PROJECT FINAL REPORT

Grant Agreement number: KBBE - 2009 - 226477

Project acronym: AENEAS

Project title: Acquired Environmental Epigenetics Advances from Arabidopsis to maize

Funding Scheme: Collaborative project

Period covered: from 1st April 2009 to 31st March 2013

Name of the scientific representative of the project's co-ordinator[1],

Prof. Serena Varotto DAFNAE University of Padova

Tel: +39 0498272858

Fax: +39 049 6272839

E-mail: serena.varotto@unipd.it

Project website address: http://www.aeneas-a2m.eu/

[1] Usually the contact person of the coordinator as specified in Art. 8.1. of the Grant Agreement.

Participant number	Participant short name	Contacts
1 coordinator	UNIPD	serena.varotto@unipd.it
2	CRAMAC	vincenzo.rossi@entecra.it
3	UNIWA	J.F.Gutierrez-Marcos@warwick.ac.uk
4	UNIGE	Jerzy.Paszkowski@unige.ch
5	JIC	Caroline.Dean@bbsrc.ac.uk
6	UCAM	dcb40@cam.ac.uk
7	MPG	detlef.weigel@tuebingen.mpg.de
8	BIO	jacques.rouster@biogemma.com

http://www.aeneas-a2m.eu/

EXECUTIVE SUMMARY

Interest in plant epigenetics is increasing both as an area of fundamental research and a source of new traits for breeding. DNA methylation, histone modifications and histone variants play a key role in regulating gene expression during the formation of epigenetic gene variants or epialleles, which can be propagated through mitosis and transmitted to the progeny, often remaining stable for several generations. Project AENEAS aimed to assess the impact of environmental conditions on epigenetic states and in the formation of epialles in the model plant Arabidopsis thaliana and then transfer the knowledge to maize (Zea mays) an important European crop.

The first general objective of AENEAS was to provide advances in understanding the detailed mechanisms of epialleles formation in response to environmental cues and their heritable maintenance in the model plant Arabidopsis, for which molecular and genetic tools were already available at the beginning of the project. Concomitantly, the constitution of an “environmental epigenetic platform” for maize has started with the

development of tools indispensable for the shift of epigenetic research from Arabidopsis to maize, to achieve the second main objective of AENEAS: the transfer of knowledge from model to maize.

Both Arabidopsis and maize research have been sustained by application of genome wide technologies and bioinformatics analysis. The project focused on three epigenetic regulatory pathway: the autonomous, the small RNA and the CpG methylation pathway. These refer to how epiregulators perceive strong environmental change, by which environmentally-induced epialleles are formed and to what extent and how the newly formed epialleles are inherited trough generations.

Outstanding scientific article on Arabidopsis epigenetics were produced and published during the project aiming at: i) understanding the effect on destabilization of A. thaliana genome induced by environmental constrains; ii) elucidating the role of two RNA binding proteins FCA and FPA in Arabidopsis autonomous pathway; iii) producing a genome wide comparison of DNA methylation among ten A. thaliana lines, derived 30 generations ago from a common ancestors. Moreover, in Arabidopsis a new working model for the establishment of DNA methylation at endogenous loci were developed, characterizing different types of sRNAs and focussing on different RNA silencing factors.

In maize, a dataset on response to various environmental stresses previously tested in Arabidopsis was collected and analysed. Plants from both the B73 maize inbred line and maize mutants of epiregulators of the three epigenetic pathways under investigations in AENEAS were then used for preparing small RNAs , RNAs, ChIP-seq and BS-seq libraries. These were sequenced and the data was analysed in order to investigate, at genome-wide level, the effect of stress application on transcription, gene regulation and some epigenetic marks. Specific maize epitargets were identified and are currently under characterization.

In parallel a series of tools for maize were developed, protocols and mutants for epiregulators, some of which were characterized at molecular level in order to determine whether the three pathways are acting in the same way as in Arabidopsis. Indeed, the results indicated that some epiregulators belonging to these pathways have different targets in the maize genome and newly-produced mutants have been characterized and are now available at the Scientific Community.

It is already clear that the novel epigenetic mechanisms present in plants constitute unique opportunities for research into the development of the crops of the future. However, the extent to which species specific epigenetic marks are related to difference in genomic organization, and how they are linked to distinct responses remains to be seen. The comparison that AENEAS have made of the behaviour of Arabidopsis and maize epitargets in response to specific stress treatments will pave the way for further identifications of the similarities and differences between two evolutionarily distinct plant species.

AENEAS come to a close in March 2013 and the Final Meeting was held in Padova, where novel and important evidence produced during the project were discuss in view of new researches towards Horizon 2020.

SUMMARY DESCRIPTION OF THE PROJECT CONTEXT AND THE MAIN OBJECTIVES

The project entitled AENEAS (Acquired Environmental Epigenetics Advances: from *Arabidopsis* to maize) aimed to investigate environmentally-induced epigenetic changes as the “new frontier” of natural and artificial variability. Epigenetics is defined as the study of heritable traits that are not encoded in the primary sequence of DNA and epigenetic mechanisms act through DNA and chromatin modifications. Environmental cues are thought to activate specific epigenetic mechanisms, which add epigenetic marks and in consequence alter spatial and temporal patterns of gene expression, destabilizes the plant genome and cause phenotypic changes that may be transmitted to the progeny, sometimes remaining stable for several generations.

The first general objective of AENEAS was to provide advances in understanding the detailed mechanisms of epiallele formation in response to environmental cues and their heritable maintenance in the model plant

Arabidopsis. To this end, the project focuses on three epigenetic regulatory pathways, which have been well characterized for their interaction with environmental signals in mediating changes into the epigenome. They are the autonomous, the small RNA and the CpG methylation pathways. Cytosine methylation is one of the best characterised epigenetic mechanisms. Methylation occurs in CpG sequence context (mCpG) both in higher plants and animals, but plants differ from animals by significant levels of methylation at symmetric CpNpG (where N can be any nucleotide) and asymmetric CpHp (where H is an A, C or T) sites. Two essential roles have been ascribed to DNA methylation: defending the genome against transposons and regulating gene expression. There is good evidence that siRNAs, which are generated by the sRNA pathway, can provide sequence specificity to guide cytosine methylation as well as other types of epigenetic modifications. In the Arabidopsis genome, there are thousands of loci for production of endogenous small RNAs that, in turn, can epigenetically modify the loci of origin. The epigenetic modifications involve DNA methylation, chromatin silencing or both. The autonomous pathway represents a good example of an interaction between environmental cues and the plant epigenome. This pathway was initially characterised because it regulates the switch from vegetative to reproductive development in Arabidopsis. However, more recent advances indicate that this pathway plays important, genome-wide roles and, in several cases, its function overlaps with other pathways mediating chromatin regulation.

Indeed, several bodies of evidence indicate that these three epigenetic regulatory pathways play a pivotal role in mediating epigenomic change in response to environmental cues. A large part of the information on the environmental-related epialleles formations is arising from studies carried out on the Arabidopsis model plant. In this model, it is well-documented that environmental cues, particularly stresses and shocks, strongly affect gene and genome activity. It is also evident that the environment, in addition to inducing genetic variability due to mutation of the DNA nucleotide sequence, also induces formation of stably inherited epialleles with relevant effects on the phenotype. Importantly, more preliminary results indicate that their function is also conserved in maize. On the initial bases described above, AENEAS research activity has led to the generation of a genetic system(s) responding to a set(s) of environmental conditions, which affect the epigenome to favor adaptive selection of new epi-variants. This objective has been pursued through activity in work package 1 (WP1). Concomitantly, the constitution of an “Environmental Epigenetic platform” for maize has started with the development of tools indispensable for the shift of epigenetic research from *Arabidopsis* to maize and this has been directly followed by the second main objective of the AENEAS project: the transfer of knowledge from model plant to maize, an important European crop (WP2 and WP3 activity, respectively). Throughout the achievement of these two general objectives of the AENEAS project, we have started to make available for the scientific community a series of deliverables that can be the “progenitors” for the “next-generation” of breeding programs, based on the exploitation of the environmentally-induced epigenetic variability. For instance, it is thought that the ability to respond to abiotic stresses represents the main constraint on maize production. Thus the future release of maize varieties with high yield and stability will largely depend on the possibility to increase tolerance to environmental stresses. In plant improvement, breeders select for a consistent phenotype that is often conditioned by many genes with incremental effects, rather than single large effect mutations. Therefore, two important questions to address were: how much of this available variation exploited in breeding programs is caused by epigenetic processes, and how it is possible to increase epigenetic diversity from which advances in crop research and development can be drawn.

Furthermore, the main objectives of AENEAS research on Arabidopsis were the characterization of the epigenetic mechanisms involved in mitotically and meiotically heritable changes of epigenetic information in response to environmental cues and the development of reliable assays to efficiently monitor environmentally induced epigenomic alterations in this model plant. Each of the three WP1’s participants

(**P4**, **P5**, and **P6**) focused their research activity on three selected pathways and applied different stress treatments, on the basis of their specific expertise and tools. They were interacting by using comparable developmental stages of plants for stress treatment and recovery, as well as comparable strategies for subsequent genomic analyses, which comprise RNA profiling, DNA methylation and siRNAs analyses. The research on Arabidopsis showed that the autonomous pathway has a general role on gene transcription termination and sRNA pathway is involved in controlling both genome stability and abiotic stress response. The next-generation sequencing of Arabidopsis and bioinformatics support for the AENEAS project were guaranteed by activity in WP4 coordinated by **P7**. **P7** has previously developed the SHORE pipeline for mapping of short reads and a graph method, GenomeMapper, for effective mapping of variant sequences and adopted GenomeMapper in the context of SHORE for the analysis of reads that come from bisulfite conversion of Arabidopsis gDNA. P7 has compared genome wide.DNA methylation among 10 A. thaliana lines, derived 30 generations ago from a common ancestor, showing that CpG methylation marks is partially maintain across generations.

A further main objective of AENEAS research was to generate the molecular and genetic tools in a crop, maize, for the exploitation of the epigenetic regulation of gene expression in response to environmental stresses.. The objective of WP2 is to identify maize transcriptomic and epigenomic changes induced by stress treatments and analyze how the above mentioned epi-regulatory pathways (autonomous, small RNA and CpG methylation) will contribute to this changes. To achieve this objective, after preliminary investigations and considering the indications from Arabidopsis that provide some evidences on the effectiveness of specific treatment in the formation of trans-generationally stable epialleles, different stresses were applied to maize B73 and to maize mutants in epi-regulators by **P1**, **P2** and **P3**. To address the role of stress-induced epigenetic gene and TE regulation **P1** has analyzed the salt and drought stresses effects on transcriptional modulation (RNA-seq), transposon activity and on associated smallRNAs (sRNA-seq) and epigenetic marks (ChIP-seq for these histone modifications H3K4me3, H3K9ac and H3K27me3 in collaboration with **P2**), both in B73 inbred line and the epiregulator mutant rmr6. A collaboration between **P3** and **P7** was set to identify stress epi-targets, using a genome-wide approach. The meristematic area (MA) from cold and heat stressed and control populations at different time points was collected by **P3** to produce Illumina libraries for sequencing: - P3 prepared and sequenced mRNA and small RNA (sRNA) libraries, **P7** prepared libraries for BIS-seq form gDNA. In the time course of the project maize mutants for environmental epi-regulators were identified and characterized: insertional lines for five epi-regulators (one mutant for each pathway) were identified by **P8** and introgressed by **P1** in B73 background. **P8** has also produced transgenic mutants of the three epigenetic pathways using a RNAi approach. Finally P2 has has deeply characterized an epiregulator mutant of the autonomous pathway and studied the effect of stress in maize genome destabilization. While significant leaps in understanding of epigenetic phenomena had been made before AENEAS' inception in 2009, several key questions remained, and the project has made solid progress towards finding workable answers. For instance, it has helped to clarify the interactions between the environment and epigenetic mechanisms and is already applying knowledge from its basic research to environmentally induced and epigenetic-related sources of variability. At the beginning of the project it was not clear whether different abiotic stresses and treatments cause distinct changes in the epigenome; since then, AENEAS has established protocols for the most effective stress treatments in epiallele formation in *Arabidopsis* and maize, where a genome-wide evaluation of their effects has produced interesting results.

Despite these successes, it is clear that there will be challenges ahead, as maize exhibits some characteristics not present in *Arabidopsis*, necessitating further characterizations and the development of unique tools. During the last part of AENEAS, the Consortium had the opportunity to start the comparison between the

information produced on the model plant (Arabidopsis) and those coming from the crop (maize). This opened an interesting discussion both from an evolutionary and applicative point of view. The Consortium stressed the importance of crossing the data produced by the participants in order to depict a picture as much complete as possible on epialle formation in maize. A further effort is now needed for understanding the hereditability of epialles in maize, for their future exploitation in breeding programs. AENEAS research into the correlation between environment and epigenetics in maize is relatively young and only few epi-regulatory pathways have been characterized at a functional level, but the project has created a valuable collection of tools and information which together constitute the maize environmental-epigenetics platform. The objectives of AENEAS WP5, the coordination WP, was to collect and coordinate all the activities related to the management of the project, the dissemination of the results, the staff exchange and the technology transfer among laboratories. P1, the Coordinator, was the responsible for maintaining the contact with the members of AENEAS Advisory Board, which was appositely constituted to guarantee the appropriate execution of the planned work within the project and to further exploit potential impacts from the project, throughout collaborations with non-EC Institutes of excellence. At the end of the project, it was a general opinion that one of the outstanding feature of AENEAS research project was the constitution of a consortium in which European groups, leaders in Arabidopsis research, were combined with groups working on maize. This combination has allowed a very fruitful collaboration making the ideas circulating among the people and driving the research on the right direction: AENEAS succeeded very well in the realization of the idea " transfer of knowledge from model to crop".

Description of the main S & T results/foregrounds

Participant number	Participant short name	Beneficiary	Name of Principle Investigator
1 coordinator	UNIPD	Università degli Studi di Padova	Serena Varotto
2	CRAMAC	CRA Unità di Ricerca per la maiscoltura	Vincenzo Rossi
3	UNIWA	University of Warwick	Josè Gutierrez Marcos
4	UNIGE	University of Geneva	Jerzy Paszkowski
5	JIC	John Innes Centre	Caroline Dean
6	UCAM	University of Cambridge	David Baulcombe
7	MPG	Max Planck Institute Tubingen	Detlef Weigel
8	BIO	BIOGEMMA	Jacques Rouster

Work package 1

Work package number	**1**	**Start date or starting event:**			Month 1	
Work package title	Epigenetic responses to environmental signals in *Arabidopsis*					
Activity Type[1]	RTD					
Participant number	4	5	6	7			
Participant short name	UNIGE WP coordinator	JIC	UCAM	MPG			

General Objectives:

The objectives of WP1 are the characterization of epigenetic mechanisms involved in mitotically and meiotically heritable changes of epigenetic information in response to environmental and development of reliable assays to efficiently monitor environmentally induced epigenomic alterations.

Specific objectives of WP1 for the period are:

1. To characterize the global transcript changes in mutants affected in the autonomous, small RNA (sRNA) and CG methylation (^{m}CpG) pathways – aiming to determine chromosomal targets affected by each of the pathways and thus the extent of their overlap. For certain mutants such as *met1,* and *fca*/*fpa* these data are available already, thus can be used directly for the comparative analysis.

2. To characterize the changes in RNA processing and in populations of small RNAs due to the deficiencies in each of the above mentioned pathways, allowing the overlaps between the pathways to be described in terms of RNA metabolism and RNA directed regulation. For certain mutants such as *fca*/*fpa* and *nrpd1a* small scale analysis have been completed and these data can be used directly for the comparative analysis.

3. To characterize genomic DNA methylation, chromatin properties and possibly direct targets of selected components of each of the pathways. For certain mutants (*met1* at a genome-wide level and *fca*/*fpa* on a small number of selected targets) these data are available already, thus can be used directly in the comparative analysis.

4. To examine the influence of environmental cues/stresses/shocks of progressive severity on chromosomal targets regulated by a specific pathway, and targets controlled by multiple pathways (selected from objectives **1, 2** and **3**). We aim to establish the most effective stress condition influencing each of the studied pathways (or several of them simultaneously). Plant material subjected to specific environmental triggers will be reassessed for genome-wide transcriptional responses, DNA methylation and chromatin analyses.

Summary of progress towards objectives:

1. Utilization of Affymetrix Whole Genome Tiling Array to investigate the genome-wide mis-expression profiles of the autonomous, mCpG and RNAi pathways. Comparative analysis of genome-wide differential expression profiles of mutants of the autonomous pathways (fcafpa double mutants), RNAi mutants (dcl1, dcl23, hyl1, serrate and rdr6) and previously published data of the Arabidipsis DNA methylation mutants (drm1-1drm2-2cmt3-11ddc triple mutant).

2. Analyses of small RNA profiles in Arabidopsis wt and fcafpa double mutants (autonomous pathway) after cold stress application and comparison with previously published ddm1 mutant (mCpG pathway).

3. Studies on the destabilization of transcriptional gene silencing (TGS) in wt Arabidopsis and in mutants compromised in mCpG and siRNA pathways.

4. Development of a variety of different stress assays (high temperature, low temperature, high salt, osmotic stress, drought, Cu^{++}, Cd^{++}) to determine the effect of stress on Arabidopsis seed germination and seedling growth. The *nrpd1a* mutant and the respective wild-type were compared in these assays.

Task 1 Discovery of epigenetic mechanisms involved in stress mediated modulations of the autonomous pathway (P5 Lead participant; other participants P4– P6– P7)

Background:

FCA and FPA have been identified as RNA binding proteins that have important roles in vegetative to floral transition in *Arabidopsis thaliana*. It has been previously shown that FCA and FPA have widespread roles in silencing transgenes and transposons as well as single copy genes.

Sub-task 11 and 1.2: Transcriptome regulation by autonomous pathway and validation of the transcriptome data.

In order to investigate the extent of FCA and FPA targets in the *Arabidopsis* genome, **P5** employed the Affymetrix Whole Genome Tiling Array Chip in collaboration with **P7**. Genome-wide transcript levels of *Arabidopsis thaliana* Columbia (Col-0) accession were compared to *fcafpa* double mutant seedlings from 7 and 17-day old plants, respectively. The results indicate a developmental difference in number of targets misregulated at two different time points. Performed analysis revealed that FCA and FPA have genome-wide targets not restricted to a certain class of genes or transcripts that belong to a single pathway/process. A number of upregulated segments from the *fcafpa* microarray analysis were selected for further verification.

Initially, **P5** focused on unannotated segments that were novel transcripts in *fcafpa* double mutants. **P5** selected a number of upregulated segments from the *fcafpa* microarray analysis for further verification. Out of 82 annotated segments, 27 were tested for verification via one-step RT-PCR and 18 were confirmed (~67%). Of these 18, 15 were further selected for analysis via quantitative RT-PCR. Particularly transcriptional changes were analyzed in the 15 selected targets in *fcafpa* double as well as *fca* and *fpa* single mutants compared to wild type Col-0. FCA and FPA show functional redundancy and they have differential interaction on different loci. **P5** analyzed the 15 verified upregulated segments either using RT-PCR followed by sequencing or 5′ and 3′ rapid amplification of cDNA ends (RACE) analysis. This revealed a complex picture, with every misexpressed segment resulting from a slightly different event at each locus. Loss of FCA and FPA led to increased transcript read-through that continued several kilobases downstream to the adjacent gene, i) through an adjacent gene, ii) into an intergenic sequence, iii) into transposon rich regions, iv) or in convergently transcribed genes. Northern analysis of the misexpressed transcripts suggested that loss of FCA and FPA resulted mainly in increased transcriptional read-through. Northern blot also revealed accumulation of high molecular weight transcripts in fpa single and fcafpa double mutants, which was not detected by the RT-PCR or 3′ RACE analyses. In summary, **P5** investigations showed that transcriptional read-through, alternative polyadenylation, and alternative splicing result in apparently unannotated genomic segments being misexpressed in the fcafpa double mutant. In some cases, the transcriptional read-through significantly reduced expression of the associated genes. Furthermore FCA/FPA dependent changes in DNA methylation were found at several loci, supporting previous associations of FCA/FPA function with chromatin modifications and suggesting that extensive read-through transcription because of defective 3′ processing is associated with chromatin changes and that FCA and FPA are involved in the interplay of these co-transcriptional mechanisms. Finally, data suggest that FCA and FPA play important roles in the A. thaliana genome in RNA 3′ processing and transcription termination, thus limiting intergenic transcription.

C Sonmez, I Bäurle, A Magusin, R Dreos, S Laubinger, D Weigel and C Dean (2011) RNA 3' processing functions of Arabidopsis FCA and FPA limit intergenic transcription. PNAS 108 (20): 8508-13

In collaboration with **P7**, **P5** employed the Affymetrix Whole Genome Tiling Array in order to do a comparative analysis of genome-wide differential expression profiles of mutants of the Autonomous pathway (i.e. *fcafpa* double mutants), RNAi mutants (i.e. *dcl1*, *dcl234*, *hyl1*, *serrate* and *rdr6*) and previously published data of the Arabidopsis DNA methylation mutants (*drm1-1drm2-2cmt3-11* (*ddc*) triple mutants) (Kurihara et. al. *Biochem Biophys Res Commun* 2008 Nov 21;376(3):553-7).

P5 and **P7** observed that the percentage of differentially expressed segments is quite similar between Autonomous, RNAi and DNA methylation pathways at 2 weeks. However, only ~9 % of the segments were misregulated in 1-week old *fcafpa* mutants. These results indicate a developmental bias in differential expression profiles of *fcafpa* mutants.

Of the total number of misregulated segments in *fcafpa* mutants, ~27% were in common with the RNAi and DNA methylation mutants irrespective of the time point. At 2-weeks, the autonomous pathway shared ~15 and 12% segments in common with the RNAi and DNA methylation pathways, respectively. When RNAi and DNA methylation pathways are compared alone, the percentage of the segments that were commonly misregulated was ~15%. These results show the extent of commonality between the Autonomous, RNAi and DNA Methylation pathways and suggest a genome-wide role of the autonomous pathway in gene silencing.

The tiling array analysis of the genome-wide targets of FCA and FPA with **P7** in wild type Columbia and fcafpa mutants at two different time points identified a series of target sequences that comprised several unannotated segments (UAs). The detailed analysis of these targets showed that upregulation of UA segments/downregulation of genes in fcafpa can partly be explained by a "switch" in polyadenylation site.
The analysis of the antisense transcript of Helitron1 (AT1TE93275) that is upregulated in fcafpa mutants demonstrated a possible overlap of the Autonomous pathway with other silencing pathways, such as the DNA methylation and the sRNA pathway. This last observation was also confirmed by the analysis of several others UAs. Furthermore it was observed that DNA Methylation at the Helitron 1 locus is perturbed in fcafpa double mutant.
To better characterized the fcafpa double mutant In collaboration with **P5** , **P4** performed RNA-seq using Illumina technology for the genome-wide analysis of cold stress on wild type Col vs fcafpa plants, sequencing both polyA and non-polyA transcripts, strand-specific and both strands.The sequencing was carried out in 2012 and repeated in 2013.
P5 focused the research on the autonomous pathway, analyzing the role of proteins involved in this pathway and the role of antisense RNA production at the FLC locus. The fcafpa mutant was shown to inhibit proper polyadenylation of mRNAs, resulting in transcriptional read-through, longer transcripts and alternative polyA tails. At the FLC locus this means the production of long antisense RNA instead of short antisense - COOLAIR – RNA, hampering the downregulation of FLC expression and thereby flowering. Analyses of polyA and non-poly RNAs showed that in fcafpa mutant, cold treatment can substitute the function of FCA and FPA proteins; cold treatment of wildtype and fcafpa mutant gave rise to similar RNA profiles. The **P5** lab also identified a homeodomain protein, AtNDX that represses expression of COOLAIR, delaying flowering. It appears to do so via stabilization of an R-loop at the COOLAIR promoter.

P5 have investigated whether higher-order chromatin structures are involved in the regulation of the Arabidopsis floral repressor gene FLC. They identified a gene loop involving the physical interaction of the 5' and 3' flanking regions of the FLC locus using chromosome conformation capture. The FLC loop is unaffected by mutations disrupting conserved chromatin regulatory pathways leading to very different expression states. However, the loop is disrupted during vernalization, the cold-induced, Polycomb-dependent epigenetic silencing of FLC. Loop disruption parallels tim-ing of the cold-induced FLC transcriptional shut-down and upregulation of FLC antisense transcripts, but does not need a cold-induced PHD protein required for the epigenetic silencing.
Pedro Crevillen, Cagla Sonmez, Zhe Wu and Caroline Dean (2013) A gene loop containing the floral repressor FLC is disrupted in the early phase of vernalization. The EMBO Journal (2013) 32, 140–148.

Sub-task 1.4: Transcriptional stress response of the FCA/FPA targets
In collaboration with the **P6**, **P5** has analysed the small RNA profile of Col seedlings before and after 2 weeks cold and fcafpa double mutants. Size-fractionated small RNA from 2 week old seedlings (in the size range 20-

30) were cloned and sequenced using 454 technology. From 208k extractable reads there were 30,332 wt; 128,536 fcafpa; 49,908 2w cold. David Studholme (Sainsbury Laboratory) analyzed the samples and found a similar profile between Col and fcafpa. However, there was a large difference in small RNA size distribution when Col before and after 2 week cold was compared. A more detailed bioinformatic analysis on this dataset was generously performed by Alexis Sarazin in Vincent Colot's group (IBENS Paris). In the cold-treated sample there was an overabundance of 21nt-long sRNAs (40% of total, 19 422 out of 48 002 sequences) compared 24nt-long sRNAs (10% of total, 11 214 out of 48 002 sequences). The distribution of matching reads was plotted for each chromosome, the density of reads was highest in the centromeric and pericentromeric regions. Those in the euchromatic arms correspond almost exclusively to miRNA sequences. An overrepresentation of the miRNA category in cold-treated seedlings explains most of the 21-mer overabundance observed in the Col0 cold-treated library. Inspection of the 21nt class indicates that most of the increase in miRNA abundance in cold-treated seedlings is due to miR169 or its priMiRNA stem loop (58,6%, or 10363 out of 14364 21nt reads; vs (0.3% or 7 out of 5131 21nt reads in our Col0 library).

In conclusion, the comparisons of genome-wide small RNA profiles in Arabidopsis mutants for autonomous, sRNA and DNA methylation pathways grown under cold stress conditions, showed that cold stress may affect transcriptional read-through, possibly through a different mechanism than fcafpa and other silencing mutants. JIC lab started a collaboration with **P7** for RNA-seq of the following materials: Col-0, fcafpa, Col-0 2w cold; fcafpa 2w cold; Ler; fcafpa in Ler.

Task 2 <u>Discovery of epigenetic mechanisms involved in stress mediated modulations of</u> the small RNAs pathway (Lead participant P6; other participants: P4– P5– P7)

A variety of different stress assays (high temperature, low temperature, high salt, osmotic stress, drought, Cu^{++}, Cd^{++}) have been developed to determine the effect of stress on Arabidopsis seed germination and seedling growth. The *nrpd1a* mutant and the respective wild-type were compared in these assays. No phenotypic variation was observed between the *nrpd1a* mutant and wild-type under these conditions. Seed from the primary stressed plants and controls, both mutant and wild-type, has been harvested in order to test the effects of the same stresses on the progeny.

Heat, cold and high salt treatments were used to analyze the effects of stress on global small RNA levels. Stress treatments have been performed on the *nrpd1a* mutant and relevant wild-type seedlings. Shoot tissue from stressed and control plants were harvested and small RNA libraries were sequenced. Bioinformatic analysis (in collaboration with Tom Hardcastle) was undertaken to determine differentially expressed sRNAs in the different samples. Very few sRNAs were differentially expressed in response to high salt conditions (approximately 3%); it is likely that this result is due to the stress treatment and not to do with the construction and analysis of the sRNA cDNA libraries. In contrast, 34% of sRNA loci were differentially expressed in response to high temperature. Of the loci that were differentially expressed, 80% were up-regulated. The majority of the up-regulated genes were 24-nucleotide in length suggesting novel epigenetic marks could be established in response to heat. However, the majority of these differentially expressed loci have similar sRNA profiles to those of floral tissue and DNA methylation is associated with these sRNA populations. Molecular analysis of the heat stressed plants have shown that known markers for the transition from vegetative to reproductive phase where up-regulated in response to high temperature suggesting that the induced 24-nucleotide sRNAs are a result of this transition. A second round of bioinformatic analysis focussed on comparing differentially expressed sRNA loci in control, heat and floral

samples. The analysis identified 21-nucleotide sRNAs that were induced in response to high temperature. Specifically, a set of sRNAs that map to VANDAL6 transposable elements were induced by heat and they persist following recovery after the stress. The VANDAL6 transposable elements are of the DNA/MuDR subfamily class and their transcription was also induced in response to heat stress. The 21-nucleotide sRNAs are likely the product of RNA silencing mechanisms that target the VANDAL6 transposable elements. Preliminary data suggests that the VANDAL6 sRNAs can also target endogenous genes that are repressed in response to high temperatures, suggesting a large heat stress network of gene regulation as a result of VANDAL6 transposable element expression. Current work is focussing on the RNA silencing components required for this response.P6 is also assembling sensor constructs in order to test transcriptional gene silencing and trans-generational inheritance of gene silencing induced by siRNAs. This construct involves a multiple cloning site placed upstream of a minimal 35s promoter driving the expression of the GFP reporter gene. Target regions of small RNAs will then be cloned into the multiple cloning site and the level of GFP expression will be used as an output for this small RNA target region – do small RNAs target genomic region in the reporter construct and cause stable silencing of *GFP* in progeny? The use of the sensor construct will allow us to detect low frequency events.

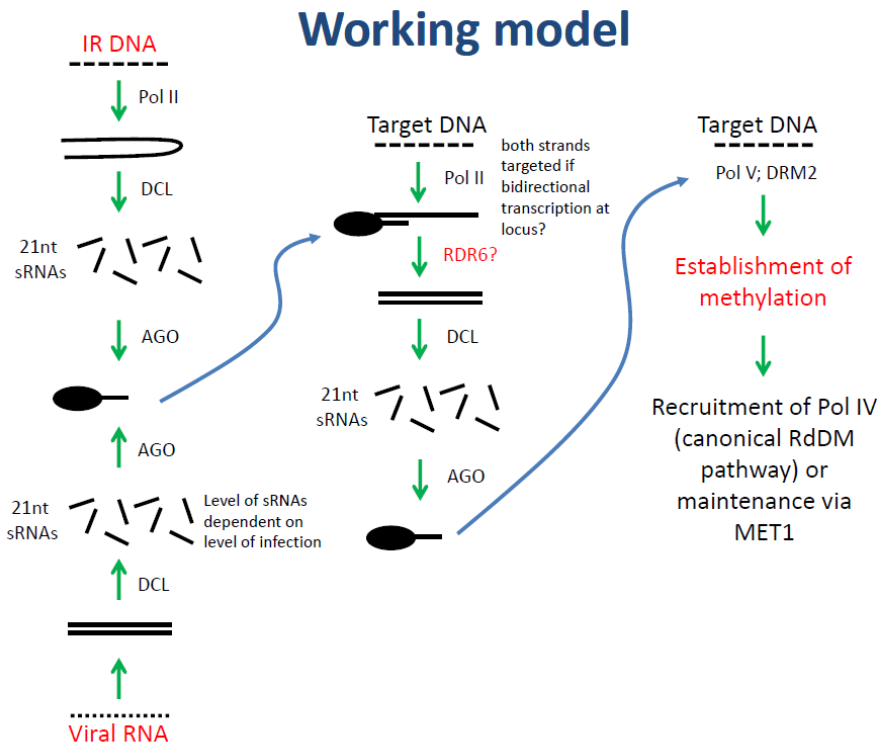
Initial constructs have been made that include cDNA fragments from the Tobacco Rattle Virus (TRV) genome. These constructs have been transformed into Arabidopsis and, once homozygous lines have been identified, they will be infected with TRV. Infections of wild-type Arabidopsis with TRV have been performed and optimized. that can direct heritable epigenetic modifications at an endogenous locus. Tobacco Rattle Virus (TRV) Virus Induced Gene Silencing (VIGS) has been used to target promoter regions of endogenous genes for silencing. No silencing of adjacent genes has been observed when VIGS was employed to target promoter regions that have never experienced silencing or DNA methylation. In contrast, VIGS was successful when the promoter of FWA was targeted in an FWA epi-mutant line. In wild-type plants, the promoter region of FWA is heavily methylated which causes the gene to be in a repressed state and as a result Arabidopsis wild-type (Col-0) plants flower early. In the FWA epi-mutant, methylation is lost at the FWA promoter which releases the repression of FWA so these plants flower late. TRV VIGS constructs containing a portion of the FWA promoter was successful in inducing methylation at the FWA promoter. This increase in methylation did not occur in the infected plants but in the progeny of plants that exhibit high levels of viral sRNAs as a result of the viral infection. The increase in methylation at the FWA promoter in the progeny plant causes an early flowering, wild-type phenotype. Various RNA silencing mutations were crossed into the FWA epi-mutant background in order to determine the factors required to establish DNA methylation.

An additional system to study the establishment DNA methylation was set-up using inverted repeat transgene constructs. The targets of these constructs were two novel epialleles in Arabidopsis: RITA and MRD1, two loci that exist in differentially methylated states within different inbred populations of Arabidopsis. Both of these loci comprise of two divergent transcripts that overlap at their 5' regions – one is a protein coding RNA and the other is a proposed non-coding RNA. Where the two transcripts overlap is the site of the differential methylation – if the overlap region is methylated both transcripts are repressed, if the overlap region is unmethlyated both transcripts are expressed. In order to determine the factors required to establish DNA methylation at these loci, inverted repeat transgenes that target the overlap regions at RITA and MRD1 were transformed into various RNA silencing mutants where RITA and/or MRD1 exist in an unmethylated state.

Establishment of DNA methylation at the different endogenous loci, using VIGS or the inverted repeat transgene, required Pol V and the de novo methyltransferase gene DRM2. However, other components of

the canonical RdDM pathway were not required. Surprisingly this included DCL3, the DICER-like enzyme responsible for the production of 24-nucleotide sRNAs from double stranded RNA precursors. This result suggested that other size classes of RNAs could be responsible for the establishment of DNA methylation at target loci. In support of this hypothesis, 21-nucleotide primary sRNAs are produced from the inverted repeat transgene and 21-nucleotide secondary sRNAs from the target loci (RITA or MRD1) in a dcl3 mutant. In addition, 21-nucleotide sRNAs are produced as a result of viral infection in the VIGS experiments that target the FWA epimutant. Current work is focussing on other RNA silencing factors, such as those involved in the trans-acting sRNA pathway, and their role in the establishment of DNA methylation.

These results have led us to produce a new working model for the establishment of DNA methylation at endogenous loci.



When sRNAs complementary to a target are supplied from an exogenous source (VIGS or an inverted repeat), the primary phase involves 21-nucleotide sRNAs that act to prime amplification of sRNAs at a target locus (a trans-acting like pathway). In order to establish methylation Pol V is required suggesting a Pol V transcript must be present at the target. The amplified 21-nucleotide sRNAs can bind to the complementary Pol V transcript which will recruit the de novo methyltransferase to methylate DNA at the target region. This will then recruit the canonical RdDM components to the locus (perhaps because Pol IV prefers methylated DNA as a template), which will lead to the production of 24-nucleotide sRNAs at the target. Maintenance of DNA methylation at target regions will depend on the context of methylation – symmetric methylation will be maintained independent of sRNAs whereas maintenance of symmetric methylation will depend on 24-nucleotide sRNAs. Final experiments for this project are being conducted to further support this model.

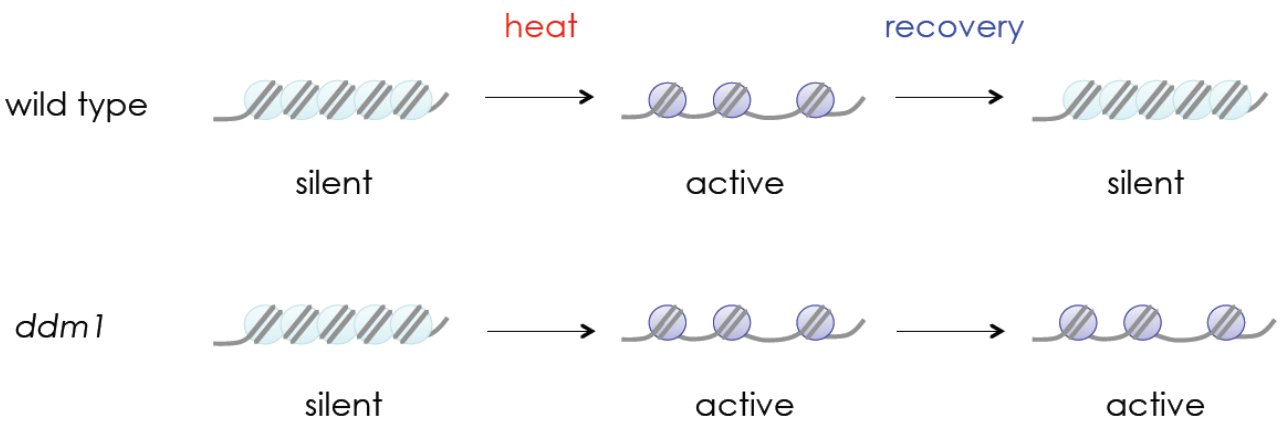
Task 3: Discovery of epigenetic mechanisms involved in stress mediated modulations of the CpG methylation pathway (Lead participant P4 other participants: P5 – P6– P7)

Background: Recently, **P4** has established particular environmental "shock" conditions that release silencing of "constitutively" silent CpG methylated heterochromatic loci.
P4 described an experimental system designed to test the influence of various environmental challenges on transcriptional suppression in Arabidopsis heterochromatin. The system exploits the well-documented observation that multicopy transgenic inserts tend to acquire properties and epigenetic marks characteristic of constitutive heterochromatin. Such silent transgenic loci can be activated in mutants affecting epigenetic regulation of endogenous targets residing in heterochromatin. **P4** applied a series of abiotic stresses to transgenic Arabidopsis plants and used the activation of an originally silent transgenic locus as readout for the destabilization of heterochromatic TGS. This approach allowed the definition of environmental stress conditions that not only destabilize transgene silencing but also result in genomewide
reactivation of endogenous heterochromatic loci. However, silencing release was mostly transient and was rapidly restored upon return to normal growth conditions. This transient activation of heterochromatic transcription occurred genome wide and was not associated with changes in DNA methylation or repressive histone modifications that were examined at a subset of reactivated loci. Intriguingly, mutations in common epigenetic gene silencing regulators, including those involved in de novo DNA methylation or H3K9me, did not prevent rapid resilencing after stress treatments.
M Tittel-Elmer, E Bucher, L Broger, O Mathieu, J Paszkowski and I Vaillant (2010) Stress-Induced Activation of Heterochromatic Transcription. Plos Genetics 6(10): e1001175.

P4 also analyzed the effect on genome destabilization in Arabidopsis seedlings subjected to heat stress and reported that a copia-type retrotransposon named ONSEN (Japanese 'hot spring') not only became transcriptionally active but also synthesized extrachromosomal DNA copies. Heat-induced ONSEN accumulation was stimulated in mutants impaired in the biogenesis of small interferingRNAs (siRNAs); however, there was no evidence of transposition occurring in vegetative tissues. After stress, both ONSEN transcripts and extrachromosomal DNAgradually decayed and were no longer detected after 20–30 days. Surprisingly, a high frequency of new ONSEN insertions was observed in the progeny of stressed plants deficient in siRNAs. Insertion patterns revealed that this transgenerational retrotransposition occurred during flower development and before gametogenesis. Therefore in plants with compromised siRNA biogenesis, memory of stress was maintained throughout development, priming ONSEN to transpose during differentiation of generative organs. Retrotransposition was not observed in the progeny of wild-type plants subjected to stress or in non-stressed mutant controls, pointing to a crucial role of the siRNApathway in restricting retrotransposition triggered by environmental stress. Finally, we found that natural and experimentally induced variants in ONSEN insertions confer heat responsiveness to nearby genes, and therefore mobility bursts may generate novel, stress-responsive regulatory gene networks.
H Ito, H Gaubert, E Bucher, M Mirouze, I Vaillant1 & Jerzy Paszkowski (2011) An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. Nature 7(472):115-9

P4 screened for mutants able to release gene silencing upon heat shock. As a model system they used a transcriptionally silenced luciferase gene in a mom1 mutant background. They found multiple mutants and selected four, two of which appeared to be ddm1 mutants. Further testing showed that ddm1 and mom1 mutants are sensitive to heat stress on their own, and the ddm1 and ddm1 mom1 double mutants prevent the luciferase gene from getting silenced again. The release of silencing was not associated with major changes in DNA methylation and they hypothesized the effect could act at the nucleosome occupancy level.

This is the current working model of P4 on DDM1 concerning the effect of stress on genome destabilization



P4 is also involved in isolation of retrotransposons activated by stress. To identify the activated transposons they focused on extrachromosomal DNA derived from retrotransposons and developed a SIRT assay (sequence-independent retro-element trapping) that can identify mobilized LTR transposons. They performed studies on both Arabidopsis and maize in collaboration with **P2**, using the SIRT assay. In Arabidopsis they observed ONSEN and EVD LTR element DNA upon heat stress. In maize (collaboration with P2) they are analyzing sequenced libraries produced from non-stressed and cold-stressed wild-type B73 and rmr6 mutant plants.

P4 also analyzed meiotic recombination in mutant plants with hypomethylated DNA and observed unexpected and counterintuitive effects of DNA methylation losses on cross-overs (CO) distribution. Recombination was further promoted in the hypomethylated chromosome arms while it was inhibited in heterochromatic regions encompassing pericentromeric DNA. Importantly, the total number of COs was not affected, implying that loss of DNA methylation led to a global redistribution of COs along chromosomes. To determine by which mechanisms altered levels of DNA methylation influence recombination—whether directly in cis or indirectly in trans by changing expression of genes encoding recombination components—we analyzed CO distribution in wild-type lines with randomly scattered and well-mapped hypomethylated chromosomal segments. The results of these experiments, supported by expression profiling data, suggest that DNA methylation affects meiotic recombination in cis.
Marie Mirouze, Michal Lieberman-Lazarovich, Riccardo Aversano, Etienne Bucher, Joël Nicolet, Jon Reinders and Jerzy Paszkowski (2012) Loss of DNA methylation affects the recombination landscape in Arabidopsis. PNAS, 2012 vol. 109 (1) 5880-5

Work package 2

Work package number	**2**	**Start date or starting event:**			Month 1	
Work package title	Molecular and genetics tools for studies of epigenetic responses to environment in maize					
Activity Type[1]	RTD					
Participant number	1	2	3	7	8		
Participant short name	UNIPD	CRAMAC WP coordinator	UNIWA	MPG	BIO		

Objectives:

The main objective of WP2 is to generate the molecular and genetic tools in maize for the exploitation of the epigenetic regulation of gene expression in response to environmental stresses. This WP has three specific objectives:

1. The identification, through genome-wide analysis, of the target sequences (hereinafter named epi-targets) of selected stress and epi-regulator mutants of three key maize environmental epi-regulators homologues of their *Arabidopsis* counterparts (*FVE*, *PolIVa*, and *MET1* genes) characterized in WP1. These environmental epi-regulators have already been characterized at a genetic level in maize, showing that they play a crucial role in the three epigenetic pathways investigated in this proposal: the autonomous pathway, the small RNA pathway (sRNA), and CpG methylation (^{m}CpG) pathway.

2. The detailed characterization of the epigenetic and transcriptional profile of a selected number (up to 25) of the maize epi-targets.

3. The phenotypic characterization and identifications of differentially expressed sequences in new and not yet characterized maize mutants, which compromise function of epi-regulators for the three epigenetic pathways mentioned above.

Summary of progress towards objectives:

1. Development of cold, and heat stress protocols in maize and sequencing of mRNA, sRNAs, ChIP and Bisulfite libraries produced to identify maize epitargets. Application of temperature shift, drought and salinity stress protocols to maize B73 and mutants.

3. Characterization of the *FVE* (nfc102 RNAi and of three nfc102 AS) mutant. Introgression of new 5 maize mutants into B73 and production of new transgenic maize mutants of three pathways.

Task 1: Identification of maize targets of selected stress and environmental epi-regulators Lead participant: P2: other participants: P1, P3, P7)

P3 applied cold and heat stress treatments to B73 maize plants. Young plants (10 – 14 days corresponding at V6 developmental stage) were grown in glasshouse conditions during the day, followed by night time incubation at 4 ˚C for 7 consecutive nights. After a recovery period of 4 days (T4) and 7 days (T7), fresh material samples consisting of the inner growing leaf and the meristem (hereinafter named “meristematic area”: MA), from stressed and control populations were isolated and frozen at -80 ˚C for RNA and DNA preparation.

The RNA extracted from MA was employed by **P3** for production and sequencing of small RNA (sRNA) and RNA-seq libraries (with polyadenylated RNA enrichment) using Illumina sequencing machine. UNIWA applied both cold and heat stresses in his lab. Both B73 and mutants from the three epigenetic pathways (rmr6;

mop1; fve and dmt1) were grown and stressed at UNIWA; for each time point, five individual plants were pooled for material collection. Plant materials were used for preparing small RNA and DGEx libraries, which were sequenced. **P1** received DNA from P3 and developed a protocol for preparing libraries to be employed in BS-seq; the first sample was paired end sequenced by **P7**. The protocol developed by **P7** to prepare BS-seq libraries in Arabidopsis has been adapted by **P1** to maize; a main point concerned the bisulphite conversion efficiency and the observation that maize chloroplastic DNA is methylated and cannot be used to evaluate conversion. Subsequently, BS-seq libraries were prepared and analysed by MPG before sequencing and data processing. **P2** optimized chromatin preparation using plants grown at UNIWA and ChIP assays from MA and prepared libraries for ChIP-seq, for cold stressed B73 materials. Libraries of chromatin immunoprecipitated with anti-H3K4me3 and with H3K27me3 were sequenced by **P7**, Depending upon the antibody used for ChIP, the mappable sequenced ranged between 25% and 65%.

P3 data analysis showed that heat seemed to have bigger effects than cold. Furthermore, at early time points after the treatment, compared to control treatment, the changes in transcriptome, small RNAs and DNA methylation (5mC) are larger than at later time points, indicating reversal of stress-induced effects. Upon stress treatment, large amounts of 'stress unlocked siRNAs (susiRNAs) derived from transposon and other repeat sequences were observed and most of these mapped to the gene-rich regions of the genome. Most of the changes in siRNAs are reset in time; some of them are still present at later time-points after the treatment. Intriguingly, part of the susiRNAs seem to direct de novo 5mC resulting in differentially methylated regions (DMRs). These changes are mostly reset again upon recovery of the temperature stress. It appears that about 10% of the DMRs retain the changed 5mC pattern. Together with the group of Detlef Weigel, they showed that in maize CpG and CHG methylation is mainly present at the centromeric region, while CHH methylation is enriched at the chromosome arms. Heat stress appears to primarily induce changes in 5mC at CHH, not at CG and CHG sites.

Task 2: Detailed molecular characterization of maize epi-targets (Lead participant P1, other participants: P2, P3)

Maize epitargets identification is the results of different stresses application to B73 inbred and selected maize epiregulator mutants followed both by the analyses of transcriptomes using RNA-seq andsmall RNAs-seq and epigenetic marks by ChIP-seq and Bi-seq. Consequently, the characterization of the epitargets has been postponed with respect to the original workplan and was performed during the final part of the project and in many cases is still ongoing. Indeed the number of epitargets that can be identified and characterized , analyzing the genome-wide data produced using the next generation sequencing technology (Illumina) is really high as consequences of the application of hightroughput technology. WP2 maintained the aim reported in Annex Ito produce results that could be then used in WP3 for comparative genomic between *Arabidopsis* and maize, but the comparison is mainly intended in genome wide terms and considering the emerging differences between the epiregulators function and the epigenetic pathways For this reason, the detailed description of the epi-targets derived both from the genome wide analysis of plants stressed with different treatments and mutants for epiregulators are reported describing the analysis of the data set obtained after stress applications (see the effect of cold and heat stress above and of the salinity stress below).

Task 3: Characterization of new maize mutants for environmental epi-regulators (Lead participantP8: other participants: P1, P2, P3)

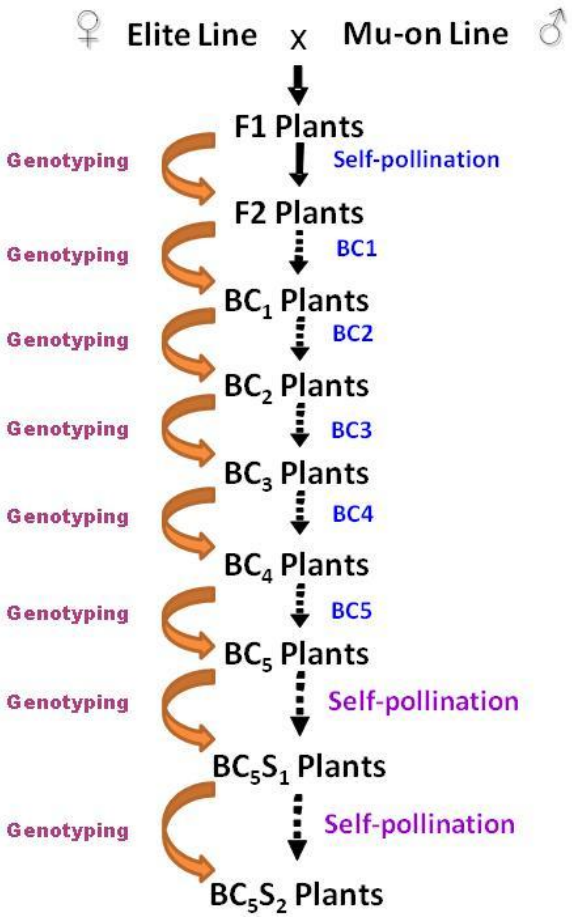
P2 characterized the maize (Zea mays) nucleosome remodeling factor complex component101 (nfc101) and nfc102, putative paralogs encoding WD-repeat proteins with homology to plant and mammalian components of various chromatin modifying complexes. They generated transgenic lines with simultaneous nfc101 and nfc102 downregulation and analyzed phenotypic alterations, along with effects on RNA levels, the binding of NFC101/NFC102, and Rpd3-type histone deacetylases (HDACs), and histone modifications at selected targets. Direct NFC101/NFC102 binding and negative correlation with mRNA levels were observed for indeterminate1 (id1) and the florigen Zea mays CENTRORADIALIS8 (ZCN8), key activators of the floral transition. In addition, the abolition of NFC101/NFC102 association with repetitive sequences of different transposable elements (TEs) resulted in tissue-specific upregulation of nonpolyadenylated RNAs produced by these regions. All direct nfc101/nfc102 targets showed histone modification patterns linked to active chromatin in nfc101/nfc102 downregulation lines. However, different mechanisms may be involved because NFC101/NFC102 proteins mediate HDAC recruitment at id1 and TE repeats but not at ZCN8. These results, along with the pleiotropic effects observed in nfc101/nfc102 downregulation lines, suggest that NFC101 and NFC102 are components of distinct chromatin modifying complexes, which operate in different pathways and influence diverse aspects of maize development.

Iride Mascheretti, Raffaella Battaglia, Davide Mainieri, Andrea Altana, Massimiliano Lauria, and Vincenzo Rossi (2013) The WD40-Repeat Proteins NFC101 and NFC102 Regulate Different Aspects of Maize Development through Chromatin Modification. The Plant Cell, Vol. 25: 404–42

In comparison to Arabidopsis only some maize mutants affecting the function of environmental epi-regulators belonging to the autonomous, sRNAs, and mCpG pathways are available for scientific community and have been characterized. In the time course of this project we characterized additional maize mutants for environmental epi-regulators belonging to the three pathways. Lines for five epi-regulators were identified by **P8** using MAG. The following mutants were sent to UNIPD for genotyping and introgression in B73 background.

Maize Mutant	Arabidopsis ortholog	Pathway	Description
hda108	AtHDA6		Hystone deacetylase.
chr120	AtMOM		Involved in the transcriptional silencing of repetitive sequences.
fpa	AtFPA	Autonomous	Putative RNA-binding protein regulating flowering time throughout the autonomous pathway and involved in transposons, retroelements and intergenic regions silencing.
dmt101	AtMET1	^{m}CpG	DNA methyltransferase involved in the propagation of ^{m}CpG patterns through DNA replication.
ago5	AtAGO5	sRNA	Slicer enzyme that uses miRNAs and siRNAs for silencing.

P1 applied the following procedure for introgression in B73 background:



P1 characterized the following mutations:

dmt101 BC2S1 homozygous mutant plants obtained as intermediate checkpoint showed no gene down-regulation so the mutant line was abandoned; *ago5* BC3S1 homozygous plants did not survive in greenhouse conditions, however **P1** are growing the BC5 plants in the field. *fpa* BC2S1 homozygous plants showed variation in gene expression and abnormal splicing: **P1** is investigating the mutation more deeply while introgression is going on. BC5S1 plants shows the complete silencing of CHR120 gene: no visible phenotype can be observed and the molecular characterization is underway. *hda108* BC5S1 homozygous plants show a strong phenotype, with reduction of plant height, leaf blade and seed germination rate: further characterization is in progress.

P8 performed RNAi transgenesis and obtained the following mutants that werephenotypically characterized.

						Phenotyping		
Pathway	Strain	Summary	Targeted genes	Nb of T1 events produced	Transferred to P03	T2 plants (field 2011)	T2 plants (Warwick)	HO T4 plants
A	T 01834	rubi3 FCA_RNAi Sac66	FCA	31	14	Flowering (7 out of 15 analysed)		
A	T 01838	rubi3 FLD_like_RNAi Sac66	HDMA102	27	18	Flowering (3 out of 15 analysed)		
sRNA	T 01837	rubi3 AGO4_like_RNAi Sac66	AGO104 / AGO105 / AGO109	32	19	Flowering (4 out of 14 analysed)	silent transgene reactivation →	?
sRNA	T 01839	rubi3 NRPD2_like_RNAi Sac66	NRPDB101 / NRPDB102	19	14	Flowering (1 out of 14 analysed)		
sRNA	T 01841	rubi3 NRPDB1b_like_RNAi Sac66	NRPDA102	15	12	Flowering (3 out of 15 analysed)		
Met	T 01840	rubi3 DDM1_like_RNAi Sac66	CHR101 / CHR106	24	16	Flowering (2 out of 15 analysed)	silent transgene reactivation →	Flowering (late)
Met	T 01846	rubi3 DRM2_like_RNAi Sac66	DMT103 / DMT107 / **CHR106**	46	21	Flowering (1 out of 15 analysed)	silent transgene reactivation	
Met	T 02405	rubi3 DRM2_like_RNAi HSP	DMT103 / DMT107 / DMT106	in progress				

P3 started using the mutants generated by Biogemma to test the effect of the downregulation of epigenetic regulators on the activation of a silenced GFP reporter locus. They did already see reactivation in a number of RNAi lines targeting DRM2, AGO4 and DDM1.

Finally, **P8** performed the transcritome analysis of selected lines and the results obtained are summarized in the following table.

	Leaf	SAM	Leaf	SAM
Macromolecules Chromatin	**AGO4_010**	AGO4_001 AGO4_010		
DNA metabolism	AGO4_001 **AGO4_010**			
Stress	**AGO4_010**	AGO4_001	**NRPDB1b_010** **NRPDB1b_012**	
Homeostasis	**AGO4_010**	AGO4_001		NRPDB1b_012
Polysaccharides		AGO4_001	NRPDB1b_012	NRPDB1b_010 NRPDB1b_012
Nitrogen			NRPDB1b_012	
Defence response				NRPDB1b_010

The phenotypic characterization and genome-wide identification of target sequences are in progress for selected lines.

Work package 3

Work package number	**3**	**Start date or starting event:**			Month 18	
Work package title	Epigenetic responses to environmental signals in maize and comparative genomics with *Arabidopsis*					
Activity Type[1]	RTD					
Participant number	1	2	3	7		
Participant short name	UNIPD	CRAMAC	UNIWA WP coordinator	MPG		
Person-months per participant:	22	21	28	2		

Objectives:
After the changes agreed by the AENEAS Consortium in the work plan of WP2, the main objective of WP3 is to evaluate the potential of maize epi-regulators for generation and inheritance of newly formed epi-alleles in response to environmental stresses. In particular, WP3 has three specific objectives:
1. To analyze epi-targets behavior under different stress conditions.
2. To analyze the trans-generational inheritance of environmentally induced epialleles
3. To explore possible maize peculiarities in the formation of epialleles in response to environmental stresses

Summary of progress towards objectives:
Different stress protocols were developed and applied to maize plants and mutants in epiregulators and materials collected to analyze the effect of the stresses on previously identified epitargets and to eventually identified new stress-specific epitargets.
Different stress protocols were developed to identify the best combination of stress and epiregulator mutation inducing Transposable Elements (TEs) mobilization.
Seed were collected from B73 plants and rmr6 mutant plants grown under salinity and drought conditions to analyze the transgenerational effects on environmental epi-targets induced by environmental stress in maize.

Task 1*: *The molecular analysis of maize epi-targets under different environmental stresses

Sub-task 1: *Analysis of selected epi-targets (from task 1 – WP2) under different stress conditions*.
P1, **P2**, and **P3** collaborate to identify the changes in transcription and epigenetic marks for a selected number of maize epi-targets in response to environmental stresses. These epi-targets were those under identification in WP2 and WP3, as the targets both of selected stresses (cold, heat, drought and salinity stress) and some maize epi-regulators belonging to one of three epi-regulatory pathways: autonomous, sRNA, and ^{m}CpG.

Sub-task 2: Development of protocols for different stress treatments in maize

The stress conditions to be applied for analysis of maize epitargets were selected among the most effectives for the induction of trans-generationally stable epigenetic changes in *Arabidopsis* mutants characterized in WP1. However, stress conditions showing prominent effects in *Arabidopsis* might be less effective in maize, due to the different genome evolution, structure and organization. Therefore, **P1** and **P3** have performed preliminary tests to assess the effect of stresses selected in WP1. On the basis of pilot experiments, **P1** and **P2** have selected different stress conditions (heat-shocks, cold, salinity and drought) which were applied to maize wild-type and mutants affecting the function of the maize homologues of *FVE* and *PolIVa*

More in detail, **P1** has developed and applied a reproducible protocol of drought and salinity stress in maize. The effect of the application of these stress conditions were monitored using functional markers, some of which indicated by **P8 that was agronomically interested on this type of stress**. The maize plants (B73 wt and *rmr6* mutant) were stressed (drought, salinity and a combination of the two stresses were applied) for 10 days and the plant material (last developed leaf) collected from a pool of four plants. Plant material was also collected after a period of recovery of four days (T4) and seven days (T7) from both B73 and rmr6. Finally, some stressed plants (S0) from both genotypes were self-pollinated and seed (S1) harvested from B73 and rmr6 mutants. Three replicates of the stress experiments were produced in two different years. To analyze the effect of the stress at genome wide level the following sequencings were done: two replicates of RNA-seq (total RNA was previously depleted from rRNA, one replicate was a simple pair-end sequencing while the second a directional one) three replicates of sRNA-seq and ChIP-seq for three histone modifications H3K27me3, H3K9 ac, H3k4me3). The total RNA-Seq strategy was used to analyze the transcriptome of wild type and mutant leaves after ten days of stress application and after 7 day of recovery period. Illumina sequenced reads were mapped to the maize reference genome (B73), to analyze the expression of annotated genes and to create a new annotation of the "stress-specific transcriptome" including TEs and long-non-coding RNAs (lncRNAs). Recent studies showed that many lncRNAs are potent cis- and trans-regulators of gene activity during development and in response to external stimuli and they can function as scaffolds for chromatin-modifying complex. Gene expression analysis revealed the modulation of many stress-related genes and the transcription of thousands of novel loci, many of these are encoding for non-annotated TEs. The characterization of these novel loci is underway and the transcriptomic data is being integrated at genome-wide level with the smallRNA dataset (sRNA-Seq) and with the distribution of chromatin marks exploring the whole epigenomic landscape of stress response and adaptation in maize. Analysis of sRNA sequencing data started from conserved miRNAs: differentially expressed miRNAs were identified employing a generalized linear model: salt and drought treatments caused slight up-regulation of five and one miRNAs, respectively, while the interaction between the two stresses did not provoke any additional effect. We are currently working on putative novel miRNAs identification and on the investigation of the relationship between microRNAs abundances and the expression levels of the corresponding targets, through the comparison with RNA sequencing data produced from the same samples.

Analysis of siRNA population confirmed previous results that siRNAs unique sequences have low expression levels and differ between samples. We are currently comparing siRNA loci abundances among samples, making distinction depending on the kind of genomic feature near to which they align: genes, transposons or other repetitive sequences, whose expression levels will be obtained from RNA sequencing data. The aim is to look for possible significant correlations between the expression levels of siRNA loci and the expression levels of their neighboring elements.

The results obtained by combining these different approaches aim to identify a robust list of sequences targets of epigenetic regulation (epitargets) that will be further studied and analyzed in the progeny of stressed plants.

Sub-task 3 Epigenetic regulation of stress induced TEs: i) to identify the best combination of stress and epi-regulator mutation inducing TEs mobilization and ii) what are the most sensitive TEs.

P2 in collaboration with **P4** investigated the epigenetic control of stress induced maize TEs, in B73, rmr6 (maize NRPD1 homolog) and nfc102 mutants. **P2** has selected specific low copies TEs from a maize TE catalogue for B73 line. Plant material harvested after the different stress treatments was used to check for the presence of extrachromosomal DNA (intermediate of retrotrasposition) and analyze up-regulation of

RNA transcripts. Materials and information will be shared with P4 and Marie Mirouze (AENEAS associate) to study the epi-genetic control of stress induced transposition in maize. **P2** developed temperature stress protocols for maize and studied the occurrence of extrachromosomal DNA (collaboration with **P4**). They could not detect such DNA in wild type or rmr6 mutant lines in response to stress. They can, however, detect such DNA in other genotypes, but also other developmental stages of the maize plant, suggesting other features than stress are more important for the production of extrachromosomal DNA. To identify activated retrotransposons **P2** and **P4** used a SIRT assay (sequence-independent retro-element trapping) that can identify mobilized LTR transposons. They performed studies on both Arabidopsis and maize using the SIRT assay. In Arabidopsis they e.g. observed ONSEN and EVD LTR element DNA upon heat stress. For maize (collaboration with **P2**) they are in the process of analyzing sequenced libraries produced from non-stressed and cold-stressed wild-type B73 and rmr6 mutant plants.

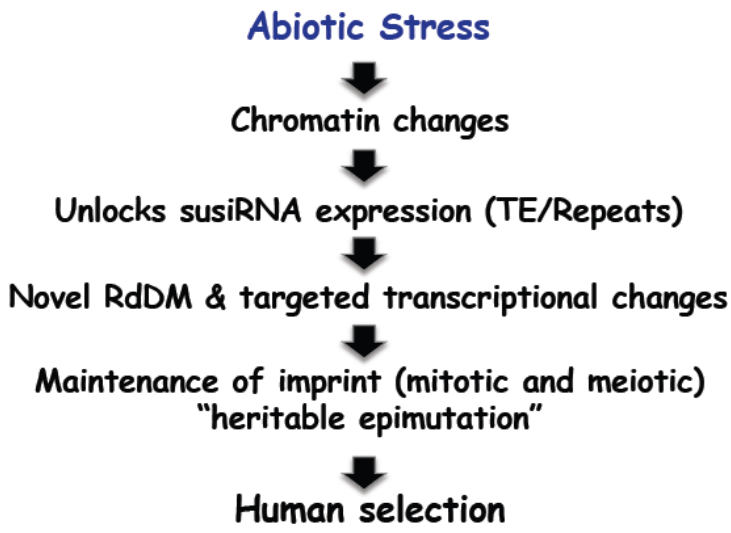
Task 2: Trans-generational effects on environmental epi-targets induced by environmental stress in maize

Different environmental stresses impact greatly on the generation of rapidly-acquired traits in maize via the creation of epialleles. However, it is not yet known whether this epigenetically-acquired information is transmitted to the progenies. To overcome this caveat in our knowledge, we aimed to analyze the impact of environment on the trans-generational inheritance of newly acquired epialleles.

P1 sowed the seed and grown the stressed plant progenies of B73 inbred line. In the next years, these progenies will produce mutation accumulation lines following a single seed descent (SSD) procedure for 5 generations. The sequencing of the lines and their transcriptome and epigenomic analyses will be used to evaluate how much the genetic mutation background is important in determining the epigenetic variation that is to separate the genetic component of variation which induces epigenetic variability from the pure epigenetic components.

To assess the stress-induced epigenetic heritability **P3** obtained the progeny of B73 WT, rmr6 mutant and mop1 mutant. In particular **P3** analyzed the Inheritance of stress-induced DMRs (Differentially methylated Regions) in the unstressed progenies of stressed plants and obtained a genome-wide hereditability maps of stress-induced DMRs.

Considering the results obtained so far both by different stresses application and response in wt B73 and in mutants of epiregulators and the first evidences on epigenetic heritability of stress response, a working model for epiallele formation by exposure to abiotic stress in maize has been produced.



This model hypothesize that maize mutants impaired in RdDM may be sensitive to stress: further elaboration of dataset from stress applications on mutants such as AGO4 and rmr6 both involved in RdDM pathway will be used to confirm this model.

Work package 4

Work package number	**4**	**Start date or starting event:**		Month 1		
Work package title	Genomics and bioinformatics platforms, website					
Activity Type[1]	RTD					
Participant number	1	2	3	4	5	6	7
Participant short name	UNIPD	CRAMAC	UNIWA	UNIGE	JIC	UCAM	MPG WP coordinator
Person-months per participant:	4	4	4	8	3	7	30

Objectives:

The objective of WP4 is to provide a genomics and bioinformatics platform for the genome-wide analysis of the epigenome carried out in the WP1, WP2, and WP3. The genomics platform is based principally on the next-generation sequencing, and where warranted, tiling microarray analyses. The data generated on these instruments will be analyzed using bioinformatics tool that need to be developed or adapted for AENEAS purposes.

In particular, the specific objectives of WP4 are:

1. The employment of a common pipeline for the bioinformatics analysis of the genome-wide data arising from activities carried out in the others WPs, to present data in a unique and common format.
2. The application of a bioinformatics-based approach for a comparative analysis of the epigenomic environmental changes identified in two evolutionary distinct species such as *Arabidopsis* and maize, and their mode of inheritance.

3. The constitution and the management of a project database for collecting and for the dissemination of the project's results.

Summary of progress towards objectives:

1. Development of a common pipeline for the analyses of Arabidopsis tiling arrays in the project.
2. Adoption of GenomeMapper in the context of SHORE for the analyses of bisulfite converted reads. GenomeMapper tools were adopted for mapping of reads from bisulfite converted DNA of *Arabidopsis thaliana* BS-seq material. Optimization for its application to maize BS-seq material is in progress Maize DNA from ChIP were analyzed by Illumina sequencing and mapped using SHOREpeak.
3. The constitution and the management of a project database for collecting and for dissemination of the project's results.

***Task 1: Adapting SHORE pipeline*.**

P7 has previously developed a pipeline for analysis of Arabidopsis tiling arrays, both for the detection of expression of annotated genes and of unannotated transcriptionally active regions (TARs). **P7** receive RNA extracted from seedlings treated with different stresses from UNIGE, JIC, and UCAM. RNA was converted to tiling array probes and hybridizations carried out. 72 samples were processed during the period. AGRONOMICS1 array platform was used for the experiments in AENEAS.

P7 has developed the SHORE pipeline for mapping of short reads and a graph method, GenomeMapper, for effective mapping of variant sequences (Schneeberger et al., 2009).

***Task 2*: *Develop statistical methods for small RNA and ChIP-seq*.**

A manuscript in which **P7** inferred indirectly patterns of DNA methylation from small RNAs in A. thaliana and its relative A. lyrata has been published: the 24-nt siRNAs were mapped to the A. thaliana and A. lyrata reference genomes by using the SHORE pipeline

Hollister, J., Smith, L. M., Guo, Y.-L., Ott, F., Weigel, D., and Gaut, B. S. (2011) On the role of transposable elements and small RNAs in driving gene expression divergence between Arabidopsis thaliana and Arabidopsis lyrata. Proc. Natl. Acad. Sci. USA 108, 2322-2327.

In the context of AENEAS, **P7** has developed a peak finder algorithm that is integrated into SHORE, SHORE peak. **P7** has provided proof of concept for the ability of SHORE peak to detect genome-wide binding site of individual transcription factors in ChIP experiments with Arabidopsis. **P7** have received from CRAMAC ChIP-seq libraries for two histone modifications (H3K4me3 and H3K27me3) from plants grown at UNIWA. Depending upon the antibody used for ChIP, the mappable sequenced ranged between 25% and 65%. using SHORE peak.

***Task 3*: *Bioinformatics analysis of bisulphite sequencing in maize*.**

P7 has adopted its GenomeMapper tools for mapping of reads from bisulfite converted DNA in Arabidopsis.

A major question of AENEAS is the transgenerational stability of epigenetic marks. To provide baseline information, MPG has compared genome-wide DNA methylation among 10 *Arabidopsis thaliana* lines, derived 30 generations ago from a common ancestor. Epimutations at individual positions were easily

detected, and close to 30,000 cytosines in each strain were differentially methylated. In contrast, larger regions of contiguous methylation were much more stable, and the frequency of changes was in the same low range as that of DNA mutations. Like individual positions, the same regions were often affected by differential methylation in independent lines, with evidence for recurrent cycles of forward and reverse mutations. Transposable elements and short interfering RNAs have been causally linked to DNA methylation8. In agreement, differentially methylated sites were farther from transposable elements and showed less association with short interfering RNA expression than invariant positions.
C Becker, J Hagmann, J Muller, D Koenig, O Stegle, K Borgwardt & D Weigel (2011) Spontaneous epigenetic variation in the Arabidopsis thaliana methylome. Nature published on line 20 Septmber 2011.

GenomeMapper tools for mapping of reads from bisulfite worked well for *A. thaliana* BS-seq material, and were optimized to maize BS-seq material in collaboration with **P3**.

P7 studied the dynamics of the DNA methylome in a natural environment. They focused on Arabidopsis populations from three different regions in the US. Arabidopsis has been introduced into the US relatively recently and these different populations therefore evolved relatively recently, providing an excellent population to study the rate by which epialleles are formed and their stability in a natural environment. The genomic mutation rate and epimutation rate appeared to occur at a similar speed in the wild as in laboratory conditions and also the genomic locations were similar. Furthermore, loci that showed to be hypervariable at the epigenomic level in the lab behaved similar in the wild, undergoing epimutations more often than other regions. A more recent study in the lab of **P7** showed that the variation in DNA methylation is more pronounced between different tissues than between control and cold-treated tissues, suggesting that the induction of epigenetic changes by environmental cues may not be as wide-spread as suggested by some published papers.

Understanding environmental effects

Professor Serena Varotto of the Univeristy of Padova offers details on her most recent project AENEAS, which explores the epigenetics of maize through *Arabidopsis thalia*

Can you outline the goals of the 'Acquired environmental epigenetics advances: from Arabidopsis to maize' (AENEAS) project?

The first general objective of the AENEAS proposal is to provide advances in understanding the detailed mechanisms of epiallele formation in response to environmental cues and their heritable maintenance in a model plant such as Arabidopsis. Concomitantly, the constitution of an 'Environmental Epigenetics platform' for maize will start with the development of tools indispensable for the shift of epigenetic research from Arabidopsis to maize. This is directly followed by the second main objective of the AENEAS project: the transfer of knowledge from model plant to maize.

What are the autonomous, the small RNA and the CpG methylation pathways and what do they regulate?

Cytosine methylation is one of the best characterised epigenetic mechanisms. Methylation occurs in CpG sequence context (mCpG) both in higher plants and animals, but plants differ from animals by significant levels of methylation at symmetric CpNpG (where N can be any nucleotide) and asymmetric CpHpH (where H is an A, C or T) sites. Two essential roles have been ascribed to DNA methylation: defending the genome against transposons and regulating gene expression. There is good evidence that siRNAs, which are generated by the sRNA pathway, can provide sequence specificity to guide cytosine methylation as well as other types of epigenetic modifications. In the Arabidopsis genome, there are thousands of loci for production of endogenous small RNAs that, in turn, can epigenetically modify the loci of origin. The epigenetic modifications involve DNA methylation, chromatin silencing or both. The autonomous pathway represents a good example of an interaction between environmental cues and the plant epigenome. This pathway was initially characterised because it regulates the switch from vegetative to reproductive development in Arabidopsis. However, more recent advances indicate that this pathway plays important, genome-wide roles and, in several cases, its function overlaps with other pathways mediating chromatin regulation.

How can a better understanding of these pathways aid knowledge of environmental factors that influence the epigenome?

Several bodies of evidence indicate that these three epigenetic regulatory pathways play a pivotal role in mediating epigenomic change in response to environmental cues. A large part of the information on the environmental-related epialleles formations is arising from studies carried out on the Arabidopsis model plant. In this model, it is well-documented that environmental cues, particularly stresses and shocks, strongly affect gene and genome activity. It is also evident that the environment, in addition to inducing genetic variability due to mutation of the DNA nucleotide sequence, also induces formation of stably inherited epialleles with relevant effects on the phenotype. Importantly, preliminary results indicate that their function is also conserved in maize.

What is the purpose of the Environmental Epigenetic platform for maize and what function will it serve?

It is our opinion that the deliverables from AENEAS Environmental Epigenetic platform for maize will be the 'progenitors' for the next generation of breeding programmes, based on the exploitation of the environmental induced epigenetics variability. For instance, it is thought that the ability to respond to abiotic stresses represents the main constraint on maize production. Thus the future release of maize varieties with high yield and stability will largely depend on the possibility to increase tolerance to environmental stresses. In plant improvement, breeders select for a consistent phenotype that is often conditioned by many genes with incremental effects, rather than single large effect mutations. Therefore, two important questions to address are: how much of this available variation exploited in breeding programs is caused by epigenetic processes, and how it is possible to increase epigenetic diversity from which advances in crop research and development can be drawn.

What progress have you made in the functional characterisation of maize mutants for epiregulators belonging to the three pathways studies in Arabidopsis?

Mutant maize lines with down regulation of the Arabidopsis FVE gene of the autonomous pathway, named nfc102, were characterised at the phenotypic and molecular level. nfc102 targets were identified using available information from Arabidopsis FVE targets and currently the attention is focused on their characterisation to shed light on nfc102 role at genome wide level. Simultaneously, another two mutants of the maize sRNA pathway are under characterisation for their response to various environmental stresses and shocks, using information derived from genome-wide characterisation of both maize B73 line and Arabidopsis. Finally, new mutants for several maize epiregulators have been introgressed in B73 maize lines, for which the genome sequence is available.

Studying stress in maize

Serena Varotto outlines efforts to translate findings on *Arabidopsis* response to environmental stresses to maize, with results that look promising for the future of genetically-enhanced crop production

What progress have you made since we last caught up with you in January 2012? What activities are you undertaking in AENEAS, which explores the epigenetics of maize through *Arabidopsis*?

In recent months, we have collected and analysed our dataset on maize's response to various environmental stresses previously tested in *Arabidopsis*. The data was obtained after the application of several stress protocols. Plant materials from both the B73 maize inbred line and maize mutants of epi-regulators of the three epigenetic pathways under investigation in AENEAS were then used for preparing small RNAs, RNAs ChIP-seq and BS-seq libraries. These were sequenced and the datasets were analysed in order to investigate the effect of stress application on transcription, gene regulation and some epigenetic marks, at genome-wide level.

Have there been many changes to the focus of your research?

The leading idea of AENEAS is to assess the effect of environmental cues on plant genome stability, focusing on the epigenetic mechanisms governing this stability. The first results were produced by the labs working with the model plant *Arabidopsis thaliana*. Information provided by the model was used to develop and apply stress protocols suitable for maize. We are now deciphering the effect of stress application on the stability of the maize genome, gathering data from both the B73 reference lines and some mutants of epi-regulators belonging to the three main epigenetic pathways – autonomous, small RNA and CpG methylation pathways.

What results have you collated on the three epigenetic regulatory pathways studied in your project?

The research on *Arabidopsis* demonstrated that the autonomous pathway has a general role in gene transcription termination; CpG methylation marks are partially maintained across generations; and the sRNA pathway is involved in controlling both genome stability and abiotic stress response. In parallel, we have developed a series of tools for maize, protocols and mutants of epi-regulators, which can determine whether the three pathways are acting in the same way as in *Arabidopsis*. Our results on maize indicate that some epi-regulators belonging to the three pathways might have different targets in the maize genome and newly-produced mutants are currently under characterisation.

Could you highlight the key features of the 'Environmental Epigenetic platform' and outline the motives behind its composition?

One of the achievements of AENEAS has been the creation of an Environmental Epigenetics Platform for maize. This platform provides information regarding the combination of environmental conditions and epi-regulator mutants able to induce the formation of epialleles in both *Arabidopsis* and maize. The platform is also being used to assess the