however, the mechanistic underpinnings remain unknown. Therefore, I decided to focus on this particular example of AGO1-dependent variation.

Specifically, *ago1-27* reveals genetic variation localized on L*er* chr5 with the coordinates 1 – 9,479,000 bp (**Figure 5.2**), adhering to the second of the above discussed scenarios. Rosette leaf number differed significantly between Col wild-type and L*er* wild type ($p = 2.596e^{-12}$, Wilcoxon test), which is consistent with the known disruption of *FLC* in L*er* wild-type (Michaels et al. 2003; Liu et al. 2004b). *FLC* is a strong repressor of flowering (Song et al. 2013; Song et al. 2014). Rosette leaf numbers also differed significantly for Col *ago1-27* and L*er ago1-27* ($p = 3.449e^{-12}$, Wilcoxon test) with L*er ago1-27* further decreasing rosette leaf number. In contrast, we observed an increase in rosette leaf number in Col-0 *ago1-27* (**Figure 5.2a**). This strong leaf number phenotype – comparable to many known early flowering mutants such as *elf3* – was highly trait-specific. Using the measure of 'days to flowering', L*er* wild type and L*er ago1-27* were undistinguishable ($p = 0.4714$, Wilcoxon

test) (**Figure 5.2b**). *AGO1* perturbation specifically affected leaf number in a background-specific manner, uncoupling these two closely correlated traits.

5.3.4 Identifying the molecular mechanism(s) that could uncouple rosette leaf number and days to flowering.

To identify the polymorphic gene(s) underlying the AGO1-dependent uncoupling of rosette leaf number and days to flowering, I focused on genes with known functions in regulating flowering time (Song et al. 2013; Spanudakis and Jackson 2014; Song et al. 2014) and polymorphisms between Col-0 and L*er* (Cao et al. 2011) (**Figure 5.2b** and **Table S5.2**).

Using qPCR, I measured expression of these candidate genes among the four genotypes; for the MIR156 genes, and MIR172B I measured expression of major target genes. As expected, *FLC* was barely detectable in L*er* wild type and Ler *ago1-27* (**Figure 5.2c**), consistent with the known disruption of Ler *FLC* (Michaels et al. 2003; Liu et al. 2004b). In contrast, *FLC* expression was increased in the Col-0 *ago1-27* relative to Col-0 wild-type, which agreed with the late flowering phenotype of Col-0 *ago1-27* in both traits. As a general trend, I observed that target genes of mir156, which are involved in ageing flowering pathways, increased in expression in L*er ago1-27* relative to Col-0 *ago1-27*.

Specifically, the *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL*) genes targeted by miR156 were upregulated in L*er ago1-27* compared to Col-0 *ago1-27* ($p = 0.003215$, T-test for SPL4), whereas the mir172b-targeted gene *SMZ* showed a slight decrease in expression (**Figure 5.2c**). Mature miR156 levels are high in young seedlings and decrease as the plants get older (Huijser and Schmid 2011; Spanudakis and Jackson 2014), allowing SPLs expression levels to increase (Huijser and Schmid 2011; Spanudakis and Jackson 2014). I hypothesized that the elevated expression levels of miR156-targeted SPLs in Ler *ago1-27* indicates differential expression or action of miR156, which in turn yields this genotype's extremely low leaf number.

The SPL gene family is comprised of 15 genes, of which 11 are targeted by miR156 (Preston and Hileman 2013). To further explore our hypothesis, I measured the expression of 10 miR156-SPL target genes (**Figure 5.3a**), and observed that *SPL5* and *SPL15* expression was increased in Ler *ago1-27* relative to Col-0 *ago1-27*. *SPL5* belongs to the same clade as *SPL3* and *SPL4*, whereas *SPL15* belongs to the same SPL clade as *SPL9* (Preston and Hileman 2013; Poethig 2013), and acts redundantly with *SPL9* (Poethig 2013; Wang et al. 2008). Of all tested SPL genes, only *SPL3*, *SPL4*, *SPL5*, *SPL9* and *SPL15* showed an increase in expression in L*er ago1-27* relative to Col-0 *ago1-27*; these genes are all miR156 targets and contribute to the aging pathway. As I describe in more detail below, of the MIR156 genes, only MIR156f is polymorphic between Col-0 and L*er*.

I also measured levels of the major floral integrator *FT*. *FT* expression was seven-fold higher in L*er ago1-27* than Col-0 *ago1-27* (**Figure S5.3**), but *FT* expression levels in L*er* wild type and Ler *ago1-27* were similar (p = 0.8308, T-test), which agreed with our finding that these genotypes did not differ in days to flowering.

TERMINAL FLOWER 1 (*TFL1*) was upregulated in L*er ago1-27* relative to Col-0 *ago1-27* (p = 0.042, T-test) (**Figure 5.2**), but *TFL1* expression was only subtly higher in L*er ago1-27* compared to L*er* wild type (p = 0.3788, T-test). This result is consistent with a recent study (Fernández-Nohales et al. 2014). *TFL1* prevents the shoot apical meristem from assuming an inflorescence meristem identity by repressing the expression of *LFY* and *AP1* (Liu et al. 2013; Wickland and Hanzawa 2015). Increased *TFL1* expression should translate into a larger number of rosette leaves, unless this effect is negated by other pathway connections.

In summary, my expression analyses suggest that the ageing pathway, and possibly polymorphism(s) in MIR156F, may be responsible for the AGO1-dependent decrease in rosette leaf number in the L*er*-background.

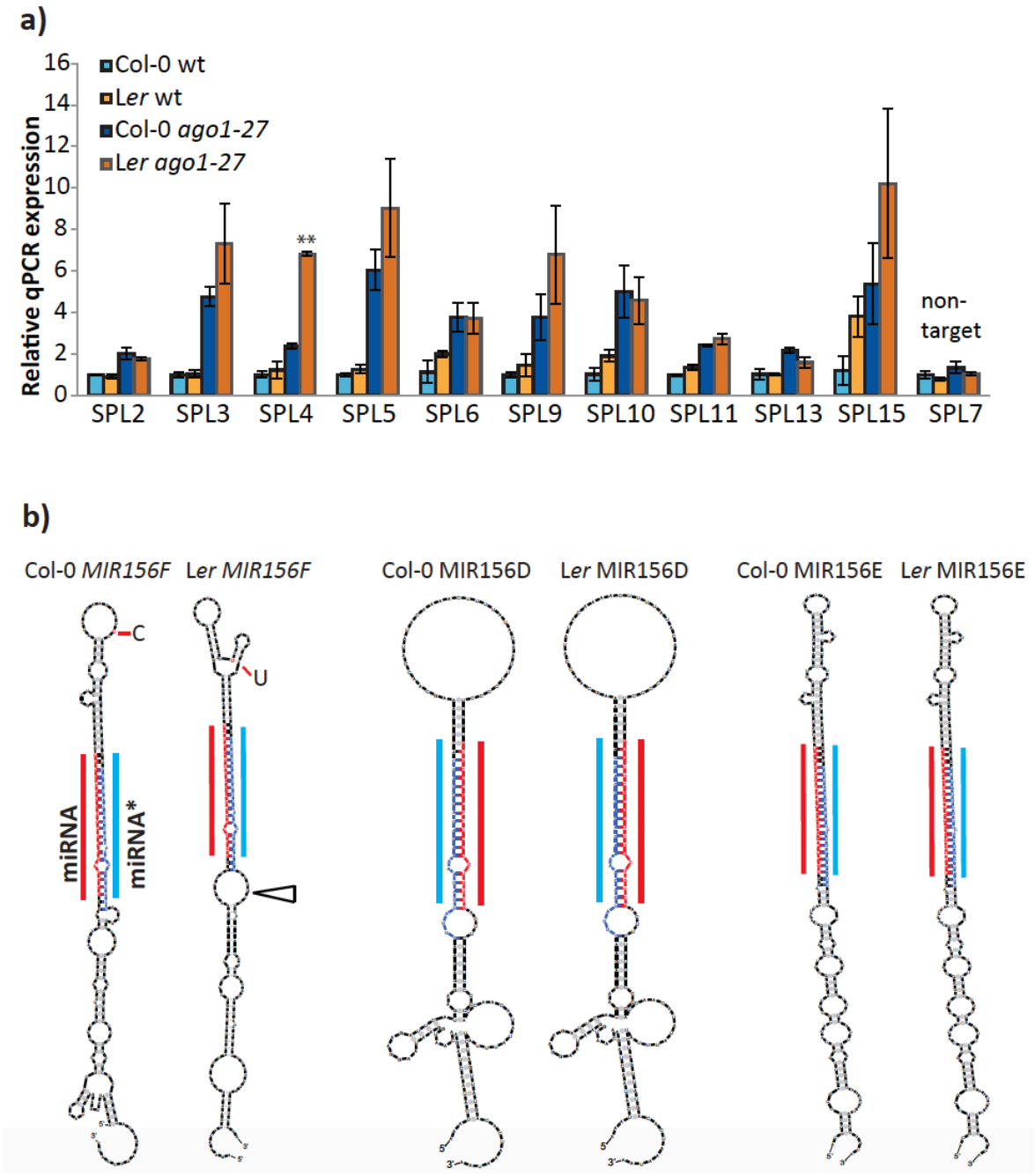


Figure 5.3. Assessing the effect of miR156 in the decoupling of days to flowering and rosette leaf number. (a) miRNA-target SPLs gene expression is higher in Col *ago1-27* than in L*er ago1-27*. SPL7 is not a miR156 target, and its expression was measured as a control. 14-day old plant tissue was collected at ZT16. Mean expression data represent two biological replicates, each with three technical replicates. ** = $p < 0.005$, T-test. **(b)** RNA structures of MIR156F, MIR156D, and MIR156E transcript sequences plus 50 nucleotides at both sides.

The structures were obtained using the program Sfold2.2, the colors were edited in Adobe Illustrator CS 2014. Red represents miR156, blue represents miR156*. The SNP between Col-0 and L*er* is indicated. The triangle represents a 14-nt deletion.

5.3.5 Genotyping of the MIR156F locus.

I sequenced Col-0, STAIRS and L*er MIR156F* to verify the annotated single nucleotide polymorphism (SNP) (**Figure S5.4**). I found that the STAIRS and L*er MIR156F* not only carried the annotated SNP, but also carried had a 14-nt deletion near the 3' end (**Figure S5.4a**). As the *MIR156* gene family is highly conserved in the plant kingdom (Cuperus et al. 2011; Luo et al. 2013); I wanted to examine the degree of its natural variation among other *A. thaliana* strains and Josh Cuperus helped me to sequence an additional 55 strains. Of all 57 strains, 42 carried the L*er*-specific C-to-T SNP, one carried a C-to-G SNP, and 32 carried the 14-nt deletion (**Figure S5.4b**). The presence or absence of the deletion was highly correlated with the presence or absence of the SNP ($R^2 = 0.3506$, $p = 0.0007$, Pearson correlation test); but not all strains carrying the SNP showed the deletion. Among the tested strains, there was no correlation of either polymorphism with rosette leaf number (data from (Lempe et al. 2005).

After discovering an unknown, but conserved deletion in MIR156F, I sequenced Col-0 and L*er MIR156D* and Josh Cuperus sequenced *MIR156E* to identify other potentially unannotated polymorphisms, which could contribute to the observed expression and flowering time phenotypes. We did not find any polymorphisms between the Col-0 and L*er* in these loci (**Figure 5.3)**.

5.3.6 MIR156F polymorphisms alter the predicted secondary structure of the miR156f precursor.

Polymorphisms residing in the hairpin structure of pre-miRNAs may alter secondary structure, which may translate into defects in miRNA biogenesis and hence altered accumulation of the mature miRNA (Bologna et al. 2013; Cuperus et al. 2010; Todesco et al. 2012). To address whether the C-to-T SNP could influence miR156f processing, I used secondary structure predictions for the miR156f precursor (Ding et al. 2004)(**Figure 5.3b**) in addition to miR156d and miR156e precursors. According to these predictions, the L*er* miR156f precursor forms a different loop structure than Col-0 miR156f (**Figure 5.3a**). Bologna *et al.* suggest that members of the miR156 family are processed from loop-to-base (Bologna et al. 2013). Therefore, I hypothesized that the altered L*er* miR156f loop structure subtly impairs miR156f processing. In combination with decreased AGO1 function, this subtle defect will decrease the overall miR156 pool, and increase miR156-targeted SPLs gene expression, which in turn will lead to decreased leaf number.

I tested this hypothesis by measuring expression of mature miR156 [miR156(a-f)] levels among the different genotypes using qPCR (Varkonyi-Gasic et al. 2007). As I hypothesized, I detected a small but significant reduction in miR156 levels between Col-0 *ago1-27* and L*er ago1-27* ($p = 0.049$, T-test) (**Figure S5.5**). Notably, miR156 levels increased in Col-0 *ago1-27* compared to Col-0 wild-type, possibly due to feed-back regulation sensing perturbed *ago1*-function (Smith et al. 2009). The L*er*-specific processing defects may interfere with this feed-back and further burden the already taxed AGO1.

5.3.7 Genes in the ageing pathway have increased mRNA levels in Ler ago1-27 relative to Col-0 ago1-27.

If the rosette leaf number phenotype in L*er ago1-27* was indeed due to altered action of miR156f through the ageing pathway, members of this pathway, and specifically meristem identity genes, should be upregulated in L*er ago1-27* relative to Col-0 *ago1-27*. To investigate this hypothesis, I measured the expression of the aging pathway genes *FUL*, *LFY*,

AP1, *AP2*, *SMZ*, *SNZ*, *TOE1*, *TOE2* as well as levels of mature miR172b (**Figure S5.6**). Consistent with my hypothesis, I found that *LEAFY* (*LFY*), *APETALA1* (*AP1*), and *FRUTIFUL* (*FUL*) (floral meristem identity genes) were highly upregulated in L*er ago1-27* relative to Col-0 *ago1-27*, whereas SUPRESSOR OF CONSTANS1 (*SOC1)* was not. As one may expect in the light of non-functional *FLC* in Ler, mature miR172b expression was upregulated in L*er ago1-27* relative to Col-0 *ago1-27* (**Figure S5.5**). However, miR172b targets *APETALA2* (*AP2*), *SCHLAFMÜTZE* (*SMZ*), *SCHNARCHZAPFEN* (*SNZ*), *TARGET OF EAT1*, (*TOE1*) and *TARGET OF EAT2* (*TOE2*) showed only a subtle decrease in expression levels, suggesting that miR172b-target genes may not play a major role in *LFY*, *AP1* and *FUL* upregulation. Therefore, I suggest that the miR156-SPL module-mediated upregulation of the floral meristem genes (**Figure S5.6**) is the most likely molecular scenario underlying the observed decrease in rosette leaf number.

5.3.8 FLC as a third epistatic factor contributing to rosette leaf number.

FLC is a major repressor of flowering, exerting its repressive function directly on the integrator gene FT as well as on other pathway members such as mir172b. *FLC* expression was upregulated in Col-0 *ago1-27* relative to L*er ago1-27*, consistent with the L*er*-specific *FLC* defect. The lack of FLC in L*er ago1-27* will increase *FT* expression, which in turn may increase expression of *SPL3* (Yamaguchi et al. 2009) (**Figure S5.6**). SPL3 promotes the expression of *AP1*, *LFY* and *FUL*. Therefore, the observed increase in SPL3 and floral meristem genes may at least in part be due to the lack of functional FLC in L*er ago1-27*. *FLC* expression is repressed when plants undergo vernalization (Andres and Coupland 2012).

To determine the extent to which FLC contributes to the observed phenotypic differences, I vernalized the Col-0 wild type, L*er* wild type, Col-0 *ago1-27* and L*er ago1-27* plants and measured rosette leaf number and days to flowering as before (**Figure 5.4**). If differences in *FLC* expression were fully responsible for the reduced Ler *ago1-27* rosette leaf number, vernalized L*er ago1-27* and Col-0 *ago1-27* plants should show similar numbers of rosette leaves. This was not the case. Although vernalization treatment subtly increased L*er*

ago1-27 rosette leaves, L*er ago1-27* plants still had significantly fewer leaves than Col-0 *ago1-27* ($p = 5.704e^{-12}$) (**Figure 5.4a**). This result indicates that *FLC* contributes to the reduced L*er ago1-27* rosette leaf number, but does not fully account for this phenotype.

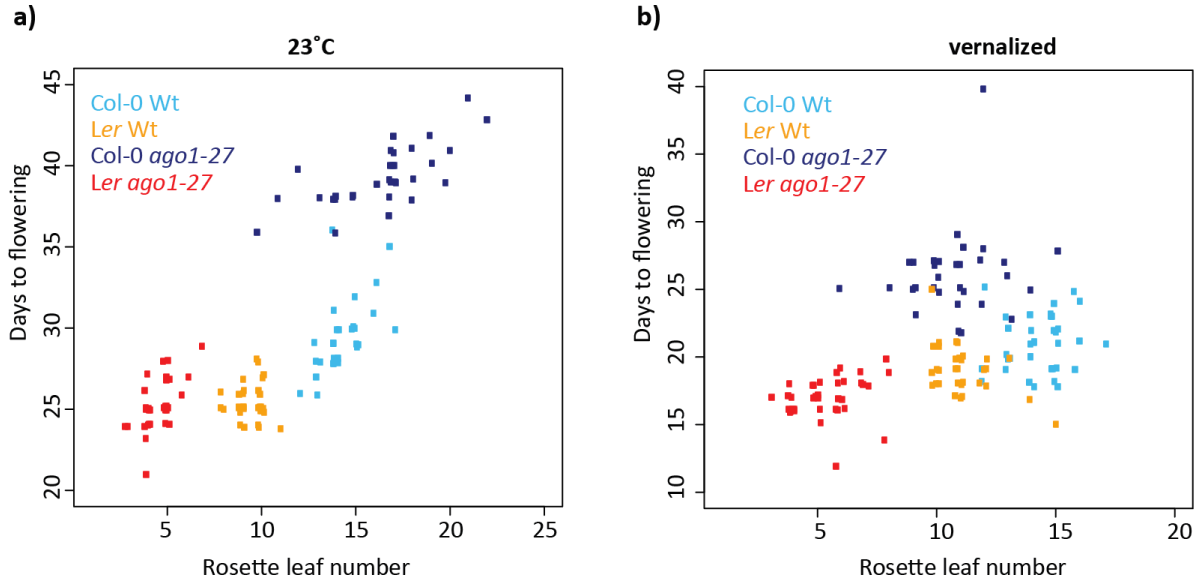


Figure 5.4. Effect of FLC in the decoupling of days to flowering and rosette leaf number. Vernalization treatment partially rescues the decoupling phenotype. **(a)** Plants were grown in LD at 23˚C. (**b**) Plants were grown for 5 days in LD at 23˚C, and then vernalized for 40 days at 4˚C. Days to flowering are the days after vernalization.

This result contrasts with my findings for Col-0. Vernalized Col-0 *ago1-27* plants showed significantly fewer rosette leaves than Col-0 wild type plants (**Figure 5.4b**), whereas in control conditions Col-0 *ago1-27* showed significantly more rosette leaves than Col-0 wild type (**Figures 5.4a,** and **5.2a**). Thus, vernalization treatment sufficed to decrease the number of rosette leaves in Col-0 *ago1-27*, suggesting a strong epistatic relationship between *FLC* and *AGO1*. As rosette leaf number and days to flowering are typically coupled (Lempe et al. 2005; Salomé et al. 2011), I would expect vernalized Col-0 *ago1-27* to flower after fewer days than Col-0 wild type. However, using days to flowering as a measure, Col-0 wild type

flowered earlier than Col-0 *ago1-27* (**Figure 5.4b**). Thus, in the Col-0 background, vernalization uncouples the traits days to flowering from rosette leaf number. In contrast, in the L*er* background, vernalization yields a small, but significant difference between L*er* wild type and Ler *ago1-27* in days to flowering ($p = 2.562e^{-06}$, Wilcoxon) (**Figure 5.3.2a**), whereas under control conditions both genotypes flower at the same time (**Figure 5.3.2b**).

Taken together, my results suggest that in Col-0 the epistasis between *FLC* and *AGO1* is a major player in determining and coupling both flowering traits, days to flowering and rosette leaf number. In L*er*, *FLC* and *AGO1* epistasis contributes to, but is a minor player in, reducing rosette leaf number in L*er ago1-27*. These results point to the involvement of at least one other locus in L*er*, which I hypothesize to be MIR156F.

5.4 DISCUSSION

I provide evidence that *AGO1* perturbation amplifies developmental noise and thereby increases phenotypic variation in isogenic seedlings similar to previous studies examining HSP90 perturbation. Prompted by this finding, I investigated whether AGO1 perturbation may affect the contribution of genetic variation to complex developmental traits. As hypothesized, I indeed observed that AGO1 perturbation reveals and conceals genetic variants; these variants appear to be common, affect several traits, and do not overlap with HSP90-dependent variation affecting the same traits (**Figure S5.2**).

5.4.1 AGO1 buffers developmental noise and genetic variation

Previous studies showed that specific miRNAs buffer noise in the pathways which they target (Sieber et al. 2007; Hilgers et al. 2010; Cassidy et al. 2013; Posadas and Carthew 2014; Siegal and Leu 2014). In contrast, here I asked whether perturbation in *AGO1*, the key enzyme of miRNA metabolism, generally acts in this way and affects multiple phenotypes. Although one may expect the observed results based on the reported specific miRNA defects, the multiple levels at which genetic networks are buffered may suffice to 'hide' the effects of

weak perturbations of *AGO1*. One such factor could be the chaperone HSP90, which is known to facilitate AGO1 functions.

Gene products that buffer developmental noise do not always buffer genetic variation (Hilgers et al. 2010). HSP90 provides robustness to both stochastic and genetic variation (Queitsch et al. 2002; Lempe et al. 2013; Masel and Siegal 2009; Siegal and Leu 2014). Similarly, *AGO1* perturbation can reveal and conceal genetic variation on L*er* chr3 and chr5 for a variety of phenotypes (**Figure S5.2**). The frequency of revealed and concealed genetic variation suggests that *AGO1*-dependent genetic variation is relatively common. In this study, I only tested the interaction of *ago1-27* with L*er* chr3 and 5 polymorphisms in five phenotypes but we assume many more AGO1-dependent loci exist in other *A. thaliana* strains. These polymorphic loci likely represent miRNA genes or their targets. In this regard, it is noteworthy that I found wide-spread variation in MIR156F across strains. Apparently MIR precursors can harbor functionally relevant natural variation, in contrast to previous assumptions (de Meaux et al. 2008; Ehrenreich and Purugganan 2008). Given that my study relied on *ago1-27* for these experiments, it would be worthwhile to test an allelic series of *ago1* mutants in the same way to investigate my speculation that dosage plays a role in the lack of overlap between HSP90-dependent and AGO1-dependent loci. Specifically, my results could be the consequence of differential sensitivity such that the identified AGO1-dependent loci are those most strongly and directly affected by AGO1 perturbation; these loci are unlikely to be the ones that are most strongly and directly affected by HSP90.

Regardless of the precise mechanistic reasons for the missing overlap between AGO1 and HSP90, our findings are relevant for understanding rules for buffering. Examples of trait-specific buffering are well-known (Hilgers et al. 2010); likewise the fact that HSP90 buffers robustness of many but not all traits is well-established. With AGO1 we have identified yet another generally acting mechanism, which shows specificity as to the revealed and concealed variants in the traits affected by HSP90. Thus, HSP90 and AGO1 maintain robustness in some of the same traits, yet do so through buffering different genetic variants. This redundancy in buffering maintains organismal robustness to some extent in the face of environmental perturbations that specifically affect one or the other network hub– at the

same time it enables such perturbations to release specific variants in response to a particular perturbation, potentially tailoring evolutionary trajectories closely to specific environmental challenges.

5.4.2 ago1-27 in Ler background uncouples days to flowering from rosette leaf number.

Days to flower and rosette leaf number are two highly correlated traits, which are often used interchangeably (Lempe et al. 2005; Salomé et al. 2011). *ago1-27* epistatically interacted with L*er* polymorphisms residing on the top of chr5 exclusively for rosette leaf number, uncoupling both traits. *A. thaliana* is an annual plant, for which it is critical to flower with sufficient resources to maximize seed amount and quality. Clearly, *Ler ago1-27* plants have fewer leaves, less biomass, and many fewer seeds than wild type *Ler* (data not shown). Col-0 *ago1-27* plants also produce fewer seeds than wild type plants (Morel et al. 2002). The AGO1-dependent variation in leaf number and flowering time and my detailed study of the participating network and complexity contributing to both traits and their relationship have deep agronomical implications for plant yield. Breeders should consider MIR genes as promising targets for trait manipulation (Wang and Wang 2015).

5.4.3 The involvement of ageing pathways in decreasing leaf number

I found that miR156-targeted SPLs involved in the ageing pathway (*SPL3*, *SPL4*, *SPL5*, *SPL9* and *SPL15*) (Spanudakis and Jackson 2014; Huijser and Schmid 2011; Poethig 2013) were upregulated in L*er ago1-27* relative to Col-0 *ago1-27*. Plants overexpressing *SPL3* and *SPL4* flower early with fewer rosette leaves (Wu and Poethig 2006; Yamaguchi et al. 2009). *SPL9* participates in the regulation of the plastochron (Wang et al. 2008), and *spl9* and *spl15* double mutants have more rosette leaves than wild type plants (Wang et al. 2008). Transgenic Arabidopsis harboring a highly expressed *MIR156F* flower later, with more rosette leaves than wild type (Wu and Poethig 2006; Wang et al. 2008). These results are consistent with my findings and conclusions.

Moreover, *FUL*, *LFY*, and *AP1* were tremendously upregulated in L*er ago1-27* compared to Col-0 *ago1-27* (**Figure S5.6**). *FUL*, *LFY*, and *AP1* are meristem identity genes (Moyroud et al. 2010; Kaufmann et al. 2010; Pajoro et al. 2014) that promote the transition of the vegetative shoot apical meristem to the reproductive inflorescence meristem (Pajoro et al. 2014; Andres and Coupland 2012). Consistent with this function, overexpression of *FUL* produces plants that flower with fewer rosette leaves (Kim et al. 2012). My extensive expression analysis supports my hypothesis that MIR156F contributes strongly to reduced leaf number in L*er ago1-27* by increasing the expression of meristem identity genes.

MIR156A and *MIR156C*, but not *MIR156F*, are associated with the sugar-regulated delay in flowering time (Yang et al. 2013; Yu et al. 2013). Therefore, *MIR156A* and *MIR156C* were hypothesized to be the main *MIR156* genes in the ageing pathway. However, *MIR156F* expression does show response to high glucose concentrations (Yang et al. 2013; Yu et al. 2013), suggesting it, too, functions in the sugar-regulated delay to flowering time. My study provides powerful evidence that MIR156F is likely a major player in the ageing pathway as subtle polymorphisms can produce significant phenotypic change in the *ago1-* mutant background. Although I have not tested the four genotypes in various sugar concentrations, this may be a worthwhile experiment for future studies.

5.4.4 Polymorphisms in MIR156F may disrupt MIR156F processing, decreasing the miR156 pool.

In addition to confirming the annotated C-to-T SNP, I found a 14-nt deletion in the 3' end of *MIR156F*. This 14-nt deletion was present in 32 of 57 strains assayed. Notably, the 14-nt deletion resides within a CT dinucleotide tandem repeat. Short tandem repeats have high mutation rates and are likely to vary among strains, unless they are subject to purifying selection (Undurraga et al. 2012; Rival et al. 2014; Press et al. 2014; Carlson et al. 2015). As I observe multiple different alleles of this CT repeat among the sequenced strains, strong purifying selection is unlikely (Rival et al. 2014; Press et al. 2014). Short tandem repeat variation is often missed with short-read sequencing (Press et al. 2014; Carlson et al. 2015).

Polymorphisms in miRNA precursors can influence miRNA biogenesis and affect levels of their mature forms (Mateos et al. 2010; Cuperus et al. 2010; Todesco et al. 2012). According to Bologna *et al.*, (Bologna et al. 2013) the *MIR156* gene family is processed from the distal loop (in which the C-to-T SNP is situated) to the stem, where the miRNA/miRNA* duplex is located. Polymorphisms in *MIR156F* could impair miR156f processing and reduce overall miR156 levels. Consistent with this hypothesis, I observed lower miR156 levels in L*er ago1-27* relative to Col *ago1-27* (**Figure S5.5)**. Future studies would include an in-depth assessment of the effects of both the SNP and the 14-nt deletion on mature miR156 processing in transgenic plants. One could construct transgenic plants with either the SNP or the 14-nt deletion to test which polymorphism has a larger effect on the phenotype.

5.4.5 FLC and MIR156F contribute to rosette leaf number in a background-specific manner.

FLC polymorphisms are major contributors to differences in flowering time, measured both in days to flowering and rosette leaf number, among different *A. thaliana* strains (Michaels et al. 2003; Werner et al. 2005; Salomé et al. 2011). I investigated the role of *FLC* in the L*er ago1-27* rosette leaf number trait by vernalizing plants of all four genotypes. Vernalization treatment represses FLC transcription (Andres and Coupland 2012). If the lack of *FLC* were solely responsible for the reduction of L*er ago1-27* rosette leaf number, Ler *ago1-27* and Col-0 *ago1-27* should have similar leaf numbers after vernalization. This was not the case (**Figure 5.4a**). The significant difference in rosette leaf number between L*er ago1-27* and Col-0 *ago1-27* persisted, although L*er ago1-27* presented on average 2-3 more rosette leaves in response to vernalization. I conclude that both the *MIR156F* and *FLC* polymorphism contribute to the drastic reduction of L*er ago1-27* rosette leaf numbers. In stark contrast, vernalization reduced leaf number in Col-0 *ago1-27* compared to wild type and uncoupled both flowering traits. My analysis reveals dramatic differences among these closely related genotypes in both genetic and environmental interactions. As others have previously observed strong background-specific effects for a different kind of polymorphism (Undurraga et al. 2012), studies in multiple strains of *A. thaliana* may be required to derive universal pathway principles.

5.5 SUPPLEMENTARY INFORMATION

Supplementary results

Genetic background affects the response of ago1-27 to different temperatures.

To determine if there were environmental conditions in which L*er* wild-type would behave like L*er ago1-27* with regard to rosette leaf number, I grew all genotypes at 16˚C, 23˚C and 27˚C. These temperatures were selected because Col-0 plants increase rosette leaf number when grown at 16˚C (Kim et al. 2012), and several miRNAs accumulate to higher levels at this temperature (Kim et al. 2012). In addition, plants growing at 27˚C tend to flower earlier, with fewer leaves (Chen 2004), and these phenotypes are correlated with increased levels of MIR172 transcripts (Lee et al. 2010).

In general, my results recapitulated these known trends: plants increased in rosette leaf number at 16˚C and decreased at 27˚C (**Figure S5.7a**). None of the temperature conditions tested significantly altered the difference between L*er* wild type and Ler *ago1-27.* However, I observed that L*er ago1-27* had a limited range of variation in rosette leaf number compared with Col-0 *ago1-27* (**Figure S5.7a**). Rosette leaf number in Col-0 *ago1-27* ranges from 7 to 29 at 16˚C, and from 8 to 22 at 27˚C, whereas *Ler ago1-27* presents a range of 5 to 10 rosette leaves across all temperatures. As I have shown, in the Col-0 background, *ago1-27* usually presents a wider phenotypic variation than wild type both in normal conditions and especially when *ago1-27* is subjected to stress conditions (Chapter 4 and this chapter). The fact that *ago1-27* in the L*er* background does not follow this behavior suggests that different genetic backgrounds and traits have different susceptibilities to AGO1 perturbation. Regarding days to flower, Col-0 *ago1-27* and Ler *ago1-27* had similar responses to temperature stresses (**Figure S5.7b**).

Supplementary figures and tables

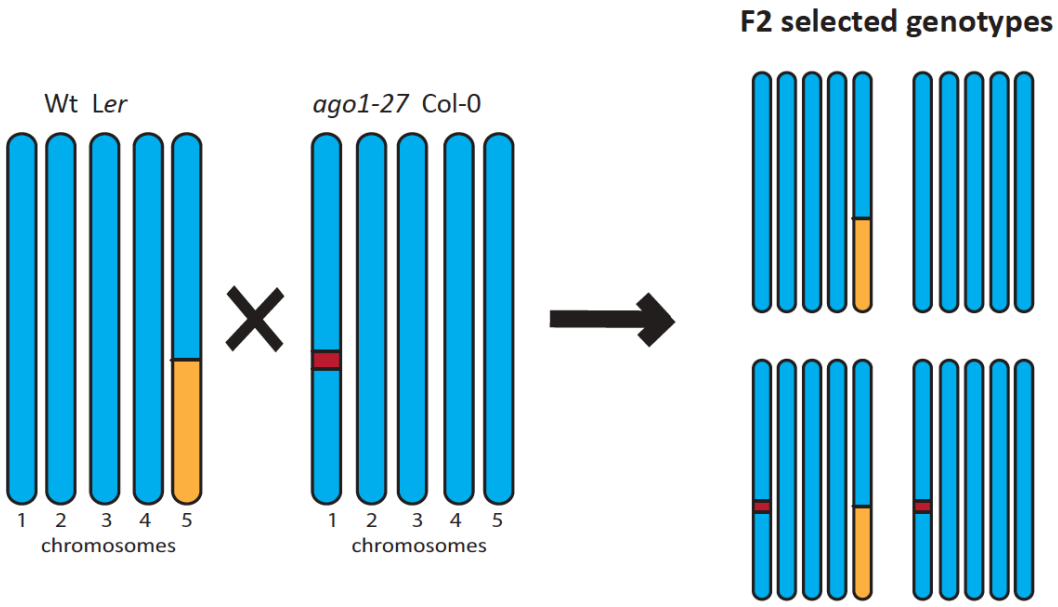


Figure S5.1. Cross scheme used to test AGO1 buffering of genetic variation. Red square represents the approximate location of the *ago1-27* allele.

a)

P -values for Wilcoxon tests

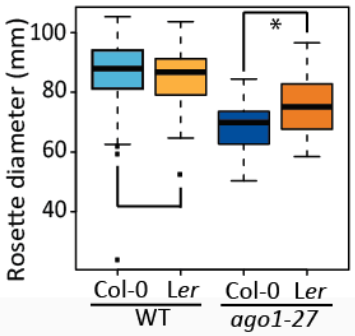
STAIRS 9448	Chr3: 18,230,000 - 23,459,830					
Phenotype	Wt comparison Col-0 vs Ler	*ago1-27* comparison Col-0 vs *Ler*	Col-0 comparison Wt vs *ago1-27*	Ler comparison Wt vs *ago1-27*	Interaction	Overlap with HSP90-loci for the selected trait?
Hypocoyl length	*1.74E-02*	*1.06E-04*	1.34E-01	*1.02E-08*	*2.36E-02*	No
Root length	*1.67E-10*	1.05E-01	*4.66E-13*	*1.55E-06*	*2.07E-06*	No
Days to flowering	***9.67E-03***	8.63E-01	***5.19E-13***	***3.66E-12***	1.15E-01	No
Rosette leaf number	*1.10E-02*	*1.93E-02*	1.83E-01	*1.76E-06*	*5.62E-04*	No
Rosette diameter	***4.06E-02***	3.56E-01	***3.84E-02***	3.68E-01	***2.90E-02***	No

STAIRS 9459	Chr3: 1 - 9,742,000					
Phenotype	Wt comparison Col-0 vs Ler	*ago1-27* comparison Col-0 vs *Ler*	Col-0 comparison Wt vs *ago1-27*	Ler comparison Wt vs *ago1-27*	Interaction	Overlap with HSP90-loci for the selected trait?
Hypocoyl length	*4.40E-14*	*2.09E-04*	*8.90E-07*	*2.35E-05*	7.91E-01	No
Root length	3.06E-01	***5.95E-02***	***< 2.2e-16***	***3.71E-11***	***1.90E-02***	No
Days to flowering	***1.03E-11***	3.45E-01	***2.34E-12***	6.14E-01	***1.73E-02***	No
Rosette leaf number	***7.55E-11***	5.09E-01	2.71E-01	***6.74E-04***	***1.48E-03***	No
Rosette diameter	4.10E-01	***7.12E-03***	***2.85E-09***	***2.33E-03***	***3.96E-02***	No

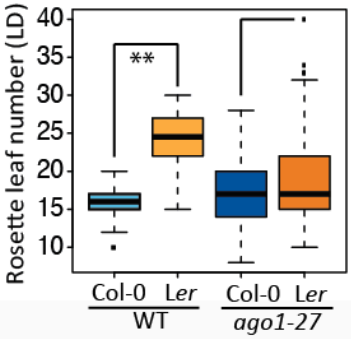
STAIRS 9472	Chr5: 1 - 9,479,000					
Phenotype	Wt comparison Col-0 vs Ler	*ago1-27* comparison Col-0 vs *Ler*	Col-0 comparison Wt vs *ago1-27*	Ler comparison Wt vs *ago1-27*	Interaction	Overlap with HSP90-loci for the selected trait?
Hypocoyl length	*5.65E-02*	*3.89E-02*	*3.29E-12*	*3.35E-10*	*4.73E-03*	No
Root length	*2.41E-08*	*1.42E-02*	*< 2.2E-16*	*3.19E-08*	*9.54E-03*	No
Days to flowering	***5.36E-11***	***4.66E-12***	5.51E-12	4.71E-01	***< 2.2 E -16***	No
Rosette leaf number	***2.60E-12***	***3.45E-12***	***1.02E-04***	***2.33E-12***	***< 2.2E-16***	No
Rosette diameter	*< 2.2E-16*	*< 2.2E-16*	*1.10E-10*	< 2.2E-16	*8.53E-04*	No

STAIRS 9501	Chr5: 9,479,000 - 26,975,502					
Phenotype	Wt comparison Col-0 vs Ler	*ago1-27* comparison Col-0 vs *Ler*	Col-0 comparison Wt vs *ago1-27*	Ler comparison Wt vs *ago1-27*	Interaction	Overlap with HSP90-loci for the selected trait?
Hypocoyl length	***4.50E-07***	3.96E-01	6.37E-01	***6.66E-05***	7.07E-02	Yes
Root length	*2.36E-04*	*3.40E-05*	*1.33E-07*	*1.46E-05*	5.58E-01	Yes
Days to flowering	*1.86E-04*	3.84E-01	*6.10E-12*	*4.06E-13*	1.49E-01	No
Rosette leaf number	***1.28E-03***	***1.08E-02***	***4.55E-04***	1.51E-01	***7.50E-05***	No
Rosette diameter	*3.82E-03*	3.22E-01	*4.48E-16*	*8.94E-12*	1.94E-01	No

b) *ago1-27* reveals on chr3



c) *ago1-27* conceals on chr3



d) L*er* chr3 is epistatic to *ago1-27*

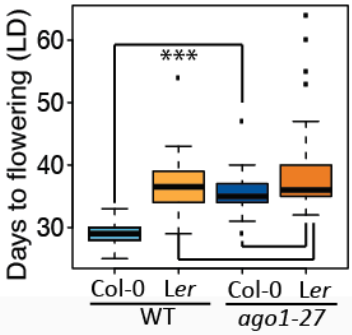


Figure S5.2. *AGO1* perturbation reveals and conceals genetic variation. (a) P-values for the comparisons per line per phenotype measured. Mann-Whitney Wilcoxon tests were used to test significance. The column "interaction" contains the p-value of the ANOVA interaction term genetic-background*mutant status. Red marks instances in which *ago1* reveals genetic variation; blue, when *ago1* conceals genetic variation; and green, cases in which the L*er* introgression was epistatic to *ago1*. For the "HSP90 overlap" column, I searched if the portions of L*er* chromosomes that harbor AGO1-dependent loci overlapped with published HSP90 –dependent loci that affect the same measured traits. All the experiments were done in Long day (LD) (16h light / 8h dark) at 23˚C. **(b)** *ago1* reveals genetic variation for rosette diameter. **(c)** *ago1* conceals genetic variation for rosette leaf number. **(d)** The L*er* introgression is epistatic to *ago1* mutation. Asterisks represent significant p-values. * = $p<0.005$, ** = $p< 0.5e^{-11}$, *** = $p<0.5e^{-12}$, Mann-Whitney Wilcoxon test.

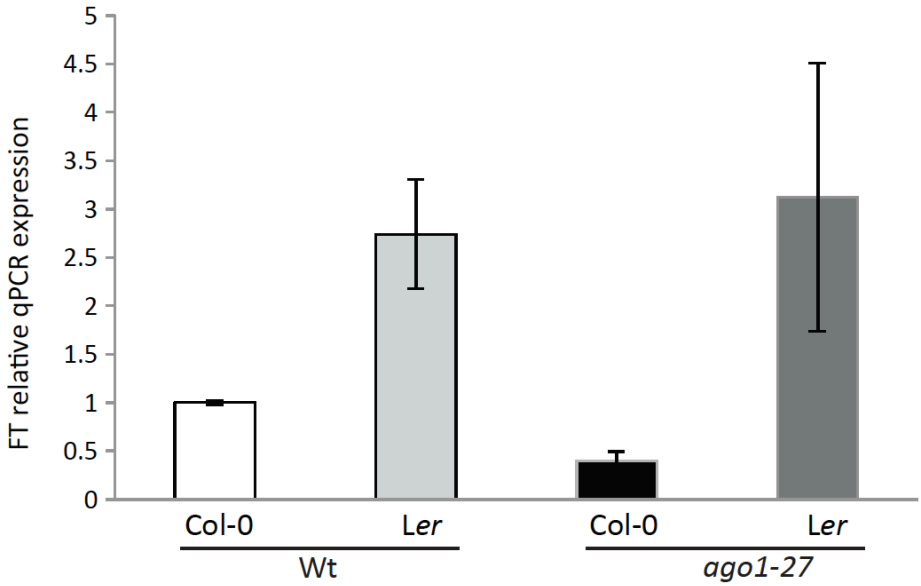


Figure S5.3. FT expression is similar between L*er* wild type L*er* *ago1*-27

FT gene expression for 14-day old seedlings is shown. Plants were grown on soil in LD at 23˚C, and tissue was collected at ZT16. Expression means represent two biological replicates, each one with three technical replicates. Error bars represent standard error.

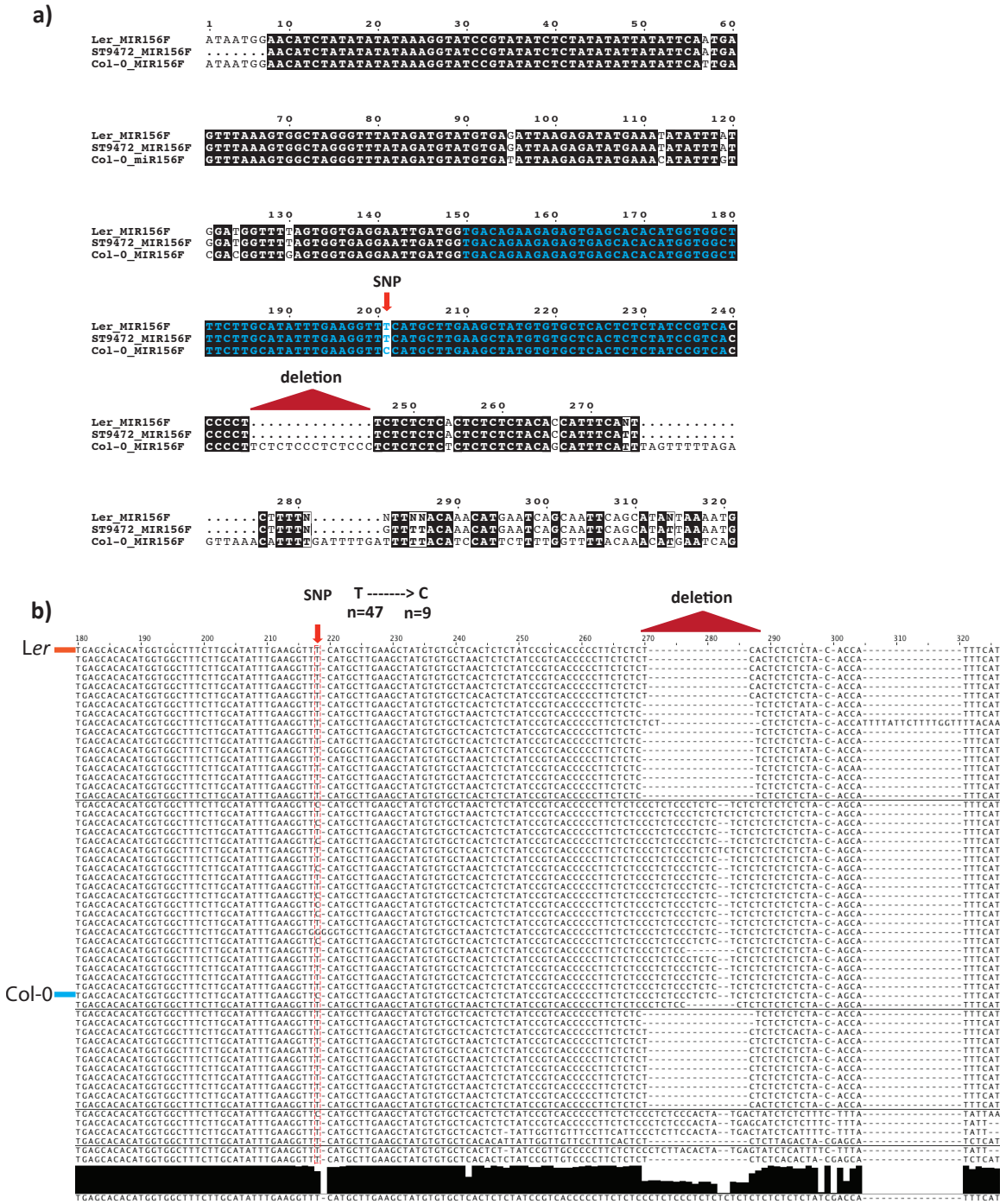


Figure S5.4. Polymorphisms in MIR156F are common. (a) Re-sequencing and alignment of MIR156F in the Col-0, L*er* and the STAIRS N9472 strains. MIR156F sequence is

indicated in blue. The red arrow indicates the position of the SNP between Col-0 and L*er*. The red triangle represents a 14-nt indel. We used T-Coffee to perform the alignments. **(b)** Alignment of MIR156F sequences for other *A. thaliana* strains. Red arrow indicates the position of the SNP. The red triangle represents the 14-nt indel. The orange line points to the L*er* sequence, and the blue line points to Col-0.

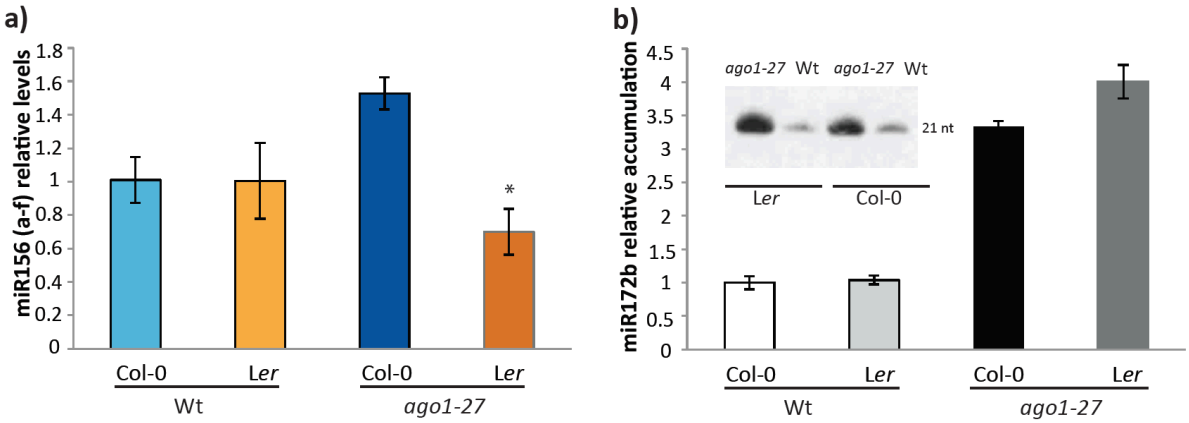


Figure S5.5. miR156 and miR172b levels. (a) miR156(a-f) levels are lower in L*er ago1-27* relative to Col-0 *ago1-27*. 14-day old plant tissue was collected at ZT16. Mean expression data represent two biological replicates, each with three technical replicates. * = $p < 0.05$, T-test. Gene expression is relative to Atsno101 RNA. **(b)** miR172b levels are up-regulated in *ago1* seedlings relative to their wild types. 14-day old plant tissue was collected at ZT16. MiR172b blots were exposed for 1 day and then quantified. Mean expression data represent two biological replicates. RNA levels were normalized to Col-0 wild type.

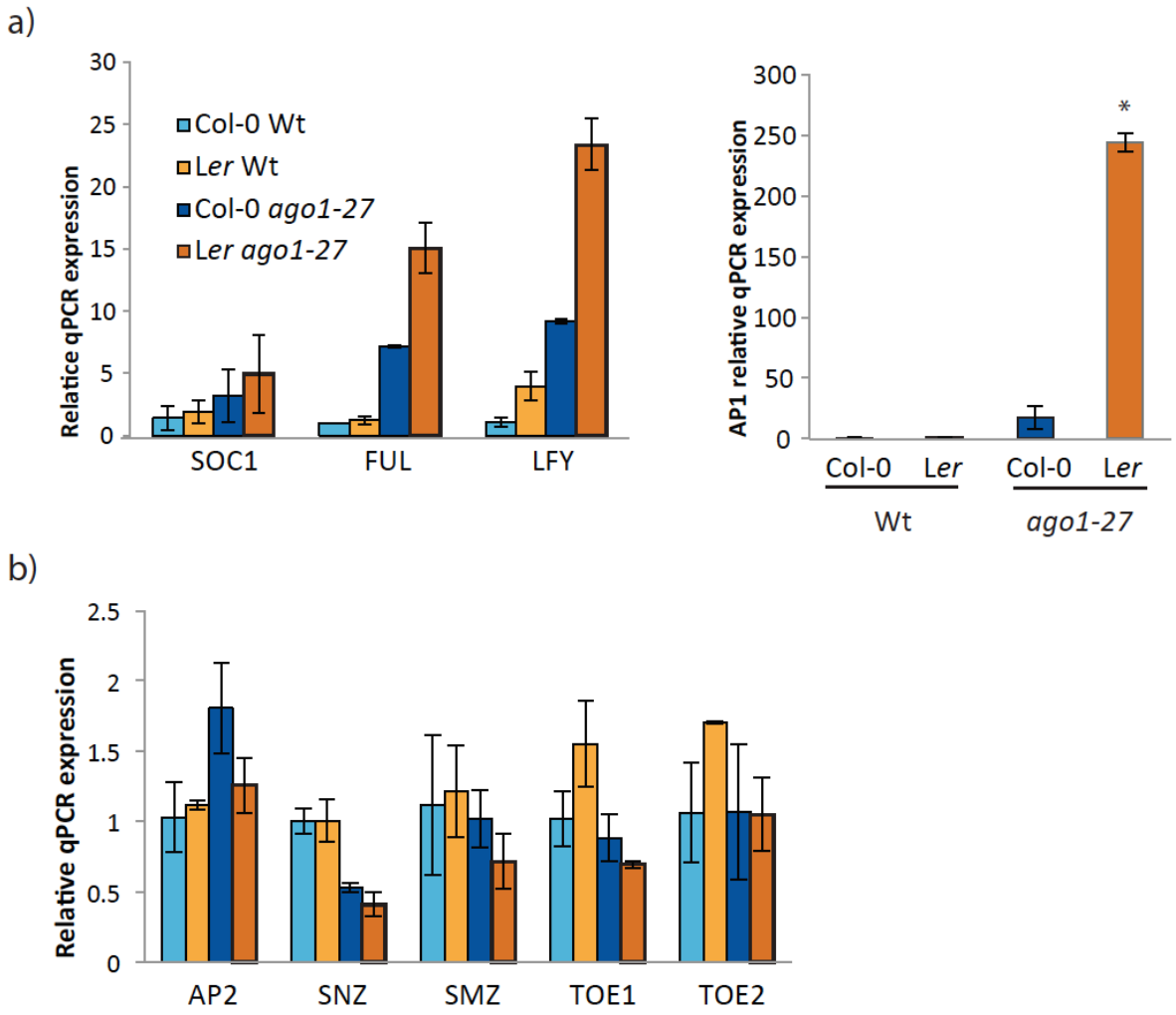


Figure S5.6. Up-regulation of genes involved in the flowering aging pathway.
(a) Gene expression of meristem identity genes is upregulated in L*er ago1-27*. **(b)** Expression of miR172 targets is similar between Col-0 *ago1-27* and L*er ago1-27*. 14-day old plant tissue was collected at ZT16. Mean expression data represent two biological replicates, each with three technical replicates. * = $p < 0.05$, T-test.

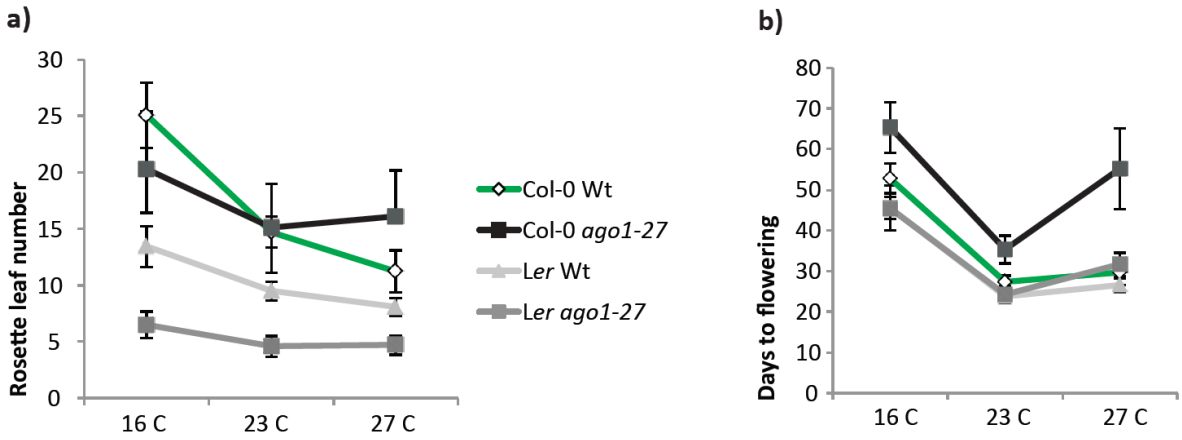


Figure S5.7. Flowering traits in response to temperature. (a) L*er* ago1-27 rosette leaf number varies little in response to different temperatures. **(b)** Days to flowering at different temperatures. Seeds were embedded in 0.1% agarose and stratified at 4˚C for five days. 36 plants per genotype were seeded on 36-pot trays. Then, they were grown at 23˚C for one day and then moved to incubators at 16˚C or 27˚C. All the experiments were performed in Long day (LD) (16h light / 8h dark) at 23˚C.

Table S5.1. Primers used in this study.

Gene	Primer Name	sequence	Purpose	Reference
MIR156F	Col_miR156f_seqFo2	acagccaatgagccaaagat	sequencing	This study
	Col_miR156f_seqRe1	gagggtgtggcgatcataaa	sequencing	This study
	Ler_miR156f_seqFo2	cgagcaatgttcctccataa	sequencing	This study
	Ler_miR156f_seqRe1	gacgcagagtgagggtgtg	sequencing	This study
MIR156D	miR156d_fow1	cacgatgggtcaacaaaatataac	sequencing	This study
	miR156d_fow2	caaaatataacccaaatggagatagac	sequencing	This study
	miR156d_rev1	gatccatcgatcaagtctttgtta	sequencing	This study
	miR156d_int1	caccttcacatacatgtgtaaag	sequencing	This study
MIR156E	miR156e_fow1	aagttgcagttcacgtggttag	sequencing	This study
	miR1563_rev1	aaccctgcataattaaaatgttagag	sequencing	This study
	miR156e_fow2	tttcttcccacatccaaagatag	sequencing	This study
	miR156_rev2	tgcttatggctagaagacacatag	sequencing	This study
MIR172B	miR172b_seq_Fow1	taacgccctaatccgtcatt	sequencing	This study
	miR172b_seq_Fow2	acgccctaatccgtcattg	sequencing	This study
	miR172b_seq_Rev1	tcccactgcttttctccttc	sequencing	This study
	miR172b_int1	ccaaagatcgatccagacttcaatc	sequencing	This study
SOC1	SOC1_fow_yama	tgaggcatactaaggatcgag	qPCR	Yamaguchi et al, 2009. Developmental Cell
	SOC1_rev_yama	gcgtctctacttcagaacttgggc	qPCR	Yamaguchi et al, 2009. Developmental Cell
CO	CO RT1 Sense	tgcagcgtaccacagac	qPCR	This study
	CO RT1 Antisense	atgaaatgtatgcgttatggttaat	qPCR	This study
ELF6	ELF6_fow1	aattccttgttatttgccataaaga	qPCR	This study
	ELF6_rev1	catcatagctgaaagggactttg	qPCR	This study
FLC	FLC_sense	cgacaagtcaccttctcca	qPCR	This study
	FLC_antisense	caaggatcttgaccaggtta	qPCR	This study
TFL1	TFL1_fow5	tcagatccttcttcactttggtg	qPCR	This study
	TFL1_rev_5	aggggtcactaggacctgga	qPCR	This study
SPL2	SPL2_Fow_Kim	acgggttggaggttgcttgagg	qPCR	Kim J.J., et al, 2012 Plant physiology
	SPL2_Rev_Kim	tttccgataccgagcacaatag	qPCR	Kim J.J., et al, 2012 Plant physiology
SPL3	SPL3_fow1	tctggtcttcaccaacgtttc	qPCR	This study
	SPL3_rev1	gcttggcttcatcaaactcac	qPCR	This study
SPL4	SPL4_fow_yama	tcaagggtagagatgacacttcctat	qPCR	Yamaguchi et al, 2009. Developmental Cell
	SPL4_rev_yama	gcgtttgcatatagctgataacat	qPCR	Yamaguchi et al, 2009. Developmental Cell
SPL5	SPL5_fow_Kim	ccagactcaagaaagaaacagggtagacag	qPCR	Kim J.J., et al, 2012 Plant physiology
	SPL5_rev_Kim	tccgtgtaggatttaataccatgacc	qPCR	Kim J.J., et al, 2012 Plant physiology
SLP9	SPL9_uncut_F	tacaagggaattggcgactc	qPCR	This study
	SPL9_uncut_R	ggcgggttcagatactgatg	qPCR	This study
SPL10	SPL10_Fow_Kim	gtgtgggagaatgctcaggagg	qPCR	Kim J.J., et al, 2012 Plant physiology
	SPL10_Rev_Kim	acgggagtgtgtttgatcccttgtg	qPCR	Kim J.J., et al, 2012 Plant physiology
SPL11	SPL11_Fow_Kim	agtccaagtttcaacttcatggcg	qPCR	Kim J.J., et al, 2012 Plant physiology

	SPL11_Rev_Kim	gaacagagtagagaaaatggctgcac	qPCR	Kim J.J., et al, 2012 Plant physiology
SPL13	SPL13_Fow_Kim	ccaatctcttcttctccaaacagtaccagaagc	qPCR	Kim J.J., et al, 2012 Plant physiology
	SPL13_Rev_Kim	gaagcaaatgagggactgacgacg	qPCR	Kim J.J., et al, 2012 Plant physiology
SPL15	SPL15_Fow_Kim	gaatgttttatcacatggaagctc	qPCR	Kim J.J., et al, 2012 Plant physiology
	SPL15_Rev_Kim	tcatcgagtcgaaaccagaagatg	qPCR	Kim J.J., et al, 2012 Plant physiology
SPL7	SPL7_fow2	caaagccctatagagtcaaaggat	qPCR	This study
	SPL7_rev2	agtctcttttccgccgctat	qPCR	This study
LFY	LFY_fow3	aacggagacgattgcaagaa	qPCR	This study
	LFY_rev3	gcaccggttcctcagataac	qPCR	This study
FUL	FUL_yama_fow	ttgcaagatcacaacaattcgcttct	qPCR	Yamaguchi et al, 2009. Developmental Cell
	FUL_yama_Rev	gagagtttggttccgtcaacgacgat	qPCR	Yamaguchi et al, 2009. Developmental Cell
AP1	AP1_Fow2	gcaagcaatgagccctaaag	qPCR	This study
	AP1_Rev2	tgctcctgtatggccttctc	qPCR	This study
FT	FT RT1 Sense	tccaagtcctagcaaccc	qPCR	This study
	FT RT1 Antisense	aaacaatataaacacgacacgatg	qPCR	This study
AP2	AP2_fow2	cgaggtcgaagctgctaga	qPCR	This study
	AP2_rev2	ctcatcgtaaatactcggatcaaa	qPCR	This study
TOE1	TOE1_fow2	caaatgtggtagatgggaagc	qPCR	This study
	TOE1_rev2	ttgattgcagccttgtcataa	qPCR	This study
TOE2	TOE2_fow2	caaccaccaccatagcttcc	qPCR	This study
	TOE2_rev2	ccaacctcagggctgtgtag	qPCR	This study
SMZ	SMZ_fow2	tcgtggatgattataggcatga	qPCR	This study
	SMZ_rev2	gccaagcctttgtatttgga	qPCR	This study
AtsnoR101	AtsnoR101_Cui_F	cttcacaggtaagttcgcttg	qPCR	Cui et al, 2014. The plant journal
	AtsnoR101_Cui_R	agcatcagcagaccagtagtt	qPCR	Cui et al, 2014. The plant journal
UBC	UBC21	gaccaagatattccatccta	qPCR	This study
	UBC21	gttaagaggactgtccg	qPCR	This study
MIR172B	miR172b probe	atgcagcatcatcaagattct	probe	This study

Table S4.2 Polymorphic genes located in chr 5 and involved in flowering time

Gene name	AT number	Chr	Start	End	SNPs	Coding	Synonymous	NonSyn	Deletions	Insertions
SPL7	AT5G18830	5	6,275,990	6,280,567	17	9	8	1	0	3
FY	AT5G13480	5	4,326,473	4,331,971	13	2	1	1	0	1
FLC	AT5G10140	5	3,173,497	3,179,448	8	1	1	0	3 - intronic	1- 3'UTR
CO	AT5G15840	5	5,171,182	5,172,758	2	0	0	0	0	0
TFL1	AT5G03840	5	1,024,641	1,025,812	1	0	0	0	0	0
miR156F	AT5G26147	5	9,136,126	9,136,215	1	0	0	0	0	0
miR156D	AT5G10945	5	3,456,647	3,456,732	0	0	0	0	0	0
miR156E	AT5G11977	5	3,867,213	3,867,308	0	0	0	0	0	0
miR162B	AT5G23065	5	7,740,610	7,740,697	3	0	0	0	1	2
miR164B	AT5G01747	5	287,586	287,734	0	0	0	0	0	0
miR164C	AT5G27807	5	9,852,664	9,852,783	0	0	0	0	0	0
miR166C	AT5G08712	5	2,838,650	2,838,757	0	0	0	0	0	0
miR166D	AT5G08717	5	2,840,628	2,840,728	0	0	0	0	0	0
miR172B	AT5G04275	5	1,188,211	1,188,299	0	0	0	0	0	0

Chapter 6

DISCUSSION AND FUTURE STUDIES

Biological systems have molecular mechanisms that allowed them to maintain a constant phenotype in the face of environmental and genetic variations. These mechanisms can take the form of master regulators of robustness, which are gene products that conceal environmental and genetic perturbations. When the function of the regulators is challenged, the concealed variation is released, and the phenotypic variation of the organism increases. The best characterized master regulator of robustness is the protein chaperone HSP90 (Rutherford and Lindquist 1998; Queitsch et al. 2002; Yeyati et al. 2007; Sangster et al. 2008a; Specchia et al. 2010; Jarosz and Lindquist 2010; Casanueva et al. 2012). Until now, in multicellular organisms, HSP90 represented the only master regulator of robustness, raising questions of whether there are more regulators, if HSP90 ability and buffering mechanisms were shared with other potential regulators, if they buffered the same traits and loci, and to what extent the different regulators overlapped in their functions (Lempe et al. 2013; Siegal and Leu 2014). Here, I show that AGO1 is another master regulator of robustness, that it buffers environmental and genetic variation, and that AGO1 and HSP90 have little overlap in their buffering functions.

6.1 IMPLICATIONS OF AGO1 BUFFERING OF ENVIRONMENTAL PERTURBATIONS

Previous studies have shown that miRNAs can buffer environmental and genetic variation, but that buffering is restricted to particular traits and is not epistatic to multiple stimuli or loci (Hilgers et al. 2010). My work shows that AGO1 perturbation increases the phenotypic variation of isogenic seedlings for two qualitative and three quantitative traits. In addition, I showed that AGO1 perturbation impairs the plants ability to perceive and respond to environmental and abiotic stresses, such as temperature and light intensity (Chapters 4 and 5). The fact that *ago1* mutant plants confound an abiotic signal for a biotic one suggest that AGO1 perturbation can change the underlying genetic network involving environmental responses and rewire it to produce the undesirable result of lesions in cotyledons. This

misinterpretation of signals resembles some instances in animals, such as autoimmune diseases, in which stimuli is misinterpreted. It may be possible that these diseases reflect a general instability of the network, in which the disease can be triggered by several genetic and environmental factors, and perhaps is one of the reasons it is difficult to identify loci involved in the process.

Normal, healthy organisms can change their phenotype to adapt to environmental stresses (phenotypic plasticity) (Lempe et al. 2013), possibly by rewiring their biological networks. AGO1 perturbation alters the organism's ability to respond to environmental stresses. When *ago1-27* plants were exposed to heat and cold stresses they explored more phenotypic space than wild type plants (Chapter 5). This increase in phenotypic variation highlights the role of AGO1 as a master regulator of robustness and may suggest that AGO1 perturbation may be a mechanism of plant adaptation.

6.2 IMPLICATIONS OF AGO1 BUFFERING OF GENETIC PERTURBATIONS

Gene products that buffer environmental variation are not always congruent, meaning that they do not confer robustness to another type of perturbation, such as genetic variation. I demonstrated that AGO1 is a congruent regulator of robustness since it buffers genetic variation, and this buffering is widespread (Chapter 5). I demonstrated that AGO1 buffers variation on two chromosomes for four different quantitative traits. When I crossed the *ago1-27* mutant into different STAIRS lines I observed that the phenotypic variation revealed was background-dependent. This result suggests that different phenotypes may have different susceptibilities to AGO1 perturbation, and therefore, to other kinds of perturbations.

Investigating which phenotypic traits have a greater variance in response to their environment is of great interest the crop industry. To date, crop improvement efforts are made under the assumption that the resulting modified seeds are going to produce plants in which the desired phenotypic trait, for example biomass, will be constant in all locations where the seeds are grown. The trait in question could be more liable to perturbations than other traits from the initial parental strain. Then, when the modified seeds undergo

environmental or biotic stresses, they may not produce the desired biomass. The modified plants could also mistake the responses to perturbations and produce undesirable compounds. It is therefore necessary to investigate which phenotypic traits have a greater variance in their response to perturbations than others.

To gain insight into which traits have more variation under perturbations, I could alter AGO1 function in different *A. thaliana* strains. I propose introducing an *ago1* mutant allele into several different Arabidopsis strains. This could be done by generating transgenic plants, or by crossing different strains with a Col-0 *ago1* mutant allele. The former method is faster, but could be difficult due the fact that AGO1 is an essential gene, and strong *ago1* mutants are infertile. An alternative approach is to evaluate the phenotypic variation of other hypomorphic *ago1* mutants generated in other strains, although to my knowledge only one other strain has characterized *ago1* mutants (Fernández-Nohales et al. 2014).

6.3 Extreme example of AGO1 revealed perturbations

As stated above, AGO1 can reveal genetic differences in the underlying genetic networks governing a particular phenotypic trait. An extreme example of a revealed phenotypic variation was the uncoupling of the closely related traits days to flowering and rosette leaf number (Chapter 5). *ago1-27* plants flower with only four leaves, but with the same days to flowering as wild type plants. The uncoupling of phenotypes suggests that particular phenotypic traits are regulated by more than one genetic network, and the influence of these networks on the phenotype can be taken apart when the overall robustness of the organism is reduced, like in the case of AGO1 perturbation. The ability of AGO1 perturbation to uncouple tightly related traits acts as a tool to solve a long-standing problem in crop improvement: the trade-off of enhancing one phenotypic trait, like biomass, but affecting other desirable phenotypic traits, such as yield (Wang and Wang 2015). Perturbation of AGO1 provides the tools to uncouple closely related traits (i.e. flowering time and leaf number) and to point out which loci might be responsible for the uncoupling, and modify it.

6.4 IMPLICATIONS OF AGO1 IN THE EVOLUTION OF GENOMES: A HYPOTHESIS

The observation that AGO1 perturbation can increase the phenotypic variation when plants undergo stress, together with the fact that AGO1-revealed genetic variation was background-specific, led me to hypothesize that AGO1 may influence the evolution of genes. In this scenario, AGO1 would conceal variation affecting primarily MIRNA and miRNA-target genes. While the miRNA site is maintained, the target gene could accumulate mutations in its coding sequence that could be masked by their miRNA–mediated regulation. When the activity of AGO1 is challenged, the accumulated variation of the gene could be revealed.

Taking advantage of the information learned from assessing the role of HSP90 in the evolution of its clients (Chapter 2 and 3), in assessing the influence of AGO1 in the evolution of miRNA-target genes, I would identify pairs of duplicated genes in which only one was targeted by a miRNA. I would also identify gene families in which only some of its members were targeted by a miRNA. I will then test if the miRNA-targeted genes had more relaxed selection that the non-targets. Some of these analyses have already been done. In a recent study, Wang and Adams (Wang and Adams 2015) analyzed gene duplicates from *Arabidopsis* and *Brassica* and found that genes that were duplicated had an increased probability of being targeted by a miRNA. This preference has also been shown for animals (Li et al. 2008a). To my knowledge, this analysis has not been done for gene families.

Measuring the nucleotide diversity among miRNA-target genes, and even miRNA genes themselves could also provide support for my hypothesis. Nucleotide diversity measures the degree of polymorphism within a population (Hernandez et al. 2011). If a gene were subject to positive or relaxed selection, one would expect the sequence to be more variable among the individuals of a population. In this case I would expect miRNA-target genes to have an increased nucleotide variation than non-target genes. I would also need to normalize for several other factors that could hinder genetic variation such as the number of miRNAs that target a particular gene, the number of tissues where the gene is expressed, number of isoforms, if all of its isoforms have a miRNA target site, and perhaps the mode of repression that the miRNA exerts upon the gene (slicing versus translational repression).

A more direct way to test AGO1's role in buffering mutations would be to mutagenize *ago1* mutant plants. For this experiment, I would make two populations of both wild type plants and *ago1-27* plants. I will then subject them to EMS mutagenesis. I will then have four groups of seeds: (i) wild type non-mutagenized, (ii) wild type mutagenized, (iii) *ago1-27* non-mutagenized, and (iv) *ago1-27* mutagenized. I would then subset of the seeds, and obtain M2 seedlings, to give the opportunity to recesive mutations to become homozygous. I will then take the four groups of seeds, grow them on agar plates or soil, and evaluate their phenotypic diversity. If AGO1 confers robustness to mutations, I would expect (i) *ago1-27* mutagenized plants to show a greater phenotypic variance than *ago1-27* non-mutagenized plants, and (ii) that the difference in the phenotypic variance between *ago1-27* plants to be greater than the difference in phenotypic variance between wild type non-mutagenized and mutagenized plants.

In addition to buffering mutations, perturbation of AGO1 may be a conserved mechanism of adaptation. Plants, as sessile organisms, have to contend with constant changes in their environment, and quickly respond to them. Perturbing AGO1 could increase the phenotypic space of the stress response and expose many genetic variants, some of which can be beneficial to overcome the stress in question. Acevedo et al. (Azevedo et al. 2010) found that depletion of AGO1 caused by plant exposure to virus can take apart the interwoven DCL1 homeostasis, and this may be important to counteract viral infection.

6.5 AGO1 AND HSP90 ACT IN DIFFERENT PATHWAYS TO CONFER ROBUSTNESS

The discovery and characterization of AGO1 as a master regulator of robustness allowed me to address the standing questions of whether master regulators of robustness overlap in their buffered traits, loci, and what is the extent of this overlap, by comparing AGO1 buffering with that of HSP90. I did not find an overlap between HSP90 and AGO1-dependent buffering of phenotypic variation for the traits and loci that I analyzed (Chapters 4 and 5). Given that AGO1 is an HSP90 client (Iki et al. 2010), this result is unexpected. A possible explanation for the result is that the HSP90 analyses and my AGO1 analyses were both done

using partial mutants, because HSP90 and AGO1 are essential genes. The partial reduction of the HSP90 and AGO1 function may change the network differently than it would in their complete absence. Developing conditional mutants for both HSP90 and AGO1 could help circumvent the problem of partial mutants. Inducible mutants would also allow testing the interaction of HSP90 and AGO1 in different tissues and at different developmental stages.

HSP90 and AGO1 could interact differently in diverse *A. thaliana strains*. For the traits of hypocotyl and root length, different strains have different responses to HSP90 inhibition (Sangster et al. 2008b) (also Chapter 2). One might speculate that in the less sensitive strains AGO1 or other master regulators of robustness could have more influence in the buffering of the phenotype of those strains.

Finally, it seems that each level of gene regulation, protein homeostasis and expression, harbors a master regulator, and based on the results of this study, the overlap between the regulators is minimal, arguing that robustness may be modular. This modularity implies that it is necessary to find more regulators to dissect the molecular mechanisms and sub-network connections underlying a particular trait or plant response to stimuli. If the trend that each level of regulation has its own master capacitor were correct, one would expect the next master regulator of robustness to be discovered among genes that regulate chromatin remodeling, epigenetic modifications, and perhaps protein degradation. An initial screening for these genes would be to identify genes that interact with many others. I could do this by searching in protein-protein interaction datasets, as well as in gene expression datasets and look for genes whose expression is correlated to many other genes. Then, I would measure if the variance of certain phenotypic traits increases in the candidate-genes mutant, and under environmental stresses. A similar approach has already been taken in yeast (Levy and Siegal 2008), where they found more than 300 genes that could buffer environmental variation. However, to my knowledge, no such approach has been done for multicellular organisms.

REFERENCES

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. Nature methods 7 (4):248-249. doi:10.1038/nmeth0410-248

Albrecht C, Russinova E, Kemmerling B, Kwaaitaal M, de Vries SC (2008) Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASE Proteins Serve Brassinosteroid-Dependent and -Independent Signaling Pathways. Plant physiology 148 (1):611-619. doi:10.1104/pp.108.123216

Anderson JP, Badruzsaufari E, Schenk PM, Manners JM, Desmond OJ, Ehlert C, Maclean DJ, Ebert PR, Kazan K (2004) Antagonistic Interaction between Abscisic Acid and Jasmonate-Ethylene Signaling Pathways Modulates Defense Gene Expression and Disease Resistance in Arabidopsis. The Plant Cell Online 16 (12):3460-3479

Andres F, Coupland G (2012) The genetic basis of flowering responses to seasonal cues. Nat Rev Genet 13 (9):627-639

Asami T, Min YK, Nagata N, Yamagishi K, Takatsuto S, Fujioka S, Murofushi N, Yamaguchi I, Yoshida S (2000) Characterization of Brassinazole, a Triazole-Type Brassinosteroid Biosynthesis Inhibitor. Plant physiology 123 (1):93-100. doi:10.1104/pp.123.1.93

Aukerman MJ, Sakai H (2003) Regulation of Flowering Time and Floral Organ Identity by a MicroRNA and Its APETALA2-Like Target Genes. The Plant cell 15 (11):2730-2741. doi:10.1105/tpc.016238

Aury JM, Jaillon O, Duret L, Noel B, Jubin C, Porcel BM, Segurens B, Daubin V, Anthouard V, Aiach N, Arnaiz O, Billaut A, Beisson J, Blanc I, Bouhouche K, Camara F, Duharcourt S, Guigo R, Gogendeau D, Katinka M, Keller AM, Kissmehl R, Klotz C, Koll F, Le Mouel A, Lepere G, Malinsky S, Nowacki M, Nowak JK, Plattner H, Poulain J, Ruiz F, Serrano V, Zagulski M, Dessen P, Betermier M, Weissenbach J, Scarpelli C, Schachter V, Sperling L, Meyer E, Cohen J, Wincker P (2006) Global trends of whole-genome duplications revealed by the ciliate Paramecium tetraurelia. Nature 444 (7116):171-178. doi:nature05230 [pii]
10.1038/nature05230

Axtell MJ (2013) Classification and comparison of small RNAs from plants. Annual review of plant biology 64:137-159. doi:10.1146/annurev-arplant-050312-120043

Azevedo J, Garcia D, Pontier D, Ohnesorge S, Yu A, Garcia S, Braun L, Bergdoll M, Hakimi MA, Lagrange T, Voinnet O (2010) Argonaute quenching and global changes in Dicer homeostasis caused by a pathogen-encoded GW repeat protein. Genes & Development 24 (9):904-915. doi:10.1101/gad.1908710

Balbi V, Devoto A (2008) Jasmonate signalling network in Arabidopsis thaliana: crucial regulatory nodes and new physiological scenarios. New Phytologist 177 (2):301-318

Ballaré CL (2011) Jasmonate-induced defenses: a tale of intelligence, collaborators and rascals. Trends in plant science 16 (5):249-257

Ballare CL, Scopel AL, Stapleton AE, Yanovsky MJ (1996) Solar Ultraviolet-B Radiation Affects Seedling Emergence, DNA Integrity, Plant Morphology, Growth Rate, and Attractiveness to Herbivore Insects in Datura ferox. Plant physiology 112 (1):161-170
Baumberger N, Baulcombe DC (2005) Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. Proceedings of the National Academy of Sciences of the United States of America 102 (33):11928-11933
Bergman A, Siegal ML (2003) Evolutionary capacitance as a general feature of complex gene networks. Nature 424 (6948):549-552. doi:http://www.nature.com/nature/journal/v424/n6948/suppinfo/nature01765_S1.html
Bielawski JP, Yang Z (2003) Maximum likelihood methods for detecting adaptive evolution after gene duplication. J Struct Func Genom 3 (1-4):201-212
Bloom JD, Adami C (2003) Apparent dependence of protein evolutionary rate on number of interactions is linked to biases in protein–protein interactions data sets. BMC evolutionary biology 3 (1):21
Bloom JD, Labthavikul ST, Otey CR, Arnold FH (2006) Protein stability promotes evolvability. Proceedings of the National Academy of Sciences 103 (15):5869-5874. doi:10.1073/pnas.0510098103
Bogumil D, Alvarez-Ponce D, Landan G, McInerney JO, Dagan T (2014) Integration of Two Ancestral Chaperone Systems into One: The Evolution of Eukaryotic Molecular Chaperones in Light of Eukaryogenesis. Molecular Biology and Evolution 31 (2):410-418. doi:10.1093/molbev/mst212
Bogumil D, Dagan T (2010) Chaperonin-Dependent Accelerated Substitution Rates in Prokaryotes. Genome Biology and Evolution 2:602-608. doi:10.1093/gbe/evq044
Bogumil D, Dagan T (2012) Cumulative Impact of Chaperone-Mediated Folding on Genome Evolution. Biochemistry 51 (50):9941-9953. doi:10.1021/bi3013643
Bogumil D, Landan G, Ilhan J, Dagan T (2012) Chaperones Divide Yeast Proteins into Classes of Expression Level and Evolutionary Rate. Genome Biology and Evolution 4 (5):618-625. doi:10.1093/gbe/evs025
Bohmert K, Camus I, Bellini C, Bouchez D, Caboche M, Benning C (1998) AGO1 defines a novel locus of Arabidopsis controlling leaf development. The EMBO journal 17 (1):170-180
Bologna NG, Schapire AL, Zhai J, Chorostecki U, Boisbouvier J, Meyers BC, Palatnik JF (2013) Multiple RNA recognition patterns during microRNA biogenesis in plants. Genome research 23 (10):1675-1689. doi:10.1101/gr.153387.112
Bologna NG, Voinnet O (2014) The Diversity, Biogenesis, and Activities of Endogenous Silencing Small RNAs in Arabidopsis. Annual Review of Plant Biology 65 (1):473-503. doi:doi:10.1146/annurev-arplant-050213-035728
Boutilier K, Offringa R, Sharma VK, Kieft H, Ouellet T, Zhang L, Hattori J, Liu C-M, Lammeren AAMv, Miki BLA, Custers JBM, Campagne MMvL (2002) Ectopic Expression of BABY BOOM Triggers a Conversion from Vegetative to Embryonic Growth. The Plant Cell Online 14 (8):1737-1749
Breunig JS, Hackett SR, Rabinowitz JD, Kruglyak L (2014) Genetic Basis of Metabolome Variation in Yeast. PLoS Genet 10 (3):e1004142. doi:10.1371/journal.pgen.1004142
Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O (2008) Widespread translational inhibition by plant miRNAs

and siRNAs. Science (New York, NY) 320 (5880):1185-1190. doi:10.1126/science.1159151

Burga A, Casanueva MO, Lehner B (2011) Predicting mutation outcome from early stochastic variation in genetic interaction partners. Nature 480 (7376):250-253. doi:http://www.nature.com/nature/journal/v480/n7376/abs/nature10665.html - supplementary-information

Cao J, Schneeberger K, Ossowski S, Gunther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, Wang X, Ott F, Muller J, Alonso-Blanco C, Borgwardt K, Schmid KJ, Weigel D (2011) Whole-genome sequencing of multiple Arabidopsis thaliana populations. Nature genetics 43 (10):956-963. doi:http://www.nature.com/ng/journal/v43/n10/abs/ng.911.html - supplementary-information

Carbonell A, Fahlgren N, Garcia-Ruiz H, Gilbert K, Montgomery T, Nguyen T, Cuperus J, Carrington J (2012) Functional Analysis of Three Arabidopsis ARGONAUTES Using Slicer-Defective Mutants. The Plant cell 24:3613–3629

Carlson KD, Sudmant PH, Press MO, Eichler EE, Shendure J, Queitsch C (2015) MIPSTR: a method for multiplex genotyping of germline and somatic STR variation across many individuals. Genome research. doi:10.1101/gr.182212.114

Casanueva MO, Burga A, Lehner B (2012) Fitness Trade-Offs and Environmentally Induced Mutation Buffering in Isogenic C. elegans. Science (New York, NY) 335 (6064):82-85. doi:10.1126/science.1213491

Cassidy JJ, Jha AR, Posadas DM, Giri R, Venken KJ, Ji J, Jiang H, Bellen HJ, White KP, Carthew RW (2013) miR-9a minimizes the phenotypic impact of genomic diversity by buffering a transcription factor. Cell 155 (7):1556-1567. doi:10.1016/j.cell.2013.10.057

Castle JC, Armour CD, Lower M, Haynor D, Biery M, Bouzek H, Chen R, Jackson S, Johnson JM, Rohl CA, Raymond CK (2010) Digital genome-wide ncRNA expression, including SnoRNAs, across 11 human tissues using polyA-neutral amplification. PLoS One 5 (7):e11779. doi:10.1371/journal.pone.0011779

Chamary JV, Parmley JL, Hurst LD (2006) Hearing silence: non-neutral evolution at synonymous sites in mammals. Nat Rev Genet 7 (2):98-108

Chen H, Xue L, Chintamanani S, Germain H, Lin H, Cui H, Cai R, Zuo J, Tang X, Li X, Guo H, Zhou J-M (2009) ETHYLENE INSENSITIVE3 and ETHYLENE INSENSITIVE3-LIKE1 repress SALICYLIC ACID INDUCTION DEFICIENT2 expression to negatively regulate plant innate immunity in Arabidopsis. The Plant cell 21 (8):2527-2540

Chen X (2004) A MicroRNA as a Translational Repressor of APETALA2 in Arabidopsis Flower Development. Science (New York, NY) 303 (5666):2022-2025. doi:10.1126/science.1088060

Chen X (2009) Small RNAs and Their Roles in Plant Development. Annual Review of Cell and Developmental Biology 25 (1):21-44. doi:doi:10.1146/annurev.cellbio.042308.113417

Chinwalla A (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420 (6915):520-562. doi:http://www.nature.com/nature/journal/v420/n6915/suppinfo/nature01262_S1.html

Chitwood DH, Nogueira FTS, Howell MD, Montgomery TA, Carrington JC, Timmermans MCP (2009) Pattern formation via small RNA mobility. Genes & Development 23 (5):549-554. doi:10.1101/gad.1770009
Citri A, Harari D, Shohat G, Ramakrishnan P, Gan J, Lavi S, Eisenstein M, Kimchi A, Wallach D, Pietrokovski S, Yarden Y (2006) Hsp90 Recognizes a Common Surface on Client Kinases. Journal of Biological Chemistry 281 (20):14361-14369. doi:10.1074/jbc.M512613200
Clarke GM (1998) The genetic basis of developmental stability. IV. Individual and population asymmetry parameters. Heredity 80 (5):553-561
Conant GC, Wolfe KH (2008) Turning a hobby into a job: How duplicated genes find new functions. Nat Rev Genet 9 (12):938-950
Cooper GM, Stone EA, Asimenos G, Green ED, Batzoglou S, Sidow A (2005) Distribution and intensity of constraint in mammalian genomic sequence. Genome research 15 (7):901-913. doi:10.1101/gr.3577405
Cuperus JT, Fahlgren N, Carrington JC (2011) Evolution and Functional Diversification of MIRNA Genes. The Plant cell 23 (2):431-442. doi:10.1105/tpc.110.082784
Cuperus JT, Montgomery TA, Fahlgren N, Burke RT, Townsend T, Sullivan CM, Carrington JC (2010) Identification of MIR390a precursor processing-defective mutants in Arabidopsis by direct genome sequencing. Proceedings of the National Academy of Sciences of the United States of America 107 (1):466-471. doi:10.1073/pnas.0913203107
Czech B, Hannon GJ (2011) Small RNA sorting: matchmaking for Argonautes. Nat Rev Genet 12 (1):19-31
Dahl CCv, Baldwin IT (2007) Deciphering the Role of Ethylene in Plant–Herbivore Interactions. Journal of Plant Growth Regulation 26 (2):201-209
Danisman S, van der Wal F, Dhondt S, Waites R, de Folter S, Bimbo A, van Dijk AD, Muino JM, Cutri L, Dornelas MC, Angenent GC, Immink RGH (2012) Arabidopsis Class I and Class II TCP Transcription Factors Regulate Jasmonic Acid Metabolism and Leaf Development Antagonistically. Plant physiology 159 (4):1511-1523. doi:10.1104/pp.112.200303
de Meaux J, Hu J-Y, Tartler U, Goebel U (2008) Structurally different alleles of the ath-MIR824 microRNA precursor are maintained at high frequency in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America 105 (26):8994-8999. doi:10.1073/pnas.0803218105
De Vos M, Van Zaanen W, Koornneef A, Korzelius JP, Dicke M, Van Loon LC, Pieterse CMJ (2006) Herbivore-induced resistance against microbial pathogens in Arabidopsis. Plant physiology 142 (1):352-363
Dean EJ, Davis JC, Davis RW, Petrov DA (2008) Pervasive and Persistent Redundancy among Duplicated Genes in Yeast. PLoS Genet 4 (7):e1000113. doi:10.1371/journal.pgen.1000113
DeLuna A, Vetsigian K, Shoresh N, Hegreness M, Colon-Gonzalez M, Chao S, Kishony R (2008) Exposing the fitness contribution of duplicated genes. Nature genetics 40 (5):676-681. doi:http://www.nature.com/ng/journal/v40/n5/suppinfo/ng.123_S1.html
Demkura P, Abdala G, Baldwin I, Ballaré C (2010) Jasmonate-Dependent and -Independent Pathways Mediate Specific Effects of Solar Ultraviolet B Radiation on Leaf Phenolics and Antiherbivore Defense. Plant physiology 152:1084–1095

Ding Y, Chan CY, Lawrence CE (2004) Sfold web server for statistical folding and rational design of nucleic acids. Nucleic Acids Research 32 (suppl 2):W135-W141. doi:10.1093/nar/gkh449

Doyle MR, Davis SJ, Bastow RM, McWatters HG, Kozma-Bognar L, Nagy F, Millar AJ, Amasino RM (2002) The ELF4 gene controls circadian rhythms and flowering time in Arabidopsis thaliana. Nature 419 (6902):74-77. doi:http://www.nature.com/nature/journal/v419/n6902/suppinfo/nature00954_S1.html

Drummond DA, Bloom JD, Adami C, Wilke CO, Arnold FH (2005) Why highly expressed proteins evolve slowly. Proceedings of the National Academy of Sciences of the United States of America 102 (40):14338-14343. doi:10.1073/pnas.0504070102

Ebert MS, Sharp PA (2012) Roles for microRNAs in conferring robustness to biological processes. Cell 149 (3):515-524. doi:10.1016/j.cell.2012.04.005

Ehrenreich IM, Purugganan MD (2008) Sequence Variation of MicroRNAs and Their Binding Sites in Arabidopsis. Plant physiology 146 (4):1974-1982. doi:10.1104/pp.108.116582

Ehrlich ES, Wang T, Luo K, Xiao Z, Niewiadomska AM, Martinez T, Xu W, Neckers L, Yu X-F (2009) Regulation of Hsp90 client proteins by a Cullin5-RING E3 ubiquitin ligase. Proceedings of the National Academy of Sciences 106 (48):20330-20335

Fares MA, Ruiz-Gonzalez MX, Moya A, Elena SF, Barrio E (2002) Endosymbiotic bacteria: GroEL buffers against deleterious mutations. Nature 417 (6887):398-398

Felsenstein J (1985) Phylogenies and the comparative method. Am Nat 125:1-15

Fernandez A, Lynch M (2011) Non-adaptive origins of interactome complexity. Nature 474 (7352):502-505. doi:http://www.nature.com/nature/journal/v474/n7352/abs/nature09992-f1.2.html - supplementary-information

Fernández-Nohales P, Domenech MJ, Martínez de Alba AE, Micol JL, Ponce MR, Madueño F (2014) AGO1 controls arabidopsis inflorescence architecture possibly by regulating TFL1 expression. Annals of Botany. doi:10.1093/aob/mcu132

Feys B, Benedetti CE, Penfold CN, Turner JG (1994) Arabidopsis Mutants Selected for Resistance to the Phytotoxin Coronatine Are Male Sterile, Insensitive to Methyl Jasmonate, and Resistant to a Bacterial Pathogen. The Plant Cell Online 6 (5):751-759

Fraser HB, Hirsh AE, Steinmetz LM, Scharfe C, Feldman MW (2002) Evolutionary rate in the protein interaction network. Science (New York, NY) 296 (5568):750-752. doi:10.1126/science.1068696

Fryer JD, Yu P, Kang H, Mandel-Brehm C, Carter AN, Crespo-Barreto J, Gao Y, Flora A, Shaw C, Orr HT, Zoghbi HY (2011) Exercise and Genetic Rescue of SCA1 via the Transcriptional Repressor Capicua. Science (New York, NY) 334 (6056):690-693

Fu J, Keurentjes JJB, Bouwmeester H, America T, Verstappen FWA, Ward JL, Beale MH, de Vos RCH, Dijkstra M, Scheltema RA, Johannes F, Koornneef M, Vreugdenhil D, Breitling R, Jansen RC (2009) System-wide molecular evidence for phenotypic buffering in Arabidopsis. Nature genetics 41 (2):166-167. doi:http://www.nature.com/ng/journal/v41/n2/suppinfo/ng.308_S1.html

Gan X, Stegle O, Behr J, Steffen JG, Drewe P, Hildebrand KL, Lyngsoe R, Schultheiss SJ, Osborne EJ, Sreedharan VT, Kahles A, Bohnert R, Jean G, Derwent P, Kersey P, Belfield EJ, Harberd NP, Kemen E, Toomajian C, Kover PX, Clark RM, Ratsch G,

Mott R (2011) Multiple reference genomes and transcriptomes for Arabidopsis thaliana. Nature 477 (7365):419-423. doi:http://www.nature.com/nature/journal/v477/n7365/abs/nature10414.html - supplementary-information

Gangaraju VK, Yin H, Weiner MM, Wang J, Huang XA, Lin H (2011) Drosophila Piwi functions in Hsp90-mediated suppression of phenotypic variation. Nature genetics 43 (2):153-158. doi:10.1038/ng.743

Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO (2000) Genomic Expression Programs in the Response of Yeast Cells to Environmental Changes. Molecular biology of the cell 11 (12):4241-4257. doi:10.1091/mbc.11.12.4241

Gfeller A, Dubugnon L, Liechti R, Farmer EE (2010) Jasmonate Biochemical Pathway. Science Signaling 3 (109):cm3

Giannini A, Bijlmakers M-J (2004) Regulation of the Src Family Kinase Lck by Hsp90 and Ubiquitination. Molecular and Cellular Biology 24 (13):5667-5676. doi:10.1128/MCB.24.13.5667-5676.2004

Guan Y, Dunham MJ, Troyanskaya OG (2007) Functional Analysis of Gene Duplications in Saccharomyces cerevisiae. Genetics 175 (2):933-943. doi:10.1534/genetics.106.064329

Guindon S, Gascuel O (2003) A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. Systematic Biology 52 (5):696-704. doi:10.1080/10635150390235520

Guo H, Li L, Ye H, Yu X, Algreen A, Yin Y (2009) Three related receptor-like kinases are required for optimal cell elongation in Arabidopsis thaliana. Proceedings of the National Academy of Sciences. doi:10.1073/pnas.0812346106

Guranowski A, Miersch O, Staswick P, Suza W, Wasternack C (2007) Substrate specificity and products of side-reactions catalyzed by jasmonate:amino acid synthetase (JAR1). FEBS Lett 581:815–820

Hahn MW (2009) Distinguishing among evolutionary models for the maintenance of gene duplicates. J Hered 100 (5):605-617. doi:10.1093/jhered/esp047

Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W (2008) GEIGER: investigating evolutionary radiations. Bioinformatics 24 (1):129-131. doi:10.1093/bioinformatics/btm538

Hartson SD, Matts RL (2012) Approaches for Defining the Hsp90-dependent Proteome. Biochimica et biophysica acta 1823 (3):656-667. doi:10.1016/j.bbamcr.2011.08.013

He JX, Gendron JM, Sun Y, Gampala SS, Gendron N, Sun CQ, Wang ZY (2005) BZR1 is a transcriptional repressor with dual roles in brassinosteroid homeostasis and growth responses. Science (New York, NY) 307 (5715):1634-1638. doi:10.1126/science.1107580

Heath M (2000) Hypersensitive response-related death. Plant Molecular Biology 44 (3):321-334. doi:10.1023/A:1026592509060

Helliwell CA, Anderssen RS, Robertson M, Finnegan EJ (2015) How is FLC repression initiated by cold? Trends in plant science 20 (2):76-82. doi:10.1016/j.tplants.2014.12.004

Hernandez RD, Kelley JL, Elyashiv E, Melton SC, Auton A, McVean G, Project G, Sella G, Przeworski M (2011) Classic Selective Sweeps Were Rare in Recent Human

Evolution. Science (New York, NY) 331 (6019):920-924. doi:10.1126/science.1198878
Hilgers V, Bushati N, Cohen SM (2010) Drosophila microRNAs 263a/b Confer Robustness during Development by Protecting Nascent Sense Organs from Apoptosis. PLoS Biol 8 (6):e1000396. doi:10.1371/journal.pbio.1000396
Hock J, Meister G (2008) The Argonaute protein family. Genome biology 9 (2):210. doi:10.1186/gb-2008-9-2-210
Hornstein E, Shomron N (2006) Canalization of development by microRNAs. Nature genetics
Hubert DA, Tornero P, Belkhadir Y, Krishna P, Takahashi A, Shirasu K, Dangl JL (2003a) Cytosolic HSP90 associates with and modulates the Arabidopsis RPM1 disease resistance protein. The EMBO journal 22 (21):5679-5689. doi:10.1093/emboj/cdg547
Hubert DA, Tornero P, Belkhadir Y, Krishna P, Takahashi A, Shirasu K, Dangl JL (2003b) Cytosolic HSP90 associates with and modulates the Arabidopsis RPM1 disease resistance protein, vol 22. vol 21. doi:10.1093/emboj/cdg547
Huijser P, Schmid M (2011) The control of developmental phase transitions in plants. Development (Cambridge, England) 138 (19):4117-4129. doi:10.1242/dev.063511
Iki T, Yoshikawa M, Nishikiori M, Jaudal MC, Matsumoto-Yokoyama E, Mitsuhara I, Meshi T, Ishikawa M (2010) In vitro assembly of plant RNA-induced silencing complexes facilitated by molecular chaperone HSP90. Molecular cell 39 (2):282-291. doi:10.1016/j.molcel.2010.05.014
Ishiguro S WY, Ito N, Nonanka H, Takeda N, Sakai T, Kanaya H, Okada K (2002) SHEPARD is the Arabidopsis GRP94 responsible for the formation of functional CLAVATA proteins. The EMBO journal 21 (5):898-908
Iwasaki S, Kobayashi M, Yoda M, Sakaguchi Y, Katsuma S, Suzuki T, Tomari Y (2010) Hsc70/Hsp90 Chaperone Machinery Mediates ATP-Dependent RISC Loading of Small RNA Duplexes. Molecular Cell 39 (2):292-299. doi:http://dx.doi.org/10.1016/j.molcel.2010.05.015
Jarosz DF, Lindquist S (2010) Hsp90 and environmental stress transform the adaptive value of natural genetic variation. Science (New York, NY) 330 (6012):1820-1824. doi:10.1126/science.1195487
Johnston M, Geoffroy MC, Sobala A, Hay R, Hutvagner G (2010) HSP90 protein stabilizes unloaded argonaute complexes and microscopic P-bodies in human cells. Molecular biology of the cell 21 (9):1462-1469. doi:10.1091/mbc.E09-10-0885
Karagoz GE, Rudiger SG (2015) Hsp90 interaction with clients. Trends in biochemical sciences 40 (2):117-125. doi:10.1016/j.tibs.2014.12.002
Kaufmann K, Wellmer F, Muino JM, Ferrier T, Wuest SE, Kumar V, Serrano-Mislata A, Madueno F, Krajewski P, Meyerowitz EM, Angenent GC, Riechmann JL (2010) Orchestration of floral initiation by APETALA1. Science (New York, NY) 328 (5974):85-89. doi:10.1126/science.1185244
Kazan K (2015) Diverse roles of jasmonates and ethylene in abiotic stress tolerance. Trends in plant science 20 (4):219-229. doi:http://dx.doi.org/10.1016/j.tplants.2015.02.001
Kim JJ, Lee JH, Kim W, Jung HS, Huijser P, Ahn JH (2012) The microRNA156-SQUAMOSA PROMOTER BINDING PROTEIN-LIKE3 module regulates ambient temperature-responsive flowering via FLOWERING LOCUS T in Arabidopsis. Plant physiology 159 (1):461-478. doi:10.1104/pp.111.192369

Kim T-W, Wang Z-Y (2010) Brassinosteroid Signal Transduction from Receptor Kinases to Transcription Factors. Annual Review of Plant Biology 61 (1):681-704. doi:doi:10.1146/annurev.arplant.043008.092057
Koo AJK, Howe GA (2009) The wound hormone jasmonate. Phytochemistry 70 (13–14):1571-1580
Koornneef A, Pieterse CMJ (2008) Cross Talk in Defense Signaling. Plant physiology 146 (3):839-844
Koumproglou R, Wilkes TM, Townson P, Wang XY, Beynon J, Pooni HS, Newbury HJ, Kearsey MJ (2002) STAIRS: a new genetic resource for functional genomic studies of Arabidopsis. The Plant Journal 31 (3):355-364. doi:10.1046/j.1365-313X.2002.01353.x
Lachowiec J, Lemus T, Thomas JH, Murphy PJ, Nemhauser JL, Queitsch C (2013) The protein chaperone HSP90 can facilitate the divergence of gene duplicates. Genetics 193 (4):1269-1277. doi:10.1534/genetics.112.148098
Lam E, Kato N, Lawton M (2001) Programmed cell death, mitochondria and the plant hypersensitive response. Nature 411 (6839):848-853
Landthaler M, Gaidatzis D, Rothballer A, Chen PY, Soll SJ, Dinic L, Ojo T, Hafner M, Zavolan M, Tuschl T (2008) Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. RNA (New York, NY) 14 (12):2580-2596
Lawton KA, Potter SL, Uknes S, Ryals J (1994) Acquired Resistance Signal Transduction in Arabidopsis Is Ethylene Independent. The Plant cell 6 (5):581-588
Le Hir H, Nott A, Moore MJ (2003) How introns influence and enhance eukaryotic gene expression. Trends in biochemical sciences 28 (4):215-220. doi:10.1016/S0968-0004(03)00052-5
Lee H, Yoo SJ, Lee JH, Kim W, Yoo SK, Fitzgerald H, Carrington JC, Ahn JH (2010) Genetic framework for flowering-time regulation by ambient temperature-responsive miRNAs in Arabidopsis. Nucleic Acids Research 38 (9):3081-3093. doi:10.1093/nar/gkp1240
Lempe J, Balasubramanian S, Sureshkumar S, Singh A, Schmid M, Weigel D (2005) Diversity of Flowering Responses in Wild <italic>Arabidopsis thaliana</italic> Strains. PLoS Genet 1 (1):e6. doi:10.1371/journal.pgen.0010006
Lempe J, Lachowiec J, Sullivan AM, Queitsch C (2013) Molecular mechanisms of robustness in plants. Current Opinion in Plant Biology 16 (1):62-69. doi:http://dx.doi.org/10.1016/j.pbi.2012.12.002
Leung AK, Sharp PA (2010) MicroRNA functions in stress responses. Molecular cell 40 (2):205-215. doi:10.1016/j.molcel.2010.09.027
Levy SF, Siegal ML (2008) Network Hubs Buffer Environmental Variation in Saccharomyces cerevisiae. PLoS Biol 6 (11):e264. doi:10.1371/journal.pbio.0060264
Li J, Musso G, Zhang Z (2008a) Preferential regulation of duplicated genes by microRNAs in mammals. Genome biology 9 (8):R132
Li L, Yu X, Thompson A, Guo M, Yoshida S, Asami T, Chory J, Yin Y (2009a) Arabidopsis MYB30 is a direct target of BES1 and cooperates with BES1 to regulate brassinosteroid-induced gene expression. The Plant Journal 58 (2):275-286. doi:10.1111/j.1365-313X.2008.03778.x

Li W, Bengtson MH, Ulbrich A, Matsuda A, Reddy VA, Orth A, Chanda SK, Batalov S, Joazeiro CA (2008b) Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling. PLoS One 3 (1):e1487. doi:10.1371/journal.pone.0001487
Li X, Cassidy JJ, Reinke CA, Fischboeck S, Carthew RW (2009b) A MicroRNA Imparts Robustness against Environmental Fluctuation during Development. Cell 137 (2):273-282. doi:10.1016/j.cell.2009.01.058
Lim WA, Pawson T (2010) Phosphotyrosine Signaling: Evolving a New Cellular Communication System. Cell 142 (5):661-667. doi:10.1016/j.cell.2010.08.023
Liu C, Teo ZW, Bi Y, Song S, Xi W, Yang X, Yin Z, Yu H (2013) A conserved genetic pathway determines inflorescence architecture in Arabidopsis and rice. Dev Cell 24 (6):612-622. doi:10.1016/j.devcel.2013.02.013
Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song J-J, Hammond SM, Joshua-Tor L, Hannon GJ (2004a) Argonaute2 is the catalytic engine of mammalian RNAi. Science (New York, NY) 305 (5689):1437-1441
Liu J, He Y, Amasino R, Chen X (2004b) siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in Arabidopsis. Genes & Development 18 (23):2873-2878. doi:10.1101/gad.1217304
Liu Y, Burch-Smith T, Schiff M, Feng S, Dinesh-Kumar SP (2004c) Molecular Chaperone Hsp90 Associates with Resistance Protein N and Its Signaling Proteins SGT1 and Rar1 to Modulate an Innate Immune Response in Plants. Journal of Biological Chemistry 279 (3):2101-2108
Lorenzo O, Chico JM, Sánchez-Serrano JJ, Solano R (2004) JASMONATE-INSENSITIVE1 Encodes a MYC Transcription Factor Essential to Discriminate between Different Jasmonate-Regulated Defense Responses in Arabidopsis. The Plant Cell Online 16 (7):1938-1950
Lu R, Malcuit I, Moffett P, Ruiz MT, Peart J, Wu A-J, Rathjen JP, Bendahmane A, Day L, Baulcombe DC (2003) High throughput virus-induced gene silencing implicates heat shock protein 90 in plant disease resistance. The EMBO journal 22 (21):5690-5699
Luo Y, Guo Z, Li L (2013) Evolutionary conservation of microRNA regulatory programs in plant flower development. Developmental Biology 380 (2):133-144. doi:http://dx.doi.org/10.1016/j.ydbio.2013.05.009
Lynch M, Conery JS (2000) The Evolutionary Fate and Consequences of Duplicate Genes. Science (New York, NY) 290 (5494):1151-1155. doi:10.1126/science.290.5494.1151
Maere S, De Bodt S, Raes J, Casneuf T, Van Montagu M, Kuiper M, Van de Peer Y (2005) Modeling gene and genome duplications in eukaryotes. Proceedings of the National Academy of Sciences of the United States of America 102 (15):5454-5459. doi:10.1073/pnas.0501102102
Maniataki E, Mourelatos Z (2005) A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. Genes & Development 19 (24):2979-2990
Manning G, Hunter T (2009) Eukaryotic kinomes: genomics and evolution
of protein kinases. In: Bradshaw RA, Dennis E, editors. Handbook of cell signaling. San Diego: Academic Press.p. 13–17
Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S (2002) The protein kinase complement of the human genome. Science (New York, NY) 298 (5600):1912-1934. doi:10.1126/science.1075762

Masel J, Siegal ML (2009) Robustness: mechanisms and consequences. Trends in genetics : TIG 25 (9):395-403. doi:10.1016/j.tig.2009.07.005
Mateos JL, Bologna NG, Chorostecki U, Palatnik JF (2010) Identification of microRNA processing determinants by random mutagenesis of Arabidopsis MIR172a precursor. Current biology : CB 20 (1):49-54. doi:10.1016/j.cub.2009.10.072
Mathieu J, Yant LJ, Mürdter F, Küttner F, Schmid M (2009) Repression of Flowering by the miR172 Target SMZ. PLoS Biol 7 (7):e1000148. doi:10.1371/journal.pbio.1000148
Mayer MP (2010) Gymnastics of Molecular Chaperones. Molecular Cell 39 (3):321-331. doi:10.1016/j.molcel.2010.07.012
Mazza CA, Zavala J, Scopel AL, Ballaré CL (1999) Perception of solar UVB radiation by phytophagous insects: Behavioral responses and ecosystem implications. Proceedings of the National Academy of Sciences 96 (3):980-985
McClellan AJ, Tam S, Kaganovich D, Frydman J (2005) Protein quality control: chaperones culling corrupt conformations. Nat Cell Biol 7 (8):736-741
McLellan CA, Turbyville TJ, Wijeratne EMK, Kerschen A, Vierling E, Queitsch C, Whitesell L, Gunatilaka AAL (2007) A Rhizosphere Fungus Enhances Arabidopsis Thermotolerance through Production of an HSP90 Inhibitor. Plant physiology 145 (1):174-182. doi:10.1104/pp.107.101808
Meiklejohn CD, Hartl DL (2002) A single mode of canalization. Trends in ecology & evolution 17 (10):468-473. doi:10.1016/S0169-5347(02)02596-X
Meinke D, Koornneef M (1997) Community standards for Arabidopsis genetics. The Plant Journal 12 (2):247-253
Memelink J (2009) Regulation of gene expression by jasmonate hormones. Phytochemistry 70 (13–14):1560-1570
Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, Chen S, Hannon GJ, Qi Y (2008) Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. Cell 133 (1):116-127. doi:10.1016/j.cell.2008.02.034
Michaels SD, He Y, Scortecci KC, Amasino RM (2003) Attenuation of FLOWERING LOCUS C activity as a mechanism for the evolution of summer-annual flowering behavior in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 100 (17):10102-10107. doi:10.1073/pnas.1531467100
Miyoshi T, Takeuchi A, Siomi H, Siomi MC (2010) A direct role for Hsp90 in pre-RISC formation in Drosophila. Nat Struct Mol Biol 17 (8):1024-1026. doi:http://www.nature.com/nsmb/journal/v17/n8/abs/nsmb.1875.html - supplementary-information
Morel J-B, Godon C, Mourrain P, Béclin C, Boutet S, Feuerbach F, Proux F, Vaucheret H (2002) Fertile Hypomorphic ARGONAUTE (ago1) Mutants Impaired in Post-Transcriptional Gene Silencing and Virus Resistance. The Plant cell 14 (3):629-639. doi:10.1105/tpc.010358
Morishima Y, Wang AM, Yu Z, Pratt WB, Osawa Y, Lieberman AP (2008) CHIP deletion reveals functional redundancy of E3 ligases in promoting degradation of both signaling proteins and expanded glutamine proteins. Human molecular genetics 17 (24):3942-3952. doi:10.1093/hmg/ddn296

Motani R, Schmitz L (2011) Phylogenetic versus functional signals in the evolution of form-function relationships in terrestrial vision. Evolution 65 (8):2245-2257. doi:10.1111/j.1558-5646.2011.01271.x
Moyroud E, Kusters E, Monniaux M, Koes R, Parcy F (2010) LEAFY blossoms. Trends in plant science 15 (6):346-352. doi:10.1016/j.tplants.2010.03.007
Mukhtar M, Nishimura M, Dangl J (2009) NPR1 in Plant Defense: It's Not over "til It"s Turned over. Cell 137:804–806
Murata S, Minami Y, Minami M, Chiba T, Tanaka K (2001) CHIP is a chaperone-dependent E3 ligase that ubiquitylates unfolded protein. EMBO reports 2 (12):1133-1138
Nathan DF, Vos MH, Lindquist S (1997) In vivo functions of the Saccharomyces cerevisiae Hsp90 chaperone. Proceedings of the National Academy of Sciences 94 (24):12949-12956
Nei M, Roychoudhury AK (1973) Probability of Fixation of Nonfunctional Genes at Duplicate Loci. The American Naturalist 107 (955):362-372. doi:10.2307/2459537
Nemhauser JL (2008) Dawning of a new era: photomorphogenesis as an integrated molecular network. Current Opinion in Plant Biology 11 (1):4-8. doi:http://dx.doi.org/10.1016/j.pbi.2007.10.005
Orozco-Cárdenas ML, Narváez-Vásquez J, Ryan CA (2001) Hydrogen Peroxide Acts as a Second Messenger for the Induction of Defense Genes in Tomato Plants in Response to Wounding, Systemin, and Methyl Jasmonate. The Plant Cell Online 13 (1):179-191
Overmyer K, Brosché M, Kangasjärvi J (2003) Reactive oxygen species and hormonal control of cell death. Trends in plant science 8 (7):335-342
Pagel M (1994) Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. Proceedings of the Royal Society of London Series B: Biological Sciences 255 (1342):37-45
Pagel M (1999) Inferring the historical patterns of biological evolution. Nature 401 (6756):877-884
Pagel M, Meade A, Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. Syst Biol 53 (5):673-684. doi:10.1080/10635150490522232
Pajoro A, Biewers S, Dougali E, Valentim FL, Mendes MA, Porri A, Coupland G, Van de Peer Y, van Dijk ADJ, Colombo L, Davies B, Angenent GC (2014) The (r)evolution of gene regulatory networks controlling Arabidopsis plant reproduction; a two decades history. Journal of Experimental Botany. doi:10.1093/jxb/eru233
Pál C, Papp B, Hurst LD (2001) Highly Expressed Genes in Yeast Evolve Slowly. Genetics 158 (2):927-931
Papp B, Pal C, Hurst LD (2003) Dosage sensitivity and the evolution of gene families in yeast. Nature 424 (6945):194-197. doi:http://www.nature.com/nature/journal/v424/n6945/suppinfo/nature01771_S1.html
Paradis E, Claude J, Strimmer K (2004) APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20 (2):289-290. doi:10.1093/bioinformatics/btg412
Peláez P, Sanchez F (2013) Small RNAs in plant defense responses during viral and bacterial interactions: similarities and differences. Frontiers in plant science 4:343. doi:10.3389/fpls.2013.00343
Pena MI, Davlieva M, Bennett MR, Olson JS, Shamoo Y (2010) Evolutionary fates within a microbial population highlight an essential role for protein folding during natural selection. Molecular systems biology 6:387. doi:msb201043 [pii]

10.1038/msb.2010.43
Peña MI, Davlieva M, Bennett MR, Olson JS, Shamoo Y (2010) Evolutionary fates within a microbial population highlight an essential role for protein folding during natural selection, vol 6. vol 1. doi:10.1038/msb.2010.43
Picard D, Khursheed B, Garabedian MJ, Fortin MG, Lindquist S, Yamamoto KR (1990) Reduced levels of hsp90 compromise steroid receptor action in vivo. Nature 348 (6297):166-168
Poethig RS (2013) Chapter Five - Vegetative Phase Change and Shoot Maturation in Plants. In: Ann ER, Michael BOC (eds) Current Topics in Developmental Biology, vol Volume 105. Academic Press, pp 125-152. doi:http://dx.doi.org/10.1016/B978-0-12-396968-2.00005-1
Posadas DM, Carthew RW (2014) MicroRNAs and their roles in developmental canalization. Current opinion in genetics & development 27:1-6. doi:10.1016/j.gde.2014.03.005
Posé D, Yant L, Schmid M (2012) The end of innocence: flowering networks explode in complexity. Current Opinion in Plant Biology 15 (1):45-50. doi:http://dx.doi.org/10.1016/j.pbi.2011.09.002
Pré M, Atallah M, Champion A, Vos MD, Pieterse CMJ, Memelink J (2008) The AP2/ERF Domain Transcription Factor ORA59 Integrates Jasmonic Acid and Ethylene Signals in Plant Defense. Plant physiology 147 (3):1347-1357
Press MO, Carlson KD, Queitsch C (2014) The overdue promise of short tandem repeat variation for heritability. Trends in Genetics 30 (11):504-512. doi:10.1016/j.tig.2014.07.008
Press MO, Li H, Creanza N, Kramer G, Queitsch C, Sourjik V, Borenstein E (2013) Genome-scale Co-evolutionary Inference Identifies Functions and Clients of Bacterial Hsp90. PLoS Genet 9 (7):e1003631. doi:10.1371/journal.pgen.1003631
Preston JC, Hileman LC (2013) Functional Evolution in the Plant SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE (SPL) Gene Family. Frontiers in plant science 4:80. doi:10.3389/fpls.2013.00080
Pulido A, Laufs P (2010) Co-ordination of developmental processes by small RNAs during leaf development. Journal of experimental botany 61 (5):1277-1291. doi:10.1093/jxb/erp397
Queitsch C, Carlson KD, Girirajan S (2012) Lessons from Model Organisms: Phenotypic Robustness and Missing Heritability in Complex Disease. PLoS genetics 8 (11):e1003041. doi:10.1371/journal.pgen.1003041
Queitsch C, Sangster TA, Lindquist S (2002) Hsp90 as a capacitor of phenotypic variation. Nature 417 (6889):618-624. doi:http://www.nature.com/nature/journal/v417/n6889/suppinfo/nature749_S1.html
Rando OJ, Verstrepen KJ (2007) Timescales of Genetic and Epigenetic Inheritance. Cell 128 (4):655-668. doi:10.1016/j.cell.2007.01.023
Rat-Genome-Sequencing-Project-Consortium (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature 428 (6982):493-521. doi:http://www.nature.com/nature/journal/v428/n6982/suppinfo/nature02426_S1.html
Reis RS, Hart-Smith G, Eamens AL, Wilkins MR, Waterhouse PM (2015) Gene regulation by translational inhibition is determined by Dicer partnering proteins. Nature Plants 1. doi:10.1038/nplants.2014.27
http://www.nature.com/articles/nplants201427 - supplementary-information

Rival P, Press MO, Bale J, Grancharova T, Undurraga SF, Queitsch C (2014) The Conserved PFT1 Tandem Repeat Is Crucial for Proper Flowering in Arabidopsis thaliana. Genetics 198 (2):747-754. doi:10.1534/genetics.114.167866
Robert-Seilaniantz A, Grant M, Jones JD (2011a) Hormone crosstalk in plant disease and defense: more than just jasmonate-salicylate antagonism. Annual review of phytopathology 49:317-343. doi:10.1146/annurev-phyto-073009-114447
Robert-Seilaniantz A, MacLean D, Jikumaru Y, Hill L, Yamaguchi S, Kamiya Y, Jones JDG (2011b) The microRNA miR393 re-directs secondary metabolite biosynthesis away from camalexin and towards glucosinolates. The Plant Journal 67 (2):218-231
Rohner N, Jarosz DF, Kowalko JE, Yoshizawa M, Jeffery WR, Borowsky RL, Lindquist S, Tabin CJ (2013) Cryptic Variation in Morphological Evolution: HSP90 as a Capacitor for Loss of Eyes in Cavefish. Science (New York, NY) 342 (6164):1372-1375. doi:10.1126/science.1240276
Rubio-Somoza I, Weigel D (2012) MicroRNA networks and developmental plasticity in plants. Trends in plant science 16 (5):258-264. doi:10.1016/j.tplants.2011.03.001
Rutherford SL, Lindquist S (1998) Hsp90 as a capacitor for morphological evolution. Nature 396 (6709):336-342. doi:http://www.nature.com/nature/journal/v396/n6709/suppinfo/396336a0_S1.html
Salathia N, Lee HN, Sangster TA, Morneau K, Landry CR, Schellenberg K, Behere AS, Gunderson KL, Cavalieri D, Jander G, Queitsch C (2007) TECHNICAL ADVANCE: Indel arrays: an affordable alternative for genotyping. The Plant Journal 51 (4):727-737. doi:10.1111/j.1365-313X.2007.03194.x
Salomé PA, Bomblies K, Laitinen RAE, Yant L, Mott R, Weigel D (2011) Genetic Architecture of Flowering-Time Variation in Arabidopsis thaliana. Genetics 188 (2):421-433. doi:10.1534/genetics.111.126607
Salverda MLM, Dellus E, Gorter FA, Debets AJM, van der Oost J, Hoekstra RF, Tawfik DS, de Visser JAGM (2011) Initial Mutations Direct Alternative Pathways of Protein Evolution. PLoS Genet 7 (3):e1001321. doi:10.1371/journal.pgen.1001321
Sangster TA, Bahrami A, Wilczek A, Watanabe E, Schellenberg K, McLellan C, Kelley A, Kong SW, Queitsch C, Lindquist S (2007a) Phenotypic Diversity and Altered Environmental Plasticity in <italic>Arabidopsis thaliana</italic> with Reduced Hsp90 Levels. PLoS ONE 2 (7):e648. doi:10.1371/journal.pone.0000648
Sangster TA, Bahrami A, Wilczek A, Watanabe E, Schellenberg K, McLellan C, Kelley A, Kong SW, Queitsch C, Lindquist S (2007b) Phenotypic diversity and altered environmental plasticity in Arabidopsis thaliana with reduced Hsp90 levels. PloS one 2 (7):e648. doi:10.1371/journal.pone.0000648
Sangster TA, Queitsch C (2005) The HSP90 chaperone complex, an emerging force in plant development and phenotypic plasticity. Current Opinion in Plant Biology 8 (1):86-92. doi:http://dx.doi.org/10.1016/j.pbi.2004.11.012
Sangster TA, Salathia N, Lee HN, Watanabe E, Schellenberg K, Morneau K, Wang H, Undurraga S, Queitsch C, Lindquist S (2008a) HSP90-buffered genetic variation is common in Arabidopsis thaliana. Proceedings of the National Academy of Sciences 105 (8):2969-2974. doi:10.1073/pnas.0712210105
Sangster TA, Salathia N, Undurraga S, Milo R, Schellenberg K, Lindquist S, Queitsch C (2008b) HSP90 affects the expression of genetic variation and developmental stability

in quantitative traits. Proceedings of the National Academy of Sciences 105 (8):2963-2968. doi:10.1073/pnas.0712200105
Schmiedel JM, Klemm SL, Zheng Y, Sahay A, Blüthgen N, Marks DS, van Oudenaarden A (2015) MicroRNA control of protein expression noise. Science (New York, NY) 348 (6230):128-132. doi:10.1126/science.aaa1738
Schommer C, Palatnik JF, Aggarwal P, Chételat A, Cubas P, Farmer EE, Nath U, Weigel D (2008) Control of Jasmonate Biosynthesis and Senescence by miR319 Targets. PLoS Biol 6 (9):e230
Sieber P, Wellmer F, Gheyselinck J, Riechmann JL, Meyerowitz EM (2007) Redundancy and specialization among plant microRNAs: role of the MIR164 family in developmental robustness. Development (Cambridge, England) 134 (6):1051-1060. doi:10.1242/dev.02817
Siegal ML, Leu J-Y (2014) On the Nature and Evolutionary Impact of Phenotypic Robustness Mechanisms. Annual Review of Ecology, Evolution, and Systematics 45 (1):495-517. doi:doi:10.1146/annurev-ecolsys-120213-091705
Smith MR, Willmann MR, Wu G, Berardini TZ, Möller B, Weijers D, Poethig RS (2009) Cyclophilin 40 is required for microRNA activity in Arabidopsis. Proceedings of the National Academy of Sciences. doi:10.1073/pnas.0812729106
Sollars V, Lu X, Xiao L, Wang X, Garfinkel MD, Ruden DM (2003) Evidence for an epigenetic mechanism by which Hsp90 acts as a capacitor for morphological evolution. Nature genetics 33 (1):70-74. doi:http://www.nature.com/ng/journal/v33/n1/suppinfo/ng1067_S1.html
Song YH, Ito S, Imaizumi T (2013) Flowering time regulation: photoperiod- and temperature-sensing in leaves. Trends in plant science 18 (10):575-583. doi:10.1016/j.tplants.2013.05.003
Song YH, Shim JS, Kinmonth-Schultz HA, Imaizumi T (2014) Photoperiodic Flowering: Time Measurement Mechanisms in Leaves. Annual Review of Plant Biology. doi:10.1146/annurev-arplant-043014-115555
Spanudakis E, Jackson S (2014) The role of microRNAs in the control of flowering time. Journal of experimental botany 65 (2):365-380. doi:10.1093/jxb/ert453
Specchia V, Piacentini L, Tritto P, Fanti L, D'Alessandro R, Palumbo G, Pimpinelli S, Bozzetti MP (2010) Hsp90 prevents phenotypic variation by suppressing the mutagenic activity of transposons. Nature 463 (7281):662-665. doi:10.1038/nature08739
Stasinopoulos TC, Hangarter RP (1990) Preventing Photochemistry in Culture Media by Long-Pass Light Filters Alters Growth of Cultured Tissues. Plant physiology 93 (4):1365-1369
Stief A, Altmann S, Hoffmann K, Pant BD, Scheible W-R, Bäurle I (2014) Arabidopsis miR156 Regulates Tolerance to Recurring Environmental Stress through SPL Transcription Factors. The Plant cell 26 (4):1792-1807. doi:10.1105/tpc.114.123851
Sun Y, Fan X-Y, Cao D-M, Tang W, He K, Zhu J-Y, He J-X, Bai M-Y, Zhu S, Oh E, Patil S, Kim T-W, Ji H, Wong WH, Rhee SY, Wang Z-Y (2010) Integration of Brassinosteroid Signal Transduction with the Transcription Network for Plant Growth Regulation in Arabidopsis. Developmental Cell 19 (5):765-777. doi:10.1016/j.devcel.2010.10.010

Sung S, Schmitz RJ, Amasino RM (2006) A PHD finger protein involved in both the vernalization and photoperiod pathways in Arabidopsis. Genes & Development 20 (23):3244-3248. doi:10.1101/gad.1493306
Sunkar R (2010) MicroRNAs with macro-effects on plant stress responses. Seminars in cell & developmental biology 21 (8):805-811. doi:10.1016/j.semcdb.2010.04.001
Sunkar R, Li Y-F, Jagadeeswaran G (2012) Functions of microRNAs in plant stress responses. Trends in plant science 17 (4):196-203. doi:10.1016/j.tplants.2012.01.010
Swarbreck D, Wilks C, Lamesch P, Berardini TZ, Garcia-Hernandez M, Foerster H, Li D, Meyer T, Muller R, Ploetz L, Radenbaugh A, Singh S, Swing V, Tissier C, Zhang P, Huala E (2008) The Arabidopsis Information Resource (TAIR): gene structure and function annotation. Nucleic Acids Research 36 (suppl 1):D1009-D1014. doi:10.1093/nar/gkm965
Taipale M, Jarosz DF, Lindquist S (2010) HSP90 at the hub of protein homeostasis: emerging mechanistic insights. Nat Rev Mol Cell Biol 11 (7):515-528
Taipale M, Krykbaeva I, Koeva M, Kayatekin C, Westover Kenneth D, Karras Georgios I, Lindquist S (2012) Quantitative Analysis of Hsp90-Client Interactions Reveals Principles of Substrate Recognition. Cell 150 (5):987-1001. doi:10.1016/j.cell.2012.06.047
Takahashi KH (2013) Multiple capacitors for natural genetic variation in Drosophila melanogaster. Molecular ecology 22 (5):1356-1365. doi:10.1111/mec.12091
Takahashi M, Morikawa H (2014) Nitrogen dioxide accelerates flowering without changing the number of leaves at flowering in Arabidopsis thaliana. Plant Signaling & Behavior 9 (10):e970433. doi:10.4161/15592316.2014.970433
Tang W, Yuan M, Wang R, Yang Y, Wang C, Oses-Prieto JA, Kim TW, Zhou HW, Deng Z, Gampala SS, Gendron JM, Jonassen EM, Lillo C, DeLong A, Burlingame AL, Sun Y, Wang ZY (2011) PP2A activates brassinosteroid-responsive gene expression and plant growth by dephosphorylating BZR1. Nature cell biology 13 (2):124-131. doi:10.1038/ncb2151
Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, McGee S, Do R, Liu X, Jun G, Kang HM, Jordan D, Leal SM, Gabriel S, Rieder MJ, Abecasis G, Altshuler D, Nickerson DA, Boerwinkle E, Sunyaev S, Bustamante CD, Bamshad MJ, Akey JM, GO B, GO S, Project obotNES (2012) Evolution and Functional Impact of Rare Coding Variation from Deep Sequencing of Human Exomes. Science (New York, NY) 337 (6090):64-69. doi:10.1126/science.1219240
Todesco M, Balasubramanian S, Cao J, Ott F, Sureshkumar S, Schneeberger K, Meyer Rhonda C, Altmann T, Weigel D (2012) Natural Variation in Biogenesis Efficiency of Individual Arabidopsis thaliana MicroRNAs. Current Biology 22 (2):166-170. doi:10.1016/j.cub.2011.11.060
Tokuriki N, Tawfik DS (2009) Chaperonin overexpression promotes genetic variation and enzyme evolution. Nature 459 (7247):668-673. doi:http://www.nature.com/nature/journal/v459/n7247/suppinfo/nature08009_S1.html
Turner JG, Ellis C, Devoto A (2002) The Jasmonate Signal Pathway. The Plant Cell Online 14 (suppl 1):S153-S164
Undurraga SF, Press MO, Legendre M, Bujdoso N, Bale J, Wang H, Davis SJ, Verstrepen KJ, Queitsch C (2012) Background-dependent effects of polyglutamine variation in the Arabidopsis thaliana gene ELF3. Proceedings of the National Academy of

Sciences of the United States of America 109 (47):19363-19367. doi:10.1073/pnas.1211021109
Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. Nucleic Acids Research 40 (15):e115-e115. doi:10.1093/nar/gks596
Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM (2009) A census of human transcription factors: function, expression and evolution. Nat Rev Genet 10 (4):252-263. doi:10.1038/nrg2538
Varkonyi-Gasic E, Wu R, Wood M, Walton E, Hellens R (2007) Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. Plant Methods 3 (1):12
Vaucheret H (2008) Plant ARGONAUTES. Trends in plant science 13 (7):350-358. doi:10.1016/j.tplants.2008.04.007
Vaughan CK, Neckers L, Piper PW (2010) Understanding of the Hsp90 molecular chaperone reaches new heights. Nat Struct Mol Biol 17 (12):1400-1404
Vert G, Chory J (2006) Downstream nuclear events in brassinosteroid signalling. Nature 441 (7089):96-100. doi:http://www.nature.com/nature/journal/v441/n7089/suppinfo/nature04681_S1.html
Waddington CH (1953) Genetic Assimilation of an Acquired Character. Evolution 7 (2):118-126
Waldman YY, Tuller T, Shlomi T, Sharan R, Ruppin E (2010) Translation efficiency in humans: tissue specificity, global optimization and differences between developmental stages. Nucleic Acids Research 38 (9):2964-2974. doi:10.1093/nar/gkq009
Wall DP, Hirsh AE, Fraser HB, Kumm J, Giaever G, Eisen MB, Feldman MW (2005) Functional genomic analysis of the rates of protein evolution. Proceedings of the National Academy of Sciences of the United States of America 102 (15):5483-5488. doi:10.1073/pnas.0501761102
Wang H, Wang H (2015) The miR156/SPL Module, a Regulatory Hub and Versatile Toolbox, Gears up Crops for Enhanced Agronomic Traits. Molecular Plant (0). doi:http://dx.doi.org/10.1016/j.molp.2015.01.008
Wang H, Zhang X, Liu J, Kiba T, Woo J, Ojo T, Hafner M, Tuschl T, Chua NH, Wang XJ (2011) Deep sequencing of small RNAs specifically associated with Arabidopsis AGO1 and AGO4 uncovers new AGO functions. The Plant journal : for cell and molecular biology 67 (2):292-304. doi:10.1111/j.1365-313X.2011.04594.x
Wang J-W (2014) Regulation of flowering time by the miR156-mediated age pathway. Journal of Experimental Botany. doi:10.1093/jxb/eru246
Wang J-W, Schwab R, Czech B, Mica E, Weigel D (2008) Dual Effects of miR156-Targeted SPL Genes and CYP78A5/KLUH on Plastochron Length and Organ Size in Arabidopsis thaliana. The Plant Cell Online 20 (5):1231-1243. doi:10.1105/tpc.108.058180
Wang S, Adams KL (2015) Duplicate Gene Divergence by Changes in MicroRNA Binding Sites in Arabidopsis and Brassica. Genome Biology and Evolution 7 (3):646-655. doi:10.1093/gbe/evv023

Wang Y, Mercier R, Hobman TC, LaPointe P (2013) Regulation of RNA interference by Hsp90 is an evolutionarily conserved process. Biochimica et biophysica acta 1833 (12):2673-2681. doi:10.1016/j.bbamcr.2013.06.017
Wang ZY, Nakano T, Gendron J, He J, Chen M, Vafeados D, Yang Y, Fujioka S, Yoshida S, Asami T, Chory J (2002) Nuclear-localized BZR1 mediates brassinosteroid-induced growth and feedback suppression of brassinosteroid biosynthesis. Developmental cell 2 (4):505-513
Wang ZY, Seto H, Fujioka S, Yoshida S, Chory J (2001) BRI1 is a critical component of a plasma-membrane receptor for plant steroids. Nature 410 (6826):380-383. doi:10.1038/35066597
35066597 [pii]
Wapinski I, Pfeffer A, Friedman N, Regev A (2007) Natural history and evolutionary principles of gene duplication in fungi. Nature 449 (7158):54-61. doi:http://www.nature.com/nature/journal/v449/n7158/suppinfo/nature06107_S1.html
Warnecke T, Hurst LD (2010) GroEL dependency affects codon usage--support for a critical role of misfolding in gene evolution. Molecular systems biology 6:340. doi:10.1038/msb.2009.94
Wasternack C (2007) Jasmonates: An Update on Biosynthesis, Signal Transduction and Action in Plant Stress Response, Growth and Development. Annals of Botany 100 (4):681-697
Wegele H, Müller L, Buchner J (2004) Hsp70 and Hsp90—a relay team for protein folding. In: Reviews of Physiology, Biochemistry and Pharmacology, vol 151. Reviews of Physiology, Biochemistry and Pharmacology. Springer Berlin Heidelberg, pp 1-44. doi:10.1007/s10254-003-0021-1
Weigel D, Glazebrook J (2002) Arabidopsis: A Laboratory Manual. CSHL Press,
Werner JD, Borevitz JO, Uhlenhaut NH, Ecker JR, Chory J, Weigel D (2005) FRIGIDA-Independent Variation in Flowering Time of Natural Arabidopsis thaliana Accessions. Genetics 170 (3):1197-1207. doi:10.1534/genetics.104.036533
Whitacre JM (2012) Biological Robustness: Paradigms, Mechanisms, and Systems Principles. Frontiers in Genetics 3:67. doi:10.3389/fgene.2012.00067
Whitesell L, Mimnaugh EG, De Costa B, Myers CE, Neckers LM (1994) Inhibition of heat shock protein HSP90-pp60v-src heteroprotein complex formation by benzoquinone ansamycins: essential role for stress proteins in oncogenic transformation. Proceedings of the National Academy of Sciences 91 (18):8324-8328
Wickland Daniel P, Hanzawa Y (2015) The FLOWERING LOCUS T/TERMINAL FLOWER 1 Gene Family: Functional Evolution and Molecular Mechanisms. Molecular Plant. doi:10.1016/j.molp.2015.01.007
Williams TA, Fares MA (2010) The Effect of Chaperonin Buffering on Protein Evolution. Genome Biology and Evolution 2:609-619. doi:10.1093/gbe/evq045
Woehrer SL, Aronica L, Suhren JH, Busch CJ, Noto T, Mochizuki K (2015) A Tetrahymena Hsp90 co-chaperone promotes siRNA loading by ATP-dependent and ATP-independent mechanisms. The EMBO journal 34 (4):559-577. doi:10.15252/embj.201490062
Workman P, Burrows F, Neckers L, Rosen N (2007) Drugging the cancer chaperone HSP90. Annals of the New York Academy of Sciences 1113 (1):202-216

Wright SI, Yau CB, Looseley M, Meyers BC (2004) Effects of gene expression on molecular evolution in Arabidopsis thaliana and Arabidopsis lyrata. Mol Biol Evol 21 (9):1719-1726. doi:10.1093/molbev/msh191
msh191 [pii]
Wu CI, Li WH (1985) Evidence for higher rates of nucleotide substitution in rodents than in man. Proceedings of the National Academy of Sciences 82 (6):1741-1745
Wu CI, Shen Y, Tang T (2009a) Evolution under canalization and the dual roles of microRNAs: a hypothesis. Genome research 19 (5):734-743. doi:10.1101/gr.084640.108
Wu G, Park MY, Conway SR, Wang J-W, Weigel D, Poethig RS (2009b) The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis. Cell 138 (4):750-759. doi:10.1016/j.cell.2009.06.031
Wu G, Poethig RS (2006) Temporal regulation of shoot development in Arabidopsis thaliana by miR156 and its target SPL3. Development (Cambridge, England) 133 (18):3539-3547. doi:10.1242/dev.02521
Wu L, Zhou H, Zhang Q, Zhang J, Ni F, Liu C, Qi Y (2010) DNA Methylation Mediated by a MicroRNA Pathway. Molecular Cell 38 (3):465-475. doi:http://dx.doi.org/10.1016/j.molcel.2010.03.008
Yamaguchi A, Wu M-F, Yang L, Wu G, Poethig RS, Wagner D (2009) The MicroRNA-Regulated SBP-Box Transcription Factor SPL3 Is a Direct Upstream Activator of LEAFY, FRUITFULL, and APETALA1. Developmental Cell 17 (2):268-278. doi:10.1016/j.devcel.2009.06.007
Yan J, Zhang C, Gu M, Bai Z, Zhang W, Qi T, Cheng Z, Peng W, Luo H, Nan F, Wang Z, Xie D (2009) The Arabidopsis CORONATINE INSENSITIVE1 Protein Is a Jasmonate Receptor. The Plant cell 21 (8):2220-2236. doi:10.1105/tpc.109.065730
Yang L, Huang W, Wang H, Cai R, Xu Y, Huang H (2006) Characterizations of a hypomorphic argonaute1 mutant reveal novel AGO1 functions in Arabidopsis lateral organ development. Plant molecular biology 61 (1-2):63-78
Yang L, Xu M, Koo Y, He J, Poethig RS (2013) Sugar promotes vegetative phase change in Arabidopsis thaliana by repressing the expression of MIR156A and MIR156C. eLife 2:e00260. doi:10.7554/eLife.00260
Yeyati PL, Bancewicz RM, Maule J, van Heyningen V (2007) Hsp90 Selectively Modulates Phenotype in Vertebrate Development. PLoS Genet 3 (3):e43. doi:10.1371/journal.pgen.0030043
Yin Y, Wang ZY, Mora-Garcia S, Li J, Yoshida S, Asami T, Chory J (2002) BES1 accumulates in the nucleus in response to brassinosteroids to regulate gene expression and promote stem elongation. Cell 109 (2):181-191
Ying BW, Taguchi H, Ueda T (2006) Co-translational binding of GroEL to nascent polypeptides is followed by post-translational encapsulation by GroES to mediate protein folding. The Journal of biological chemistry 281 (31):21813-21819. doi:10.1074/jbc.M603091200
Yu S, Cao L, Zhou C-M, Zhang T-Q, Lian H, Sun Y, Wu J, Huang J, Wang G, Wang J-W (2013) Sugar is an endogenous cue for juvenile-to-adult phase transition in plants. eLife 2:e00269. doi:10.7554/eLife.00269
Yu X, Li L, Zola J, Aluru M, Ye H, Foudree A, Guo H, Anderson S, Aluru S, Liu P, Rodermel S, Yin Y (2011) A brassinosteroid transcriptional network revealed by

genome-wide identification of BESI target genes in Arabidopsis thaliana. The Plant Journal 65 (4):634-646. doi:10.1111/j.1365-313X.2010.04449.x

Zhang W, Gao S, Zhou X, Chellappan P, Chen Z, Zhou X, Zhang X, Fromuth N, Coutino G, Coffey M, Jin H (2011) Bacteria-responsive microRNAs regulate plant innate immunity by modulating plant hormone networks. Plant Molecular Biology 75 (1-2):93-105. doi:10.1007/s11103-010-9710-8

Zhao R, Davey M, Hsu YC, Kaplanek P, Tong A, Parsons AB, Krogan N, Cagney G, Mai D, Greenblatt J, Boone C, Emili A, Houry WA (2005) Navigating the chaperone network: an integrative map of physical and genetic interactions mediated by the hsp90 chaperone. Cell 120 (5):715-727. doi:S0092-8674(04)01249-8 [pii]
10.1016/j.cell.2004.12.024

Zheng X, Zhu J, Kapoor A, Zhu JK (2007) Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. The EMBO journal 26 (6):1691-1701. doi:10.1038/sj.emboj.7601603

Zhou X, Wang G, Zhang W (2007) UV-B responsive microRNA genes in Arabidopsis thaliana. Molecular systems biology 3 (103)

Zhu H, Zhou Y, Castillo-González C, Lu A, Ge C, Zhao Y-T, Duan L, Li Z, Axtell MJ, Wang X-J, Zhang X (2013) Bidirectional processing of pri-miRNAs with branched terminal loops by Arabidopsis Dicer-like1. Nat Struct Mol Biol 20 (9):1106-1115. doi:10.1038/nsmb.2646

http://www.nature.com/nsmb/journal/v20/n9/abs/nsmb.2646.html - supplementary-information

Zou J, Guo Y, Guettouche T, Smith DF, Voellmy R (1998) Repression of Heat Shock Transcription Factor HSF1 Activation by HSP90 (HSP90 Complex) that Forms a Stress-Sensitive Complex with HSF1. Cell 94 (4):471-480. doi:10.1016/S0092-8674(00)81588-3

VITA

Tzitziki J. Lemus Vergara was born in Colima, Mexico. During her childhood, she enjoyed mixing showering and cleaning products in the quest for the perfect bubble soap. Then, she proceeded to mix perfumes and essential oils, just to see what the smells would be. When she was 12, she helped a grad student with his DNA extraction, and though that working with liquid nitrogen was the coolest thing one can do. Therefore, she decided to study science. She did her High School education in her hometown and then moved to Cuernavaca, Morelos, to study her undergrad in Genome Sciences, at the National and Autonomous University of Mexico (UNAM). After graduating, she moved to Seattle, Washington, to further continue her education in Genome Sciences. During the course of her graduate work she did science outreach, ran a half marathon, made good friends, and became more knowledgeable about tea, wine, coffee, and chocolate.