

UNIVERSITÀ DEGLI STUDI DI PERUGIA



DIPARTIMENTO DI CHIMICA, BIOLOGIA e BIOTECNOLOGIE

Corso di Laurea in SCIENZE BIOLOGICHE

**PROVA FINALE**

***Effetti dell'innesto sull'epigenoma del pomodoro  
(*Solanum lycopersicum*)***

LAUREANDO

Thomas Sirchi

RELATORE

Prof. Emidio Albertini

Correlatore

Dr. Marco Di Marsico

## SOMMARIO

<b>1. ABSTRACT.....</b>	<b>5</b>
<b>2. INTRODUZIONE .....</b>	<b>8</b>
2.1 EPIGENETICA .....	9
2.1.1 Meccanismi epigenetici.....	10
2.1.2 Epialleli.....	18
2.1.3 Stress.....	21
2.1.4 Analisi di metilazione .....	24
2.2 INNESTO .....	26
2.2.1 Storia.....	27
2.2.2 Tecniche dell'innesto.....	28
2.2.3 Benefici .....	29
2.2.4 Conseguenze epigenetiche.....	32
2.3 POMODORO.....	36
2.3.1 Genoma del pomodoro .....	36
2.3.2 L'innesto in pomodoro .....	38
2.3.3 Studi sulla metilazione del DNA in pomodoro innestato.....	39
<b>3. SCOPO DELLA TESI .....</b>	<b>40</b>
<b>4. MATERIALI E METODI .....</b>	<b>42</b>
4.1 MATERIALE VEGETALE .....	43
4.2 PROTOCOLLO DI LABORATORIO .....	43
4.2.1 ESTRAZIONE DEL DNA E PREPARAZIONE DEGLI ADATTATORI.....	43
4.2.2 Preparazione delle librerie .....	44
4.3 ANALISI BIOINFORMATICA: TECNICA MCSEED .....	52
4.3.1 Preparazione del file counts.....	52
4.3.2 Da counts al formato Methykit .....	53
4.3.3 Identificazione DMPs con Methykit .....	54
4.3.4 Identificazione DMRs .....	55
4.3.5 Identificazione delle DMGs .....	56
4.4 ANALISI GENE ONTOLOGY CON PANTHER - GENE LIST ANALYSIS .....	57
4.4.1 Panther (Protein analysis through evolutionary relationships) .....	57
<b>5. RISULTATI .....</b>	<b>58</b>
5.1 IDENTIFICAZIONE DELLE DMPs .....	59
5.1.1 PCA (principal component analysis).....	62
5.2 IDENTIFICAZIONE DELLE DMRS .....	66

5.2.1	<i>DMRs</i> .....	66
5.3	DMG .....	74
5.4	GENE ONTOLOGY.....	75
5.5	ANALISI DEI GENI.....	77
5.5.1	<i>I12</i> .....	79
5.5.2	<i>I14</i> .....	79
5.5.3	<i>I15</i> .....	80
5.5.4	<i>O3</i> .....	80
5.5.5	<i>OPT</i> .....	81
5.5.6	<i>Self</i> .....	81
6.	CONCLUSIONI .....	83
7	RINGRAZIAMENTI.....	86
	BIBLIOGRAFIA .....	88

## 1. ABSTRACT

L'innesto è una tecnica di coltura con radici profonde nella storia. Ad oggi è largamente utilizzata nella agricoltura moderna, portando un vasto numero di vantaggi documentati per gli agricoltori e la pianta stessa. Esempi di questi vantaggi sono l'aumento del raccolto, maggior assorbimento di nutrienti, tolleranza allo stress biotico ed abiotico. Su quest'ultimi, per molto tempo si sono concentrati gli studi e le ricerche riguardanti l'innesto. I risultati di queste ricerche, insieme ad una crescente attenzione economica, ha portato l'innesto a costituire una possibile alternativa all'uso di agrofarmaci. Nel pomodoro (*Solanum lycopersicum*), parte della famiglia delle *Solanaceae*, la tecnica dell'innesto è ampiamente utilizzata per fini commerciali e di salute della pianta. Le osservazioni compiute sull'effetto dell'innesto in pomodoro riguardano principalmente: la resistenza a patogeni, l'aumento del raccolto, le variazioni nelle proprietà organolettiche, qualità e presentazione del frutto per fini commerciali. Diversi studi volti a mostrare gli effetti dell'innesto sulla pianta hanno portato una minor attenzione all'aspetto epigenetico privilegiando invece l'osservazione delle proprietà del frutto. La tecnica dell'innesto, in quanto atto di fusione tra parti di una pianta, rappresenta una grande fonte di stress alla quale la nuova pianta deve adattarsi per sopravvivere. L'adattamento necessario per la sopravvivenza comporta una serie di modificazioni reversibili, spesso ereditabili, nell'espressione genica e talvolta nel fenotipo della pianta, di fatto mostrando un collegamento tra l'innesto e le modificazioni dell'epigenoma. L'epigenomica si propone di studiare i meccanismi dietro l'organizzazione dell'epigenoma come le modificazioni istoniche e i cambiamenti di metilazione delle citosine. In questo lavoro di tesi l'attenzione è stata focalizzata proprio sull'ultimo aspetto. Le differenze nella metilazione sono state individuate in dei confronti tra innesti effettuati con cinque tipi di portainnesto diverso (I12, I14, I15, O3 e OPT) ed un ultimo innesto prodotto con nesto e portainnesto della stessa varietà ed i loro relativi controlli. Nei confronti tra trattati e controlli è stata valutata la metilazione relativa a carico dei context CG, CHG e CHH. Queste sono sequenze diffuse in tutto il genoma il cui tasso di metilazione va a regolare l'espressione genica, silenziando o attivando, specifici geni e pathway metabolici. L'analisi è stata effettuata tramite la metodica MCSed, la quale sfrutta enzimi di restrizione sensibili alla metilazione, successivamente attraverso l'utilizzo di diversi strumenti bioinformatici per elaborare le informazioni. L'obiettivo è stato quello di valutare le differenze nella metilazione relativa tra i confronti e identificare i geni differenzialmente metilati. L'identificazione di quest'ultimi è stata effettuata grazie all'analisi delle regioni differenzialmente metilate (DMRs) ed il loro confronto con il genoma di riferimento di pomodoro. Al fine di comprendere l'ampiezza delle modificazioni dell'epigenoma di pomodoro e l'influenza delle stesse sul frutto è stato essenziale individuare e studiare i geni differenzialmente metilati.

Grafting is an agricultural technique with ancient origin. Nowadays, it is widely used in modern agriculture, bringing many documented advantages for farmers and for the plant itself. Examples of these advantages are increased harvesting, increased nutrient absorption, tolerance to biotic and abiotic stresses. For a long time, the studies concerning grafting have been focused on the stresses. The results of these studies, together with an increasing economic attention, led grafting to be a possible alternative to the use of pesticides. In tomato (*Solanum lycopersicum*), part of the family of the Solanaceae, the grafting technique is widely used for commercial purposes and plant health. The research made on the effect of the grafting in tomato mainly concern the resistance to pathogens, the increase of the harvesting, the variations in the nutritional properties, quality and presentation of the fruits for commercial purposes. Grafting, as an act of fusion between parts of a plant, represents a great source of stress to which the new produced plant must adapt to survive. The necessary adaptation involves a series of reversible modifications, often heritable, in the gene expression and sometimes in the phenotype of the plant, showing a link between the grafting and the modifications of the epigenome. Epigenomics aims to study the mechanisms behind the organization of the epigenome such as histone modifications and methylation changes of cytosines. In this work the attention was focused on this last epigenetic change. The differences in methylation have been identified through comparisons between grafted plants produced with five different types of rootstocks (I12, I14, I15, O3 and OPT) and one last graft produced with scion and rootstock of the same variety and their controls. The relative methylation of the CG, CHG and CHH contexts was evaluated in the comparisons between treatments and controls. These are sequences spread through the whole genome and their methylation rate regulates gene expression by silencing or activating specific genes and metabolic pathways. The analysis was done through the MCSed methodology, which utilizes endonucleases that are sensible to DNA methylation, then using various bioinformatic tools to process the produced information. The goal of the analysis was to assess the differences in the DNA methylation between the comparisons and to identify the differentially methylated genes. Their identification was performed thanks to the analysis of the differentially methylated regions (DMRs) and their comparison with the tomato reference genome. In order to understand the extent of the alteration in the tomato epigenome and their influence on the fruit, it was essential to identify and study the differentially methylated genes.

## 2. INTRODUZIONE

## 2.1 Epigenetica

Il genoma degli organismi viventi è un oggetto di studio molto importante nel panorama scientifico attuale, sul quale si continua a effettuare ricerca per comprendere a pieno tutti i meccanismi legati ad esso e alla sua funzione nell'organismo. La genomica, branca della biologia molecolare, basandosi sulla bioinformatica e i suoi strumenti, gestisce l'enorme quantità di dati prodotti riuscendo ad occuparsi dello studio della struttura, del contenuto e della funzione evolutiva del genoma. Con il termine struttura si intende una vasta gamma di componenti sui quali è incentrato il nostro interesse di studio, poiché proprio da questi è possibile valutare meccanismi, interazioni ed effetti sul genoma e sull'organismo di cui fanno parte. Grazie alle tecniche di sequenziamento del DNA ed alla più recente bioinformatica, branca della scienza che ha visto un notevole sviluppo, è possibile ricostruire, nucleotide per nucleotide, intere sequenze di DNA. Dando la possibilità di capire come interazioni e modificazioni della sequenza nucleotidica, possano influire sul fenotipo degli esseri viventi e come l'ambiente esterno abbia effetti su di esso. L'ambiente che circonda un organismo, con i suoi cambiamenti, influenza lo stesso portandolo ad attivare dei meccanismi complessi per potersi adattare alle nuove condizioni. Questi meccanismi di risposta allo stress coinvolgono il genoma, non solo in quanto fonte di istruzioni per tali cambiamenti, ma come oggetto stesso di tali modificazioni. Tutto questo porta ad una condizione di adattamento all'ambiente circostante costante e dinamico. Lo stress nelle piante viene definito come condizione sfavorevole o sostanza che possa provocare un effetto negativo o un drastico blocco nel metabolismo, nella crescita e nello sviluppo della pianta stessa (Lichtenthaler & Hartmut, 1998). Questo fenomeno induce le piante ad attuare dei meccanismi di adattamento all'ambiente, strategie che permettono la sopravvivenza dell'individuo e della specie. Tali strategie, volte alla minimizzazione degli effetti negativi, possono coinvolgere cambiamenti fisiologici dei pathway metabolici e la modifica dell'espressione genica. Epigenetica è il termine utilizzato per definire lo studio di questi cambiamenti o modificazioni nell'espressione genica che non apportano alterazioni al DNA stesso. Pertanto, l'epigenetica studia i cambiamenti ereditabili nell'espressione genica che si verificano senza alterazioni nella sequenza del DNA (Bottaccioli *et al.*, 2012). Questi cambiamenti sono definiti ereditabili poiché capaci di essere trasmessi alle generazioni future, ma soprattutto reversibili, in quanto non vanno a modificare la sequenza del DNA stesso. L'epigenetica agisce come un codice "al di sopra" (dal greco *ἐπί*) della sequenza nucleotidica (Romani, 2019), nello specifico viene definita come un adattamento strutturale delle regioni cromosomiche così da registrare, segnalare o promuovere uno stato di alterazione (Bird, 2007). Lo studio di tali alterazioni ci permette di comprendere i meccanismi di questi eventi regolativi, che possono essere naturali o artificiali.



### 2.1.1 Meccanismi epigenetici

Nel corso dell'evoluzione, le cellule animale e vegetale hanno sviluppato un sofisticato e complesso sistema multicomponente e interconnesso per la decodifica dell'informazione genetica (Romani, 2019). Tra i componenti dei meccanismi epigenetici troviamo: la modifica degli istoni con la sua conseguenza sullo stato della cromatina, l'attività degli RNA non codificanti, la metilazione e demetilazione del DNA. Tutti questi meccanismi richiedono un complesso e ridondante sistema enzimatico di modificatori epigenetici.

#### *Modificazioni istoniche*

Gli istoni sono delle proteine basiche che si aggregano formando ottameri in grado di legare il DNA determinando la produzione del nucleosoma. Il DNA forma una spirale intorno all'ottamero istonico per 146pb. L'ottamero è composto dagli istoni H2A, H2B, H3 e H4. Oltre a questi è presente l'istone H1, detto istone linker, il quale forma un ponte tra tratti del filamento di DNA per mantenerlo in posizione. Tra due nucleosomi vi è una distanza di 80pb di DNA alle quali sono legate proteine non istoniche. L'istone H1 e le proteine non istoniche facilitano il ripiegamento dei nucleosomi andando a formare la cromatina, il complesso di DNA, RNA e proteine dinamicamente associate tra loro a compattezza variabile. Il termine eterocromatina si riferisce ad un livello di impacchettamento del DNA accessibile da proteine e fattori trascrizionali. Nell'eucromatina l'impacchettamento diviene serrato rendendo il DNA non accessibile. La carica di queste proteine istoniche è molto importante, infatti, essendo cariche positivamente, permettono l'interazione con la molecola di DNA carica negativamente e l'avvolgimento. Le code degli istoni possiedono amminoacidi positivi e grazie alla loro carica possono essere sottoposti a modificazioni post-traduzionali come la metilazione, acetilazione, ubiquitinazione, fosforilazione e ribosilazione (FIG. 1). Questo tipo di modifiche può provocare cambiamenti nella cromatina e nella sua compattezza, andando così ad alterare l'espressione genica e regolare il processo di trascrizione.

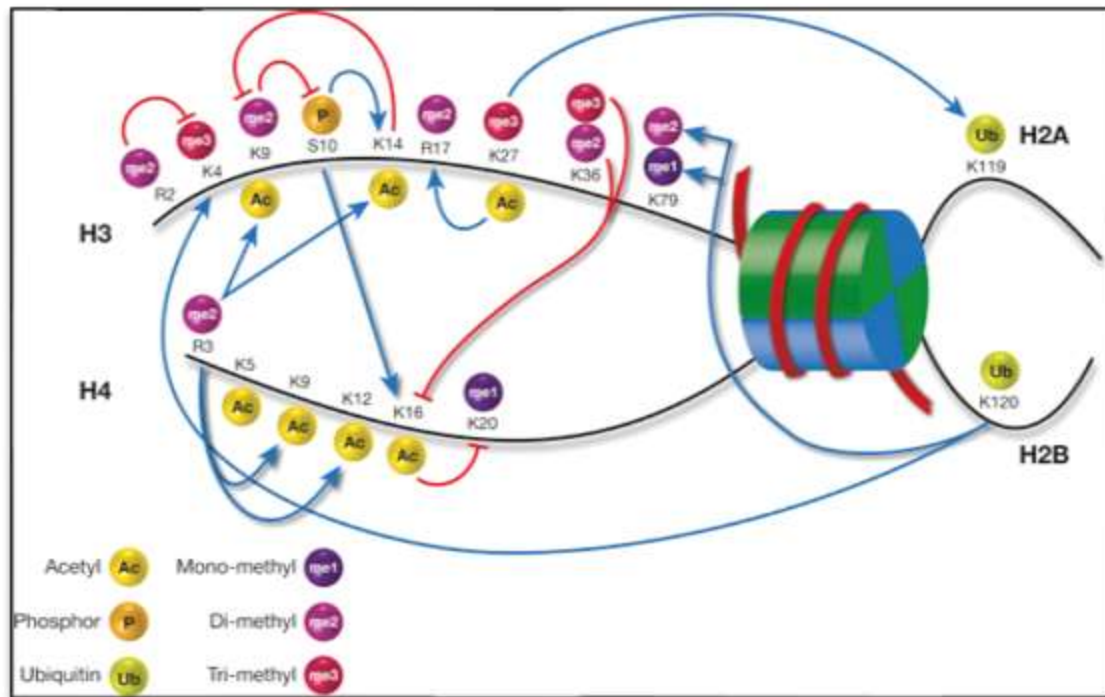


Figura 1: possibili modificazioni post-traduzioni su code istoniche (Simonet et al., 2016)

### Small RNA (sRNA)

Gli small RNA sono una parte fondamentale della regolazione genica nelle piante ed animali: questi filamenti di RNA non codificanti danno luogo al fenomeno del RNA silencing (Dunoyer *et al.*, 2013) oltre ad essere coinvolti nella formazione della eterocromatina e la difesa contro i virus. Gli small RNA hanno la capacità di muoversi e diffondersi dalla cellula che li ha prodotti ad una vicina. Questo fenomeno è stato studiato anche attraverso la tecnica dell'innesto che vedremo successivamente. Esperimenti sull'innesto sono stati utilizzati per scoprire: proteine mobili, mRNAs, smallRNA e piccole molecole che possono avere un impatto sulla morfologia della pianta innestata (Thomas *et al.*, 2019). Differenti classi di sRNA sono state scoperte nelle piante fino ad ora, e sono associate con il silenziamento post trascrizionale dei geni (Post Transcriptional Gene Silencing, PTGS) o con il silenziamento di geni trascrizionalmente attivi (Transcriptional Gene Silencing, TGS) (Berger *et al.*, 2018). I primi coinvolti nel PTGS sono i micro-RNA (miRNA), questi sono lunghi 21 nucleotidi e responsabili, a seguito del coinvolgimento di altre proteine specifiche, della degradazione di mRNA omologhi e l'inibizione della trascrizione. I secondi, coinvolti nella TGS, sono gli small interfering RNA (siRNA), i quali esistono in forme da 21 nucleotidi o 24 nucleotidi e si distinguono in base alle loro funzioni. I siRNA con una lunghezza di 21 nucleotidi sono coinvolti sia nella degradazione di mRNA specifici tramite PTGS che nel silenziamento TGS in loci specifici per trasposoni. Gli altri, lunghi 24

nucleotidi, sono coinvolti nella metilazione del DNA mediata da RNA (RNA-directed DNA methylation, RdDM) (Berger), che sarà spiegata in seguito approfondendo il meccanismo della metilazione del DNA.

### La metilazione del DNA

La metilazione del DNA è una tipologia di modificazione epigenetica, molto importante per la regolazione di geni e la stabilità del genoma stesso (Zhang *et al.*, 2018). Nelle piante la metilazione del DNA è categorizzata in classi distinte, basate sulla sequenza del context nel quale la citosina metilata (mC) è presente (Bewick *et al.*, 2017). Questi context sono: CG, CHG e CHH, dove H sta ad indicare un qualsiasi nucleotide tra A, T e C. La metilazione nel context CG è la più abbondante forma di mC nel genoma delle piante (Bewick, 2007). La metilazione del DNA si distingue in metilazione di mantenimento e *de novo* (FIG. 2). Il mantenimento consiste nel metilare residui di citosina non metilati nel DNA di nuova sintesi, dove la metilazione persiste in uno solo dei due filamenti, così conservare la modificazione (Unoki, 2019). La metilazione *de novo* è l'attività mediante la quale i gruppi metilici sono aggiunti alla citosina non metilata, il pathway più studiato di questa tipologia è la RdDM.

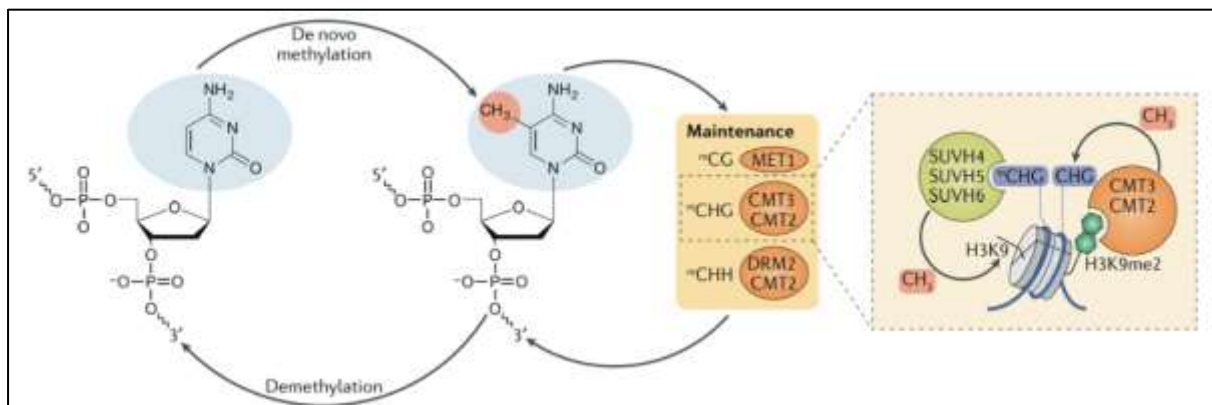


Figura 2: Regolazione dinamica della metilazione del DNA nelle piante. (Zhang, 2018)

### Mantenimento della metilazione

Prendendo come esempio *Arabidopsis thaliana*, la metilazione CG (mCG) è catalizzata e mantenuta dalla proteina METHYLTRANSFERASE 1 (MET1). Questa, a seguito della replicazione del DNA, va a metilare le citosine non modificate nel nuovo filamento (Zhang *et al.*, 2018). La MET1 è reclutata dalla famiglia delle proteine VIM (variation in methylation) necessarie per il mantenimento della metilazione CG. La metilazione di CHG (mCHG) è mantenuta da CHROMOMETHYLASE 3 (CMT3) ed è catalizzata da KRYPTONITE (KYP)/SU(VAR)3-9 HOMOLOG 4 (SUVH4), SUVH5 ed anche SUVH6 LISINA METILTRANSFERASI (Bewick; Finnegan *et al.*, 1996; Lindroth *et al.*, 2001; Cao & Jacobsen, 2002).

Infine, la metilazione CHH (mCHH) è mantenuta da DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) oppure CMT2, dipendentemente dalla regione del genoma dove si trova la modifica. La DRM2 mantiene le metilazioni CHH nelle regioni RdDM che sono collocate principalmente nei trasposoni corti, in sequenze ripetute a livello dell'eucromatina, oppure ai vertici dei lunghi trasposoni nell'eterocromatina. Invece, CMT2 catalizza la metilazione di CHH a livello dell'istone H1 nella cromatina, dove la RdDM è inibita. Anche se la metilazione di CHH può essere mantenuta solo da DRM2 e CMT2, questi enzimi possono anche metilare citosine negli altri due context (Zhang, 2018).

#### *La metilazione de novo del DNA*

Nelle piante, la metilazione *de novo* è mediata dalla RNA-directed DNA methylation (RdDM) (Zhang). La RdDM comprende l'impiego di siRNAs e scaffold RNAs (scRNA) in aggiunta ad una vasta gamma di proteine responsabile del processo. Questo pathway necessita di tre RNA polimerasi specifiche per le piante: la polimerasi IV (POL IV), la quale ha il compito di iniziare la biogenesi di siRNA e la polimerasi V (POL V) che sintetizza trascritti scRNA per reclutare altri fattori della RdDM (Gao *et al.*, 2010) e la polimerasi II che promuove una metilazione al di fuori del pathway canonico della RdDM. Inoltre, la POL IV e la POL V possono essere reclutate per la RdDM grazie a modificazioni preesistenti della cromatina, che ne faciliterebbe l'arrivo ai loci interessati.

#### *SAWADEE HOMEODOMAIN HOMOLOG 1 e la Polimerasi IV*

La POL IV è reclutata dalla proteina SAWADEE HOMEODOMAIN HOMOLOG 1 (SHH1), la quale lega l'istone H3 demetilato. Il dominio SAWADEE di questa proteina è un modulo con la capacità di legare la cromatina, il quale funziona come un lettore a due lisine, sondando sia per le modificazioni della lisina in posizione 4 (K4) non metilata e K9 metilato nella coda dell'istone H3 (Law *et al.*, 2013; Law *et al.*, 2011; Liu *et al.*, 2011). Questo è fondamentale non solo per il processo di reclutamento della POL IV, ma per la capacità di questo dominio di discriminare lo stato delle modificazioni sulla coda dell'istone e mostrare la specifica preferenza per le modificazioni repressive nella posizione H3K9 metilata. La discriminazione svolta dalla SHH1, in particolare il suo dominio SAWADEE è la chiave alla sua importanza nella regolazione dell'espressione genica.

#### *Polimerasi V*

La POL V richiede un complesso proteico per associarsi alla cromatina e sintetizzare scRNA: questo è noto come complesso DDR. Il DDR comprende anche la proteina per il modellamento della cromatina DEFECTIVE IN RNADIRECTED DNA METHYLATION 1 (DRD1), fondamentale per il mantenimento

strutturale dei cromosomi DEFECTIVE IN MERISTEM SILENCING 3 e infine la proteina RDM1. Il complesso DDR interagisce con due proteine appartenenti alla famiglia dell'istone metiltransferasi, la SUPPRESSOR OF VARIATION 3-9 HOMOLOG PROTEIN 2 (SUVH2) e la SUVH9. Queste possono riconoscere citosine non metilate nel DNA e reclutare la POL V.

### *Polimerasi II*

La POLIMERASI II (POL II) è una polimerasi in grado di sintetizzare siRNA che non fanno parte del pathway canonico per la RdDM (Zhang *et al.*, 2009). I trascritti mediati dalla POL II possono sia portare alla produzione di siRNA e scRNA a 24 nucleotidi ma possono anche reclutare la POL IV e POL V promuovendo la sintesi di siRNA a specifici loci (Zheng *et al.*, 2009). La POL II si associa alla *RNA-dependent RNA polymerase 6* (RDR6) di fatto costituendo una via alternativa alla più comune RdDM della POL IV e RDR2. Questa RdDM alternativa, mediata da RDR6, può passare attraverso siRNA di 21 nucleotidi o di 24 nucleotidi prodotti dalle proteine Dicer-like 2 (DLC2) e Dicer-like 4 (DLC4) oppure attraverso siRNA di solo 24 nucleotidi prodotti dalla proteina Dicer-like 3 (DLC3) (FIG. 3) (McCue *et al.*, 2015; Marí-Ordóñez *et al.*, 2013).

### *La RNA-directed DNA methylation (RdDM)*

Tenendo presente tutte le proteine appena illustrate, si prende come esempio la RdDM in *A. thaliana*. Questo processo inizia con la produzione di un siRNA di 24 nucleotidi tramite la trascrizione effettuata dalla POLIMERASI IV (POL IV) la quale è subito seguita dalla RNA dipendente RNA polimerasi 2 (RDRP2) e dalla proteina Dicer-like proteine 3 (DCL3). La prima produce una copia del trascritto sotto forma di un RNA a doppio filamento (dsRNA) e la seconda produce, dal dsRNA, filamenti di siRNA. Con la produzione dei siRNA viene attivato un gruppo di proteine chiamate Argonaute (AGO). A questo punto, i siRNA vengono caricati principalmente nelle AGO4 ed AGO6 e successivamente appaiati con scRNA di nuova produzione dalla POL V (Zhang). Questo scRNA prodotto è mantenuto sulla cromatina anche grazie alla presenza della proteina RRP6L1, analoga alla ribosomal RNA-processing (RRP6) del lievito e mammiferi. L'associazione della sequenza specifica tra siRNA, scRNA e AGO 4 è resa possibile dal segmento AGO hook, cioè un motivo proteico che media l'interazione con le proteine AGO 4 e 6, contenuto nella subunità più grande della POL V (He *et al.*, 2009; Bies-Etheve *et al.*, 2009).

Il processo di appaiamento tra siRNA e scRNA è stabilizzato dal complesso proteico IDN2-IDP che lega RNA ed interagisce con il complesso rimodellante della cromatina SWI/SNF (FIG. 4), in particolare la subunità SWI3B. Questa subunità è molto importante per la POL V, in quanto partecipa al silenziamento mediato dalla trascrizione della polimerasi, effettuando un'alterazione nella posizione dei nucleosomi nella cromatina. A questo punto avviene l'interazione tra AGO4/AGO6 e la DNA metil-

transferasi DOMAINS REARRANGED METHYLASE 2 (DRM2) la quale catalizza la metilazione de novo sul DNA (Zhong *et al.*, 2014).

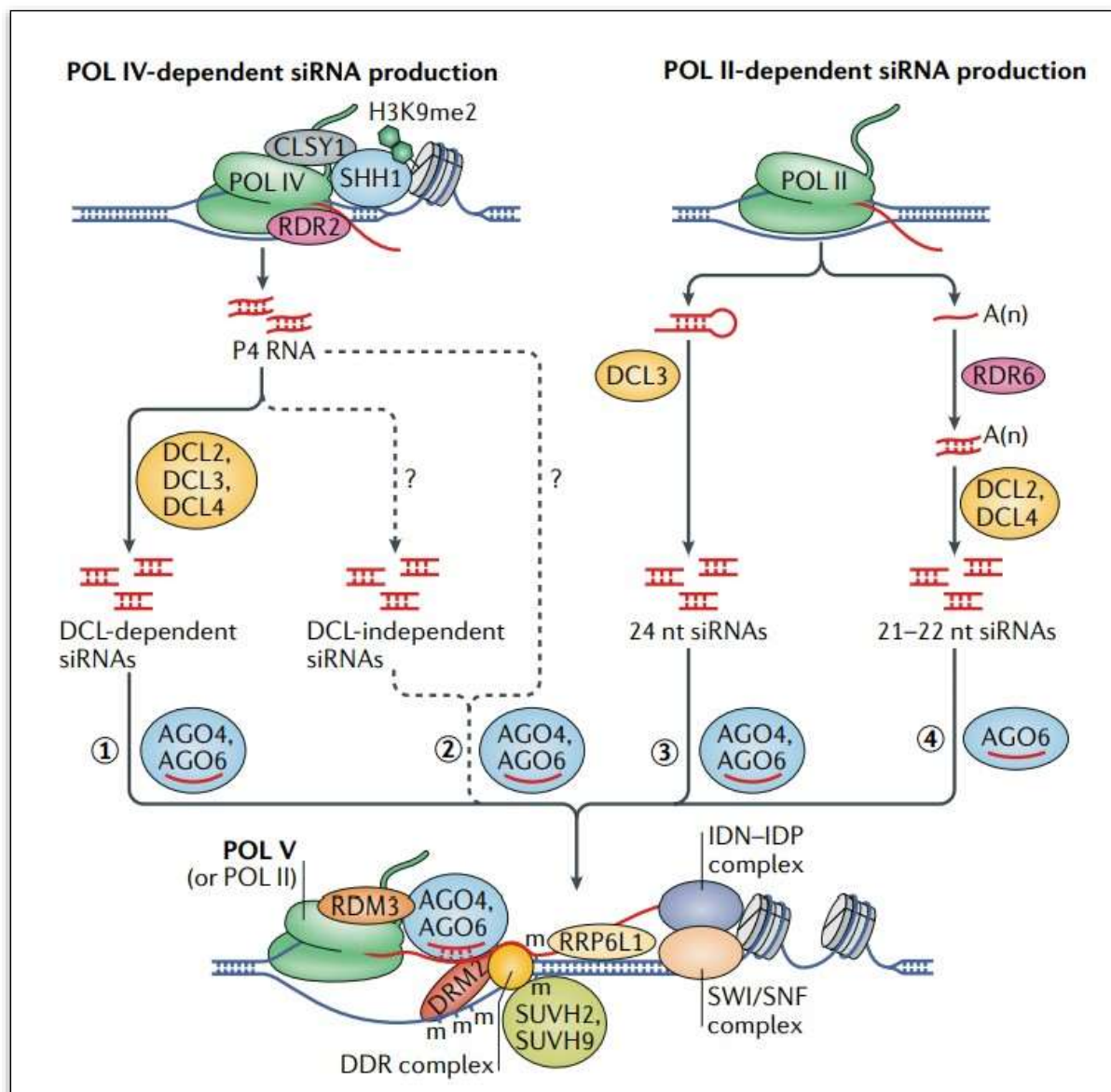


Figura 3: RdDM in *A. thaliana* (Zhang)

### Demetilazione del DNA

La demetilazione del DNA è un processo chiave per la regolazione dell'espressione genica. Esistono due categorie principali di demetilazione: una passiva caratterizzata dalla carenza di metilazione o impossibilità nel mantenimento della stessa (Rocha *et al.*, 2005; Zhou *et al.*, 2013) dovuto ad una carenza di gruppi metilici al momento della replicazione, l'altra, chiamata attiva, è mediata dalla proteina REPRESSOR OF SILENCING 1 (ROS1) (FIG. 4) (Zhu, 2009; Liu *et al.*, 2020).

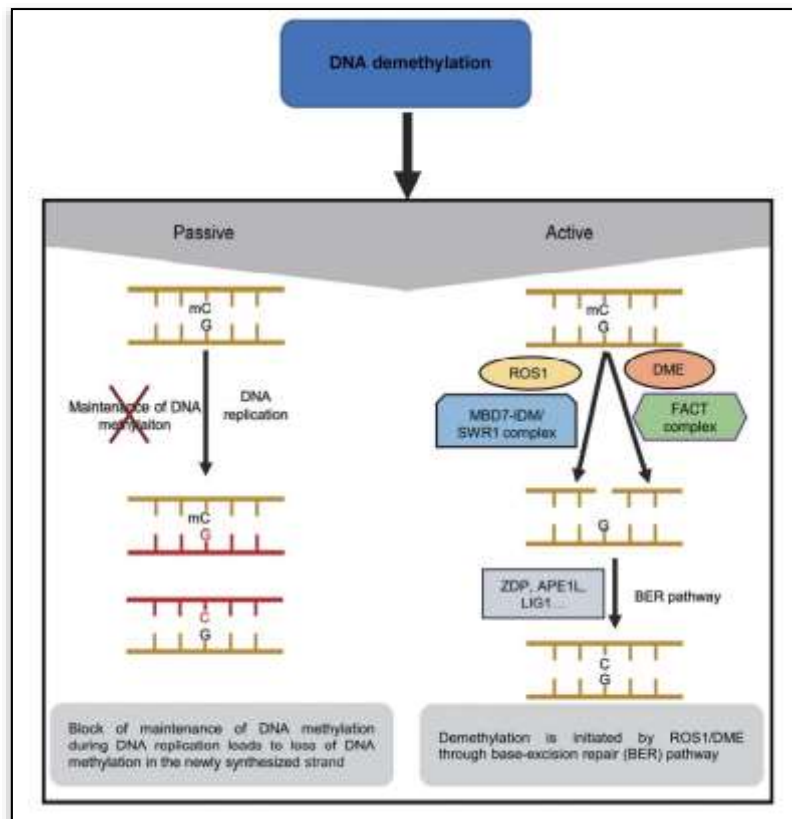


Figura 4: Meccanismo di demetilazione nelle piante (Liu et al., 2020)

Il livello di metilazione è controllato dinamicamente tramite l'inserimento, il mantenimento e la rimozione di citosine metilate. La demetilazione attiva del DNA è possibile grazie all'azione di un gruppo di enzimi bifunzionali capaci di rimuovere facilmente la metilazione. Nelle piante, il processo è iniziato da una famiglia di enzimi bifunzionali 5-mC DNA glicosilasi-apuriniche o apirimidiniche in grado di riconoscere e rimuovere direttamente il gruppo metile dalla citosina utilizzando un sistema di riparazione per escissione di basi (FIG. 6) (Gong et al., 2002; Gehring et al., 2006). In *A. thaliana* è presente una famiglia di quattro DNA glicosilasi bifunzionali: REPRESSOR OF SILENCING 1 (ROS1), DEMETER-LIKE PROTEIN 2 (DML2), DML3 e l'ultima proteina trascrptional activator demeter (DME). Le prime due sono proteine espresse in tutti i tessuti mentre l'ultima è espressa solamente a livello delle cellule compagne nei gameti femminili e maschili (Ortega-Galisteo et al., 2008). Queste proteine possono eseguire l'escissione della citosina in tutti i context di metilazione. Durante la demetilazione del DNA, le glicosilasi effettuano l'escissione della base idrolizzando il legame glicosidico tra la base ed il suo deossiribosio, tagliando così il DNA e creando un sito privo di base. Questo processo va a rimuovere le citosine metilate indipendentemente dalla sequenza e dal context di metilazione, generando un *gap* di un singolo nucleotide (FIG. 5) (Choi et al., 2002; McCullough et al., 1999; Morales-Ruiz et al., 2006). L'escissione è seguita da una reazione di  $\beta$ -eliminazione o  $\beta,\delta$ -eliminazione. A questo punto due endonucleasi lavorano dopo l'eliminazione per sintetizzare un gruppo OH al 3' terminale.

Rispettivamente la DNA liasi (APE1L) per la  $\beta$ -eliminazione e la DNA fosfatasi polinucleotide 3'-fosfatasi (ZDP). Il *gap*, adesso appositamente modificato, può essere riempito da una DNA polimerasi. Grazie ad una DNA ligasi si ha il ripristino della struttura del DNA con il completamento della riparazione (FIG. 5) (Lee *et al.*, 2014; Li *et al.*, 2015)

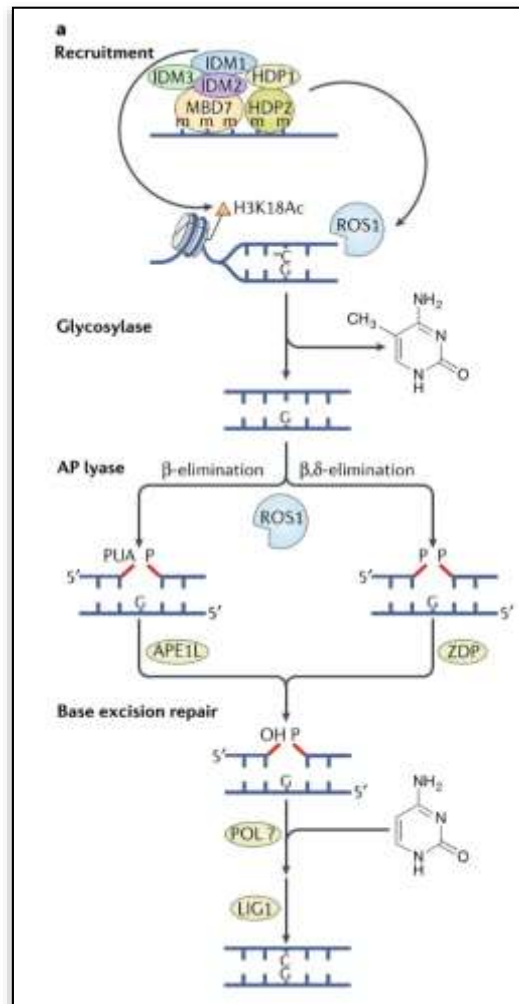


Figura 5: La ROS1-mediata demetilazione del DNA in *Arabidopsis thaliana*. (Zhang)

La proteina ROS1 ha dimostrato di possedere una specificità per particolari siti nel DNA. Questi siti sono definiti dalle caratteristiche della cromatina e dalle proteine di reclutamento che li identificano. Nel caso della DME, la demetilazione avviene in regioni contenenti trasposoni ricchi in sequenze AT in regioni di eucromatina, portando ad un'alterazione nell'espressione dei geni vicini (Gehring *et al.*, 2009). La demetilazione mediata da ROS1, oltre a coinvolgere trasposoni, aiuta a stabilire i confini tra gli stessi e i geni adiacenti. Facendo in modo di evitare la diffusione della metilazione del DNA e del silenziamento dell'espressione genica dai trasposoni ai geni vicini. Infine, ROS1 e la sua azione di



demetilazione sono molto importanti per contrastare l'espansione della metilazione, agendo in contrasto con il pathway della RdDM (FIG. 6) (Gong; Tang *et al.*, 2016).

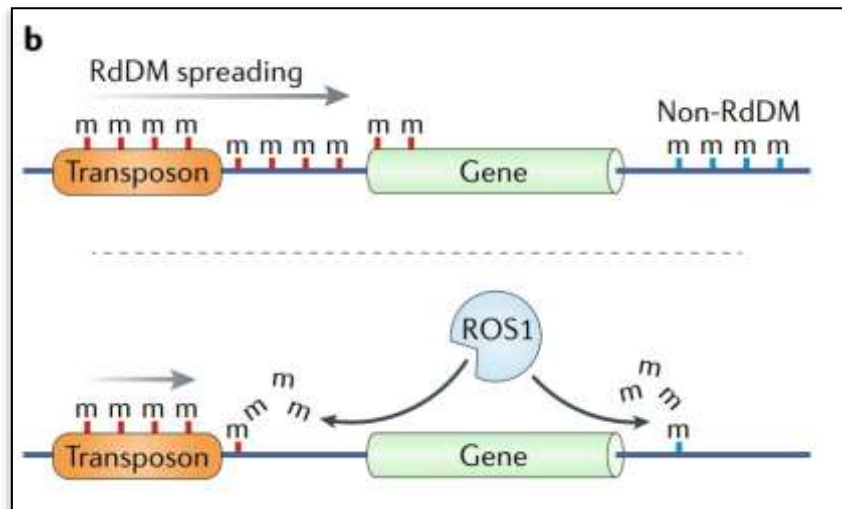


Figura 6: Demetilazione mediata da ROS1, aiuta a definire i limiti tra trasposoni e geni, prevenendo l'espansione della RdDM. (Zhang, 2018)

### 2.1.2 Epialleli

Nelle piante, metilazione e demetilazione concorrono a definire l'espressione genica e di conseguenza il fenotipo. Metilazioni spontanee possono alterare l'espressione genica portando alla formazione di alleli epigenetici o epialleli, questi differiscono nel numero e nella distribuzione di nucleotidi metilati in specifiche sequenze. Piante con epialleli differenti possono esibire, come detto, fenotipi distinti. Molti di questi epialleli sono solitamente stabili dopo una mitosi ed in molti casi anche dopo meiosi (Kakutani *et al.*, 2002; Kakutani *et al.*, 1999). Questa stabilità conduce alla possibilità di trasmissione alla progenie, poiché durante la replicazione di sequenze di DNA metilato si ha la sintesi di una sequenza emi-metilata, dove un solo filamento di DNA presenta la metilazione del filamento stampo. La cellula, pertanto, è capace di ripristinare il livello di metilazione del DNA di origine. Questo va a costituire un solido meccanismo per la trasmissione di schemi di metilazione tra cellule di uno stesso organismo e anche attraverso le generazioni (Kalisz *et al.*, 2004). Molti di tali epialleli sono influenzati da variazioni genetiche, principalmente da varianti strutturali e da attivazione di trasposoni, inoltre la formazione di epialleli indipendentemente da variazione genetica è stata osservata in molte specie (Srikant *et al.*, 2021; Kalisz *et al.*). Un'analisi del metiloma in *A. thaliana*, che ha coinvolto 30 generazioni della pianta, ha rivelato che le mutazioni epigenetiche avvengono con una frequenza molto più alta in confronto alle mutazioni geniche (Ossowski *et al.*, 2010; Van Der Graaf *et al.*, 2015).

Inoltre, un altro studio, ha riportato che per ogni generazione prodotta tramite riproduzione sessuale, gli epialleli spontanei si manifestano in regioni differenzialmente metilate (DMR) biologicamente significative in meno del 0.003% del metiloma, mentre il restante 99.997% delle regioni metilate è stato passato alla progenie fedelmente (Hofmeister *et al.*, 2017). Le DMR sono un concetto fondamentale per dare un significato biologico alla metilazione del DNA. Ci indicano regioni genomiche con un differente stato di metilazione tra confronti biologici e sono considerate zone di alto interesse per la loro possibile influenza nella regolazione genica. La formazione di nuove DMR può dipendere da nuove mutazioni sotto forma di epialleli. Tali mutazioni hanno dimostrato un grande collegamento con l'origine geografica ed il clima della pianta nella quale avvengono, mostrando dunque il loro potenziale coinvolgimento nei meccanismi di risposta e adattamento della pianta (Kawakatsu *et al.*, 2016), oltre ad essere coinvolti in varie funzioni biologiche: nel controllo della senescenza fogliare, nel metabolismo dell'amido, nella biosintesi del triptofano, nella resistenza della pianta a patogeni e nello sviluppo floreale (He *et al.*, 2018; Silveira *et al.*, 2013; Srikant). Un esempio di importanza biologica di epialleli ereditabili è evidente nel pomodoro, precisamente nel mutante *crn*, nel quale la trascrizione del gene COLOURLESS NON-RIPENING (*crn*) è repressa da una ipermetilazione del DNA a livello del promotore, mostrando come risultato dei frutti privi del tipico colore rosso e con maturazione inibita (FIG. 7) (Manning & *et al.*, 2006).

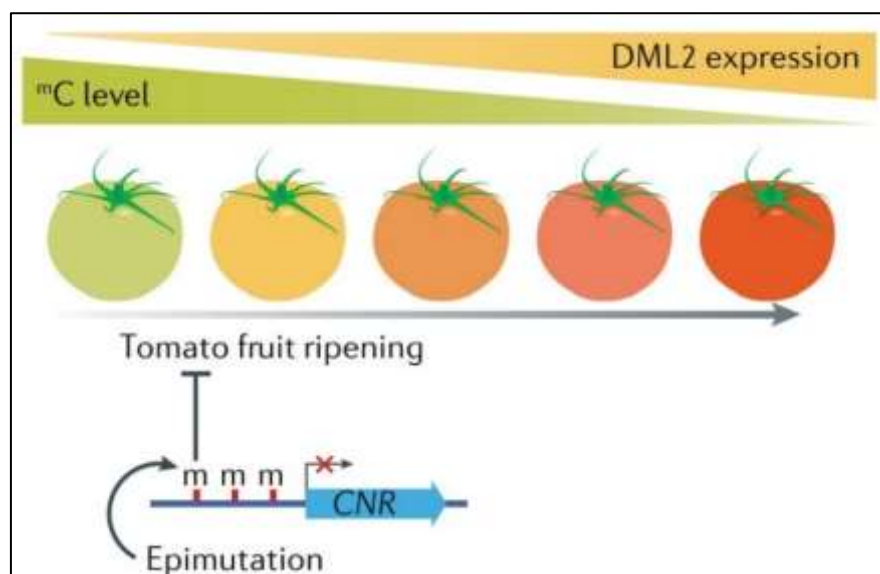


Figura 7: Pomodoro durante la maturazione a seguito di una epimutazione a livello del gene CRN. (Zhang, 2018)

Questo tipo di epimutazioni spontanee avviene con una bassa frequenza rispetto a quella con cui mutazioni genetiche possono avvenire a livello della stessa citosina (Johannes *et al.*, 2019). Considerando però che la maggior parte delle regioni metilate sono conservate e fedelmente trasmesse alla progenie attraverso le generazioni, anche queste nuove mutazioni hanno il potenziale

di essere mantenute e divenire ereditabili, degli epialleli funzionali (Srikant, 2021). In breve, gli epialleli possono essere impressi, alleli puri, spontanei. Si definiscono impressi quando sono generati durante lo sviluppo riproduttivo in piante wild-type (WT), in cui vengono rimossi i segni di metilazione specifici dell'allele parentale, avviando la trascrizione genica (FIG. 8A). “Epialleli puri” sono cambiamenti spontanei della metilazione, che possono verificarsi indipendentemente dalla variazione genetica sottostante e possono anche essere soggetti a fissazione evolutiva (FIG. 8B). Infine gli epialleli spontanei si formano in seguito a mutazioni nei geni codificanti proteine, compresi i geni coinvolti nel mantenimento della metilazione del DNA (metiltransferasi) e la mobilitazione di elementi trasponibili (TE). Portando alla formazione di epialleli, poiché possono influenzare il livello di metilazione e la trascrizione del gene prossimale (FIG. 8C). Insieme, questi cambiamenti genetici ed epigenetici appena acquisiti possono fornire un vantaggio adattivo per la pianta (Srikant, 2021).

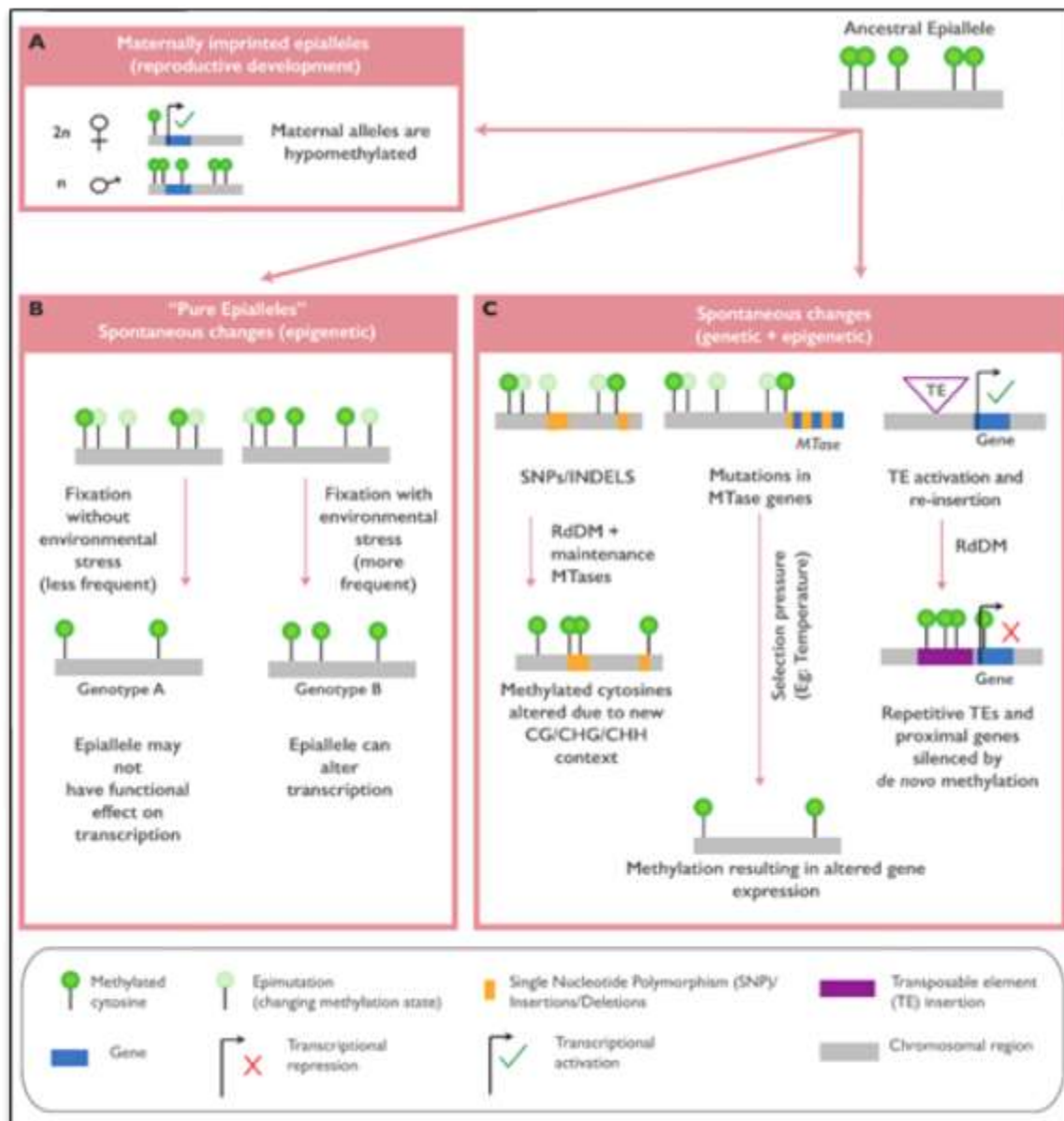


Figura 8: Origine degli epialleli naturali. (Srikant et al., 2021)

### 2.1.3 Stress

L'ambiente va a rappresentare, per la pianta, una continua fonte di cambiamento, il quale può presentarsi sotto forma di stress. Esso, come abbiamo visto parlando di epialleli, costituisce uno stimolo molto importante nel contesto dei cambiamenti epigenetici e in particolar modo nella metilazione del DNA. Quest'ultima agisce come una risposta da parte della pianta ai vari stimoli abiotici e biotici presentati dall'ambiente (Zhang, 2018). In natura, le piante sono esposte a parassiti, fitofagi e microrganismi, che sono i principali agenti di stress biotico. In presenza di questo tipo di

stimoli, come infezioni da potenziali patogeni oppure una colonizzazione di microrganismi simbiotici, le piante hanno dimostrato la capacità di rispondere attivando una serie di meccanismi, tra cui la metilazione del DNA. Il processo comune di formazione e sviluppo di un nodulo, presente a livello della radice, consiste nella formazione di cellule ingrossate e dall'accumulo di amiloplasti tra le cellule meristematiche della radice, il tutto è dovuto ad un processo simbiotico di un batterio. Il nodulo può arrivare a modificare la metilazione in centinaia di DMR, incluse alcune specifiche per il suo sviluppo (Nagyimihály *et al.*, 2017). Alterazioni ulteriori al processo di metilazione, come modifiche nella regolazione, possono determinare modifiche alla suscettibilità della pianta a patogeni (Yu *et al.*, 2013). Lo stress di tipo abiotico si compone di una vasta gamma di condizioni sfavorevoli che l'ambiente può presentare alla pianta. Ad esempio: calore, freddo, siccità, elevata salinità, stress osmotico, carenza di nutrienti, anossia e molti altri. In molti studi è stata dimostrata la correlazione tra stress e risposta della pianta per mezzo di metilazione o demetilazione del DNA, a livello di tutto il genoma e/o in loci specifici che fungono da importante strumento di adattamento (Yong-Villalobos *et al.*, 2015). La metilazione del DNA permette quindi di regolare l'espressione genica in risposta a stimoli esterni, a seguito di una esposizione persistente e continua, andando a determinare una memoria dello stress. Tale memoria è in grado di essere riprodotta dalle generazioni successive se è presente una continuità nell'esposizione (FIG. 9) (Sanchez *et al.*, 2014). Un esempio di modificazione epigenetica indotta da uno stress abiotico è la condizione in cui, nei frutti di pomodoro, l'esposizione a basse temperature può portare ad una perdita di aroma. Le basse temperature causano una riduzione dell'espressione della DNA metilasi DML2, per cui provocano una ipermetilazione a livello dei promotori e relativo silenziamento dei geni responsabili della biosintesi dei composti che ne definiscono l'aroma (Zhang *et al.*, 2016). Alla luce di tutto questo possiamo quindi dire che la metilazione è correlata con il silenziamento dei geni e degli elementi trasponibili, mentre la demetilazione è correlata con l'attivazione della trascrizione.

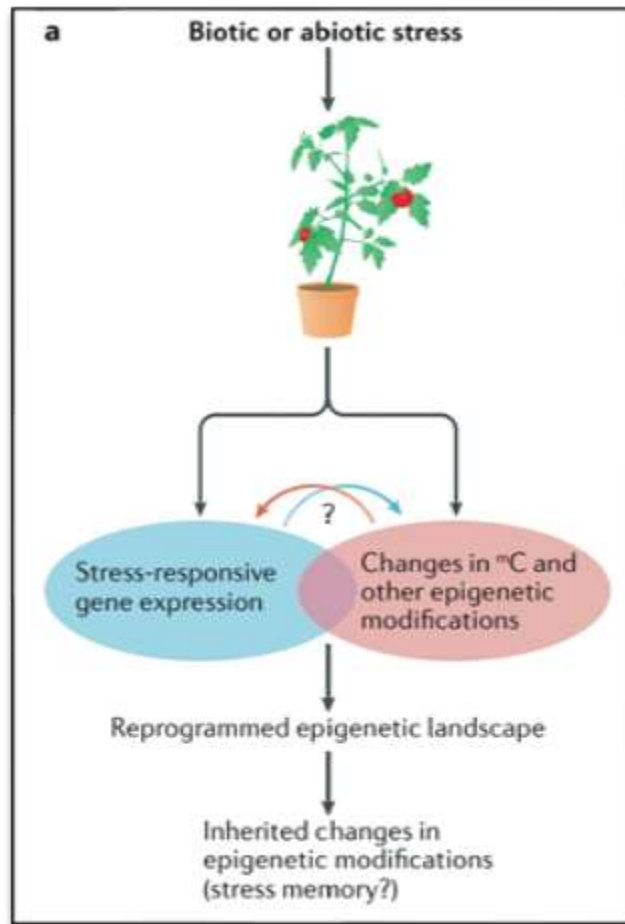


Figura 9: Modificazioni epigenetiche causate da esposizione a stress biotico o abiotico e possibile ereditabilità.

Uno degli strumenti per combattere lo stress abiotico e biotico è la tecnica dell'innesto. Questo strumento, molto diffuso in ambito agricolo in tutto il mondo, rappresenta una valida soluzione per aumentare la resistenza a stress termici e salini (Rivero *et al.*, 2003). Tuttavia, l'innesto sottopone la pianta a stress multipli che possono portare a modificazioni nella metilazione del DNA, tra cui lo stress meccanico generato dalla ferita per il taglio iniziale (Berger *et al.*, 2018). A sostegno di questo, uno studio ha trovato in *A. thaliana* una parte importante di small RNA, soprattutto siRNA, con capacità di influenzare il processo di RdDM e di conseguenza la metilazione del DNA. Il siRNA può muoversi attraverso l'innesto ed arrivare ad influenzare altre cellule (Molnar *et al.*, 2010). Questo dimostra come la pratica dell'innesto possa portare a cambiamenti epigenetici e fenotipici.

#### 2.1.4 Analisi di metilazione

Per studiare la metilazione del DNA e le eventuali modificazioni a suo carico, è stato necessario lo sviluppo di tecniche in grado di permettere l'identificazione dello stato di metilazione di geni e la quantificazione delle citosine metilate su tutto il genoma. Questo ci permette di valutare lo stato di metilazione e la differenza tra due o più confronti. Tra le tecniche disponibili la Gold Standard è di certo la "whole-genome bisulfite sequencing" (WGBS), seguita dalla "reduced-representation bisulfite sequencing" (RRBS). Queste sfruttano il bisolfito di sodio per convertire le citosine non metilate in uracile, mentre quelle metilate restano invariate. Consentendo quindi una misurazione del livello di metilazione del DNA, il che le pone alla base delle analisi di metilazione.

##### *Whole-genome bisulfite sequencing (WGBS)*

Tramite la tecnica WGBS è possibile verificare lo stato di metilazione delle singole citosine lungo tutto il genoma, ma il costo elevato di tale pratica la rende inadeguata per studi con un limitato numero di confronti (Suzuki *et al.*, 2018) e per genomi troppo grandi. Questa metodica prevede: la purificazione e rottura del genoma, l'arricchimento dei frammenti prodotti con adenine nella parte terminale 3' e il legame di adattatori metilati. Successivamente il DNA di interesse viene trattato con bisolfito di sodio, un composto chimico in grado di deaminare le citosine non metilate in uracili, i quali sono in seguito convertiti in timine. Al contrario, le citosine metilate non vengono convertite (Hayatsu & Hikoya, 2008). Grazie a questa reazione è possibile determinare l'esatta posizione genomica delle citosine metilate tramite il confronto tra la sequenza del campione trattato e quella del controllo non trattato oppure con la sequenza del genoma di riferimento disponibile per la specie studiata.

##### *Reduced-representation bisulfite sequencing (RRBS)*

La tecnica RRBS è un'alternativa efficace alla WGBS per l'analisi di grandi genomi. RRBS è in fatti capace di fornire dati di alta qualità per analisi di genomi di grandi dimensioni (Meissner *et al.*, 2005). Questa metodica aggiunge al trattamento con bisolfito di sodio, una digestione enzimatica a cura dell'enzima *MspI*. Il processo di selezione della lunghezza avviene tra segmenti compresi tra 40 e 220bp arrivando a coprire circa 85% delle isole CG, ed in particolare sui promotori. Questa tecnica di analisi della metilazione del DNA permette un risparmio notevole in confronto al WGBS, in quanto si concentra sulle regioni ricche del context CG in prossimità dei siti di restrizione. Ma determina anche una perdita di coverage a livello delle regioni intergeniche ed elementi regolatori distali (Yong *et al.*, 2016). Presenta però sempre un costo elevato e poco accessibile.

### *Methylation Context Sensitive Enzyme ddRAD, MCSeEd*

Per superare alcune delle limitazioni presenti nelle precedenti tecniche, è stata sviluppata la tecnica MCSeEd. Questa presenta un'estensione molto semplice, un'estesa scalabilità, una buona efficienza nei costi ed inoltre permette la determinazione dei cambiamenti nella metilazione nei context CG, CHG, CHH e 6mA (Marconi *et al.*, 2019). Tra i vantaggi di questa tecnica c'è la possibilità di sviluppare delle librerie a rappresentazione ridotta, sfruttando un approccio scalare molto più efficiente ed economico rispetto ai metodi con bisolfito. Infine offre la possibilità di mappare loci utilizzando un genoma di riferimento oppure condurre un'analisi priva di riferimento come nell'approccio genome-independent. La tecnica MCSeEd porta come innovazione l'utilizzo di due enzimi di restrizione per la digestione del confronto. La digestione avviene in parallelo da parte di un enzima insensibile alla metilazione (*MseI*) ed uno sensibile alla metilazione delle citosine (es. *AclI*) (FIG. 10). Il taglio da parte di questi due enzimi produrrà frammenti con le estremità adesive alle quali saranno legati due adattatori. Un adattatore *barcode* specifico per l'estremità tagliate dall'enzima sensibile alla metilazione, utile per identificare i confronti nella successiva fase di sequenziamento. L'altro adattatore è detto ad "Y", specifico per i siti tagliati dall'enzima insensibile alla metilazione. Successivamente viene effettuata l'amplificazione dei frammenti utilizzando primer specifici per i due adattatori. Dopodiché si procede con la size selection, cioè una selezione dei campioni in base alle loro dimensioni. I campioni vengono valutati con una elettroforesi su gel e si selezionano frammenti di DNA nell'intervallo da 250bp a 600 bp. La peculiare natura dell'adattatore ad "Y" lo porta a svolgere un ruolo chiave in questa fase, infatti permette l'amplificazione solo di quei frammenti in possesso di entrambi gli adattatori. A seguito si procederà con un sequenziamento NGS Illumina, ed una analisi bioinformatica utilizzando pipeline appropriate per la tecnica (FIG. 10).



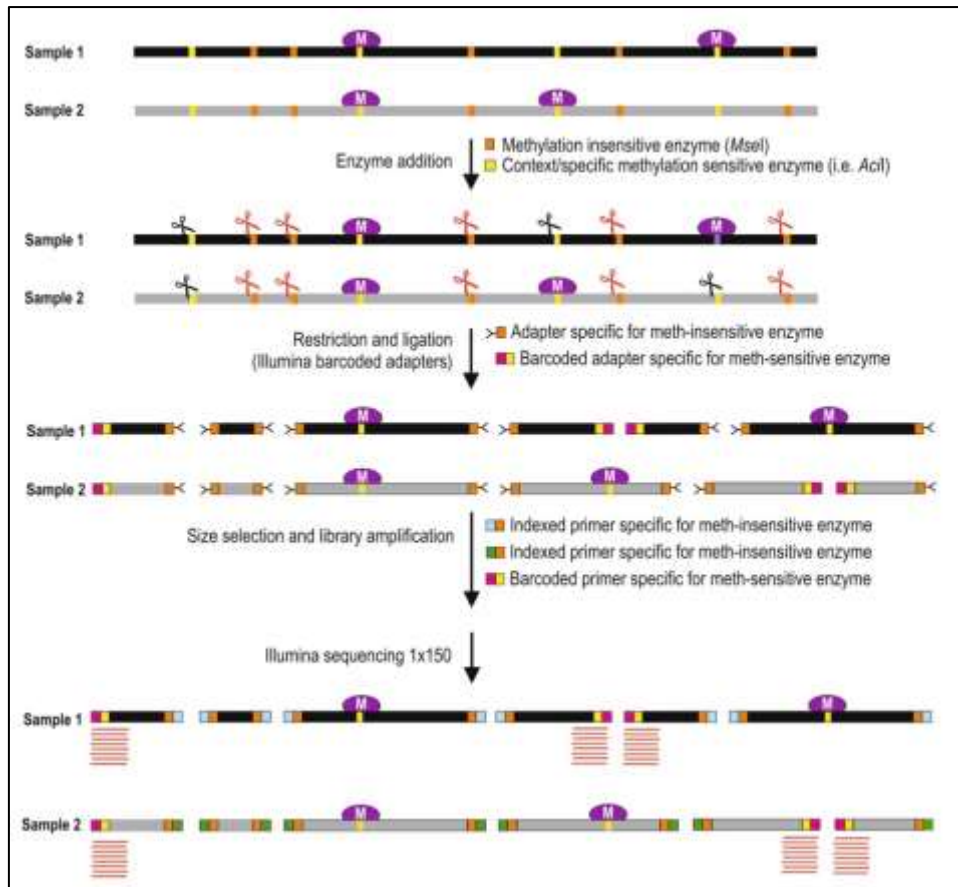


Figura 10: Rappresentazione schematica di MCSeEd a due confronti (Marconi *et al.*, 2019)

## 2.2 INNESTO

La pratica dell'innesto può essere definita come il naturale o artificiale atto di fusione tra parti di una pianta, in modo che sia presente una continuità vascolare tra le parti (Pina *et al.*, 2005) e che l'organismo risultante possa continuare a funzionare come una singola pianta. In generale, due piante differenti precedentemente intatte possono essere innestate insieme naturalmente o intenzionalmente. L'innesto intenzionale, svolto dall'uomo, consiste nell'inserire la parte superiore di una pianta, precedentemente tagliata, chiamata nesto (scion) all'interno di un'apertura in un'altra pianta con un proprio sistema radicale sviluppato, chiamato portainnesto (rootstock). Con il termine nesto ci si riferisce alla parte della pianta tagliata dal donatore, questa a seguito dell'innesto crescerà come la parte superiore della pianta innestata. Il portainnesto è la parte che riceve il nesto, e diverrà la parte inferiore della pianta innestata. Esso può presentarsi in diverse configurazioni a seconda se comprende solamente il sistema radicale della pianta oppure anche una breve porzione del fusto subito sovrastante le radici. La pratica più comune di innesto comprende una singola unione di un nesto ad un portainnesto, è presente anche una variante chiamata "double working" (Mudge *et al.*,

2009) che consiste in una unione tra tre piante, il risultato è una nuova pianta innestata composta da tre parti geneticamente distinte (FIG. 11).

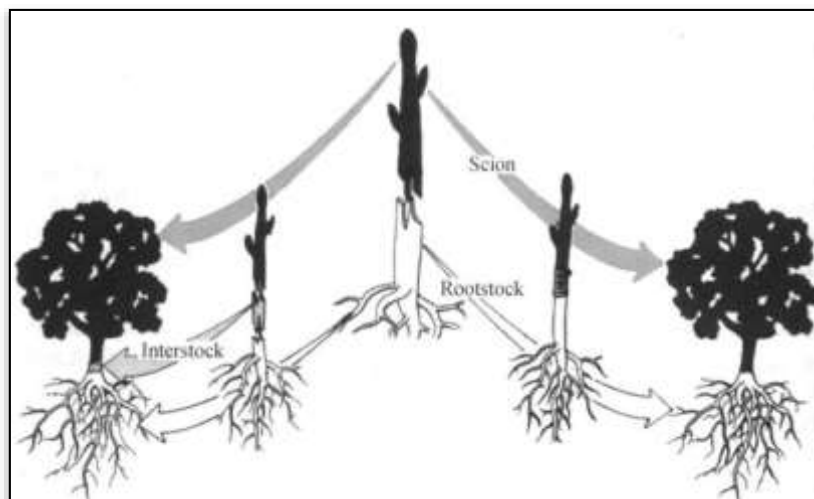


Figura 11: Illustrazione delle possibilità della tecnica dell'innesto (Mudge et al., 2009).

Una delle considerazioni più importanti da effettuare è quella della compatibilità tra le piante ed i suoi limiti. Questa pone le sue basi sulla componente genetica, ma si avvale anche della differenza tra la tassonomia dei soggetti sottoposti all'innesto.

### 2.2.1 Storia

L'innesto ha origini molto antiche. La tecnica è stata utilizzata in piante da frutto per millenni, ma solo recentemente ha ricevuto le attenzioni mondiali ed un impiego su larga scala per il commercio (Sakata *et al.*, 2007). Vecchie testimonianze possono essere trovate in antichi testi cinesi, coreani e giapponesi. L'innesto è noto alle popolazioni cinesi sin dal 1560 a.C. Ulteriori testimonianze dell'utilizzo della tecnica sono state trovate in alcuni testi biblici, nella letteratura greca a cura di Aristotele nei suoi scritti per agricoltura (384-322 a.C.) e da Theophrastus (371-287 a.C.) (Lee & Jung-Myung, 2003). Anche se gode di antiche origini, la tecnica dell'innesto è solo recentemente divenuta di utilizzo commerciale. L'opportunità per questa espansione è arrivata grazie all'estensiva diffusione e produzione di plastica nei tardi anni '50 che ha permesso lo sviluppo di nuovi strumenti. La prima introduzione per fini commerciali si deve al Giappone ed alla Corea all'inizio degli anni '90.

Successivamente fu introdotta anche dalle nazioni occidentali e quindi entrò in uso su scala globale. Questa diffusione è dovuta anche ai numerosi benefici che questa tecnica porta con sé, rendendola valida nel ridurre problemi associati a successivi raccolti e tolleranza allo stress. Inoltre, l'aumento della consapevolezza nei consumatori, i quali cercano sempre con maggior richiesta un'alimentazione sicura, funzionale e rispettosa dell'ambiente, comporta un grande incentivo per la tecnica.

### 2.2.2 Tecniche dell'innesto

La tecnica dell'innesto comprende determinati passaggi per una sua esecuzione corretta: la scelta della specie per il portainnesto e il nesto, creazione di una unione fisica tra le due parti, la guarigione dall'unione e l'acclimatazione della nuova pianta innestata (Lee *et al.*, 2010). La tecnica utilizzata, così come la tipologia di strumenti, varia molto in base alla pianta sulla quale si andrà a lavorare, all'esperienza dell'operatore, alle preferenze nel metodo di coltura, al numero di innesti da eseguire e allo scopo stesso dell'innesto se eseguito per fini commerciali o per ricerca. Ad oggi, a livello commerciale, l'innesto eseguito a mano rimane la via di produzione più diffusa, soprattutto per le piante erbacee, seppure siano presenti molti macchinari in grado di eseguire tutti i compiti necessari (Lee, 2010). Per questa ricerca è stata utilizzata la più comune tecnica dell'innesto, per la produzione di piante di pomodoro innestate, *Tongue approach grafting* (TAG). Questa tecnica è una delle più diffuse per la sua semplicità di esecuzione e la più alta possibilità di successo e sopravvivenza della pianta, specialmente per le piante erbacee ed appartenenti alla famiglia delle *Solanacee* (Lee & Jung-Myung, 2003). Questa tecnica di innesto ha maggior successo quando le piante sono molto giovani e sia il nesto che il portainnesto hanno una lunghezza simile. Le parti in crescita del portainnesto devono essere rimosse con attenzione prima di procedere, così da evitare sprechi nei nutrienti e velocizzare il processo di recupero dall'innesto. Il taglio viene effettuato ad un angolo di circa 30°-40° perpendicolare agli assi, profondo abbastanza da permettere la fusione di più canali vascolari possibili, ascendente per il nesto e discendente per il portainnesto. Dopodiché si posizionano delle specifiche pinze per tenere in posizione l'innesto (FIG. 12). Al completamento le nuove piante innestate sono tenute sotto controllo in un ambiente adatto per 1-2 giorni prima di essere immesse in serra. Alla fine, dopo un periodo che varia dai 10 ai 12 giorni, se l'innesto è avvenuto con successo, si procede alla pulizia delle radici del nesto effettuando un taglio il più vicino possibile al punto di unione. La pinza sarà poi tolta al momento del trapianto della pianta (Lee *et al.*, 2010).

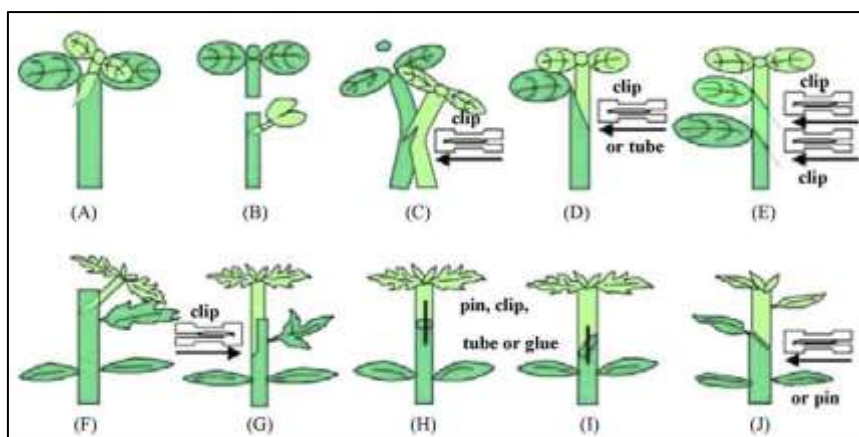


Figura 12: Principali metodi di innesto nelle Cucurbitacee e nelle Solanacee: innesto (A e B); (C) Tongue approach grafting (TAG); (D, E e J) innesto di giunzione; (F, G) cleft grafting (H e I) (Lee et al., 2010).

### Acclimatazione

L'acclimatazione è fondamentale per assicurare la sopravvivenza delle piante innestate. Questa procedura coinvolge la riparazione del taglio provocato dall'innesto fino a che la pianta non è pronta per essere spostata in una normale condizione di serra (Lee, 2010). La tecnica, nella maggior parte dei casi, comprende l'utilizzo di coperture o involucri di plastica di colore nero. Le piante possono essere sia avvolte nell'involucro di plastica finché l'unione non si è formata, ed evitare la produzione di eccessivo calore. In aggiunta, tenere la pianta in una condizione ombreggiata può rivelarsi utile per ridurre il calore, per questo scopo sono state sviluppate molte camere di climatizzazione. Esse sono in grado di creare l'ambiente perfetto per sostenere la pianta fino alla completa guarigione dell'innesto (Lee et al., 2008).

### 2.2.3 Benefici

La pratica dell'innesto è ad oggi largamente utilizzata in tutto il mondo grazie ai suoi ovvi benefici, specialmente nelle piante della famiglia delle *Solanacee* (Cui et al., 2021). Come si può osservare nella Tab. 1, questi benefici sono molto importanti per i coltivatori e per la ricerca. Essi variano in base alla specie di pianta ed al portainnesto scelto per l'unione, aprendo le porte a un crescente numero di combinazioni che garantiscono la massima efficienza e adattabilità ad ogni tipo di coltura, in campo ed in serra. I più importanti e ricercati tra questi sono la: tolleranza a stress abiotico e biotico, l'incremento del raccolto, l'aumento di vigore nella pianta e gli effetti sulla qualità del frutto (Lee, 2010).

Tabella 1 Elenco dei vari benefici che si possono avere con l'innesto. (Lee et al., 2010)

Vantaggi	Svantaggi
Aumento del raccolto	Semi aggiuntivi per il portainnesto
Maggior crescita del fusto	Esperienza richiesta
Resistenza alle malattie	Selezione della combinazione nesto/portainnesto migliore
Resistenza e tolleranza a nematodi	Combinazioni differenti per la stagione di raccolta
Resistenza ad alte temperature	Combinazioni differenti a seconda del metodo di raccolta
Resistenza a basse temperature	Prezzi maggiori
Maggiore assorbimento di nutrienti	Maggior occorrenza di malattie del seme
Maggiore assorbimento d'acqua	Crescita eccessiva
Tolleranza a stress salino	La raccolta dei frutti potrebbe essere posticipata
Tolleranza a suolo bagnato	Possibile qualità inferiore del frutto
Tolleranza a metalli pesanti ed inquinamento organico	Maggior possibilità di disordini fisiologici della pianta
Cambiamenti di qualità	Eventuali sintomi di incompatibilità a stadi avanzati dell'innesto
Periodo di raccolta esteso	Utilizzo di diverse tradizioni culturali
Raccolti multipli o consecutivi permessi	Prezzi aumentati per le piante innestate
Produzione di scarti organici convenienti	
Valore ornamentale anche per educazione	

#### *Tolleranza da stress biotici*

Quando si parla di stress biotico, ci si riferisce all'attacco della pianta dovuta da agenti patogeni residenti nel suolo. In questo caso l'innesto ha provato di essere una pratica soluzione al problema, riducendo l'utilizzo di agrofarmaci. I portainnesti vigorosi hanno dimostrato di possedere una eccellente resistenza agli stress biotici (Morra & Bilotto, 2006), conferendo la possibilità di proteggere i raccolti. La resistenza della nuova pianta innestata dipende dal portainnesto, poiché le vulnerabilità del nesto non vengono trasmesse (Lee, 2010)

### *Tolleranza da stress abiotici*

Le piante innestate hanno presentato una maggiore resistenza a molti tipi di stress abiotici, a seconda del portainnesto scelto la pianta risultante può acquisire la resistenza desiderata. La tecnica viene utilizzata per indurre la resistenza contro basse ed alte temperature (Venema *et al.*, 2008), molto importante per piante *Solanacee* e *Cucurbitacee*. Inoltre viene utilizzata per aumentare la resistenza a stress idrico, diminuire l'assorbimento di inquinanti dal suolo, aumentare l'importo di nutrienti e la resistenza a stress osmotici (Rivero *et al.*, 2003; Schwarz *et al.*, 2010).

### *Aumento di vigore*

Con il termine aumento del vigore della pianta si intende la capacità della stessa di svilupparsi. Questo può essere stimolato e migliorato tramite l'utilizzo dell'innesto. Le radici nel portainnesto sono di solito molto più grandi e sviluppate di quelle della pianta da cui è preso il nesto e ciò porta ad un assorbimento di nutrienti e di acqua molto più efficiente (Lee, 2010). Inoltre, permette un utilizzo di agrofarmaci e fungicidi minore, non essendo più necessari nelle stesse quantità grazie all'effetto vigoroso sulla pianta.

### *Incremento del raccolto*

L'innesto è associato ad un incremento del raccolto, oltre a garantire una maggior resistenza a patogeni e stress, la quale porta una maggior sopravvivenza della pianta. I portainnesti vigorosi possono anche portare ad un incremento generale del raccolto. In pomodoro è stato riscontrato un aumento del raccolto in condizioni di serra e con sufficiente acqua e nutrienti. Sono anche aumentati il peso del frutto ed il numero di frutti per pianta (Djidonou *et al.*, 2017). Simili aumenti di raccolto sono stati registrati in molte altre colture.

### *Qualità del frutto*

L'influenza dell'innesto sulla qualità del frutto è un campo ancora non molto esplorato. Le sue meccaniche sono tuttora sotto studio per decifrare i complessi meccanismi coinvolti. La qualità può essere brevemente definita come la combinazione di stimoli visivi, forma, colore, aroma, contenuti solidi solubili (SSC) e acidità titolabile (TA) (Flores *et al.*, 2010). Queste proprietà possono essere influenzate dall'innesto, in particolare dal portainnesto. Una diversa combinazione di nesto e portainnesto porterà a differenti risultati riguardanti il frutto, un aumento dei valori di minerali e grandezza del frutto. Una generale aumento della qualità è stato osservato in studi riguardanti

pomodoro con portainnesti ibridi (Djidonou, 2017). Va tenuto in considerazione che i miglioramenti possono essere trascurabili rispetto alla coltura originale. Talvolta si possono ottenere effetti indesiderati e deleteri andando a ridurre gli SSC e provocando una degradazione interna del frutto durante la maturazione (Lee, 2010).

#### 2.2.4 Conseguenze epigenetiche

Quando l'innesto è avvenuto con successo, esso può dunque portare a cambiamenti fenotipici, vere e proprie variazioni che possono dipendere da modificazioni epigenetiche. Questo fenomeno osservato in molti studi ci porta ad ipotizzare, in accordo con i dati prodotti, che la metilazione del DNA e il trasferimento di mRNA, miRNA e siRNA tra nesto e portainnesto, siano parti fondamentali in un più grande dialogo epigenetico (Berger *et al.*, 2018). L'innesto è in grado di stabilire un linguaggio di comunicazione a lunga distanza, tra le due parti, così da definire quelle che saranno le caratteristiche della nuova pianta innestata.

##### *Risposta ad una ferita causata dall'innesto*

Il processo a cui va in contro la pianta sottoposta all'innesto causa un gravissimo stress su di essa, in quanto si va a tagliare fisicamente parte della pianta interrompendo i flussi ormonali e metabolici. Questo mette in moto delle risposte specifiche, che possono includere una modifica della metilazione del DNA. Analisi trascrittomiche di piante innestate hanno confermato che, poco dopo l'innesto, delle risposte contro lo stress sono attivate in *Arabidopsis*, e che simili risposte sono attive in piante tagliate ma non innestate (Melnik *et al.*, 2018). Dopo il taglio e la conseguente perdita di integrità della pianta, viene generata una risposta, localmente a livello della giunzione dell'innesto, ma potenzialmente lungo tutta la pianta grazie ad una comunicazione a lunga distanza. L'interruzione dei flussi metabolici ed ormonali pone la pianta in una situazione critica a livello del taglio, causando l'accumulo di zucchero nel nesto ma non nel portainnesto (FIG. 13A). Vista la natura dell'innesto e la complicità delle modificazioni epigenetiche in situazioni di stress, è possibile ipotizzare una connessione tra le due.

### Callogenesi e differenziamento cellulare

Una dei primi eventi che avviene dopo l'inserimento dell'innesto è la callogenesi, essa consiste nella formazione di un callo (FIG. 13B). Questo è un complesso processo che necessita di modificazioni a carico di cellule somatiche adulte, le quali regrediscono ad uno stadio meno differenziato e riacquistano le loro capacità proliferative (Ikeuchi *et al.*, 2013), andando in contro ad una riprogrammazione epigenetica. Un modello per studiare la formazione di un callo ed il differenziamento cellulare ad esso associato, in vitro, sono le colture con la presenza del modello "Callus inducing medium" (CIM) (Berger). Grazie a questo sono state trovate prove del coinvolgimento della metilazione del DNA e di modificazioni istoniche in mutanti di *A. thaliana*. Nei mutanti *met1* è stata registrata una quasi completa eliminazione di metilazione nel context CG (Berdasco *et al.*, 2008) e nei mutanti *kyp-2* è stata rilevata una riduzione di metilazione nell'istone H3 nella lisina 9 (H3K9me2) e del context CHG (Jackson *et al.*, 2004). Questi ed altri studi evidenziano come un coinvolgimento epigenetico sia molto presente nei processi di guarigione della pianta. (FIG. 13C) Le dinamiche con le quali l'epigenetica influenzi la rigenerazione dei tessuti sono ancora sconosciute (Berger).

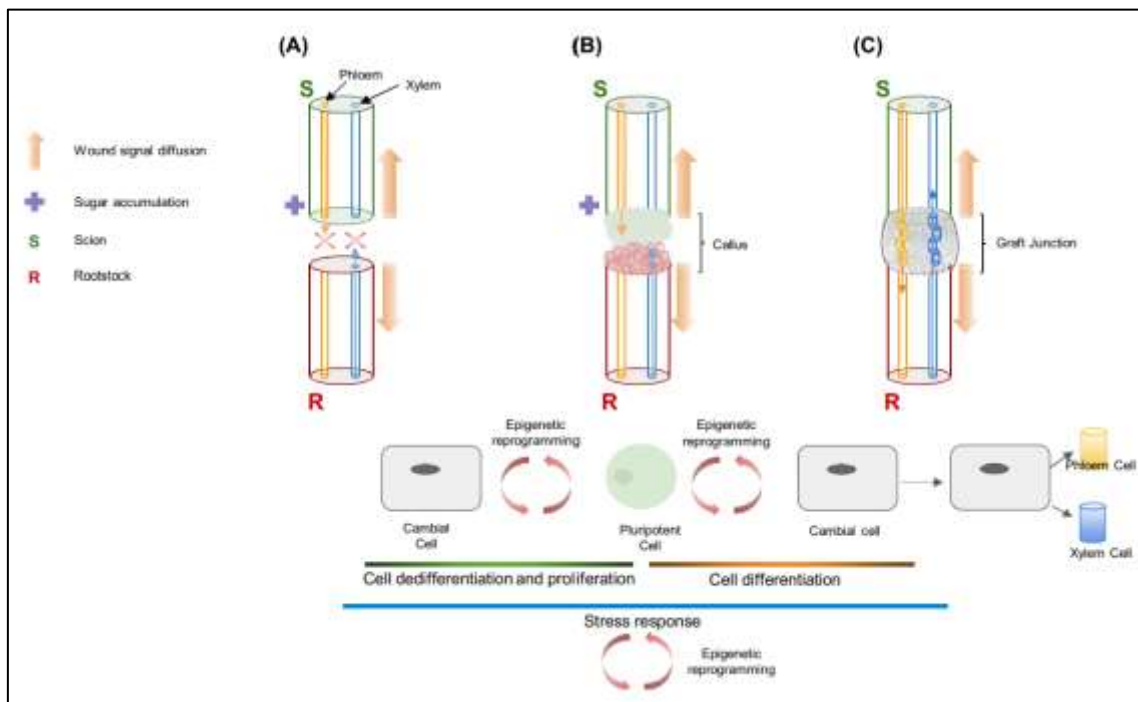


Figura 13: A Situazione iniziale, la pianta ferita e perdita dell'integrità. B Callogenesi. C Riparazione delle vie vascolari e guarigione della pianta (Berger)



### Scambio di small RNA tra le due parti della pianta

Ad oggi sono presenti molte evidenze che sottolineano il movimento a lungo raggio degli small RNA attraverso la pianta, questi sono i responsabili del silenziamento di geni e di conseguenza influiscono sul fenotipo della pianta (FIG. 14) (Pyott *et al.*, 2015). Lo scambio di questi small RNA (21-24 nt) avviene tra il nesto ed il portainnesto. La maggior efficienza che è presente nel trasporto dal portainnesto al nesto potrebbe essere la causa della grande influenza che esso ha sulla pianta finale. I miRNA, coinvolti nella regolazione dell'espressione genica post trascrizionale, possono muoversi attraverso il floema della pianta innestata. I siRNA invece sono coinvolti nella metilazione del DNA, questo porta il loro trasporto lungo la pianta a svolgere il compito di regolatore per la metilazione de novo nei tessuti che li ricevono, attraverso la via RdDM. Questo perché in situazione di innesto tra due specie differenti o genotipi diversi (hetero-graft) i siRNA possono trovare loci contenenti citosine non metilate, dando origine ad eventi di metilazione del DNA con il potenziale di influenzare il fenotipo. Queste modificazioni epigenetiche possono divenire stabili ed essere trasmesse tramite divisione cellulare formando delle stabili modificazioni fenotipiche.

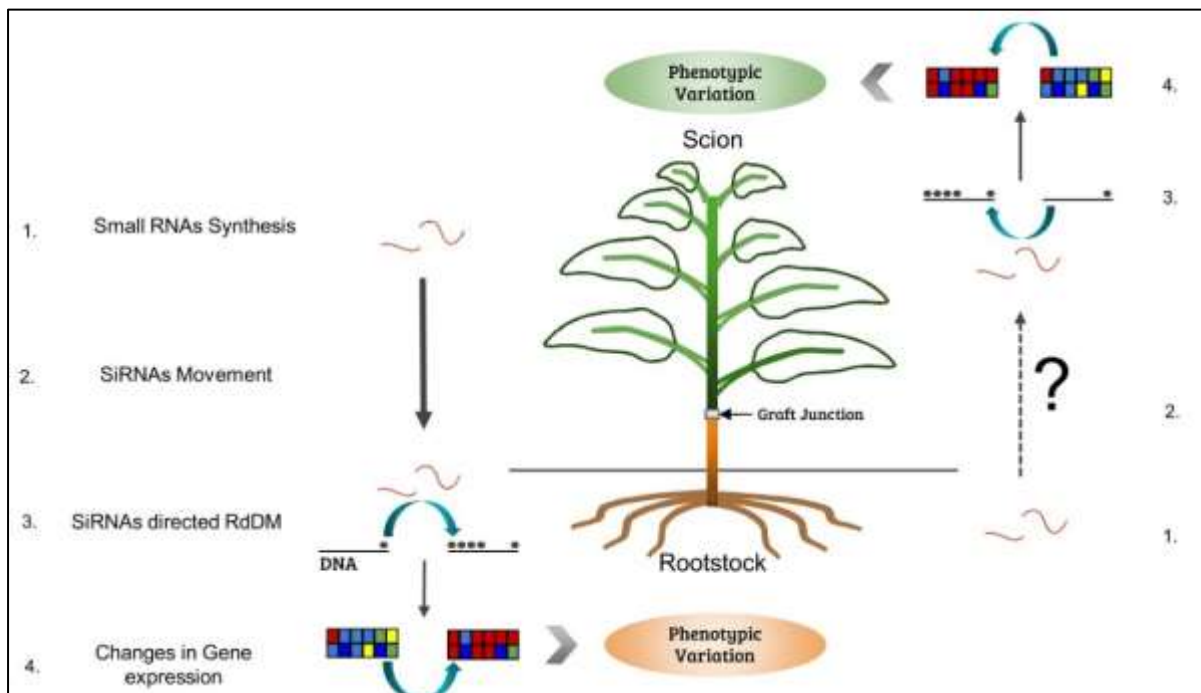


Figura 14: Movimento di small RNA ed il loro effetto sulla pianta (Berger)

### *Riprogrammazione trascrizionale*

L'espressione genica è quindi influenzata dall'innesto in tutte le sue combinazioni, tuttavia, è solo in piante innestate hetero-graft che si possono osservare le maggiori influenze riguardanti l'epigenetica. Tenendo conto del fatto che lo stato fisiologico della pianta (Yamamuro *et al.*, 2016) e lo stato ormonale (Shen *et al.*, 2016) hanno una grande influenza sulle modificazioni epigenetiche, la pratica dell'innesto, che le tocca entrambe, si pone come un ottimo strumento per provocare queste reazioni. Dunque, sia per merito dell'innesto stesso o per l'influenza che la trasmissione di siRNA ha sulla metilazione de novo, le modificazioni epigenetiche sono al centro dell'interesse per comprendere appieno tutti gli aspetti di questa tecnica. Tali modificazioni della metilazione del DNA possono provocare cambiamenti fenotipici stabili ed ereditabili, come quelle osservate nel pomodoro *crn* (Manning, 2006).

## 2.3 Pomodoro

Il pomodoro coltivato (*Solanum lycopersicum*) appartiene alla famiglia delle *Solanaceae* (Angelini *et al.*, 2010) (FIG. 15). Questa contiene molte altre specie conosciute come: tabacco, patate, peperone e melanzana. Il pomodoro deve le sue origini in Sud America ed è stato portato in Europa dai conquistatori spagnoli nel sedicesimo secolo. La pianta presenta un tempo di coltura relativamente breve ed un alto raccolto, che la rendono appetibile economicamente per un vasto impegno. L'importanza alimentare del pomodoro è nota su scala globale, portando il frutto ad essere uno dei più consumati nel pianeta. Dal punto di vista alimentare esso contribuisce ad una dieta sana e ben bilanciata poiché ricco di minerali come ferro e fosforo, vitamine B e C, amminoacidi essenziali, zuccheri e fibre (Naika *et al.*, 2005).



Figura 15: *Solanum lycopersicum*

### 2.3.1 Genoma del pomodoro

All'interno del progetto "The tomato genome" scritto dal Tomato Genome Consortium, una squadra formata da ricercatori di 14 nazioni, ha offerto uno sguardo più da vicino al genoma di pomodoro.

Sono state sequenziate, la variante *Heinz 1706* e la specie *Solanum pimpinellifolium*. La varietà Heinz 1706 è stata sequenziata utilizzando una combinazione del metodo Sanger e di tecniche di nuova generazione (NGS), ottenendo un genoma della grandezza di circa 900 megabasi (Mb) assegnate in 12 cromosomi. Il genoma della specie *S. pimpinellifolium* LA1589 è stato sequenziato ed assemblato utilizzando la piattaforma NGS Illumina, ottenendo la grandezza di 739 Mb (Knapp, 2012). In queste analisi è stato osservato che i cromosomi di pomodoro consistono in porzioni di eterocromatina pericentrica ed euromatina distale, con ripetizioni di questa disposizione dentro ed intorno i centromeri, nei cromomeri e telomeri. Sono state osservate zone con maggior densità di ricombinazione e geni nell'euromatina. Le inserzioni cloroplastiche e zone conservate di geni codificanti miRNA sono distribuite più equamente su tutto il genoma. Inoltre, nel genoma nucleare del pomodoro è stato determinato un contenuto di guanina (G) e citosina (C) di circa il 37%, il quale è tra i più bassi riportati nel mondo vegetale (Messeguer *et al.*, 1991). Grazie a questa analisi è stato possibile ottenere una topografia dei cromosomi (FIG. 16).

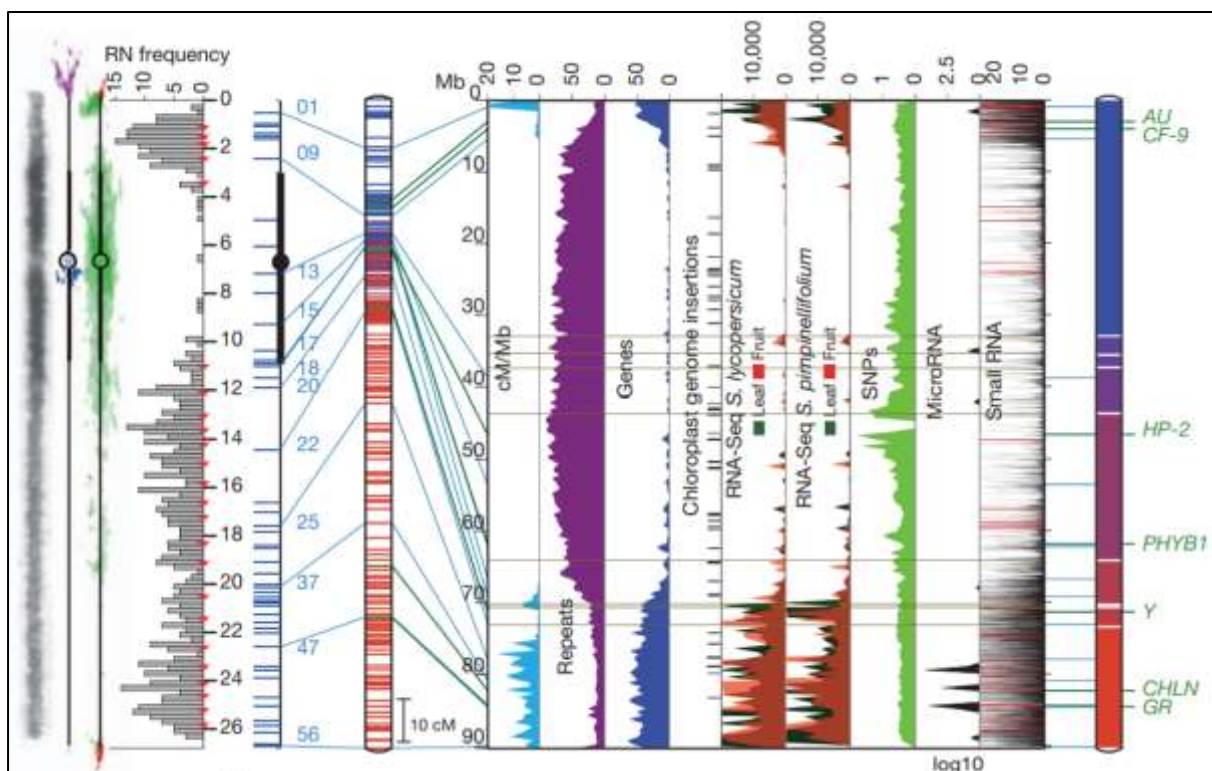


Figura 16: topografia multidimensionale del cromosoma 1 di pomodoro (Knapp & Tomato Genome Consortium, 2012).

### *Metilazione*

La metilazione delle citosine, nel genoma di pomodoro, è stata osservata in media sul 23% dei residui di citosine. Questa varia da tessuto immaturo con una media del 20%, a maturo con una media del 25%. La metilazione del DNA è stata osservata anche in polline e semi. Nei primi si attesta intorno alla media del 22% e nei secondi si osserva il dato più alto con una media del 27%, suggerendo la presenza di metilazione de novo successiva all'impollinazione (Messeguer, 1991). Inoltre, è stato stimato che circa il 55% dei siti con context CG e 85% con context CHG sono metilati (Messeguer, 1991). Questo non è dovuto al caso, ma al fatto che i siti non metilati sono distribuiti in gruppi che ricordano le isole GC presenti in mais e tabacco. Il basso contenuto di G e C e l'alto livello di metilazione può essere spiegato dall'ipotesi di transizione di metil citosina in timina (T). A supporto di ciò sono stati rilevati bassi livelli di G e C in porzioni del DNA soggette a minor controllo, dove queste mutazioni sono più tollerate e dalla mancanza di loci suscettibili alla metilazione in regioni non codificanti (Messeguer, 1991).

#### 2.3.2 L'innesto in pomodoro

Nel contesto della coltivazione del pomodoro, la pratica dell'innesto è comune e diffusa in tutto il mondo in quanto sono presenti numerose malattie che possono infettare la pianta e potenzialmente rovinare il raccolto. Per questi motivi è una pratica consigliata per delle colture in serra. Ad oggi sono disponibili molti portainnesti vigorosi capaci di resistere alle malattie. Inoltre l'innesto è consigliato grazie anche ai vantaggi che comporta: l'aumento di vigore, un frutto più ricco di nutrienti, la garanzia di un raccolto maggiore con più frutti vendibili e più stabili (Djidonou, 2017). La recente introduzione nel mercato di tecnologie e macchinari in grado di superare le difficoltà legate alla tecnica, soprattutto alle ridotte dimensioni della pianta ed il livello di esperienza e laboriosità richiesto, ha reso possibile anche a piccole aziende agricole di applicare la tecnica e produrre piante innestate di qualità (Lee, 2010). Al giorno d'oggi la tecnica utilizzata nell'innesto non è più solo la TAG, ma può variare in base alla preferenza e l'esperienza dell'operatore.

### *Portainnesto*

Negli anni i portainnesti più comunemente utilizzati per il pomodoro sono stati ricavati da melanzana, oggi sono presenti molte varietà di portainnesto per le esigenze più diverse (Lee, 2010). Recentemente

la melanzana è caduta in disuso se non per alcune situazioni specifiche. Molto diffuse per i benefici, ma anche per la ricerca scientifica, l'utilizzo di portainnesti di varietà differenti di pomodoro.

### 2.3.3 Studi sulla metilazione del DNA in pomodoro innestato

Al momento della scrittura di questa tesi, molti studi sono stati effettuati riguardo l'importanza della metilazione del DNA nelle piante. La tecnica dell'innesto ha sempre affascinato agricoltori e ricercatori per le sue proprietà e benefici, dei quali ad oggi non comprendiamo a pieno i complessi meccanismi. Una delle possibili direzioni da esplorare è la metilazione del DNA legata all'innesto e come questo possa provocare modificazioni epigenetiche. Esse portano a cambiamenti fenotipici nella pianta ibrida e nella progenie.

Da uno studio effettuato utilizzando due varianti di pomodoro, la 1221 come portainnesto, con proprietà di resistenza alle malattie e la variante vulnerabile Money maker (MM) come nesto, sono state riportate delle differenze nella metilazione. A seguito dell'innesto è stata effettuata una analisi sulla metilazione del DNA nei context CG e CHG ed i suoi cambiamenti sul tutto il genoma, utilizzando il metodo MSAP (methylation sensitive amplification polymorphism). Da questo sono state osservate delle variazioni nello stato di metilazione. Nello specifico i risultati mostrano una variazione nel grado di metilazione in MM coltivato dove nel context CG è stato rilevato il valore più alto di metilazione al 49.9% ed il più basso al 11.4%. Nella variante 1221 è stato invece rilevato un più alto grado di metilazione in CHG. Nell'ibrido innestato è stata osservata una diminuzione nella metilazione in CG al 31.4% ed un aumento della metilazione in CHG al 28.6%. Questi dati sono stati rivelati in concordanza con le medie delle due varietà coltivate separatamente, dimostrando quindi una variazione statisticamente significativa nella metilazione del DNA (Xu *et al.*, 2016). In un altro studio, è stato osservato che l'innesto effettuato tra individui della stessa specie può causare grandi alterazioni nella metilazione del DNA in loci specifici nel nesto. Inoltre, sono state rilevate delle alterazioni a livello dei context CG e CHG, a livello dei siti 5'-CCGG analizzabili con la tecnica MSAP, seppur mantenendo il livello di metilazione generale intatto. Gli studi citati sostengono la tesi dell'esistenza di modifiche epigenetiche correlate direttamente con l'innesto (Wu *et al.*, 2013). La metilazione del DNA rappresenta un oggetto di studio attuale per il suo coinvolgimento nell'adattamento alle condizioni avverse e nelle modificazioni epigenetiche. Esse, vista la loro importanza biologica nel silenziamento dell'espressione genica, possono produrre modificazioni funzionali del fenotipo, potenzialmente misurabili ed utilizzabili per migliorare i raccolti e la qualità del frutto.

### 3. SCOPO DELLA TESI

L'obiettivo di questa analisi è stato identificare gli effetti dell'innesto sull'epigenoma di pomodoro, precisamente i cambiamenti di metilazione del DNA e le conseguenze sulla qualità del frutto. Le differenze di metilazione sono state osservate in sei innesti, effettuati con un'unica varietà di pomodoro e cinque tipi di portainnesto (I12, I14, I15, O3 e OPT). Un ulteriore innesto è stato prodotto con entrambe le parti della stessa varietà di pomodoro. Per poter identificare tali cambiamenti è stata usata prima la metodica MCSeEd, la quale sfrutta enzimi di restrizione sensibili alla metilazione, successivamente diversi strumenti bioinformatici volti all'identificazione di differenze nella metilazione relativa tra i confronti. Argomento di questa tesi è dunque la differente metilazione tra i confronti e le conseguenze nell'espressione genica, che può portare possibili conseguenze sul frutto di pomodoro.



#### 4. MATERIALI E METODI

## 4.1 Materiale vegetale

I materiali vegetali che sono stati utilizzati per questa analisi sono dei frutti di pomodoro prelevati dal nesto di varietà *Creativo* e cresciuti in serra negli stabilimenti di una azienda siciliana. Essi provengono da sei diversi portainnesti ed un controllo non innestato. Il progetto sperimentale prevede la produzione di piante innestate, utilizzando come nesto la stessa varietà di pomodoro e come portainnesto degli ibridi proprietari dell'azienda: I12, I14, I15, O3, OPT e Self. Sono state prodotte tre repliche biologiche di ciascun campione, così da avere una statistica più stabile ed evitare la produzione di falsi positivi. I portainnesti I12, I14, I15 sono delle varietà sperimentali in fase di sviluppo per fini commerciali, OPT e O3 sono due varianti già in commercio. Da ognuno di essi è stato raccolto il frutto per effettuare l'estrazione del DNA.

## 4.2 Protocollo di laboratorio

### 4.2.1 Estrazione del DNA e preparazione degli adattatori

Il primo passaggio è stato quello di raccogliere i campioni e successivamente congelarli nella camera dell'autoclave a  $-80^{\circ}\text{C}$ . Sezioni di tessuto di frutto sono state prelevate e grindate con mortaio e pestello. Il tessuto grindato è stato utilizzato per l'estrazione del DNA, impiegando un apposito kit di estrazione e purificazione. Alla fine del processo di estrazione, il DNA è stato eluito in  $150\ \mu\text{L}$  di acqua ultra-pura. La quantificazione del DNA è avvenuta con NanoDrop utilizzando  $2\ \mu\text{L}$  di DNA. Al fine di valutare la buona qualità del DNA genomico estratto e valutare la presenza di possibili degradazioni e/o contaminazioni,  $5\ \mu\text{L}$  di ogni confronto sono stati caricati su gel di agarosio all'  $1.2\%$  e visualizzati al transilluminatore. Le librerie sono state prodotte utilizzando degli adattatori con *barcode* a doppio filamento che andranno a legarsi ai frammenti di DNA. Per prima cosa è stata effettuata la diluizione degli adattatori liofilizzati a  $100\ \mu\text{M}$  con acqua ultra-pura. Successivamente sono stati miscelati gli oligopeptidi (FIG 17) a singolo filamento con un rapporto 1:1, per poi farli appaiare attraverso il termociclatore seguendo il profilo di temperatura riportato in Tab. 2.

Figura 17: oligopeptidi a singolo filamento

Nome	Adp Info	Sequenza
P7	ACI_TOP_7	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCCACCGT
	ACI_BOT_7	[Phos]CGACGGTGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P8	ACI_TOP_8	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGCGA
	ACI_BOT_8	[Phos]CGTCGCACGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P9	ACI_TOP_9	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAACAAT
	ACI_BOT_9	[Phos]CGATTGTTTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P10	ECO_TOP_10	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCCGATATTGCA
	ECO_BOT_10	[Phos]ATATCGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P11	ECO_TOP_11	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGACACGTTGCA
	ECO_BOT_11	[Phos]ACGTGTTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P12	ECO_TOP_12	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCGGTTCTGCA
	ECO_BOT_12	[Phos]GAACCGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1	PST_TOP_1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTACAATGCA
	PST_BOT_1	[Phos]TTGTAGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P2	PST_TOP_2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCATCATGCA
	PST_BOT_2	[Phos]TGATGCTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P3	PST_TOP_3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCCATCCATGCA
	PST_BOT_3	[Phos]TGGATGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

Tabella 2: preparazione degli adattatori

Denaturazione	Ramp	Annealing	Ramp	Annealing	Hold
95 °C x 3'	0,5 °C ogni 10"	70 °C x 10'	0,5 °C ogni 10"	" 20 °C x 20'	10 4 °C

La verifica è stata svolta facendo correre gli adattatori appaiati con un gel di agarosio ad alta concentrazione (4% di agarosio)

#### 4.2.2 Preparazione delle librerie

##### 4.2.2.1 Restrizione e ligazione

Il DNA è stato digerito da enzimi sensibili e insensibili alla metilazione. I frammenti prodotti dalla digestione sono stati ligati dall'adattatore barcode nel sito tagliato dall'enzima sensibile alla metilazione e da un adattatore comune nel sito tagliato dall'enzima insensibile. Per la reazione di restrizione/ligazione, è stato necessario preparare un buffer appropriato. Sono stati aggiunti 500 µL di buffer CutSmart 10X a 25 µL di DTT 1M e 475 µL di acqua ultra. A questo punto il DNA genomico è stato digerito con un enzima sensibile alla metilazione (Acil, PstI, o EcoT22I) e con l'enzima insensibile

alla metilazione *MseI* (Tab. 3). Ad esempio, come mostrato nella Tab. 4, per *Acil*, a 200 ng di DNA genomico sono state aggiunte 4 U di enzima *Acil*, 4 U di enzima *MseI*, 2µM di adattatore barcode, 2µM di adattatore Y comune, 1 U di DNA ligasi T4 (ThermoFisher), 0.2 mM di ATP e 1x buffer RL (Tab. 5) per un volume finale di 50 µL. Infine, sono stati incubati i confronti in un termociclatore per 4 ore a 37°C per assicurare la reazione di restrizione e di ligazione.

*Tabella 3: enzimi di restrizione utilizzati*

Enzima	Sito di riconoscimento	Sito di taglio	Sensibile metilazione	C metilata	Context metilazione
<b><i>Acil</i></b>	CCGG/GCGG	C'CGC/G/CGG	Si	mCGC/G 5mCGG	CG
<b><i>PstI</i></b>	CTGCAG	CTGCA'G	Si	CTG 5mCAG	CHG
<b><i>EcoT22I</i></b>	ATGCAT	ATGCA'T	Si	ATG 5mCAT	CHH
<b><i>MseI</i></b>	TTAA	T'TAA	No	Non sensibile	“-“

Tabella 4: mix di reazione per la produzione delle library

Enzimi	<i>MseI/Aci</i> mix (μL)	<i>MseI/Eco</i> mix (μL)	<i>MseI/PstI</i> mix (μL)
Restr. Ligation 10X	10	10	10
<i>MseI</i> 10U/μL (4U)	0,4	0,4	0,4
<i>AciI</i> 10U/μL (4U)	0,4	X	X
<i>EcoT22I</i> 10U/μL (4U)	X	0,4	X
<i>PstI</i> 20U/μL (4U)	X	X	0,3
<i>Mse</i> Adp (50μM)	2	2	2
T4 DNA ligase 5U/μL	0,2	0,2	0,2
ATP 10 mM	1	1	1
H <sub>2</sub> O	12	12	12,4
tot MIX	26	26	26
Barcode Adp	2	2	2
NA (200ng)	22	22	22
Tot MIX RL	50	50	50
Temp reaction	37°C 4hrs	37°C 4hrs	37°C 4hrs
	4°C	4°C	4°C

Tabella 5: preparazione del buffer RL

Reagente	Diluizione	Volume (μL)	Conc. finale
CutSmart Buffer (B7204S)	10x	500	5x
DTT	1M	25	25mM
H <sub>2</sub> O	-	475	-
Volume finale	-	1000	-

#### 4.2.2.2 Precipitazione tramite PEG delle librerie genomiche

I frammenti ristretti-ligati (RL) dei diversi confronti sono stati uniti per generare un pool di individui da amplificare con un indice comune (Tab. 6)

Tabella 6: individui con indice comune per amplificazione

	Acil	EcoT221	PstI	Confronti
1	1A_P7	1A_P10	1A_P1	Self
2	1B_P8	1B_P11	1B_P2	
3	1C_P9	1C_P12	1C_P3	
4	2A_P7	2A_P10	2A_P1	O3
5	2B_P8	2B_P11	2B_P2	
6	2C_P9	2C_P12	2C_P3	
7	3A_P7	3A_P10	3A_P1	OPT
8	3B_P8	3B_P11	3B_P2	
9	3C_P9	3C_P12	3C_P3	
10	4A_P7	4A_P10	4A_P1	Ungrafted
11	4B_P8	4B_P11	4B_P2	
12	4C_P9	4C_P12	4C_P3	
13	A_P7	A_P10	A_P1	I12
14	B_P8	B_P11	B_P2	
15	C_P9	C_P12	C_P3	
16	A_P7	A_P10	A_P1	I14
17	B_P8	B_P11	B_P2	
18	C_P9	C_P12	C_P3	
19	A_P7	A_P10	A_P1	I15
20	B_P8	B_P11	B_P2	
21	C_P9	C_P12	C_P3	

Successivamente è stata preparata una soluzione di PEG8000/30 mM  $MgCl_2$  fresco al 30%, aggiungendo 300  $\mu L$  di  $MgCl_2$  a 6 mL di acqua ultra-pura in una provetta Falcon. Quindi sono stati aggiunti 3 g di PEG8000. Al fine di raggiungere il volume di 10 mL è stata aggiunta acqua ultra-pura sterile. Divenuta limpida la soluzione è stata filtrarla con un filtro da 0,2  $\mu m$ . Dopo aver preparato la soluzione si è dispensato 0,5x di soluzione PEG per ciascun confronto e pipettato il contenuto su e giù almeno 20 volte. Successivamente i campioni sono stati centrifugati a 10.000g per 20 minuti a temperatura ambiente e per rimuovere rapidamente la fase liquida. Alla fine il pellet di DNA è stato risospeso in 30  $\mu L$  di acqua ultra-pura.

#### 4.2.2.3 *Purificazione con beads Agilent AMPure (1,1X)*

Le beads Agilent AMPure sono servite per purificare la nostra soluzione di tutti i componenti non desiderati e lasciare solo il DNA di interesse sul quale hanno agito gli enzimi di restrizione. All'inizio è stato utilizzato il vortex sui campioni per qualche secondo e poi sono stati aggiunti 33  $\mu L$  di soluzione AMPure a 30  $\mu L$  di campione. Dopodiché la miscela è stata pipettata almeno una decina di volte e i campioni sono stati incubati per cinque minuti a temperatura ambiente. In questo intervallo di tempo, il DNA ha aderito alle beads. Poi i campioni sono stati collocati su piastra magnetica per almeno cinque minuti, in modo da separare le beads dalla soluzione ed eliminare il surnatante in eccesso. A questo punto sono stati pipettati 100  $\mu L$  di etanolo al 70% in ogni provetta e i campioni incubati per 30 secondi a temperatura ambiente. Il passaggio è stato ripetuto due. Dopo il secondo lavaggio, le beads sono state essiccate nella piastra magnetica per 8-10 minuti, fino a quando tutto l'etanolo non è evaporato. I campioni sono stati rimossi dalla piastra magnetica e sono stati aggiunti 20  $\mu L$  di acqua ultra-pura, risospesando le beads in soluzione. La soluzione è stata lasciata ad incubare per altri cinque minuti a temperatura ambiente. Infine, i campioni sono stati collocati sulla piastra magnetica, fino alla separazione della soluzione acquosa contenente DNA dalle beads, andando poi a prelevare il surnatante trasferendolo in una nuova provetta.

#### 4.2.2.4 *Corsa su gel e selezione delle dimensioni da 250bp a 600bp*

Considerando che per il sequenziamento Illumina è necessario utilizzare frammenti di DNA con una dimensione compresa tra 250 e 600 bp, è stato opportuno effettuare una selezione dei frammenti prodotti. Per tal fine, i ristretti-ligati purificati sono stati corsi su di un gel di agarosio all'1,2% per 30

minuti con un voltaggio a 50V. Al termine della corsa elettroforetica, sono state prelevate dal gel solo le sezioni contenenti frammenti di DNA con dimensione compresa tra 250bp e 600 bp. Quindi è stata eseguita la purificazione del gel con il kit di estrazione QIAquick seguendo le istruzioni del produttore, utilizzando il tampone di eluizione riscaldato a 50 °C anziché a temperatura ambiente. Per l'eluizione, il tampone è stato pipettato al centro della membrana, lasciando incubare per un minuto a temperatura ambiente, per poi centrifugare a 16.000 g per un minuto.

#### *4.2.2.5 Purificazione con beads Agilent AMPure (0.8X) delle librerie*

Durante la seconda purificazione con le beads Agilent AMPure, preventivamente vortexate per qualche secondo, sono stati aggiunti 38.4 µL di soluzione AMPure a 50 µL di confronto. Poi pipettata la miscela almeno una decina di volte e i campioni sono stati incubati per cinque minuti a temperatura ambiente. In questo intervallo di tempo, il DNA ha aderito alle beads. Successivamente, sono stati collocati i campioni su piastra magnetica per cinque minuti, in modo da separare le beads dalla soluzione ed eliminare il surnatante in eccesso. A questo punto sono stati pipettati 100 µL di etanolo al 70% in ogni provetta e lasciati incubare per 30 secondi a temperatura ambiente. Il passaggio è stato ripetuto due. Dopo il secondo lavaggio, le beads sono state essiccate nella piastra magnetica per 8-10 minuti, fino a quando tutto l'etanolo non è evaporato. I campioni sono stati rimossi dalla piastra magnetica e risospesi aggiungendo 30 µL di acqua ultra-pura. La soluzione quindi è stata messa ad incubare per altri cinque minuti a temperatura ambiente. Infine, i campioni sono stati posti sulla piastra magnetica, attenendo la separazione della soluzione acquosa contenente DNA dalle beads. Come in precedenza è stato prelevato il surnatante e trasferito in una nuova provetta.

#### *4.2.2.6 Quantificazione al Qubit delle librerie*

La quantificazione con Qubit, in ogni campione, ha richiesto 199 µL di tampone HS più 1 µL di reagente HS. Due standard sono stati preparati per costruire la retta di taratura. In ognuno dei due standard, sono stati aliquotati 190 µL di soluzione HS e 10 µL di standard (rispettivamente a concentrazione 0ng/µL e 10ng/µL). Per la quantificazione delle librerie, invece, sono stati necessari 198 µL di soluzione HS e 2 µL di DNA purificato. Ciascuna provetta è stata vortexata per circa 2-3 secondi e lasciata incubare a temperatura ambiente per 2 minuti. Il Qubit è dunque stato attivato, procedendo al



caricamento dei due standard per la produzione della retta di taratura. Infine sono stati caricati in successione i vari confronti per la quantificazione.

#### 4.2.2.7 *PCR enrichment delle librerie*

Questo passaggio è stato effettuato per arricchire i frammenti dei campioni che sono stati legati ad entrambi gli adattatori, sia all'adattatore ad Y comune e sia all'adattatore con barcode. Infatti, solo se entrambi gli adattatori sono legati ai frammenti, la PCR di arricchimento produrrà la libreria arricchita. Tramite questo procedimento è stata aggiunta una sequenza, chiamata index, a livello dell'adattatore Mse. La PCR è stata effettuata seguendo il protocollo Phusion Taq DNA polimerasi e dei oligo specifici (Tab. 7) 15 ng di DNA sono stati aggiunti alla mix di reazione riportata in Tab. 8. Sono stati effettuati 12 cicli di PCR seguendo il profilo riportato nella Tab. 9. Per valutare l'avvenuto arricchimento delle librerie, 5 µL di ciascuna libreria è stata caricata su gel di agarosio all'1,5%.

*Tabella 7: Oligo utilizzati nella PCR enrichment*

Primer	Sequenza
PCR1_ddRAD	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACG
PCR2_idx_2	CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGC
PCR2_idx_4	CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGC
PCR2_idx_5	CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGC
PCR2_idx_6	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGC
PCR2_idx_12	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGC
PCR2_idx_14	CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGC
PCR2_idx_15	CAAGCAGAAGACGGCATACGAGATTGACATGTGACTGGAGTTCAGACGTGTGC

Tabella 8: Protocollo PCR del kit

Componenti	Per 50 µl di reazione
H2O	15,4 µl
5x Buffer	10 µl
5 µM dNTP	2 µl
10 µM PCR1ddRAD	1 µl
10 µM PCRIidx2/idx4	1 µl
Phusion DN polimerasi	0,6 µl
DNA template	15 ng / µl

Tabella 9: Profilo termico dei 12 cicli di PCR

Passaggi	Temperature(°C)	Tempo
Denaturazione iniziale	98°C	2'
Denaturazione	38°C	20''
Appaiamento dei primers	56°C	30''
Allungamento	72°C	30''
Allunamento finale	72°C	5'
Mantenimento	4°	-

#### 4.2.2.8 Purificazione con beads Agilent AMPure (1X) delle librerie arricchite

L'ultima purificazione con le beads Agilent AMPure, preventivamente vortexate per qualche secondo, ha previsto l'aggiunta di 45 µL di soluzione AMPure a 45 µL di confronto. Dopodiché la miscela è stata pipettata almeno una decina di volte e i campioni incubati per cinque minuti a temperatura ambiente. In questo intervallo di tempo, il DNA ha aderito alle beads. In seguito, i campioni sono stati posti sulla piastra magnetica per cinque minuti, in modo da separare le beads dalla soluzione e poi eliminare il

surnatante in eccesso. A questo punto 100  $\mu$ L di etanolo al 70% sono stati aggiunti in ogni provetta e 30 secondi a temperatura ambiente sono stati incubati. Il passaggio è stato ripetuto due.

Dopo il secondo lavaggio, le beads sono state lasciate ad essiccare nella piastra magnetica per 8-10 minuti, fino a quando tutto l'etanolo non è evaporato. I campioni sono stati prelevati dalla piastra magnetica e sospesi con 30  $\mu$ L di acqua ultra-pura. La soluzione è stata lasciata ad incubare per cinque minuti a temperatura ambiente. Per ultimo, i campioni sono stati posti sulla piastra magnetica, attendendo la separazione della soluzione acquosa contenente DNA dalle beads, prelevando il surnatante e traferendolo in una nuova provetta. Un'ultima quantificazione con il Qubit delle librerie arricchite è stata necessaria per spedire i campioni al sequenziamento Illumina HiSeq 150bp paired-end.

### 4.3 Analisi bioinformatica: Tecnica MCSeEd

I dati di partenza per questa analisi derivano dall'utilizzo del metodo MCSeEd sul frutto del pomodoro. Questa tecnica, oltre a identificare le differenze di metilazione permette di identificare eventuali SNP nelle librerie generate dall'analisi dei confronti. Grazie a MCSeEd è possibile effettuare l'analisi utilizzando due approcci differenti: un approccio genome-independent ed un approccio genome-dependent. Per questa analisi è stato utilizzato un approccio genome-dependent, ossia i loci sequenziati sono stati mappati su di un genoma di riferimento di pomodoro (ITAG4.0; Sol Genomics Network).

#### 4.3.1 Preparazione del file counts

L'approccio genome-dependent ha previsto il *demultiplexing* delle reads prodotte dal sequenziamento utilizzando lo strumento `process_radtags`. Questo ha identificato e assegnato le reads a ciascun individuo in base alle sequenze del barcode ad esso legato. Quindi le sequenze di ciascun individuo sono state mappate sul genoma di riferimento con l'algoritmo `bwa mem`. Dopodiché sono stati prodotti con Samtools dei file bam, ordinati e indicizzati, contenenti reads mappate in modo univoco. I suddetti bam sono stati uniti in un'unica matrice di conteggio in cui le colonne rappresentano le librerie e le righe la posizione nel genoma. Ciò è stato importante per memorizzare le posizioni genomiche sequenziate e per registrare il numero di reads mappate in modo univoco in ciascuna posizione genomica. La suite bedtools è stata utilizzata per comprimere e ordinare le coordinate ridondanti e unire gli intervalli sovrapposti. Le informazioni ottenute sono state convertite

in un file GFF che può essere utilizzato come file di input in featureCounts insieme ai file bam descritti in precedenza, ottenendo un file di testo contenente le nostre reads, chiamato *count\_genome\_dependent* (Tab. 10).

Tabella 10: Esempio del contenuto del file *count\_genome\_dependent* da I12 context CG.

Geneid	Chr	Start	End	Strand	Length	ungr_aci_P 7	ungr_aci_P 8	ungr_aci_P 9	i12_aci_P7	i12_aci_P8	i12_aci_P9
SL4.0:9649-9792	Ch0	9649	9792	+	144	1	0	0	2	0	2
SL4.0:12223-12335	Ch0	12223	12335	+	1	0	0	1	4	1	1
SL4.0:14811-14938	Ch0	14811	14938	+	128	0	0	1	0	0	0
SL4.0:17877-18003	Ch0	17877	18003	+	127	1	0	0	0	0	0
SL4.0:23932-24075	Ch0	23932	24075	-	144	2	0	0	1	0	0
SL4.0:26758-26819	Ch0	26758	26819	-	62	17	4	15	12	7	5

#### 4.3.2 Da counts al formato Methykit

Il file *counts\_genome\_dependent* generato precedentemente contiene le seguenti informazioni: numero identificativo del locus, cromosoma dove le read sono state mappate, lunghezza e i count delle read. La mia analisi è partita dal file di testo *counts\_genome\_dependent* dal quale è stato ottenuto un file compatibile con il pacchetto R Methykit. Un pacchetto del programma R contenuto in Bioconductor, con il quale è possibile analizzare le posizioni differenzialmente metilate e ottenerne la posizione nel genoma di studio. I dati contenuti nel file sono stati normalizzati utilizzando un metodo RPKM-like considerando la numerosità delle reads per ogni libreria creata. Il fattore di correzione da utilizzare è stato ottenuto risolvendo l'equazione:  $10^6 / X_n$  ( $X_n$  numero di reads totali per libreria,  $n$  numero delle librerie) e successivamente moltiplicato per il raw data di ogni locus così da ricavare il valore normalizzato. La formula utilizzata è stata:

$$RPKM_{(r)} = \frac{10^6}{X_n} \times C_r$$

$C_r$ : conteggio iniziale di reads di ogni locus (raw data)

$X_n$ : numero di reads totali per libreria

$n$ : numero delle librerie

$10^6$ : fattore di correzione

Questo processo è avvenuto grazie ad uno script nel linguaggio Perl. Questo linguaggio è molto utilizzato nella bioinformatica e ricco di funzioni utili per la gestione dei dati ed analisi statistica, eseguito da terminale. A questo punto è stato effettuato un calcolo del livello di metilazione relativo per locus. Alla fine del processo sono stati ottenuti sei file in totale, di questi tre da trattati e tre dai controlli. Tali sono stati salvati in formato di testo pronti per analisi con MethyKit come nella Tab. 11.

*Tabella 11: esempio del contenuto del file methylKit. Le colonne indicano in ordine: ID identificativo, cromosoma, posizione, strand, coverage, livello relativo di metilazione e demetilazione*

Geneid	Chr	posizione	strand	coverage	metilazione	demetilazione
SL4.0:242394	ch00	242394	F	24	29.17	70.83
SL4.0:331049	ch00	331049	F	29	44.83	55.17
SL4.0:383342	ch00	383342	F	7	42.86	57.14
SL4.0:385134	ch00	385134	F	12	33.33	66.67
SL4.0:385713	ch00	385713	R	16	50.00	50.00
SL4.0:389779	ch00	389779	R	8	50.00	50.00

#### 4.3.3 Identificazione DMPs con MethyKit

##### *Software RStudio*

RStudio è un software gratuito, basato sul linguaggio R, specializzato nell'analisi statistica di dati, ricco di funzioni ed intuitivo. Il programma si avvale di pacchetti specifici, sviluppati dalla community, che ne ampliano le funzionalità rendendolo un programma completo. RStudio è molto utilizzato in ambito bioinformatico grazie alla possibilità di gestire dati e svolgere analisi avanzate. I pacchetti riguardanti la bioinformatica sono gestiti dal repository BioConductor, la quale offre un'ottima quantità di strumenti per l'analisi. In questa analisi svolge un ruolo importante grazie al tool MethyKit ed alla possibilità di creare grafici (ad esempio heatmap, hierarchical cluster, boxplot) e compiere analisi statistiche (T-test e analisi delle componenti principali).

### *Identificazione delle DMPs in ambiente RStudio*

Passando in ambiente R, il pacchetto MethyKit ci ha permesso di identificare le DMPs, cioè le posizioni differenzialmente metilate. Sono stati caricati i file del passaggio precedente in R, creando l'oggetto `myobj` e su quest'ultimo si è utilizzato tre comandi specifici. In ordine: `filterByCoverage` per filtrare le read sotto il un preciso valore di coverage, `calculateDiffMeth` per poter calcolare le differenze di metilazione tra i gruppi innestato e non innestato, settando la soglia del q-value a 0.05. In ultimo, il comando `getMethylDiff` che permette di valutare la differenza di metilazione maggiore del 5% e secondo il livello di significatività espresso dal q-value minore di 0.05.

#### 4.3.4 Identificazione DMRs

Dal file contenente le informazioni sulle DMP è stata possibile l'identificazione delle windows, ossia finestre di paia di basi all'interno delle quali possono essere presenti regioni differenzialmente metilate. La determinazione di questo intervallo è stata necessaria per un corretto calcolo delle DMR lungo tutto il genoma. Utilizzando il comando Unix *for loop* in combinazione con uno script *Perl* è stato possibile identificare le windows partendo da un valore cluster di 100bp ed aumentando progressivamente fino a 2000bp. I risultati sono stati esportati in un file Excel dove è stato considerato il rapporto numero di cluster/window più alto per determinare la migliore window, in ogni context, per l'identificazione delle DMR nei confronti (FIG. 17)

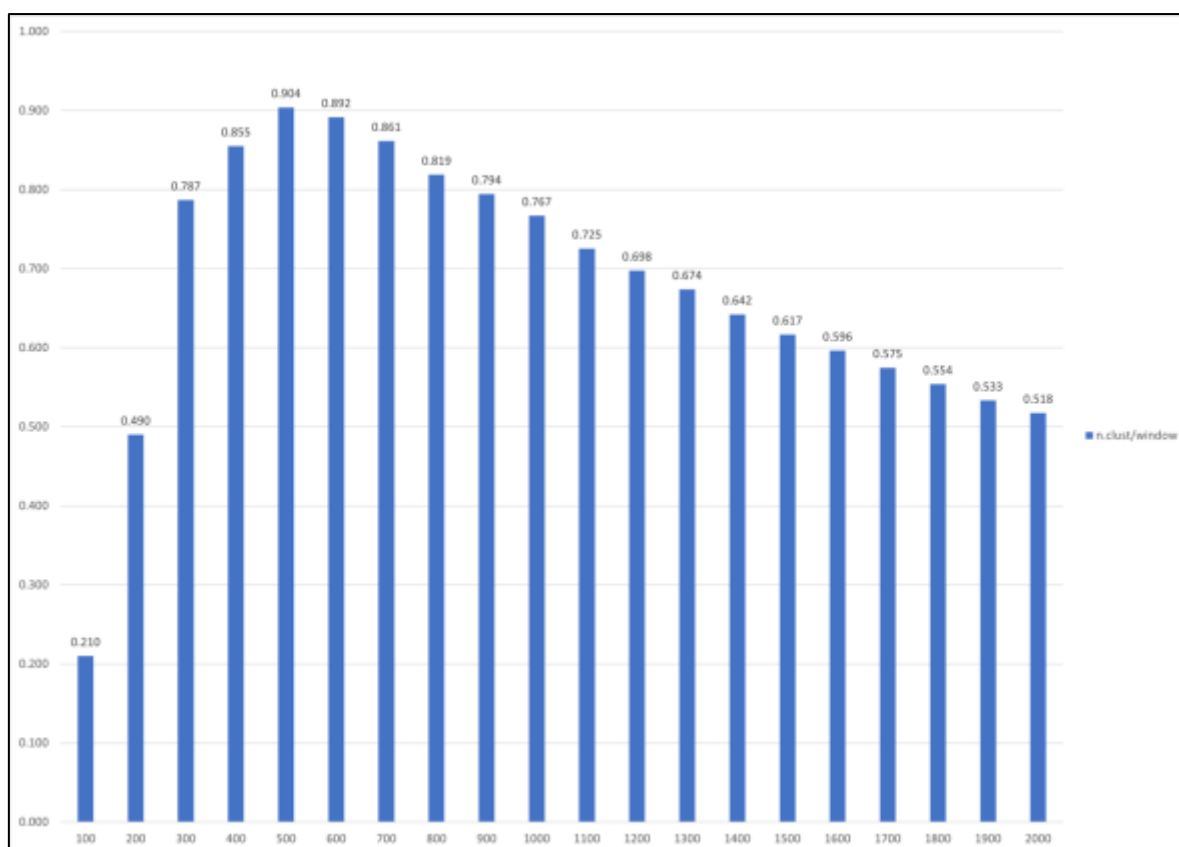


Figura 17: Rappresentazione grafica del rapporto  $n. \text{ clust}/\text{window}$  calcolato per l'identificazione della finestra migliore context-specifica di 112 context CG, in questo caso 500bp

Selezionata la window migliore è stato utilizzato un secondo script Perl, nel quale l'impostazione da modificare è stata solo l'ampiezza della window in continuità con quella selezionata precedentemente, per identificare le coordinate delle DMRs e il numero di DMPs al loro interno.

#### 4.3.5 Identificazione delle DMGs

Per mezzo delle DMR identificate precedentemente è stato possibile individuare i geni differenzialmente metilati. Questa analisi è stata svolta utilizzando un toolset per UNIX chiamato Bedtools. Questo contiene un'ampia gamma di strumenti utili per l'analisi genomica. Sebbene ogni singolo strumento sia progettato per svolgere un'attività relativamente semplice, è possibile condurre analisi piuttosto sofisticate combinando più operazioni (R. Quinlan & Kindlon, 2019). Da questo toolset sono stati utilizzati due tools specifici: intersect e window. Intersect consente di esaminare le sovrapposizioni tra due insiemi di informazioni genomiche. Il primo insieme è relativo alle coordinate

delle regioni differenzialmente metilate ed il loro differenziale di metilazione, prodotto grazie alla formattazione del precedente file. Il secondo insieme di informazioni genomiche contiene le annotazioni ufficiali di pomodoro (ITAG 4). Da questo comando, si ottengono inoltre le informazioni su dove le DMR intersecano precisamente sull'annotazione ufficiale. Con il tool Window è possibile analizzare uno specifico intervallo di paia di basi upstream e downstream dai geni annotati e pertanto ottenere informazioni riguardanti la metilazione nelle regioni regolative a monte o a valle dei geni. Dai geni differenzialmente metilati (DMGs) ottenuti grazie ai due tools, è possibile dividere i geni in ipermetilati e demetilati per effettuare le successive analisi.

## 4.4 Analisi Gene ontology con PANTHER - Gene List Analysis

L'ultima operazione è stata quella di analisi delle Gene Ontology (GO). L'analisi è svolta sul grande database GO. Questa è la fonte di informazioni più completa al mondo sulla funzione dei geni e dei prodotti genici (proteine e RNA non codificanti). GO è stato creato ed è gestito da un ampio consorzio di biologi esperti, data scientists e ingegneri informatici, con il prezioso contributo di scienziati che collaborano in aree specifiche della biologia. Le informazioni contenute in questo progetto riguardano i geni di molti organismi diversi di tutti i regni della vita, così come anche di virus (Nucleic Acids Research, 2021).

### 4.4.1 Panther (Protein analysis through evolutionary relationships)

I geni differenzialmente metilati ottenuti dai passaggi precedenti sono stati formattati in liste di geni da inserire nel motore di ricerca di PANTHER. Un sistema di classificazione che è stato progettato per classificare le proteine (e i loro geni), facilitandone la loro analisi (Thomas, 2021). La lista va incontro ad analisi secondo il test di over-representation, il quale tramite la statistica binomiale andrà a stimare il livello di arricchimento per processi biologici (BP), funzioni molecolari (MF), compartimenti cellulari (CC) e classi di proteine PANTHER. Inoltre, è stato possibile visualizzare i risultati secondo il *false discovery rate* (FDR), cioè un metodo per concettualizzare il tasso di errori di tipo I nei test di ipotesi nulle quando si effettuano confronti multipli. I dati ottenuti sono stati poi trascritti in un file Excel, divisi per categoria e selezionati in base alla loro pertinenza con lo scopo dell'analisi.



## 5. RISULTATI

Grazie all'utilizzo del metodo MCSeEd e all'analisi genome-dependent, che utilizza un genoma di riferimento ed attraverso differenti tool bioinformatici, è stato possibile ricavare dai nostri dati le informazioni relative alle differenze di metilazione nei confronti. Le informazioni sono state utilizzate per l'identificazione di geni di interesse riguardanti la pianta ed il frutto del pomodoro. Le tappe per l'interpretazione dei risultati sono state l'individuazione delle posizioni differenzialmente metilate (DMPs) e le regioni differenzialmente metilate (DMRs), con relativi grafici per visualizzare i risultati. Infine, è stata effettuata un'ulteriore analisi gene ontology (GO) per determinare la classe di appartenenza dei geni d'interesse e le loro caratteristiche, nonché il loro ruolo nel genoma di pomodoro. Tutto questo è stato valutato per i confronti di I12, I14, I15, O3, OPT e Self. Per ciascun confronto si valutano tre context di metilazione tipici delle piante: CG, CHG e CHH.

## 5.1 Identificazione delle DMPs

In ambiente R, grazie all'utilizzo di MethyKit, sono state identificate le posizioni differenzialmente metilate (DMPs). Queste rappresentano i loci che possiedono una frequenza di metilazione significativamente alterata tra controllo e trattato. I dati sono stati organizzati per il confronto e divisi per i tre context di metilazione. Inoltre, è stata effettuata un'altra distinzione tra le DMPs identificate, prendendo in considerazione il loro livello di metilazione relativo. Questo ha portato alla distinzione tra le DMPs che presentano una ipermetilazione (Hyper) oppure ipometilazione (Hypo) rispetto al controllo. Tutti i dati, compreso il numero totale delle DMPs, sono stati riportati nella Tab. 12 *a seguire*.

Tabella 12: DMPs identificate con methylKit in R.

Confronti	CG	CHG	CHH
<b>I12 Hyper</b>	8100	7373	15992
<b>I12 Hypo</b>	6404	7420	7364
<b>I12 Total</b>	14504	14793	23356
<b>I14 Hyper</b>	4928	7685	8891
<b>I14 Hypo</b>	5482	5848	9553
<b>I14 Total</b>	10464	13534	18444
<b>I15 Hyper</b>	8232	7792	14507
<b>I15 Hypo</b>	5547	7311	7036
<b>I15 Total</b>	13779	15103	21543
<b>O3 Hyper</b>	6844	6983	12659
<b>O3 Hypo</b>	6306	7707	6778
<b>O3 Total</b>	13150	14690	19437
<b>OPT Hyper</b>	5416	4763	7964
<b>OPT Hypo</b>	6788	6713	3781
<b>OPT Total</b>	12204	11476	11745
<b>SELF Hyper</b>	4268	5725	7422
<b>SELF Hypo</b>	5330	6630	2919
<b>SELF Total</b>	9598	12355	10341

Da questi dati sono stati costruiti dei grafici (FIG. 18) per spiegare immediatamente quanto presente nella Tab. 12.

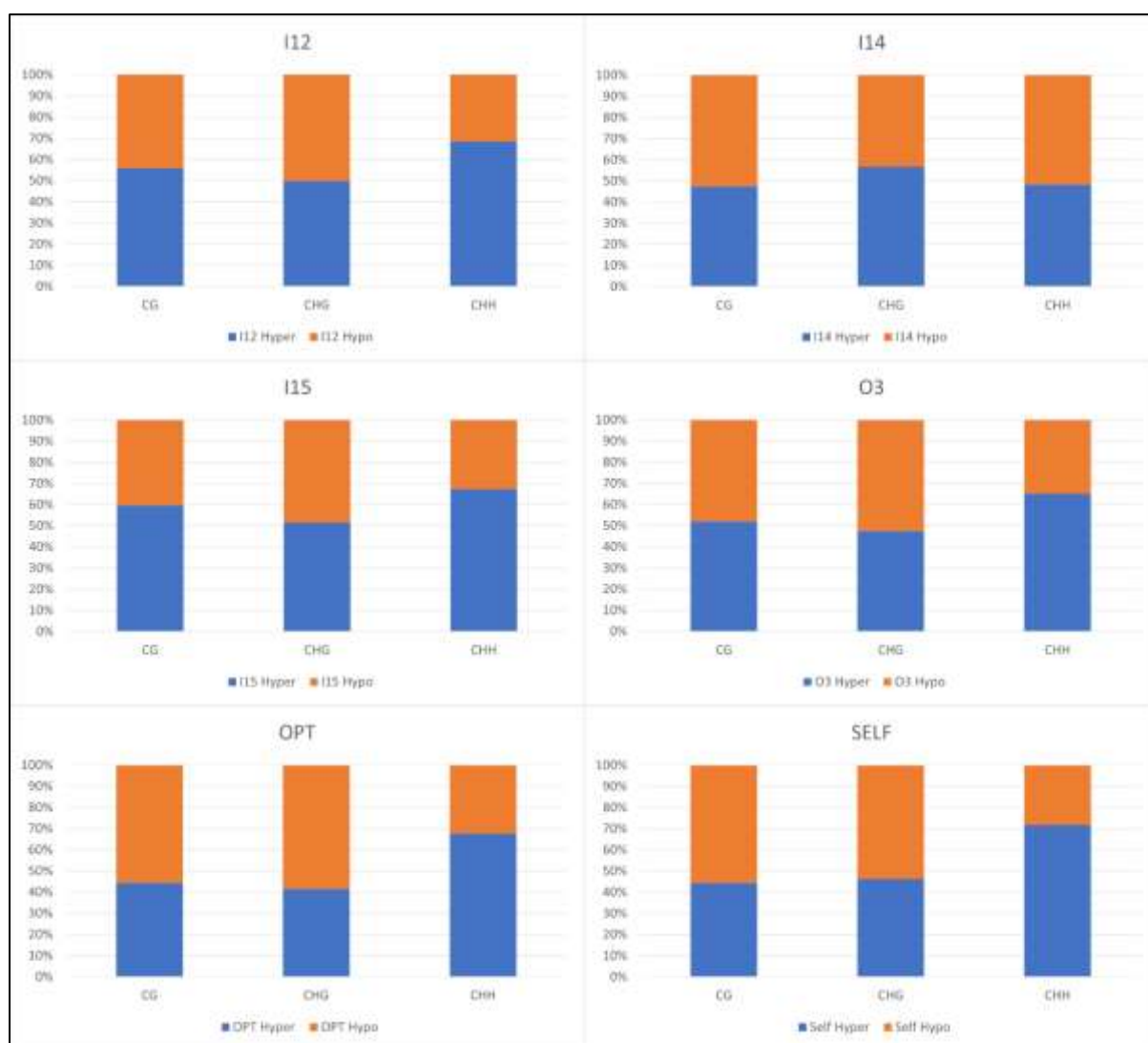


Figura 18: Andamento del livello relativo di metilazione di DNA nei context a seguito dell'innesto. I dati sono riportati in percentuale, separati sia per context di metilazione che per confronto.

La **Figura 18** ci mostra il rapporto ipometilazione/ipermetilazione relativa dei confronti riportati in percentuale. Nel caso di I12 è possibile notare che nei context CG e CHG i livelli di metilazione e demetilazione relativi sono pressoché simili, mentre in CHH i loci ipermetilati risultano essere il doppio dei demetilati. In I14 osserviamo una lieve maggioranza dei loci demetilati rispetto ai metilati; in CHG, invece, è presente una maggioranza di loci ipermetilati. Nel confronto I15 si vede l'opposto di quanto osservato con I14, cioè una lieve prevalenza di loci ipermetilati nel context CG e CHH, mentre nel context CHG è presente una lieve maggioranza di loci demetilati. Nel grafico di O3, il context CG presenta un livello di loci demetilati relativi lievemente maggiore rispetto ai non metilati, mentre il context CHG mostra un livello di ipometilazione maggiore. Il context CHH mostra un livello relativo di

ipermetilazione due volte maggiore di quello di ipometilazione. Il confronto di OTP mostra un livello relativo di ipermetilazione simile per i context CG e CHG, mentre per il CHH i loci ipermetilati sono di circa due volte superiori. Infine, nel confronto di Self si osserva che nei context CG e CHG i livelli relativi di ipermetilazione sono simili ed in CHH i loci ipermetilati sono il doppio rispetto ai ipometilati.

La presenza di DMPs in tutti i confronti evidenzia l'effetto dell'innesto e sottolinea la necessita di ulteriore investigazione attraverso strumenti statistici.

#### 5.1.1 PCA (principal component analysis)

Dai dati ottenuti dall'analisi con MethyKit riguardanti le DMPs, sono stati prodotti dei grafici rappresentativi attraverso la PCA (principal component analysis). L'analisi dei componenti principali ha permesso di semplificare la visualizzazione dei dati utilizzati attraverso delle trasformazioni lineari. Questa tecnica riduce le dimensioni di una matrice multivariata a 2 o 3 dimensioni, conservando comunque il massimo dell'informazione contenuta nei dati. Attraverso la PCA si riesce a sintetizzare la variabilità complessiva del sistema ed evidenziare eventuali strutture latenti contenute nella matrice dati. I grafici della PCA riportati di seguito sono degli *score plot* che considerano le prime due dimensioni. Sono riportate di seguito le PCA relative ai campioni (FIG. 19; FIG.20; FIG. 21).

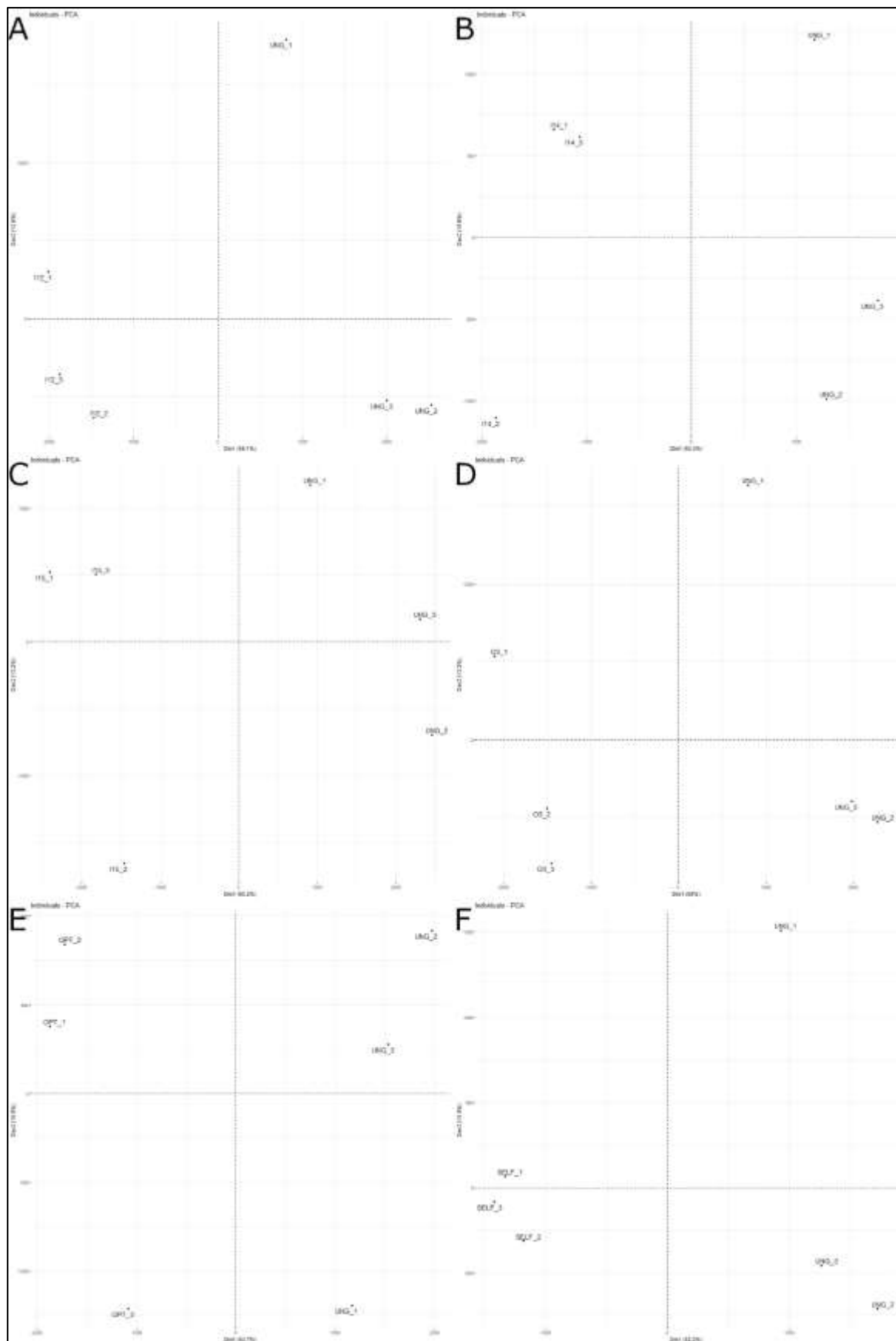


Figura 19: PCA dei confronti nel context CG. A I12; B I14; C I15; D O3; E OPT; F Self.

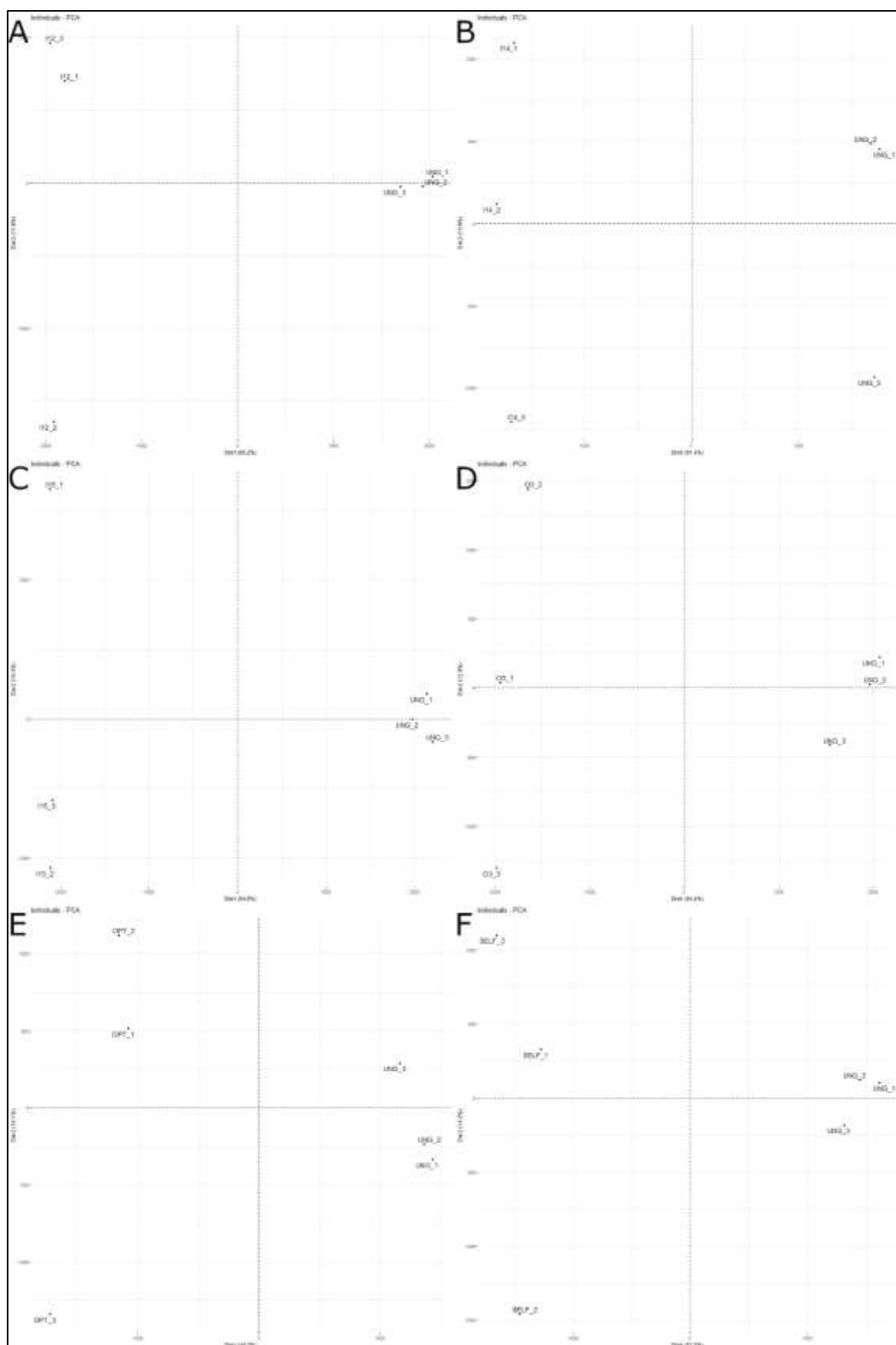


Figura 20: PCA dei confronti nel context CHG. A I12; B I14; C I15; D O3; E OPT; F Self.

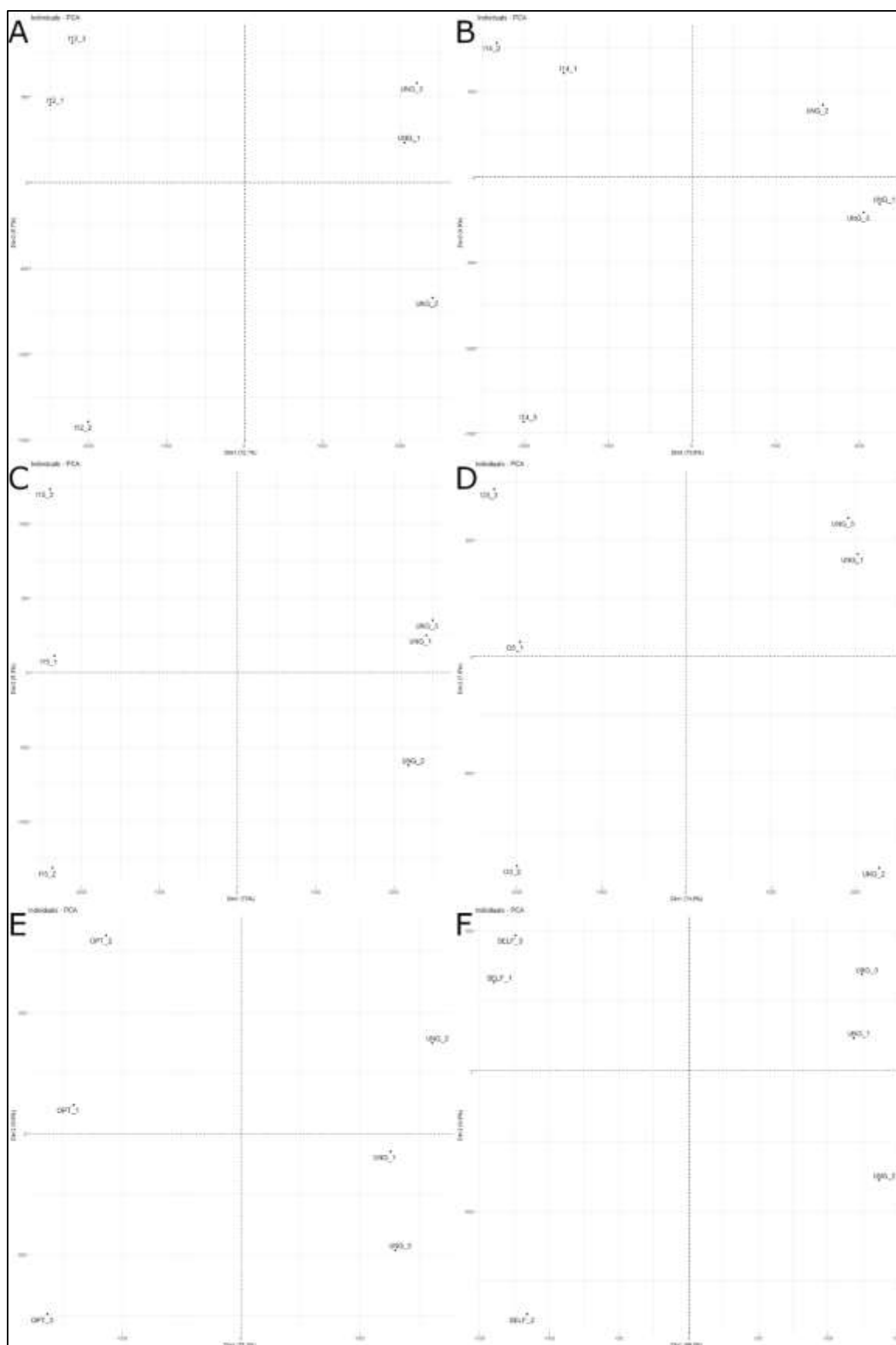


Figura 21: PCA dei confronti nel context CHH. A I12; B I14; C I15; D O3; E OPT; F Self.



Nei grafici PCA possiamo osservare come i campioni innestati si separino sensibilmente dai controlli non innestati lungo la prima componente. Questo ci dice che le differenze di metilazione tra i due gruppi sono dovute alla tecnica dell'innesto.

## 5.2 Identificazione delle DMRs

Grazie all'identificazione delle DMPs sono state calcolate le finestre entro le quali è possibile identificare le regioni differenzialmente metilate (DMRs). La Tab. 13 contiene il numero totale delle DMRs identificate per ogni confronto nei tre context di metilazione.

*Tabella 13: DMRs identificate e divise per context e per portainnesto*

Confronto	CG	CHG	CHH
<b>I12 Total</b>	452	456	464
<b>I14 Total</b>	393	361	435
<b>I15 Total</b>	526	942	575
<b>O3 Total</b>	435	500	372
<b>OPT Total</b>	434	377	194
<b>SELF Total</b>	306	375	184

### 5.2.1 DMRs

Tutti i dati provenienti dall'identificazione delle DMRs sono stati analizzati ulteriormente tramite librerie di R, grazie alle quali è stato possibile produrre grafici boxplot e heatmap.

Il *boxplot* è una rappresentazione grafica di dati numerici, che attraverso la divisione in quartili permette di descrivere la loro distribuzione grazie ad indici di posizione. Sono presenti nel grafico altre tre componenti: le linee che si estendono verticalmente andando a prendere il nome di “baffi” (detti whiskers), i quali indicano la variabilità all'esterno dei quartili più alti e più bassi, gli outlier ossia i dati che sono significativamente distanti dagli altri osservati rappresentati come punti a sé stanti e la linea centrale che rappresenta la mediana.

Le heatmap sono una tipologia di grafico nel quale l'asse delle x e delle y rappresentano due variabili, che definiscono delle categorie nei dati, ed una terza variabile definita da una gradazione di colore. In

questa ricerca la gradazione colorimetrica rappresenta il grado di metilazione relativo e le differenze tra controlli e trattati, dove si passa da uno stato di ipometilazione (in rosso) ad uno di ipermetilazione (in giallo). Inoltre, la clusterizzazione in cima rappresenta i controlli ed i trattati, a lato invece i singoli loci analizzati. Sono stati riportati di seguito boxplot ed heatmap ordinati per context e per innesto (FIG. 22; FIG. 23; FIG. 24; FIG. 25; FIG. 26; FIG. 27).

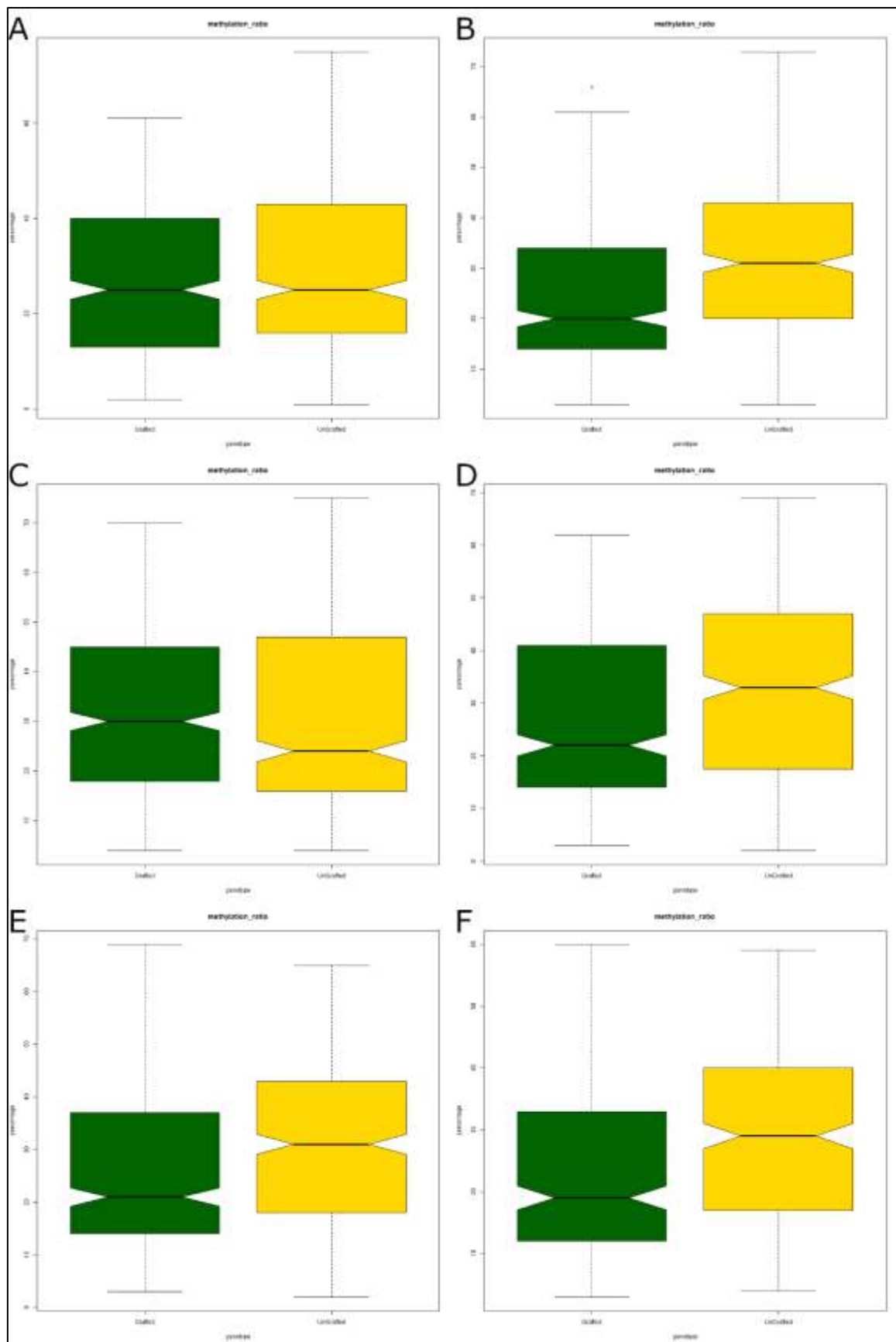


Figura 22: boxplot per context CG. A I12; B I14; C I15; D O3; E OPT; F Self.

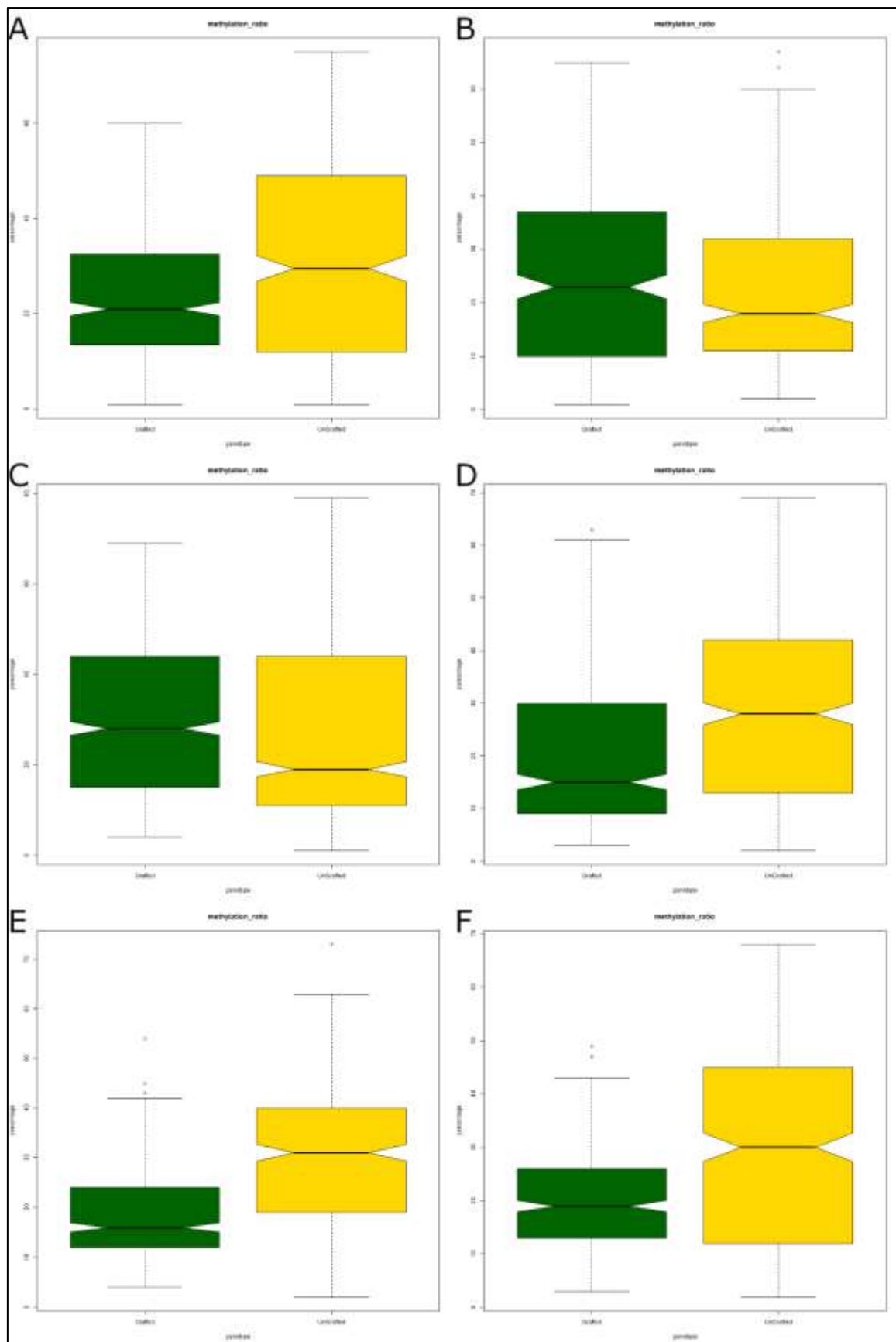


Figura 23: boxplot per context CHG. A I12; B I14; C I15; D O3; E OPT; F Self.

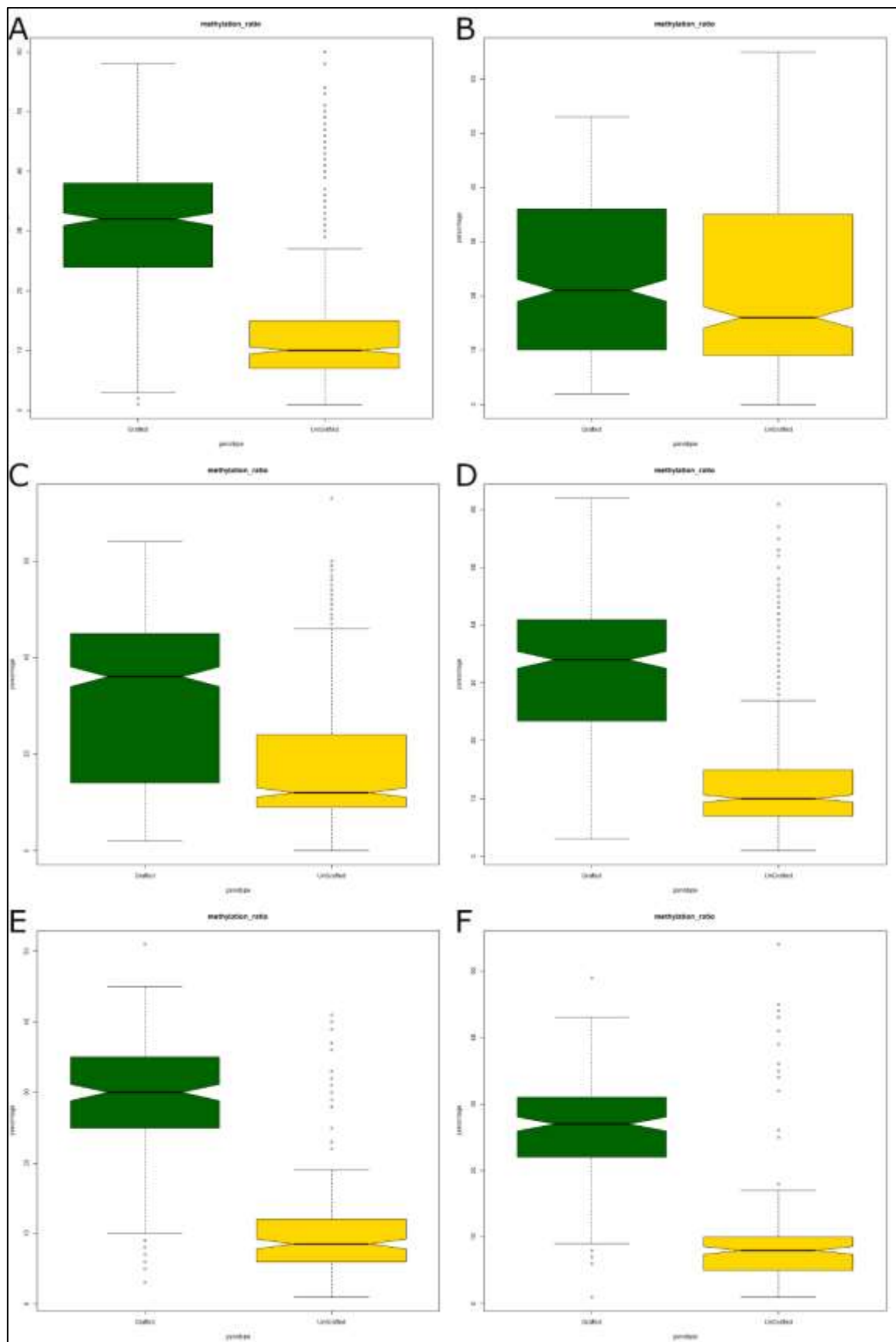


Figura 24: boxplot per context CHH. A I12; B I14; C I15; D O3; E OPT; F Self.

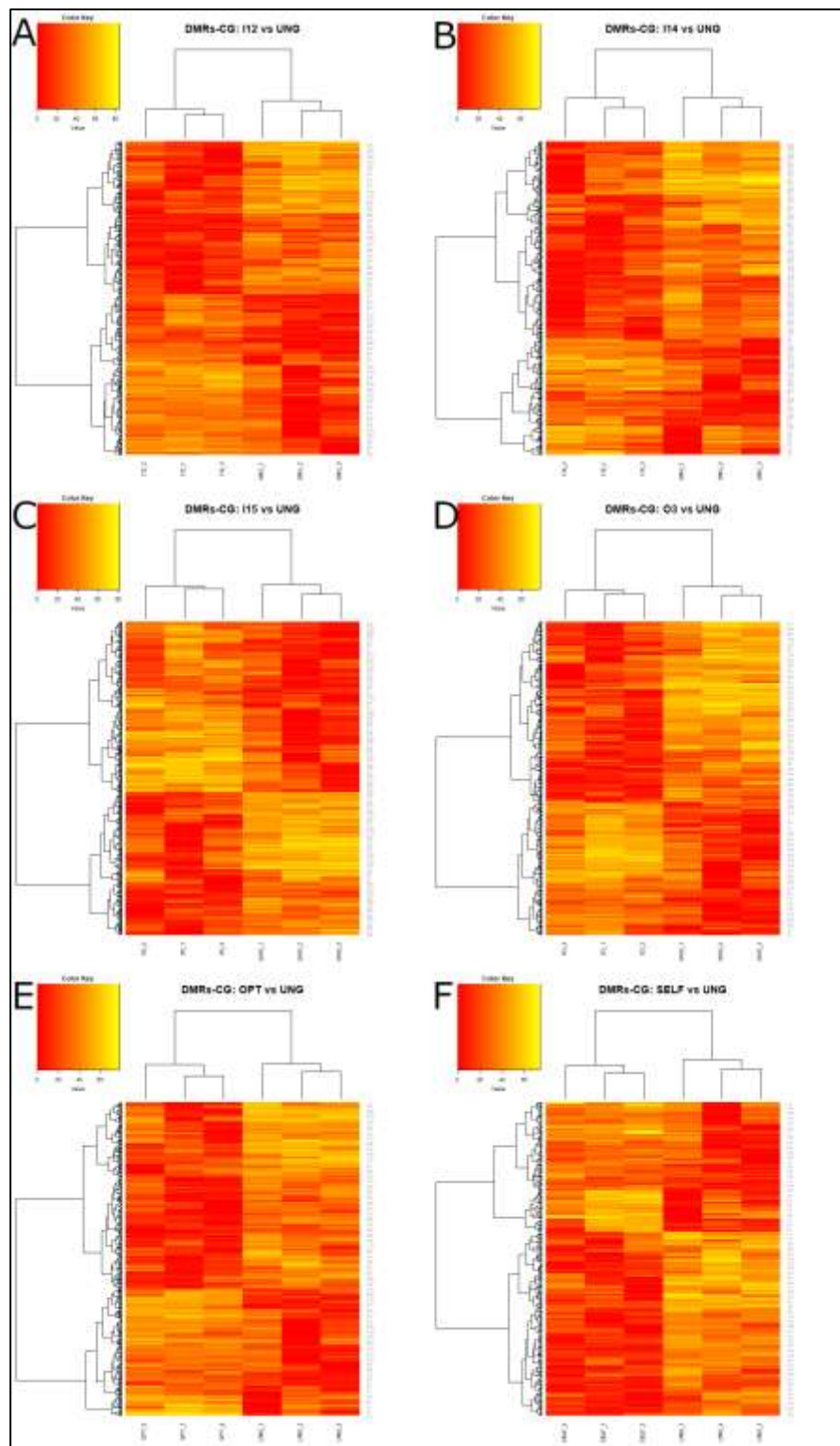


Figura 25: Heatmap per context CG. A I12; B I14; C I15; D O3; E OPT; F Self.



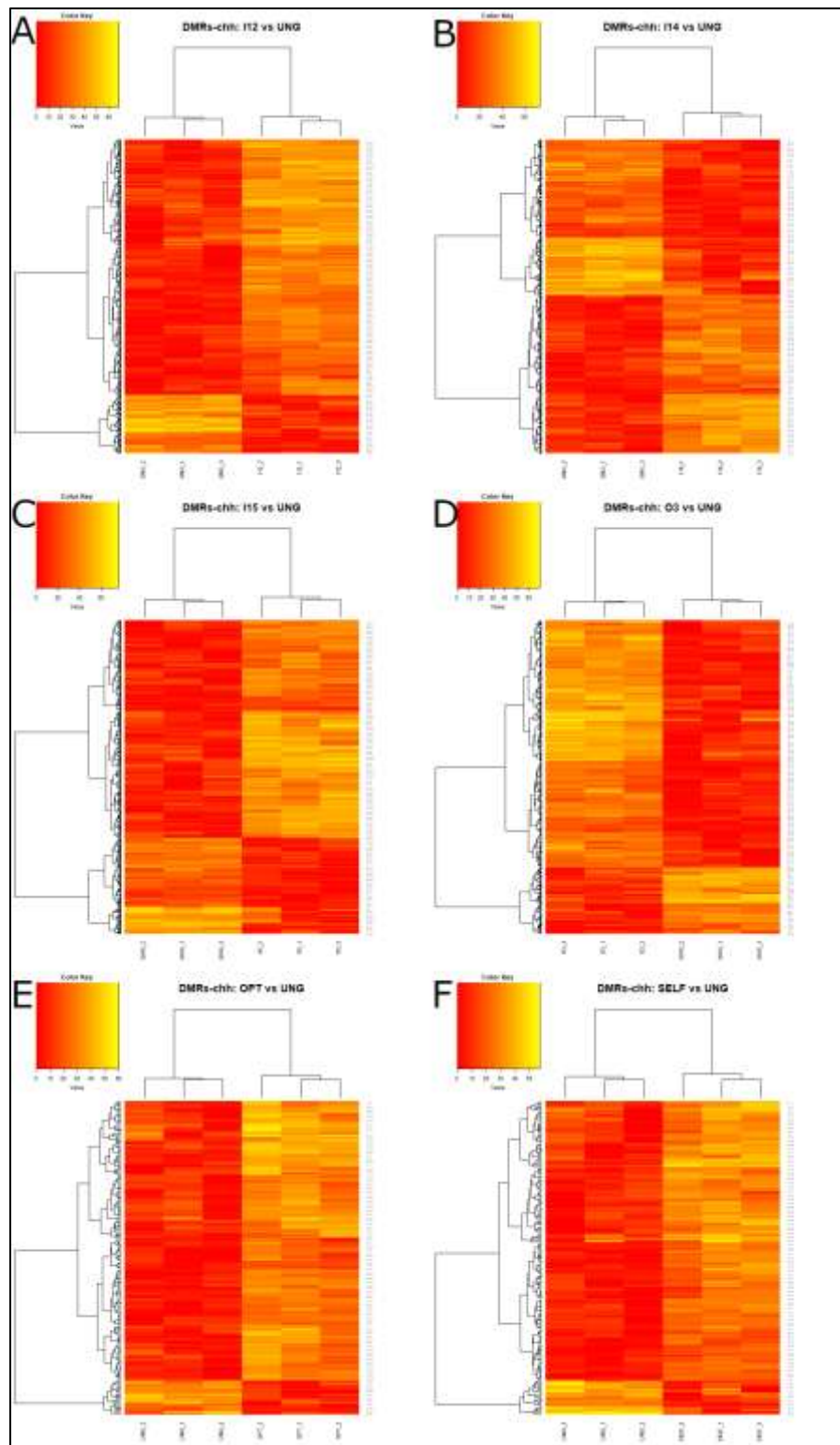


Figura 27: Heatmap per context CHH. A I12; B I14; C I15; D O3; E OPT; F Self.



Dai grafici è possibile osservare le differenze di metilazione relativa nei vari confronti. Il context CG presenta un lieve andamento verso l'ipometilazione rispetto ai controlli, simile in tutti i confronti, indicando che l'innesto ha determinato, per questo context, una possibile demetilazione. Nel context CHG si possono vedere delle differenze relative ad ogni campione. I confronti I12, O3, OPT e SELF presentano una evidente ipometilazione, il confronto I14 presenta una ipermetilazione rispetto ai controlli e I15 non vede differenze significative tra trattato e controllo. Riguardo al context CHH possiamo vedere dei risultati più marcati rispetto agli altri context, in quanto abbiamo osservato un'ipermetilazione relativa su tutti i confronti. In I12, I15, O3, OPT e Self si osserva una marcata tendenza all'ipermetilazione rispetto ai controlli, invece, in I14 non sono presenti differenze significative dal controllo. Le differenze nella metilazione delle DMRs concorrono a sottolineare l'influenza dell'innesto nella regolazione epigenetica della pianta di pomodoro. I diversi valori di ipermetilazione e ipometilazione tra i trattati ed i controlli potrebbero pertanto tradursi in una differente espressione genica, precisamente dei geni associati alle DMRs che presentano differenze significative. Il passaggio successivo di questa analisi è stato l'identificazione dei geni, importante per valutare l'effetto delle differenze di metilazione relativa nell'espressione genica dei confronti.

### 5.3 DMG

Utilizzando l'annotazione ufficiale di pomodoro (ITAG 4) è stato possibile intersecare le coordinate delle DMRs trovate con la posizione dei geni nel genoma di pomodoro. Tale analisi è stata effettuata per identificare le DMRs all'interno dei corpi genici (GB). Inoltre, le DMRs che sono state trovate a 2kb a monte o a valle da un gene sono state considerate come appartenenti alle regioni regolative (Reg) dello stesso. Nella Tab. 14, suddivisi per confronto di appartenenza e context, sono presentati i numeri di DMG per gene body, regolative e numero totale.

*Tabella 14: Numero di DMG gene body (GB), regulative (Reg). Suddivisi per confronto di appartenenza e context*

	I12		I14		I15		O3		OPT		Self	
	GB	Reg	GB	Reg	GB	Reg	GB	Reg	GB	Reg	GB	Reg
<b>CG</b>	343	197	325	162	407	228	326	164	337	184	232	149
<b>CHG</b>	428	162	339	106	454	183	468	174	360	127	361	124
<b>CHH</b>	47	29	65	31	81	45	40	32	18	18	15	13

## 5.4 Gene Ontology

I geni identificati sono stati sottoposti ad analisi con Panther, dai cui sono emerse dalle GO interessanti che coprono aspetti importanti dello sviluppo del frutto e della pianta di pomodoro. Tra le GO identificate quelle maggiormente significative e correlate ad effetti fenotipici e fisiologici sono state riportate nel testo. Queste possono coinvolgere importanti aspetti della pianta come: lo sviluppo, la biosintesi ormonale, l'assorbimento di nutrienti e la qualità del frutto. La Tab. 15 contiene le GO più interessanti e il confronto in cui sono state trovate.

Tabella 15: Gene Ontology di tutti i confronti.

Gene Ontology	Confronti
Production of miRNAs involved in gene silencing by miRNA	I12 - O3 - OPT
Auxin polar transport	I12 - I14 - I15 - O3 - SELF
Auxin polar transport	I12 - I15 - O3 – OPT - SELF
Ent-kaurene oxidation to kaurenoic acid	I14 - I15 - O3 - OPT
Negative regulation of gibberellic acid mediated signaling pathway	I14 - 15
Cellular glucose homeostasis	O3 – OPT - SELF
Abscisic acid biosynthetic process	I14 - O3
Glycerol-3-phosphate biosynthetic process	I15 - O3
UDP-glucosyltransferase activity	I12
Brassinosteroid homeostasis	I12
Abscisic acid-activated signaling pathway	I14
Abscisic acid-activated signaling pathway	I14
Cellular macromolecule metabolic process	I15
Strigolactone biosynthetic process	I15
Asymmetric cell division	O3
abscisic acid biosynthetic process	O3
Negative regulation of abscisic acid-activated signaling pathway	OPT
Negative regulation of gibberellic acid mediated signaling pathway	OPT
Regulation of signal transduction	OPT
6-phosphofructokinase activity	OPT
6-phosphofructokinase activity	OPT
Farnesyl diphosphate biosynthetic process, mevalonate pathway	Self
Pigment biosynthetic process	Self

## 5.5 Analisi dei geni

I geni appartenenti alle GO più importanti sono stati approfonditi grazie alle rispettive annotazioni e descrizione del loro ruolo all'interno del genoma di pomodoro. Nei 6 confronti di questa analisi sono stati trovati dei geni ricorrenti, questi hanno dimostrato un enrichment superiore ad altri in più confronti.

Per i confronti I12, I15, O3, OPT e Self nel context CHG è stato trovato il gene ipometilato "*Solyc03g118740.3*" a livello del gene-body. Il gene codifica per la proteina AUXIN EFFLUX CARRIER COMPONENT 1 (SIPIN1), la quale è coinvolta come componente nel trasporto di auxina. Inoltre, è parte del trasporto basipetale di auxina e media la formazione del gradiente auxina necessario per garantire la corretta organogenesi della pianta.

Il gene "*Solyc02g089260.4*" è stato trovato ipermetilato a livello del gene-body, nel context CHG, in I12, I15, O3, Self e ipometilato in I14 a livello del gene-body. Questo gene codifica per una Auxin transport protein BIG, necessaria per il trasporto di auxina che va ad influenzare lo sviluppo della pianta. L'omologo in *A. thaliana* svolge diverse funzioni, tra cui troviamo: il controllo dell'allungamento dei pedicelli e degli internodi dello stelo, la modulazione dell'espressione di geni regolati dalla luce, la repressione l'espressione dei geni CAB1 e CAB3, aumento della sensibilità agli inibitori del trasporto di auxina e inoltre influenza la distribuzione subcellulare del trasportatore di auxina PIN1 in risposta agli inibitori del trasporto di auxina. Inoltre, è anche coinvolto nella regolazione delle risposte ai fitormoni come auxina, citochina, etilene e acido gibberellico (GA), in particolare durante gli stimoli luminosi (Kanyuka *et al.*, 2003).

Per i confronti I12, O3 e OPT nel context CHG, ipermetilato a livello delle regioni regolative, è stato trovato il gene "*Solyc01g009090.3*" il quale codifica per la proteina Serrate RNA effector molecule. L'omologo in *A. thaliana* funge da mediatore tra il cap-binding complex (CBC) e le proteine per lo splicing dei pre-mRNA e dei miRNA, quindi necessario per la corretta preparazione dei principali miRNA coinvolti nel silenziamento genico RdDM. La proteina agisce anche da repressore dei geni coinvolti nello sviluppo fogliare, inoltre, è coinvolta nella regolazione generale del processamento dei miRNA, favorendo l'accumulo dei trans-acting small interfering RNA (ta-siRNA) necessari per splicing pre-mRNA (Laubinger *et al.*, 2008).

Il gene “*Solyc04g083160.2*” è stato trovato nel context CHG dei confronti I14, I15 e O3 ipermetilato a livello del gene-body e nel confronto OPT ipometilato a livello del gene-body. Questo gene codifica per una proteina ossidasi. L’omologo in *A. thaliana* codifica per la proteina Ent-kaurene ossidasi (cytochrome P450), la quale catalizza tre ossidazioni successive del gruppo 4-metile dell’ent-kaurene dando come risultato l’acido kaurenoico. Questo è un passo chiave nella biosintesi delle gibberelline (GA) (Morrone *et al.*, 2010).

Il gene “*Solyc10g086370.1*” è stato trovato nei confronti I14 e 15 nel context CHG e ipometilato al livello del gene-body. Questo gene codifica per una proteina contenente un dominio GRAS, il quale è presente in un'importante famiglia di proteine vegetali che prende il nome dai primi sui tre membri (GIBBERELLIC-ACID INSENSITIVE (GAI), REPRESSOR of GAI (RGA) and SCARECROW (SCR). Queste proteine svolgono diversi ruoli nello sviluppo di radici e germogli, nella regolazione del fitormone acido gibberellico (GA) e la trasduzione del segnale a carico del fitocromo A (Hirsch *et al.*, 2009).

Nei confronti O3, OPT e SELF, nel context CHH, è stato trovato il gene “*Solyc02g091830.3*” ipermetilato a livello della zona regolativa. Questo gene codifica per una proteina fosfotransferasi la quale è coinvolta nel metabolismo dei composti esosi, parte della glicolisi e del metabolismo dei carboidrati. In particolare, catalizza il primo step del subpathway che sintetizza D-gliceraldeide 3-fosfato e diidrossiacetone fosfato da D-glucosio.

Nel context CHG dei confronti I15 e O3 è stato identificato il gene “*Solyc03g117490.4*” il quale è ipometilato a livello del gene-body. Esso codifica per una glicerolo chinasi, una proteina coinvolta nel pathway che sintetizza sn-glicerolo 3-fosfato (G3P) dal glicerolo e parte della via di degradazione del glicerolo attraverso la via della glicerolo chinasi, che è a sua volta parte del metabolismo dei polioli. Questi sono dei carboidrati idrogenati importanti per la loro funzione dolcificante ed effetti benefici sulla salute umana (Livesey *et al.*, 2003).

Nei confronti I14 e O3, nel context CHG, è stato trovato il gene “*Solyc04g071940.3*” il quale codifica per l’enzima xantossina deidrogenasi. Questo enzima è coinvolto nella biosintesi dell’acido abscissico andando a sintetizzare l’aldeide abscissica il diretto precursore dell’ABA. La deidrogenasi catalizza la reazione: xantossina + NAD<sup>+</sup> ⇌ aldeide abscissica + NADH + H<sup>+</sup>.

### 5.5.1 I12

Il confronto di pomodoro I12 ha presentato due geni interessanti. Il primo nel context CG è “*Solyc02g081300.3*” ipermetilato a livello della regione regolativa. Questo gene codifica per un enzima che scinde il saccarosio andando a produrre fruttosio e UDP-glucosio, molecola ad alta energia che prende parte alla sintesi del glicogeno per varie vie metaboliche. Il secondo, “*Solyc12g006860.2*” è stato trovato nel context CHH e ipermetilato a livello del gene-body. Questo gene codifica per un enzima citocromo P450 che svolge la idrossilasi dei brassinosteroidi. L’enzima è coinvolto nella omeostasi e inattivazione di questi fitormoni. I brassinosteroidi (Brs) sono ormoni vegetali endogeni essenziali per la corretta regolazione di molti processi fisiologici necessari per la crescita e lo sviluppo delle piante, tra questi anche la regolazione l'espressione di numerosi geni. I Brs hanno un impatto sull'attività delle vie metaboliche più complesse e contribuiscono alla regolazione della divisione e differenziazione cellulare (Steven, 2011). Questi fitormoni, alla luce di ricerche sulla loro inibizione, sono risultati coinvolti in allungamento, flessione, divisione cellulare, sviluppo riproduttivo e vascolare, polarizzazione della membrana, e modulazione dello stress. In *A. thaliana* e nel pomodoro sono stati osservati dei mutanti dei Brs, in cui le giovani piante presentano un fenotipo deetiolato, un colore scuro, un ipocotile corto e dei cotiledoni aperti, dimostrando un ruolo importante di Brs nella fotomorfogenesi e nella morfogenesi delle foglie (Steven & Sasse, 1998).

### 5.5.2 I14

Il confronto I14 ha mostrato nel context CG due geni “*Solyc08g076960.1*” e “*Solyc10g085310.1*”, ipometilati a livello del gene-body. Questi codificano rispettivamente per PYL1 e PYL5, due ricettori dell’acido abscissico (ABA), un fitormone sesquiterpenoide coinvolto nel controllo di un gran numero di funzioni fisiologiche. I ricettori sono necessari per l’attuazione di risposte dell’organismo mediate dal fitormone stesso. Gli omologi in *A. thaliana* hanno la funzione inibitoria verso l'attività delle fosfatasi proteiche di gruppo A di tipo 2C (Pp2cs) quando attivate da ABA e promuovono la resistenza allo stress idrico (Zhang *et al.*, 2013).

### 5.5.3 I15

Nel confronto I15 nel context CG sono stati identificati due geni: “*Solyc01g057210.3*” ipermetilato e “*Solyc04g077860.4*” ipometilato, entrambi a livello del gene-body. Il primo gene codifica per una proteina chinasi, la quale inibisce il complesso mitocondriale piruvato deidrogenasi (PDH) tramite la fosforilazione della subunità alfa E1 sui residui di Ser, andando così a contribuire alla regolazione del metabolismo del glucosio. Il secondo codifica per una idrolasi, con dominio  $\alpha/\beta$ , coinvolta nel processo di sintesi dei fitormoni strigolattoni. Gli strigolattoni sono ormoni che inibiscono l'accrescimento e la ramificazione dei germogli attraverso la via MAX-dipendente, inoltre contribuiscono a regolare la risposta della pianta a condizioni fosforo limitanti e potenziano l'interazione con funghi simbionti per aumentare la capacità di assorbimento dei nutrienti. Il gene omologo in *A. thaliana* codifica per la Strigolactone esterasi DWARF 14, la quale funziona a valle della sintesi dei stringolactoni, come componente della segnalazione ormonale e come enzima che partecipa alla conversione degli strigolactoni nella forma bioattiva (Chevalier *et al.*, 2014).

### 5.5.4 O3

Il confronto O3 ha esibito dei geni interessanti in due dei 3 context, precisamente in CG e CHG. Nel context CG è stato trovato il gene “*Solyc10g074680.2*” ipometilato a livello del gene-body. Questo codifica per una proteina appartenente al gruppo delle SCARECROW. L'omologo in *A. thaliana* agisce come fattore di trascrizione per la divisione cellulare asimmetrica coinvolta nella formazione di pattern radiale nelle radici, organi assiali ed il mantenimento delle cellule staminali. Inoltre, risulta essenziale per lo svolgimento della divisione cellulare e il gravitropismo (Wysocka-Diller *et al.*, 2000; Fukaki *et al.*, 1998). Nel context CHG è stato trovato il seguente gene, ipometilato a livello del gene-body, “*Solyc12g056450.2*” che codifica per l'enzima 3-idrossi-3-metilglutaril coenzima A sintasi. Questo enzima condensa acetil-CoA con acetoacetil-CoA per formare HMG-CoA, che è il substrato per HMG-CoA reduttasi o via metabolica dell'acido mevalonico. La via del mevalonato è una via anabolica che a partire dall'Acetil-Coenzima A porta alla sintesi di composti organici essenziali per la pianta come poli-isoprenoidi e steroli.

#### 5.5.5 OPT

Nel confronto OPT solo il context CG ha dimostrato un enrichment di geni interessanti per l'analisi. In tutto sono stati trovati cinque geni: "*Solyc05g008060.4*", "*Solyc11g011260.1*", "*Solyc01g096490.3*", "*Solyc04g015200.3*" e "*Solyc04g072580.1*". Il primo gene "*Solyc05g008060.4*" risulta ipermetilato a livello del gene-body e codifica per la proteina auxin efflux carrier component 4 (SIPIN4), un componente del trasporto dell'auxina coinvolta nella crescita della pianta. Il silenziamento dell'espressione genica di SIPIN4, proteina prevalentemente espressa in germoglio di fiore di pomodoro e giovani frutti in via di sviluppo, può portare a frutti partenocarpici a causa dello sviluppo precoce prima della fecondazione (Mounet *et al.*, 2012). Il secondo gene "*Solyc11g011260.1*", ipometilato a livello del gene-body, codifica per una proteina appartenente al gruppo "DELLA protein GAI", che partecipa alla repressione della via di segnalazione della Gibberellina (GA). Il gene "*Solyc01g096490.3*" codifica per una E3 ubiquitin-protein ligase KEG, l'omologo in *A. thaliana* media l'ubiquitinazione proteica E2-dipendente. Inoltre, agisce come regolatore negativo della segnalazione dell'acido abscisico (ABA) ed è necessaria per la degradazione di ABI5 mediando la sua ubiquitinazione. Questa proteina, insieme a EDR1, può regolare il traffico endocitico e la formazione di complessi di segnalazione sulla rete trans-Golgi (TGN) durante le risposte allo stress (Liu *et al.*, 2010). Infine, gli ultimi due geni "*Solyc04g015200.3*" e "*Solyc04g072580.1*", entrambi ipermetilati a livello del gene-body, codificano per l'enzima ATP-dipendente 6-fosfofruttichinasi, il quale catalizza la fosforilazione del D-fruttosio 6-fosfato a fruttosio 1,6-disfosfato consumando ATP.

#### 5.5.6 Self

Nel confronto Self nel context CG sono stati trovati due geni. Il primo è "*Solyc06g066310.3*" ipermetilato a livello della zona regolativa e codifica per una fosfomevalonato chinasi. Questa proteina è coinvolta nella via metabolica dell'acido mevalonico o del HMG-CoA reduttasi. Il secondo è il gene "*Solyc03g123760.3*" ipometilato a livello della zona regolativa. Questo gene codifica per la proteina 15-cis-phytoene desaturase, coinvolta nella biosintesi del licopene il quale fa parte della biosintesi dei carotenoidi e catalizza la conversione fitoene in zeta-carotene. Il licopene è un idrocarburo isomero aciclico del beta-carotene molto importante per il frutto del pomodoro. Nel pomodoro alti livelli di licopene sono spesso associati a frutti con una maggior possibilità di vendita ed una qualità del frutto migliore, andando ad aumentare fattori come: colore, contenuti solidi solubili (SSC) e acidità titolabile (TA). Questi sono chiari indicatori di una maggior qualità del frutto che risulta in una preferenza di



acquisto da parte del consumatore (Ilahy *et al.*, 2018). Il licopene ha inoltre delle proprietà positive per la salute umana, esso infatti rappresenta uno dei più potenti antiossidanti naturali efficace nella prevenzione delle malattie croniche (Rao *et al.*, 2006).

Questi descritti sono alcuni dei geni evidenziati dall'analisi tramite PANTHER, con un grado di enrichment elevato rispetto agli attesi. La pratica dell'innesto e le modifiche al livello di metilazione che ne consegue, coinvolgono tutta la pianta e con essa il suo genoma, modificando l'espressione genica di molti geni. Questi coprono una vasta gamma di aspetti biologici della pianta, partendo dalla sintesi di ormoni, passando ai geni coinvolti nella glicolisi, fino ai geni codificanti per lo splicing di miRNA coinvolti nella stessa RdDM. Vista la mole e la varietà dei geni interessati da modificazioni nel livello di metilazione gli effetti dell'innesto sul epigenoma di pomodoro sono evidenti e misurabili. Una maggiore ricerca è comunque necessaria per individuare le conseguenze esatte di questi cambiamenti, così da poter strumentalizzare questa tecnica non solo nel campo della resistenza a stress biotici ed abiotici ma come una soluzione per il miglioramento del raccolto e delle qualità del frutto stesso. I risultati proposti da questa analisi sono promettenti dimostrando la possibilità dei geni coinvolti nello sviluppo del frutto, di essere influenzati dalla tecnica dell'innesto.

## 6. CONCLUSIONI

La tecnica dell'innesto ha rappresentato nel tempo una valida soluzione per aumentare la resistenza delle piante a stress abiotici e biotici, andando a produrre vantaggi importanti per le colture. Questi vantaggi, dal miglioramento della qualità del frutto fino all' aumento del raccolto hanno portato all'aumento dell'utilizzo della tecnica in agricoltura. L'innesto pone le piante in una condizione di stress che rende la tecnica soggetto di studi riguardanti le modificazioni epigenetiche. L'epigenetica e le sue modificazioni sono, in questi ultimi anni, una branca della genomica sempre più approfondita ed in rapida evoluzione. Il focus di questa ricerca è stata la modificazione della metilazione del DNA, in particolare, è stata valutata la metilazione delle citosine in tre context, il CG e due esclusivi alle piante CHG e CHH. Questi tre context sono strettamente collegati al compattamento della cromatina, alla regolazione dell'espressione genica, dello splicing e al silenziamento o attivazione dei trasposoni. Le differenze nella metilazione, tra trattati e controlli, sono state individuate in innesti effettuati con cinque tipi di portainnesto commerciale diverso (I12, I14, I15, O3 e OPT) ed un ultimo innesto prodotto con nesto e portainnesto della stessa varietà chiamato Self. L'analisi è stata effettuata a partire dai risultati della metodologia MCSed e utilizzando successivamente strumenti bioinformatici. Questo ci ha permesso di valutare in primo luogo che le differenze di metilazione tra controlli e trattati sono state causate dall'innesto. Proseguendo nell'analisi è stato possibile identificare le DMPs e le DMRs nei confronti. Le DMPs hanno mostrato, nei due context CG e CHG, andamenti misti nella metilazione relativa. Nel primo context CG è stato trovato un andamento verso l'ipermetilazione, con l'unica eccezione del confronto i14 che presenta una ipometilazione. Il secondo context CHG presenta invece la situazione opposta con un andamento generale verso l'ipometilazione, i due confronti i14 ed i15 presentano una tendenza all'ipermetilazione. Per quanto riguarda il context CHH è stata osservata una ipermetilazione comune a tutti i confronti. Le DMRs hanno mostrato nel context CG risultati contrastanti rispetto alle DMPs, una tendenza all'ipometilazione generale di tutti i confronti. Gli altri due context si allineano con i risultati trovati nelle DMPs. Per cui il context CHG ha mostrato una tendenza all'ipometilazione, i14 ed i15 invece presentano un'ipermetilazione rispetto ai controlli. Il context CHH ha presentato una ipermetilazione generale su tutti i confronti. Alla luce di queste analisi le Gene ontology ci hanno permesso di approfondire i geni individuati nelle DMRs e di capire la loro funzione e descrizione. La mole di geni identificati è stata notevole, ha riguardato le più disparate funzioni biologiche della pianta e del frutto, partendo dalla sintesi di fitormoni, nell'espressione dei geni coinvolti nella glicolisi, nei pathway che definiscono il colore del frutto, fino ai geni coinvolti nello splicing di miRNA all'interno della stessa RdDM, che regola la metilazione del DNA. Sulla base di questi risultati una maggiore ricerca è stata necessaria per individuare le conseguenze esatte di questi cambiamenti nella pianta e nel frutto. L'epigenoma di pomodoro, dunque, si è mostrato suscettibile alla tecnica dell'innesto sotto forma di variazioni nella metilazione relativa. La differente metilazione

è ricaduta su alcune delle zone regolative o sui corpi dei geni, provocandone possibilmente una differente espressione. Le conseguenze sulla metilazione del DNA vanno ad aggiungere un altro strato di complessità alla tecnica dell'innesto. Vista non più al pari di uno strumento per la protezione o il benessere della pianta, ma come una tecnica con le potenzialità di influenzare le quantità dei raccolti e la qualità dei frutti. Questi due parametri sono essenziali per una agricoltura migliore ed economicamente vantaggiosa nel rispetto dell'ambiente, utilizzando di fatto una proprietà indotta nelle piante per ottenere il meglio da esse. I risultati proposti da questa analisi mostrano l'influenza dell'innesto sull'epigenoma di pomodoro e la possibilità dei geni coinvolti da tali modificazioni di influenzare lo sviluppo e le qualità del frutto. Un ulteriore approfondimento nella ricerca è comunque necessario per poter fornire alla tecnica applicazioni pratiche nel controllo del frutto, attraverso l'influenza dell'innesto sulla metilazione del DNA.

## 7 RINGRAZIAMENTI

Vorrei ringraziare tutte le persone che mi hanno permesso di arrivare fin qui e di portare a termine questo lavoro di tesi. Innanzitutto, un sentito grazie al mio relatore Prof. Emidio Albertini per l'opportunità che mi ha concesso. Cimentarmi nella stesura di una tesi sperimentale, è stata per me e le mie conoscenze una grande sfida soprattutto alla fine di un percorso triennale. Devo a lui, perciò la possibilità di aver affrontato nel dettaglio argomenti complessi, affascinanti e lontani dai miei precedenti studi come l'epigenetica e la bioinformatica. Campi scientifici che saranno protagonisti del mio futuro. Grazie, inoltre, per la disponibilità e tempestività ad ogni mia richiesta. Un ringraziamento speciale va al mio correlatore Dr. Marco Di Marsico per i suoi indispensabili consigli, gioiosa severità e per avermi suggerito puntualmente le giuste modifiche da apportare alla mia tesi. La sua immensa pazienza e la sua complicità nella stesura di ogni capitolo dell'elaborato sono state fondamentali. Grazie per avermi fornito gli strumenti per la mia ricerca, devo a te le conoscenze e le competenze che ho acquisito durante tutto il percorso di tesi. Sei stato una figura importante per me dal primo momento in laboratorio fino all'ultima e-mail di revisione. Ti auguro il meglio. L'energia e la passione che mi hai trasmesso non andrà affatto sprecata. Non sarei mai potuto arrivare fin qui senza il supporto morale dei miei genitori, che mi hanno sempre sostenuto, appoggiato e supportato lungo tutto il percorso di studi. Vi devo molto per essermi stati accanto in questo periodo intenso, soprattutto nei momenti difficili e di sconforto, è solo grazie a voi se sono riuscito a superarli. Adesso spero di poter gioire insieme dei traguardi raggiunti. Un ringraziamento particolare va a mia sorella che ha contribuito con le sue idee e le sue intuizioni a dare un tocco di originalità a questo elaborato. Per i miei amici, siete stati la mia luce nei momenti più bui, grazie per aver reso le poche pause concesse la mia salvezza. Per ultimo ringrazio me stesso, per la tenacia e la testardaggine che mi hanno permesso di arrivare alla fine di questo percorso.

## BIBLIOGRAFIA

1. Angelini, Renzo, and V. Magnifico. "Pomodoro." ART Servizi Editoriali SpA, Italia (2010).
2. Bai, Yuling, and Pim Lindhout. "Domestication and breeding of tomatoes: what have we gained and what can we gain in the future?." *Annals of botany* 100.5 (2007): 1085-1094.
3. Berdasco, María, *et al.* "Promoter DNA hypermethylation and gene repression in undifferentiated Arabidopsis cells." *PloS one* 3.10 (2008): e3306.
4. Berger, Margot MJ, Philippe Gallusci, and Emeline Teyssier. "Roles of epigenetic mechanisms in grafting and possible applications." *Advances in Botanical Research* 88 (2018): 203-246.
5. Bewick, Adam J., and Robert J. Schmitz. "Gene body DNA methylation in plants." *Current opinion in plant biology* 36 (2017): 103-110.
6. Bies-Etheve, Natacha, *et al.*, "RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family." *EMBO reports* 10.6 (2009): 649-654.
7. Bird, Adrian. "Perceptions of epigenetics." *Nature* 447.7143 (2007): 396.
8. Bottaccioli, Francesco, and Anna Giulia Bottaccioli. "La rivoluzione epigenetica." *Sistema salute. Rivista di educazione sanitaria e di promozione della* (2012): 319-335.
9. Cao, Xiaofeng, and Steven E. Jacobsen. "Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes." *Proceedings of the National Academy of Sciences* 99.suppl 4 (2002): 16491-16498.
10. Carthew, Richard W., and Erik J. Sontheimer. "Origins and mechanisms of miRNAs and siRNAs." *Cell* 136.4 (2009): 642-655.
11. Cheng, Jingfei, *et al.*, "Downregulation of RdDM during strawberry fruit ripening." *Genome biology* 19.1 (2018): 1-14.
12. Chevalier, Florian, *et al.* "Strigolactone promotes degradation of DWARF14, an  $\alpha/\beta$  hydrolase essential for strigolactone signaling in Arabidopsis." *The Plant Cell* 26.3 (2014): 1134-1150.
13. Choi, Yeonhee, *et al.* "DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in Arabidopsis." *Cell* 110.1 (2002): 33-42.
14. Clouse, Steven D. "Brassinosteroids." *The Arabidopsis Book/American society of plant biologists* 9 (2011).
15. Clouse, Steven D., and Jenneth M. Sasse. "Brassinosteroids: essential regulators of plant growth and development." *Annual review of plant biology* 49.1 (1998): 427-451.
16. Cui, Qingqing, *et al.* "Stage-specific events in tomato graft formation and the regulatory effects of auxin and cytokinin." *Plant Science* 304 (2021): 110803.



17. Djidonou, Desire, *et al.* "Influence of interspecific hybrid rootstocks on tomato growth, nutrient accumulation, yield, and fruit composition under greenhouse conditions." *HortTechnology* 27.6 (2017): 868-877.
18. Dunoyer, Patrice, *et al.*, "Plant mobile small RNAs." *Cold Spring Harbor perspectives in biology* 5.7 (2013): a017897.
19. Finnegan, E. Jean, W. James Peacock, and Elizabeth S. Dennis. "Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development." *Proceedings of the National Academy of Sciences* 93.16 (1996): 8449-8454.
20. Flores, Francisco B., *et al.* "The effectiveness of grafting to improve tomato fruit quality." *Scientia horticulturae* 125.3 (2010): 211-217.
21. Fukaki, Hidehiro, *et al.* "Genetic evidence that the endodermis is essential for shoot gravitropism in *Arabidopsis thaliana*." *The Plant Journal* 14.4 (1998): 425-430.
22. Gehring, Mary, *et al.*, "DEMETER DNA glycosylase establishes MEDEA polycomb gene Self-imprinting by allele-specific demethylation." *cell* 124.3 (2006): 495-506.
23. Gehring, Mary, Kerry L. Bubb, and Steven Henikoff. "Extensive demethylation of repetitive elements during seed development underlies gene imprinting." *Science* 324.5933 (2009): 1447-1451.
24. Gong, Zhizhong, *et al.*, "ROS1, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase." *Cell* 111.6 (2002): 803-814.
25. Hayatsu, Hikoya. "Discovery of bisulfite-mediated cytosine conversion to uracil, the key reaction for DNA methylation analysis—a personal account." *Proceedings of the Japan Academy, Series B* 84.8 (2008): 321-330.
26. He, Li, *et al.* "A naturally occurring epiallele associates with leaf senescence and local climate adaptation in *Arabidopsis* accessions." *Nature communications* 9.1 (2018): 1-11.
27. He, Xin-Jian, *et al.*, "An effector of RNA-directed DNA methylation in *Arabidopsis* is an ARGONAUTE 4-and RNA-binding protein." *Cell* 137.3 (2009): 498-508.
28. Hirsch, Sibylle, and Giles ED Oldroyd. "GRAS-domain transcription factors that regulate plant development." *Plant signaling & behavior* 4.8 (2009): 698-700.
29. Hofmeister, Brigitte T., *et al.* "Stable inheritance of DNA methylation allows creation of epigenotype maps and the study of epiallele inheritance patterns in the absence of genetic variation." *Genome biology* 18.1 (2017): 1-16.
30. Ikeuchi, Momoko, Keiko Sugimoto, and Akira Iwase. "Plant callus: mechanisms of induction and repression." *The plant cell* 25.9 (2013): 3159-3173.

31. Ilahy, Riadh, *et al.* "When color really matters: horticultural performance and functional quality of high-lycopene tomatoes." *Critical Reviews in Plant Sciences* 37.1 (2018): 15-53.
32. Jackson, James P., *et al.* "Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in *Arabidopsis thaliana*." *Chromosoma* 112.6 (2004): 308-315.
33. Johannes, Frank, and Robert J. Schmitz. "Spontaneous epimutations in plants." *New Phytologist* 221.3 (2019): 1253-1259.
34. Kakutani, Tetsuji, *et al.* "Meiotically and mitotically stable inheritance of DNA hypomethylation induced by *ddm1* mutation of *Arabidopsis thaliana*." *Genetics* 151.2 (1999): 831-838.
35. Kakutani, Tetsuji. "Epi-alleles in plants: inheritance of epigenetic information over generations." *Plant and cell physiology* 43.10 (2002): 1106-1111.
36. Kalisz, Susan, and Michael D. Purugganan. "Epialleles via DNA methylation: consequences for plant evolution." *Trends in Ecology & Evolution* 19.6 (2004): 309-314.
37. Kanyuka, Konstantin, *et al.* "Mutations in the huge *Arabidopsis* gene *BIG* affect a range of hormone and light responses." *The Plant Journal* 35.1 (2003): 57-70.
38. Kawakatsu, Taiji, *et al.* "Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions." *Cell* 166.2 (2016): 492-505.
39. Knapp, S., and Tomato Genome Consortium. "The tomato genome sequence provides insights into fleshy fruit evolution." (2012).
40. Laubinger, Sascha, *et al.* "Dual roles of the nuclear cap-binding complex and *SERRATE* in pre-mRNA splicing and microRNA processing in *Arabidopsis thaliana*." *Proceedings of the National Academy of Sciences* 105.25 (2008): 8795-8800.
41. Law, Julie A., *et al.*, "Polymerase IV occupancy at RNA-directed DNA methylation sites requires *SHH1*." *Nature* 498.7454 (2013): 385-389.
42. Law, Julie A., *et al.*, "*SHH1*, a homeodomain protein required for DNA methylation, as well as *RDR2*, *RDM4*, and chromatin remodeling factors, associate with RNA polymerase IV." *PLoS genetics* 7.7 (2011): e1002195.
43. Lee, J. M. "Vegetable grafting: a powerful aid for cultivation of environmentally-friendly produce." *Modern Science and Technology* 2008.4 (2008): 68-85.
44. Lee, Jiyeon, *et al.* "AP endonucleases process 5-methylcytosine excision intermediates during active DNA demethylation in *Arabidopsis*." *Nucleic acids research* 42.18 (2014): 11408-11418.

45. Lee, Jung-Myung, *et al.* "Current status of vegetable grafting: Diffusion, grafting techniques, automation." *Scientia Horticulturae* 127.2 (2010): 93-105.
46. Lee, Jung-Myung. "Grafting of herbaceous vegetable and ornamental crops." *Hortic. Rev.* 28 (2003): 61-124.
47. Li, Yan, *et al.* "An AP endonuclease functions in active DNA demethylation and gene imprinting in Arabidopsis." *PLoS genetics* 11.1 (2015): e1004905.
48. Li, Yan, Suresh Kumar, and Weiqiang Qian. "Active DNA demethylation: mechanism and role in plant development." *Plant cell reports* 37.1 (2018): 77-85.
49. Lichtenthaler, Hartmut K. "The stress concept in plants: an introduction." *Annals of the new York Academy of sciences* 851.1 (1998): 187-198.
50. Lindroth, Anders M., *et al.*, "Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation." *Science* 292.5524 (2001): 2077-2080.
51. Liu, Hongxia, and Sophia L. Stone. "Abscisic acid increases Arabidopsis ABI5 transcription factor levels by promoting KEG E3 ligase self-ubiquitination and proteasomal degradation." *The Plant Cell* 22.8 (2010): 2630-2641.
52. Liu, Jun, *et al.*, "An atypical component of RNA-directed DNA methylation machinery has both DNA methylation-dependent and-independent roles in locus-specific transcriptional gene silencing." *Cell research* 21.12 (2011): 1691-1700.
53. Livesey, Geoffrey. "Health potential of polyols as sugar replacers, with emphasis on low glycaemic properties." *Nutrition Research Reviews* 16.2 (2003): 163-191.
54. Manning, Kenneth, *et al.* "A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening." *Nature genetics* 38.8 (2006): 948-952.
55. Marí-Ordóñez, Arturo, *et al.*, "Reconstructing de novo silencing of an active plant retrotransposon." *Nature genetics* 45.9 (2013): 1029-1039.
56. Marsico, Marco Di, *et al.* "MCSeEd (Methylation Context Sensitive Enzyme ddRAD): A New Method to Analyze DNA Methylation." *Plant Epigenetics and Epigenomics*. Humana, New York, NY, 2020. 47-64.
57. Martínez-Macías, María Isabel, *et al.* "A DNA 3' phosphatase functions in active DNA demethylation in Arabidopsis." *Molecular cell* 45.3 (2012): 357-370.
58. McCue, Andrea D., *et al.*, "ARGONAUTE 6 bridges transposable element m RNA-derived siRNA s to the establishment of DNA methylation." *The EMBO journal* 34.1 (2015): 20-35.

59. McCullough, Amanda K., M. L. Dodson, and R. Stephen Lloyd. "Initiation of base excision repair: glycosylase mechanisms and structures." *Annual review of biochemistry* 68.1 (1999): 255-285.
60. Meissner, Alexander, *et al.* "Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis." *Nucleic acids research* 33.18 (2005): 5868-5877.
61. Melnyk, Charles W., *et al.* "Transcriptome dynamics at Arabidopsis graft junctions reveal an intertissue recognition mechanism that activates vascular regeneration." *Proceedings of the National Academy of Sciences* 115.10 (2018): E2447-E2456.
62. Messeguer, Ramon, *et al.* "Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA." *Plant molecular biology* 16.5 (1991): 753-770.
63. Molnar, Attila, *et al.* "Small silencing RNAs in plants are mobile and direct epigenetic modification in recipient cells." *science* 328.5980 (2010): 872-875.
64. Morales-Ruiz T, Ortega-Galisteo AP, Ponferrada-Marin MI, MartinezMacias MI, Ariza RR, Roldan-Arjona T (2006) DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. *Proc Natl Acad Sci USA* 103:6853–6858
65. Morra, L., and M. Bilotto. "Evaluation of new rootstocks for resistance to soil-borne pathogens and productive behaviour of pepper (*Capsicum annuum* L.)." *The Journal of Horticultural Science and Biotechnology* 81.3 (2006): 518-524.
66. Morrone, Dana, *et al.* "Characterization of the kaurene oxidase CYP701A3, a multifunctional cytochrome P450 from gibberellin biosynthesis." *Biochemical Journal* 431.3 (2010): 337-347.
67. Mounet, Fabien, *et al.* "Down-regulation of a single auxin efflux transport protein in tomato induces precocious fruit development." *Journal of experimental botany* 63.13 (2012): 4901-4917.
68. Mudge, Ken, *et al.* "A history of grafting." (2009).
69. Nagymihály, Marianna, *et al.* "Ploidy-dependent changes in the epigenome of symbiotic cells correlate with specific patterns of gene expression." *Proceedings of the National Academy of Sciences* 114.17 (2017): 4543-4548.
70. Naika, Shankara, *et al.* "Cultivation of tomato." *Production, processing and marketing, Agromisa/CTA. Revised edition* (2005).

71. Ortega-Galisteo, Ana Pilar, *et al.*, "Arabidopsis DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks." *Plant molecular biology* 67.6 (2008): 671-681.
72. Ossowski, Stephan, *et al.* "The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*." *science* 327.5961 (2010): 92-94.
73. Pina, Ana, and Pilar Errea. "A review of new advances in mechanism of graft compatibility–incompatibility." *Scientia Horticulturae* 106.1 (2005): 1-11.
74. Pyott, Douglas E., and Attila Molnar. "Going mobile: non-cell-autonomous small RNAs shape the genetic landscape of plants." *Plant biotechnology journal* 13.3 (2015): 306-318.
75. R. Quinlan, A. and Kindlon, N., 2019. *bedtools: a powerful toolset for genome arithmetic — bedtools 2.30.0 documentation*. [online] *Bedtools.readthedocs.io*. Available at: <<https://bedtools.readthedocs.io/en/latest/index.html>> [Accessed 14 December 2021].
76. Rao, A. Venket, M. R. Ray, and L. G. Rao. "Lycopene." *Advances in food and nutrition research* 51 (2006): 99-164.
77. Rivero, Rosa M., Juan M. Ruiz, and Luis Romero. "Role of grafting in horticultural plants under stress conditions." *Journal of food agriculture and environment* 1 (2003): 70-74.
78. Rivero, Rosa M., Juan M. Ruiz, and Luis Romero. "Role of grafting in horticultural plants under stress conditions." *Journal of food agriculture and environment* 1 (2003): 70-74.
79. Rivero, Rosa M., Juan M. Ruiz, and Luis Romero. "Role of grafting in horticultural plants under stress conditions." *Journal of food agriculture and environment* 1 (2003): 70-74.
80. Rocha, Pedro SCF, *et al.*, "The Arabidopsis HOMOLOG-DEPENDENT GENE SILENCING1 gene codes for an S-adenosyl-L-homocysteine hydrolase required for DNA methylation-dependent gene silencing." *The Plant Cell* 17.2 (2005): 404-417.
81. Romani, Massimo. "Epigenetica." (2019).
82. Sakata, Y., T. Ohara, and M. Sugiyama. "The history and present state of the grafting of cucurbitaceous vegetables in Japan." *III International Symposium on Cucurbits* 731. 2005.
83. Sanchez, Diego H., and Jerzy Paszkowski. "Heat-induced release of epigenetic silencing reveals the concealed role of an imprinted plant gene." *PLoS genetics* 10.11 (2014): e1004806.
84. Sanchez, Diego H., and Jerzy Paszkowski. "Heat-induced release of epigenetic silencing reveals the concealed role of an imprinted plant gene." *PLoS genetics* 10.11 (2014): e1004806.
85. Savvides, Andreas, *et al.*, "Chemical priming of plants against multiple abiotic stresses: mission possible?." *Trends in plant science* 21.4 (2016): 329-340.

86. Schwarz, Dietmar, *et al.* "Grafting as a tool to improve tolerance of vegetables to abiotic stresses: Thermal stress, water stress and organic pollutants." *Scientia Horticulturae* 127.2 (2010): 162-171.
87. Shen, Yuan, Emmanuelle Issakidis-Bourguet, and Dao-Xiu Zhou. "Perspectives on the interactions between metabolism, redox, and epigenetics in plants." *Journal of Experimental Botany* 67.18 (2016): 5291-5300.
88. Silveira, Amanda Bortolini, *et al.* "Extensive natural epigenetic variation at a de novo originated gene." *PLoS genetics* 9.4 (2013): e1003437.
89. Simonet, Nicolas G., George Rasti, and Alejandro Vaquero. "The Histone Code and Disease: Posttranslational Modifications as Potential Prognostic Factors for Clinical Diagnosis." *Epigenetic Biomarkers and Diagnostics*. Academic Press, 2016. 417-445.
90. Srikant, Thanvi, and Anjar Tri Wibowo. "The Underlying Nature of Epigenetic Variation: Origin, Establishment, and Regulatory Function of Plant Epialleles." *International Journal of Molecular Sciences* 22.16 (2021): 8618.
91. Suzuki, Masako, *et al.* "Whole-genome bisulfite sequencing with improved accuracy and cost." *Genome research* 28.9 (2018): 1364-1371.
92. Tang, Kai, *et al.* "The DNA demethylase ROS1 targets genomic regions with distinct chromatin modifications." *Nature plants* 2.11 (2016): 1-10.
93. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Research*, 2021, 49.D1: D325-D334.
94. Thomas, Hannah R., and Margaret H. Frank. "Connecting the pieces: uncovering the molecular basis for long-distance communication through plant grafting." *New Phytologist* 223.2 (2019): 582-589.
95. Thomas, P., 2021. *PantherDB.org*. [online] Pantherdb.org. Available at: <<http://www.pantherdb.org/about.jsp>> [Accessed 14 December 2021].
96. Unoki, Motoko. "Recent Insights into the mechanisms of de novo and maintenance of DNA methylation in mammals." *DNA methylation mechanism* (2019).
97. Van Der Graaf, Adriaan, *et al.* "Rate, spectrum, and evolutionary dynamics of spontaneous epimutations." *Proceedings of the National Academy of Sciences* 112.21 (2015): 6676-6681.
98. Venema, Jan Henk, *et al.* "Grafting tomato (*Solanum lycopersicum*) onto the rootstock of a high-altitude accession of *Solanum habrochaites* improves suboptimal-temperature tolerance." *Environmental and Experimental Botany* 63.1-3 (2008): 359-367.

99. Wu, Rui, *et al.* "Inter-species grafting caused extensive and heritable alterations of DNA methylation in Solanaceae plants." *PLoS One* 8.4 (2013): e61995.
100. Wysocka-Diller, Joanna W., *et al.* "Molecular analysis of SCARECROW function reveals a radial patterning mechanism common to root and shoot." *Development* 127.3 (2000): 595-603.
101. Xu, XinCheng, *et al.* "Alterations of DNA methylation in diverse grafted hybrid tomatoes (*Solanum lycopersicum* L.)." (2016).
102. Yamamuro, Chizuko, Jian-Kang Zhu, and Zhenbiao Yang. "Epigenetic modifications and plant hormone action." *Molecular plant* 9.1 (2016): 57-70.
103. Yong, Wai-Shin, Fei-Man Hsu, and Pao-Yang Chen. "Profiling genome-wide DNA methylation." *Epigenetics & chromatin* 9.1 (2016): 1-16.
104. Yong-Villalobos, Lenin, *et al.* "Methylome analysis reveals an important role for epigenetic changes in the regulation of the Arabidopsis response to phosphate starvation." *Proceedings of the National Academy of Sciences* 112.52 (2015): E7293-E7302.
105. Yu, Agnès, *et al.* "Dynamics and biological relevance of DNA demethylation in Arabidopsis antibacterial defense." *Proceedings of the National Academy of Sciences* 110.6 (2013): 2389-2394.
106. Zhang, Bo, *et al.* "Chilling-induced tomato flavor loss is associated with altered volatile synthesis and transient changes in DNA methylation." *Proceedings of the National Academy of Sciences* 113.44 (2016): 12580-12585.
107. Zhang, Huiming, Zhaobo Lang, and Jian-Kang Zhu. "Dynamics and function of DNA methylation in plants." *Nature reviews Molecular cell biology* 19.8 (2018): 489-506.
108. Zhang, Xiaoyu, *et al.*, "Genome-wide analysis of mono-, di-and trimethylation of histone H3 lysine 4 in Arabidopsis thaliana." *Genome biology* 10.6 (2009): 1-14.
109. Zhang, Xingliang, *et al.* "Structural insights into the abscisic acid stereospecificity by the ABA receptors PYR/PYL/RCAR." *PloS one* 8.7 (2013): e67477.
110. Zheng, Binglian, *et al.*, "Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in Arabidopsis." *Genes & development* 23.24 (2009): 2850-2860.
111. Zhong, Xuehua, *et al.*, "Molecular mechanism of action of plant DRM de novo DNA methyltransferases." *Cell* 157.5 (2014): 1050-1060.
112. Zhou, Hao-Ran, *et al.*, "Folate polyglutamylation is involved in chromatin silencing by maintaining global DNA methylation and histone H3K9 dimethylation in Arabidopsis." *The Plant Cell* 25.7 (2013): 2545-2559.

113. Zhu, Jian-Kang. "Active DNA demethylation mediated by DNA glycosylases." *Annual review of genetics* 43 (2009): 143-166.