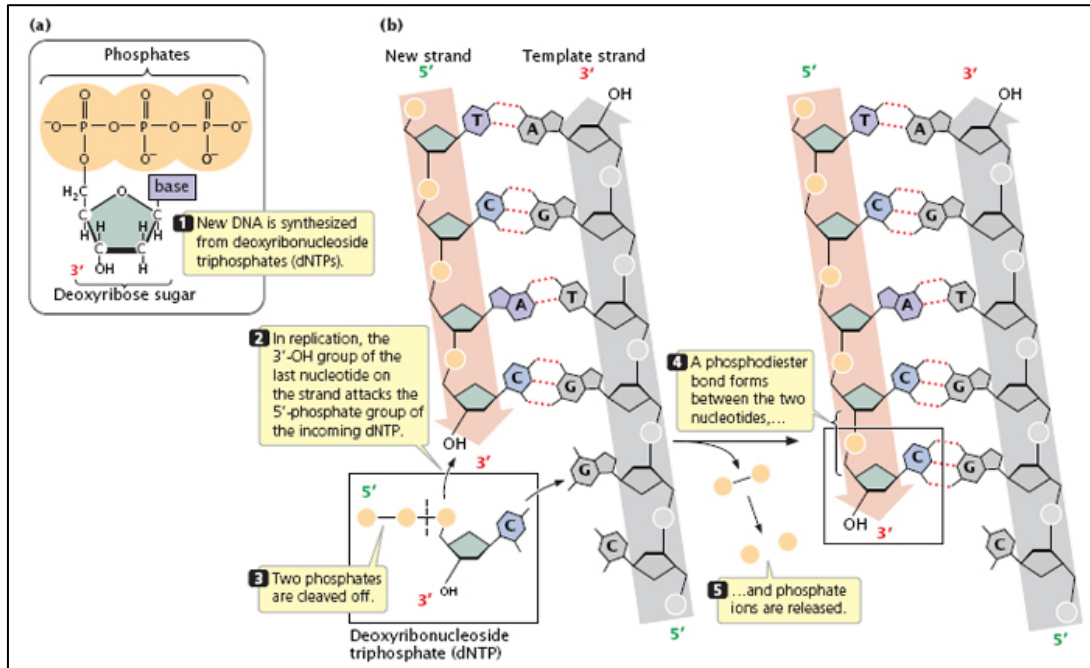


## PCR, Polymerase Chain Reaction (Reazione a catena della DNA polimerasi)

La PCR è una tecnica di biologia molecolare per replicare ripetutamente (amplificare), in modo estremamente selettivo, un tratto definito di DNA del quale si conoscano le sequenze nucleotidiche iniziali e terminali, partendo da una soluzione di DNA (nucleare, organellare, plasmidico, virale ecc.) in cui questo tratto di DNA è presente in singola copia o in un numero esiguo di copie. In questo modo è possibile isolare e studiare un qualsiasi tratto di DNA a partire da un campione biologico da cui è possibile recuperare tracce di DNA. Attraverso questa reazione, a partire da una singola molecola di DNA che fa da stampo, si possono generare miliardi di frammenti identici di uno stesso tratto di DNA in poco più di un'ora!

Tale metodica fu ideata nel 1983 da Kary B. Mullis il quale ottenne, per questo, il Premio Nobel per la chimica (1993).

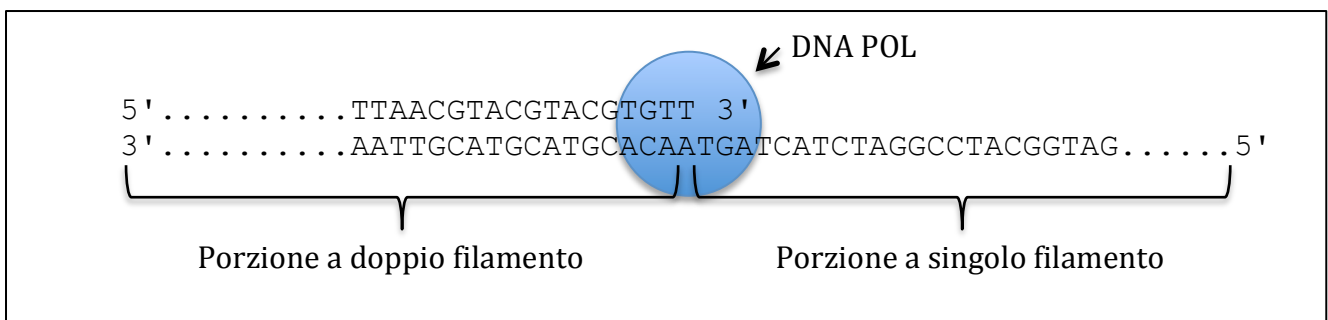
La reazione è catalizzata dall'enzima **DNA polimerasi** che, *in vivo*, è responsabile della replicazione del DNA che avviene durante la mitosi. La PCR ricostruisce *in vitro* uno specifico passaggio della duplicazione cellulare: la ricostituzione (sintesi) di un segmento di DNA "completo" (a doppia elica) a partire da un filamento a singola elica. Il filamento mancante viene ricostruito a partire da una serie di nucleotidi (i "mattoni" elementari che costituiscono gli acidi nucleici) che vengono disposti nella corretta sequenza, complementare a quella del DNA interessato (v. fig seguente).



Schema che raffigura l'incorporazione di un nuovo nucleotide da parte della DNA Polimerasi.

Per iniziare la reazione occorre che il DNA stampo venga denaturato, cioè le due singole eliche che costituiscono il DNA devono essere completamente separate; la DNA polimerasi per poter funzionare ha bisogno di un innesco, ovvero di una regione a doppio filamento seguita da una regione a singolo filamento (v. fig. seguente).

L'enzima, a partire dall'innesco, incorpora un nucleotide dopo l'altro, secondo la complementarità delle basi dell'elica a singolo filamento (che fa da stampo), muovendosi in direzione 5' > 3'.



Esempio di innesco per la DNA Polimerasi



determina un allungamento dei primers legati, utilizzando come stampo il filamento singolo di DNA (fase di polimerizzazione o di prolungamento).

Poiché molte DNA polimerasi, tra cui anche quella umana, non possono resistere alle temperature necessarie per la denaturazione si fa ricorso alle polimerasi appartenenti a organismi termofili che non sono inattivate dalle alte temperature, ad esempio la Taq polimerasi proveniente dal batterio termofilo *Thermus aquaticus*.

Il ciclo descritto viene ripetuto generalmente per circa 25-35 volte.

Dopo il primo ciclo da una singola molecola di DNA se ne originano 2, dopo il secondo ciclo da queste se ne originano 4, dopo il terzo da queste se ne originano 8 e così via, in modo geometrico. Dopo n cicli da una singola molecola di DNA si otterranno  $2^n$  frammenti di DNA compresi tra i due oligonucleotidi. Dopo 30 cicli teoricamente si possono ottenere  $2^{30}$  ovvero oltre un miliardo di molecole di DNA identiche!

L'intera procedura è assai rapida poiché ciascun ciclo dura solo pochi minuti, e per essere realizzata si utilizza un *Termociclatore* ovvero uno strumento in grado di programmare nel tempo, in modo automatico, la temperatura.

Una reazione di PCR avviene all'interno di microprovette da 200  $\mu$ l che vengono inserite all'interno del termociclatore.

All'interno di ciascuna microprovetta vengono inseriti i seguenti "ingredienti" in soluzione:

- una quantità, anche minima, del segmento di DNA che si desidera riprodurre;
- una quantità opportuna di nucleotidi liberi (desossiribonucleotidi trifosfati, dNTP) per costituire i nuovi filamenti;
- opportuni primers, costituiti da brevi sequenze di DNA (oligonucleotidi) complementari agli estremi 3' dei due filamenti del segmento da riprodurre;
- una DNA polimerasi termo-resistente
- un Buffer che serve a mantenere il pH stabile (tampone) e necessario per costituire l'ambiente adatto alla reazione;
- altri elementi di supporto (ad es. ioni magnesio) indispensabili per il corretto funzionamento della DNA polimerasi;
- acqua per portare a volume la soluzione.

In questo sito è possibile visualizzare un'animazione della PCR:  
[http://www.karymullis.com/pcr\\_vid.swf](http://www.karymullis.com/pcr_vid.swf)

### **Possibili applicazioni della PCR:**

-In caso di malattie genetiche o tumorali mediante la PCR è possibile amplificare il gene responsabile di tali stati patologici (ovviamente il gene in questione deve essere stato già riconosciuto, mentre in caso di malattie infettive si possono amplificare geni del microrganismo in questione che codifichino per funzioni vitali essenziali o per fattori di virulenza.

- Dal momento che molte sequenze geniche presentano una elevata conservazione di sequenza in specie differenti, sintetizzando gli oligonucleotidi in modo che questi si possano appaiare a due regioni conservate del gene, è possibile isolare un gene omologo in una specie in cui quel gene non è stato ancora caratterizzato.

- E' possibile amplificare porzioni di DNA "antico" (isolato da reperti fossili, mummie ecc.) in modo da effettuare confronti con campioni contemporanei di DNA per studi evoluzionistici.

- E' possibile amplificare numerose porzioni del DNA genomico di un genotipo (una specie, una varietà o un individuo) per mettere a confronto genotipi diversi, ottenendo per ciascuno una "impronta digitale molecolare", utile ad esempio in medicina legale, diagnostica, tassonomia, riconoscimento varietale.

- Il bersaglio da amplificare può anche essere una molecola di RNA (come, ad esempio, nel caso di alcuni virus o di trascritti genici) la quale deve essere, come primo passo, sottoposta ad una reazione di retrotrascrizione in cui a partire da molecole di RNA si possono ottenere molecole di cDNA che possono essere amplificate tramite la PCR (RT-PCR); ciò consente di stabilire se un determinato gene è trascritto, ovvero è espresso, e in che misura, nonché di conoscerne la sequenza codificante.

### **Bibliografia:**

Mullis KB, Faloona FA. "Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction". *Methods Enzymol.* 1987;155:335-50.

Russell P.J. "Genetica" EdiSES 1994 pp.473-475.

**In pratica:****Mix PCR (conc. finali):**

4 ng/ $\mu$ l DNA stampo  
 1X Buffer (fornito con la Taq)  
 2.5 mM MgCl<sub>2</sub>  
 0.2 mM ciascun dNTP  
 1  $\mu$ l oligonucleotide A  
 1  $\mu$ l oligonucleotide B  
 0.05 U/ $\mu$ l Taq (enzima DNA polimerasi)

**Mix PCR da 25  $\mu$ l:**

0.5  $\mu$ l DNA stampo (200 ng/ $\mu$ l)  
 2.5  $\mu$ l (Buffer 10X)  
 2.5  $\mu$ l (25 mM)  
 2  $\mu$ l (2.5 mM ciascuno)  
 0.25  $\mu$ l (100  $\mu$ M)  
 0.25  $\mu$ l (100  $\mu$ M)  
 0.25  $\mu$ l (5 U/ $\mu$ l)  
 16.75  $\mu$ l H<sub>2</sub>O

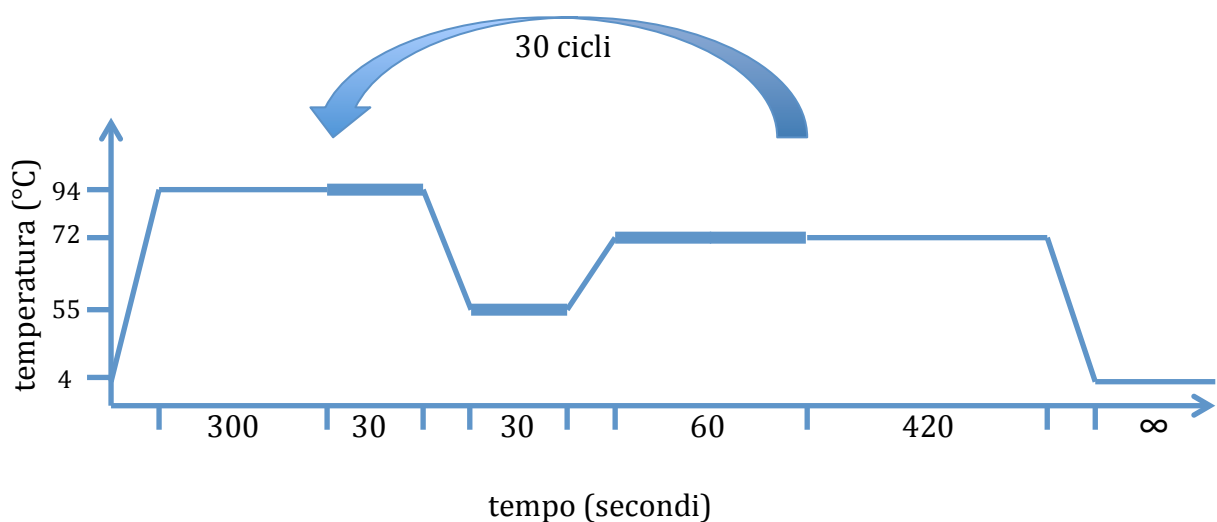
**ESEMPIO DI PROGRAMMAZIONE PCR**

-Denaturazione iniziale: 5 min a 94°C (serve per denaturare completamente il DNA genomico)

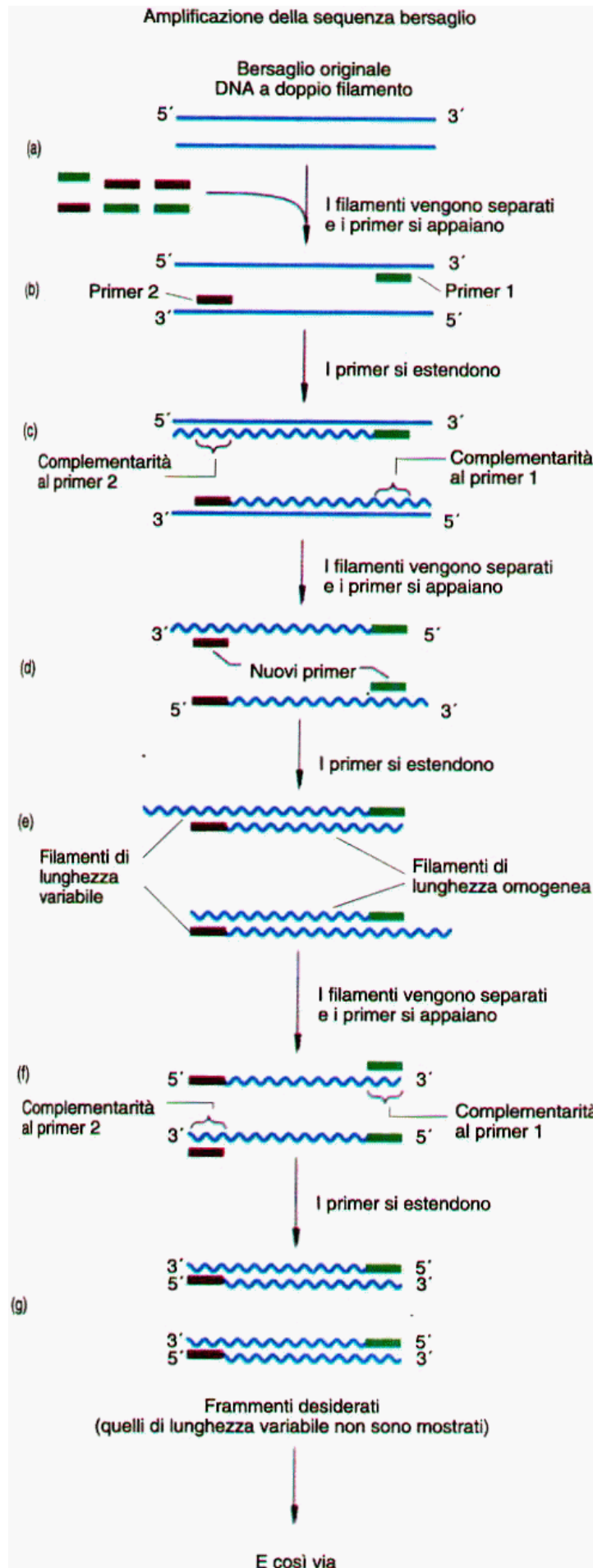
-30 cicli:

- 1) Denaturazione a 94°C 30 sec.
- 2) Appaiamento primers a 55°C per 30 sec.
- 3) Polimerizzazione a 72°C per 60 sec. (60 sec. ogni 1000 bp)

-Polimerizzazione finale a 72°C per 7 min. (per consentire di ultimare tutti i frammenti amplificati).





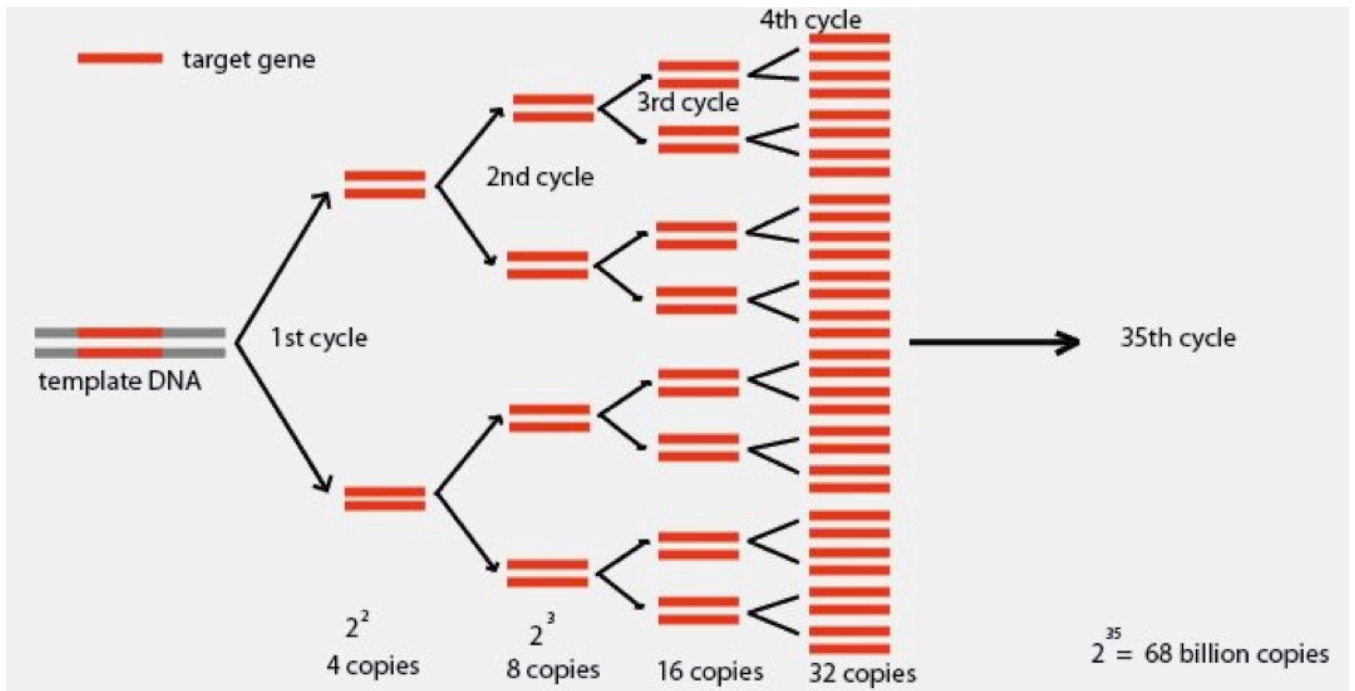


ciclo 1

ciclo 2

ciclo 3





## Caso pratico: vogliamo isolare, tramite PCR, un frammento di gene codificante una deidrina di girasole (*Helianthus annuus*)

Le deidrine sono proteine ubiquitarie delle piante che vengono sintetizzate principalmente in condizioni di deficit idrico. Ad esempio si accumulano nelle foglie di piante sottoposte a stress idrico (mancanza di irrigazione) o nel seme durante la progressiva disidratazione che avviene nelle fasi tardive della sua maturazione. Sono proteine idrofili che si legano a macromolecole cellulari (enzimi, acidi nucleici, membrane) e ne preservano la struttura nel corso della carenza idrica che si registra a livello cellulare.

1) Attraverso una ricerca per parole chiave ("dehydrin" + specie) si ricercano sequenze proteiche depositate presso il database dell'NCBI (National Center for Biotechnology Information <http://www.ncbi.nlm.nih.gov/>)

The screenshot shows the NCBI website interface. At the top, there is a search bar with the text "Protein" and "dehydrin daucus carota" entered. Below the search bar, there is a "Welcome to NCBI" section with a navigation menu on the left. The menu includes items like "NCBI Home", "Resource List (A-Z)", "All Resources", "Chemicals & Bioassays", "Data & Software", "DNA & RNA", "Domains & Structures", "Genes & Expression", "Genetics & Medicine", "Genomes & Maps", "Homology", "Literature", "Proteins", "Sequence Analysis", "Taxonomy", "Training & Tutorials", and "Variation". The main content area features a "Get Started" section with links to "Tools", "Downloads", "How-To's", and "Submissions". There is also a "Genomic Structural Variation" section with an image of a corn cob. On the right side, there are "Popular Resources" and "NCBI Announcements" sections.

2) Si copiano diverse sequenze proteiche in formato fasta ( >nome accapo SEQUANZA)

### dehydrin protein [Daucus carota]

```

GenBank: BAD06644.1
FASTA  Graphics

Go to:

LOCUS      BAD06644                149 aa            linear   PLN 12-JAN-2005
DEFINITION dehydrin protein [Daucus carota].
ACCESSION  BAD06644
VERSION    BAD06644.1 GI:57566540
DBSOURCE   accession AB105039.1
KEYWORDS   .
SOURCE     Daucus carota (carrot)
  ORGANISM  Daucus carota
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; Gunneridae;
            Pentapetales; asterids; campanulids; Apiales; Apiaceae; Apioidae;
            Scandiceae; Daucinae; Daucus.
REFERENCE  1
AUTHORS    Shiota,H., Yang,G., Shen,S., Eun,C.H., Matabe,K., Tanaka,I. and
            Kanada,H.
TITLE      Isolation and characterization of six abscisic acid-inducible genes
            from carrot somatic embryos
JOURNAL    Plant Biotechnol. 21, 309-314 (2004)
REFERENCE  2 (residues 1 to 149)
AUTHORS    Shiota,H.
TITLE      Direct Submission
JOURNAL    Submitted (05-MAR-2003) Hajime Shiota, Yokohama City University,
            Graduate School of Integrated Science; Seto 22-2, Kanazawa-ku,
            Yokohama, Kanagawa 214-0021, Japan
            (E-mail:shiota@yokohama-cu.ac.jp, Tel:01-45-787-2318,
            Fax:01-45-787-2413)
FEATURES   Location/Qualifiers
            source          1..149
                        /organism="Daucus carota"
                        /cultivar="DE-Harusakigomun"
                        /db_xref="taconid:639"
                        /tissue_type="somatic embryo"
                        /dev_stage="embryo"
            Protein         1..149
                        /product="dehydrin protein"
            Region          18..146
                        /region_name="Dehydrin"
                        /note="Dehydrin; pI=50257"
                        /db_xref="CDD:248731"
            CDS             1..149
                        /gene="CAISE1"
                        /coded_by="AB105039.1:69..518"
ORIGIN
1      1aaagagggg qkhhqsdedy gnpvrcldy gnpvqhttg tgnhtgat gskpvlst
61      fdgprlra gaxaxasad dngprckkv gskqkikgl pghkhdqgt qttipwagp
121     fhatgtawp skkqindbi kdkppgh
//

```

```

>Coffea canephoraDQ323987.1
maqygaeygn qksqydeygn pvrqtdeygn parhggmtgd ygttggtgay ggttgahgty      61 atgttgtgtt gayatqpgtd
vgkehhlglg mhrsgsgss ssseddgqgg rrrkkmkeki      121 keklpgghke aqpgqeyssa taapgyggeg vqhekkgimd
kikeklpggh hn
>Nicotiana tabacumAB049336.1
mshydnqfsa ggalqtdeyg npirqtdeyg npvhhtggtm gdygttggtg aythagga      61 ghttgilgge hrpghehgtl
ggmlhrsgss ssssssedd gqgrrkkkg mkekikeklp      121 gghkddqths tatttttgyg megehhekk gimdkikekl
pghhgpggh
>Daucus carotaAB105039.1
masygegyg gqkhhgsdey gnpvrqtdey gnpvqhttg tmhgtgatt gahgvglstg      61 sdgqprlnrs gsssssssed
dgmgrrrkkv gmkqkikgml pggkhdqqt qtttvmggs      121 ygatgtgeqp eekkgimdki kdklpgggh
>Oryza sativaNM_001050530.1
maehatgvyg hpyprvdqyg npvpvdqyg npvdepapr dtaagyvapp dpavstgdyg      61 lagaeaphph esavmsgaaa
aavapggeay trdgggvvpp agektfayeg tvsaagvtga      121 sqqlqpttre eghttlgetl rrrsgksssss sssseddgqg
grrkkksike kikeklpgsh      181 kqeeqkqagh tapaagtgtg tgthaagkhe kkgivekike klpghghh
>Hordeum vulgareAF181461.1
madyggygh pyprvdeygn pppvdqygn piprepgqap aygsagavtp aeygagvvp      61 gyglgavhp hesvvggavp
psgaayaheg avsgglapge ttayayegav sgglapgett      121 ayayegmvs gigagagdri qpakeehttl gealrrsgss
sssssssed dgqgrqrkk      181 kikekikek lpgnqkheeh kaghaapaa gtgthekkgi mekikeklpg hh
>Triticum aestivumX59133.1
mefqgghdnp anrvdeygnp fplagawger trsrhrravp gpqgraqdrw ilhrsgssss      61 sssseddmg grrkkmkek
ikeklpgghk dnqghmatgt gtggaygpgt gtggaygqqg      121 htgmagagtg tgekkgimd kikeklpggh
>Glycine maxNM_001250385.1
maeaqlrdqh gnpvpltdqy gnpviltder gnpvqltva ttatgtagsg fgsygtgag      61 ggasattvad llatqprsr
earelrrsss ssssssedd gqgrrkkgvk dkikeklpgv      121 gggnnkeha httapttat nhpadqhekk gimerikekl pghhth
>Brassica napusAY303803.1
madlkdergn pihltdehgn pvqltdefgn pmhitgvass apqykesvtg niqeyrtaap      61 pagvaagtgv aattaagvat
getttgqqqh heslgehlrr sgsssssse ddgqgrrkk      121 gmkdikekl sggkhkdeqt pstatattgpt tttgaaaadq
hhekkgilek ikeklpgghn      181 hhp
    
```

3) Si compie un multiallineamento delle sequenze proteiche attraverso il **ClustalW** (<http://www.genome.jp/tools/clustalw/>)



**Multiple Sequence Alignment by CLUSTALW**

CLUSTALW	MAFFT	PRRN
<a href="#">Help</a>		
<b>General Setting Parameters:</b>		
Output Format: <input type="text" value="CLUSTAL"/>		
Pairwise Alignment: <input type="radio"/> FAST/APPROXIMATE <input checked="" type="radio"/> SLOW/ACCURATE		
Enter your sequences (with labels) below (copy & paste): <input checked="" type="radio"/> PROTEIN <input type="radio"/> DNA		
Support Formats: FASTA (Pearson), NBRF/PIR, EMBL/Swiss Prot, GDE, CLUSTAL, and GCG/MSF		
>Coffea canephoraDQ323987.1 maqygaeygn qksqydeygn pvrqtdeygn parhggmtgd ygttggtgay ggttgahgty 61 atgttgtgtt gayatqpgtd vqkehhlglg mhrsgsgss ssseddgqgg rrrkkmkeki 121 keklpgghke aqpgqeyssa taapgyggeg vqhekkgimd kikeklpggh hn >Nicotiana tabacumAB049336.1		
<b>Or give the file name containing your query</b>		
<input type="button" value="Scegli file"/> Nessun file selezionato		
<input type="button" value="Execute Multiple Alignment"/> <input type="button" value="Reset"/>		

**Allineamento, tramite ClustalW, di alcune sequenze aminoacidiche codificanti deidrine**

```

Coffea_canephora      -MAQYGAHEYGN-QKSQYDEYGNPVVRQTDEYGNPPARHGG-TMGDYGTTGTT
Nicotiana_tabacum    -MSHYDNQFSAGQALQDEYGNPIRQTDEYGNPVHHTGGTMGDYGTTGTG
Daucus_carota        MASYGEGGYGGQKHHGDEYGNPVVRQTDEYGNPVQHTTGGTGMHGTGATT
Oryza_sativa         -MAEHATGVYGHYPYPRVDQYGNPVVPVDQYGNPVVPDEP--APRDTAAGYV
Hordeum_vulgare     --MADYGGEYGHYPYPRVDEYGNPVVPVDQYGNPI PREPGQAPAYGSAGAV
Triticum_aestivum   ---MEFQGQHNDNPANRVDEYGNPFPF LAGAWGERTRSRHRRAVPGPQGRAQ
Glycine_max__soybean_ MAEAQLRDQHGNPVPVLTDQYGNPVLTDERGNPVQLTGVATTATGTAGSG
_Brassica_napus     --MADLKDERGNPIHLTDEHGNPVLTDEFGNPMHITGVASSAPQYKESV
                    *::*** .. *:
    
```

```

Coffea_canephora      G-----
Nicotiana_tabacum    G-----
Daucus_carota        G-----
Oryza_sativa         APPDPAVSTGDYGLAGAEAPHPHESAVMSGAAAAAVAPGGE-----
Hordeum_vulgare     TPAAEYGAGVKPGYGLGGAVHPHESVVGGAVPPSGAAYAHEGAVSGLAP
Triticum_aestivum   DR-----
Glycine_max__soybean_ FG-----
_Brassica_napus     TG-----
    
```

```

Coffea_canephora      -----AYGGTTGTAGAHGTYATGTTGTTGTGAYATQPGTDVGKE
Nicotiana_tabacum    -----AYGTHAGGGAGHTTGILGGEHRP-----GHE
Daucus_carota        -----AHGVGLSTGSDGQP-----
Oryza_sativa         -----AYTRDGGGVVPAGEKTFAYEGTVSAAGVTGASGQLQPTTREEG
Hordeum_vulgare     GETTAYAYEGGAVSGLAPGETTAYAYEGGMVSGIGAGAGDRIQP--AKEE
Triticum_aestivum   -----
Glycine_max__soybean_ -----SYGTG-----AYGGGASATTVADLLATQPRSGRE--
_Brassica_napus     -----NIQEYRTAAPPAGVAAGTGVAATTAAGVATGETTTGQQQH
    
```

```

Coffea_canephora      RHGLGMLHRSGSGSSSSSS--EDDGQGGR-RKK-GMKEKIKELPGGHK
Nicotiana_tabacum    HGTLGLMLHRSGSSSSSSSS--EDDGQGGR-RKKKGMKEKIKELPGGHK
Daucus_carota        -----RLNRSGSSSSSSSS--EDDGMGG-RKKVGMKQKIKGMLPGGHK
Oryza_sativa         HTTLGETLRRSGSSSSSSSSSEDDGQGGR-RKKKSIKEKIKELPGSHK
Hordeum_vulgare     HTTLGEALRRSGSSSSSSSSSEDDGQGGRQRKKKGIKEKIKELPGNQK
Triticum_aestivum   -----WILHRSGSSSSSSSS--EDDGMGG--RKKGMKEKIKELPGGHK
Glycine_max__soybean_ -----ARELRRSGSSSSSSSS--EDDGQGGR--RKKGVKDKIKELPGVGG
_Brassica_napus     HESLGEHLRRSGSSSSSSSS--EDDGQGGR--RKKGMKDKIKELSGGKH
                    *.**...*****   **** **  :* .:*** **
    
```

```

Coffea_canephora      EAQPGQEYSSATAAPG-----YGGEGEQHEKKGIMLK
Nicotiana_tabacum    DDQTHSTATTTTTGYG-----MEG-EHHHEKKGIMDF
Daucus_carota        EDQQTQTTTTVPMGGSYGA-----TGTGEQHEKKGIMDK
Oryza_sativa         QEEQKQAGHTAPAAGTGTGT-----GTHAAGHEKKGIVEK
Hordeum_vulgare     HEEHKAGHAAAPAAGTGT-----HEKKGIMEK
Triticum_aestivum   DNQQHMATGTGTGGAYGPGTGTGGAYGQQGHTGMAGAGTGTHEKKGIMDK
Glycine_max__soybean_ GNNNKEHAHTTTAPTAT-----NHPADQHEKKGIMER
_Brassica_napus     KDEQTPSTATTTGPTTTTGA-----AAADQHHEKKGILEK
                    : .
                    *****:
    
```

```

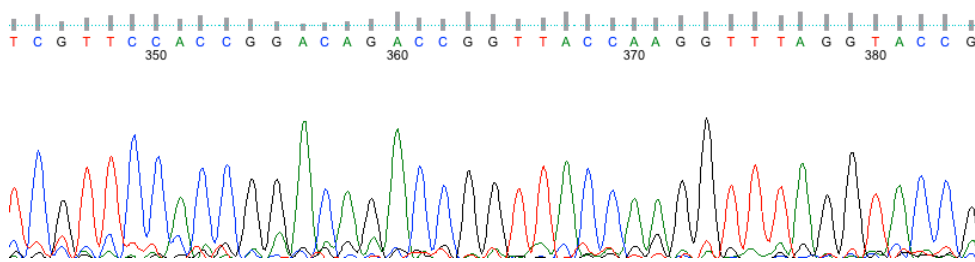
Coffea_canephora      IKEKLPGGHRN---
Nicotiana_tabacum    IKEKLPPGHHGPGHH
Daucus_carota        IKDKLPPGGHH----
Oryza_sativa         IKEKLPPGHGH---
Hordeum_vulgare     IKEKLPPGHH----
Triticum_aestivum   IKEKLPQG-----
Glycine_max__soybean_ IKEKLPGHHTH---
_Brassica_napus     IKEKLPGHNNHHP-
                    **:****
    
```



Si identificano possibili regioni su cui progettare i primers da utilizzare in PCR, utilizzando come stampo DNA genomico di girasole.

6) Si progettano i primers e si effettua una PCR. Si visualizza su gel di agarosio il prodotto di PCR ottenuto.

7) Il frammento di DNA relativo al prodotto di PCR viene sequenziato. La sequenza viene visualizzata attraverso un elettroferogramma (v. fig. seguente).

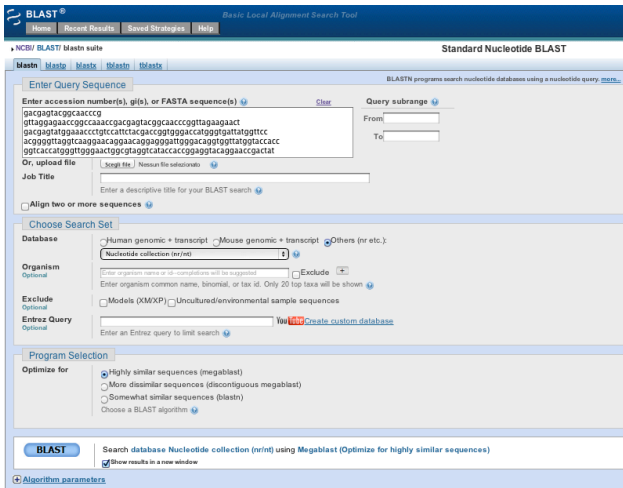


Sequenza nucleotidica parziale di un gene codificante una deidrina di girasole (782 bp) ottenuta, tramite PCR, utilizzando oligonucleotidi progettati su sequenze conservate individuate dal multi-allineamento di sequenze provenienti da altre specie.

**gacgagtacggcaaccg**

gttaggagaaccggcceaaccgacgagtacggcaaccgggttagaagaactgacgagtatggaaaccctgtccattctacgac  
 cggtaggaccatgggtgattatggtccacggggttaggtcaaggaacaggaacaggaggattgggacaggtgggtatggta  
 ccaccggtcaccatgggttgggaactggcgttaggtcataaccacggagggtacaggaaccgactatacctccgggtggtcgttc  
 accggacagaccggttaccaaggttaggtaccgagctctgagttcgggtggaagaactggaaccttccaaaaccaaccagtg  
 cacacctgtaggcggtggtgggtgagctcaggaactggggcgggtggtggagggtacaggaactggaacagggcattctgcatc  
 gttctggaagtagcagctcaagctcg**gtaagtatacaagaaaaaaaacttgctattatTTTTTatatacaaggagatagaga**  
**ggctgcttttagttaaatTTTTTaatagTTTTTgaaaaaaaaaccgTTTatgTTTTTgTTTaatTTTatgatctcattgattga**  
**tatgtag**tcggaggatgatgggcaaggaggaaggaggaagaagaagggtgatgcaaaagatcaaagagaagcttccgggtg  
 gtcaacagacaagaggagcagtatcagagccagacgactactactacgacgggtggcggggcgggttatggtgaaaccac  
**gagaagaaagggatgatgga**

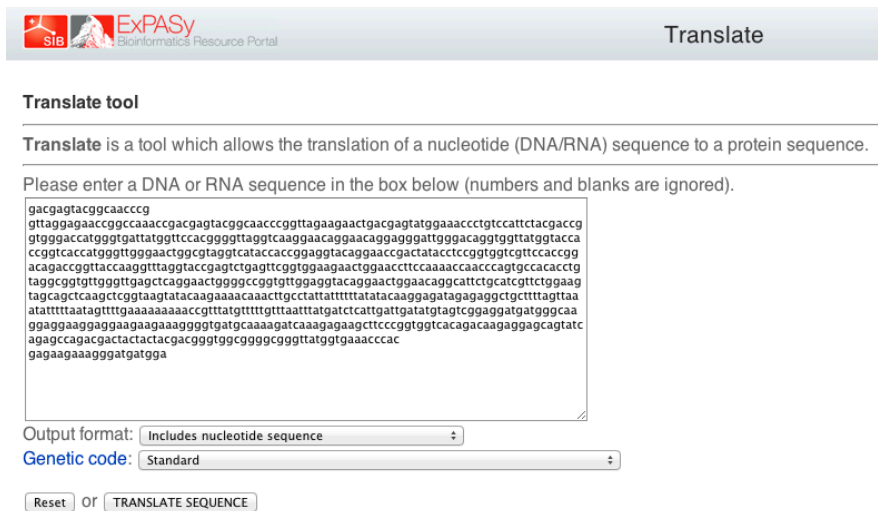
8) La sequenza nucleotidica viene analizzata con **BLASTN** (Basic Local Alignment Search Tool, <http://blast.ncbi.nlm.nih.gov/Blast.cgi>) per verificare che il frammento corrisponda effettivamente al gene desiderato.



Helianthus debilis partial sDhn1 gene for putative dehydrin, exons 1-2 subsp. silvestris  
Sequence ID: [embIAJ249710.1](#) Length: 1020 Number of Matches: 1

Range 1: 105 to 825		GenBank	Graphics	Next Match	Previous Match
Score	Expect	Identities	Gaps	Strand	
815 bits(441)	0.0	663/761 (87%)	51/761 (6%)	Plus/Plus	
Query 33	CCAAACCGACGAGTACCGCAACCCGGTTAGAAAGACTGACGAGTATGGAAACCTGTCCA				92
Sbjct 105	CCAAACTGACGAGTACCGCAACCCAGTTAGGAGAACCGACGAGTATGGAAACCTGTCCA				164
Query 93	TTCTACGACCGGTGGGACATGGGTGATTATGGTTCCAGGGTTAGGTCAAGGAAACAGG				152
Sbjct 165	TTCTACGACCGGTGGGACATGGGTGATTATGGTTCCAGGGTTAGG-C---CG---AGG				218
Query 153	AACAGGAGGATGGGACCGTGGTATGGTACACCGTCCACATGGGTGGGAACTGG				212
Sbjct 219	AACAGGAGGATGGGACCGTGGTATGGTACACCGTCCACATGGGTGGGAACTGG				278
Query 213	CGTAGGTACATACCAGGAGTACAGGAACCGACTATACCTCCGGTGGTTCACACGG				272
Sbjct 279	CATAGGTACATACCAGGAGTACAGGAACCGACTATACCTCCGGTGGTTCACACGG				338
Query 273	ACAGACCGTTACCAAGGTTTAGTACCGAGTCTGAGTCCGGTGGGAACTGGAACCTT				332
Sbjct 339	ACAGACCGTTACCAAGGTTTAGTACCGAGTCTGAGTCCGGTGGGAACTGGAACCTT				386
Query 333	CCAAAACCAACCCAGTGGCCACCTGTAGGCGGTGTGGTGGTGGTCCAGGAACTGGGGC				392
Sbjct 387	CCAAAACCAACCCAGTGGCCACCTGTAGGCGGTGTGGTGGTGGTCCAGGAACTGGGGC				446
Query 393	CGGTGTGGAGGTAC---A---GGAACTGGAAAGGCAATTCCTGCATGTTCCGAGATAG				446
Sbjct 447	CGGTGTGGAGGTACCGGCATGGAAAGGCAAGGCAATTCCTGCATGTTCCGAGATAG				506
Query 447	CAGCTCAGCTCGGTAGTATACAGGAAACCAACTGCTATATTTTTATATACAG				506
Sbjct 507	CAGCTCAGCTCGGTAGT-TA-ATTA---C---TT-C---AT-ATCCT-----AG				544
Query 507	GAGTAGAGAGGCTGCTTTAGTAAATATTTTAA-A-TAGTPTTGGAAAAAACCCGTT				564
Sbjct 545	GAGTAGAGAGGCTGCTTTAGTAAATATTTTAA-TAGTPTTGGAAAAAACCCGTT				604
Query 565	TATGTTTTTGTAAATT--T-ATGATCTCATGATGATGATGATGATGATGATGATG				621
Sbjct 605	TATGTTTTTGTAAATTAAATATGATCTCATGATGATGATGATGATGATGATGATG				664
Query 622	CAAGGAGGAGGAGGAGGAAAGGGGTGATGCAAAAGATCAAAAGGAAAGCTCCCGGT				681
Sbjct 665	CAAGGAGGAGGAGGAGGAAAGGGGTGATGCAAAAGATCAAAAGGAAAGCTCCCGGT				724
Query 682	GTCACAGACAGAGGAGGAGTATCAGAGGACGACTACTACTACTACTACTACTACTACT				741
Sbjct 725	GTCACAGTCAAGAGGACAGTATCAGAGGACGACTACTACTACTACTACTACTACTACT				784
Query 742	CGGGTGTATGTTGAAACCCAGAGAAAGGGATGATGGA 782				
Sbjct 785	CGGGTGTACTGTGAAACCCAGAGAAAGGGATGATGGA 825				

Inoltre attraverso il software **Expasy** (<http://web.expasy.org/translate/>) è possibile anche ricavare la sequenza aminoacidica a partire da quella nucleotidica.



Sequenza nucleotidica con, in basso, la traduzione ovvero la sequenza aminoacidica corrispondente ottenuta con Expasy .

**gacgagtacggcaaccg**  
**D E Y G N P**  
 gttaggagaaccggccaaccgacgagtacggcaaccgggttagaagaact  
 V R R T G Q T D E Y G N P V R R T  
 gacgagtatggaaccctgtccattctacgaccggggaccatgggtgattatggttcc  
 D E Y G N P V H S T T G G T M G D Y G S  
 acggggttagtcaaggaacaggaacaggaggattgggacagggtttaggtaccacc  
 T G L G Q G T G T G G I G T G G Y G T T  
 ggtcaccatgggttgggaactggcgtaggtcataccaccggaggtacaggaaccgactat  
 G H H G L G T G V G H T T G G T G T D Y  
 acctcgggtgctggtccaccggacagaccggttaccaaggttaggtaccgagtctgag  
 T S G G R S T G Q T G Y Q G L G T E S E  
 ttcgttgaagaactggaaacttccaaaaccaaccagtgccacacctgtaggcgggtt  
 F G G R T G T F Q N Q P S A T P V G G V  
 gggttgagctcaggaactgggcccgggtggttgaggtacaggaactggaacaggcattctg  
 G L S S G T G A G V G G T G T G T G I L  
 catcgttctggaagtagcagctcaagctcg**gtaagtatacaagaaaacaacttgcctat**  
 H R S G **S S S S S S** V S I Q E N K L A Y  
**tat**ttttatatacaaggagatagagaggctgttttagttaaatatTTTTaatagtttt  
 Y F L Y T R R - R G C F - L N I F N S F  
**g**aaaaaaaaaccggttatgtttttgtttaatttatgatctcattgattgat**gtag**tcg  
 E K K T V Y V F V - F M I S L I D M - **S**  
 gaggatgatgggcaaggaggaaggaggaagaagaaaggggtgatgcaaaagatcaaagag  
 E D D G Q G G R R K K K G V M Q K I K E  
 aagcttcccgggtgtcacagacaagaggagcagtatcagagccagacgactactactacg  
 K L P G G H R Q E E Q Y Q S Q T T T T T  
 acgggtggcggggcgggttatggtgaaaccac**gagaagaaagggatgatgga**  
 T G G G A G Y G E T H **E K K G M M D**

In **rosso** sequenza di un introne.

### Sequenza aminoacidica parziale di una deidrina di girasole

**DEYGNP**VRRRTGQTDEYGNPVRRTDEYGNPVHSTTGGTMGDYGSTGLGQGTGTGGIGTGGYGTGHHGLG  
 TGVGHTTGGTGTDTSGGRSTGQTGYQGLGTESEFFGGRTGTFQNPQPSATPVGGVGLSSGTGAGVGGTGT  
 GTGILHRSGSSSSSSSEDDGQGGRRKKKGVMQKIKEKLPGGHRQEEQYQSQTTTTTTGGGAGYGETH**EK**  
**KGMMD**

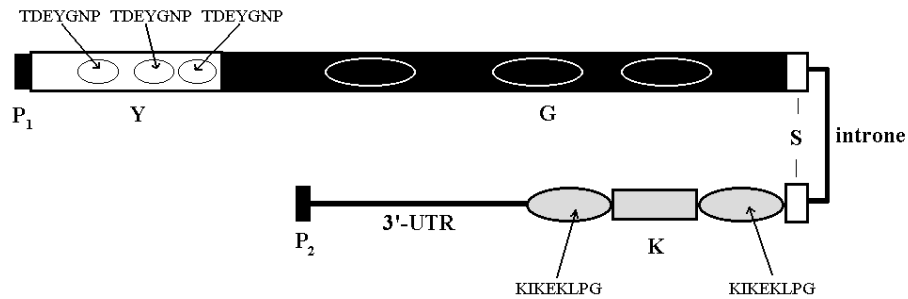


### A cosa può servire isolare un gene? Ecco un esempio.

### Isolare particolari forme del medesimo gene che potrebbero essere utili nel miglioramento genetico

Attraverso questa metodologia sono stati isolati frammenti genici codificanti deidrine da numerosi genotipi di girasole selvatico di svariate località. In particolare in due genotipi, dal Colorado e da Durango (Messico), sono stati identificati due frammenti genici che portano a due importanti modificazioni nella sequenza aminoacidica.

Schema dei domini proteici codificati dal gene della deidrina di girasole



HCM (linea di girasole coltivato)	259 aa
Durango	246 aa
Colorado	236 aa

HCM	MANYGGDKQYGRETRHTGDYESIPIHSTGGQYEQEVLTDEYGNPVRRTGOTDEYGNPVRR	60
Durango	MANYGGDKQYGRETRHTGDYENPIHSTGGQYEQEVLPQ-----TDEYGNPVRR	47
Colorado	MANYGGDKQYGRETRHTGDYENPIHSTGGQYEQEVLTDEYGNPVRRTGOTDEYGNPVRR	60
	*****.***** * ***** *	

HCM	TDEYGNPVHSTTGGTMGDYGSTGLGQGTGTGGIGTGGYGTGHHGLGTGVGHTTGGTGT	120
Durango	TDEYGNPVHSTTGGTMGDYGSTGLGQGTGTGGIGTGGYGTGHHGLGTGVGHTTGGTGT	107
Colorado	TDEYGNPVHSTAGGTMGDYGSTGLGQGTG--GIGTGGYGTGHHGLGTGIGHTTGGTGT	118
	*****:***** *****:*****	

HCM	YTSGGRSTGTGYQGLGTESEDFGKKTGTFFONQPSATPVGGVGLNSGTGAGVGGTGTGTGI	180
Durango	YTSGGRSTGTGYQGLGTEPEFGKKTGTFFONQPSATPVCCVCISSGTGAGVGGTGTGTGI	167
Colorado	YTSGGRSTGTGYQGLGTESEFRKKTGT-----GVC GTGTGTGI	157
	*****:* *****	

HCM	LHRSGSSSSSSSEDDGQGGRRKKKGVMOKIKEKLPGGHRQEEQYQSQTTTTTTGGGAGYG	240
Durango	LHRSGSSSSSSSEDDGQGGRRKKKGVMOKIKEKLPGGHRQEEQYQSQTTTTTTGGGAGYG	227
Colorado	LHRSGSSSSSSSEDDGQGGRRKKKGVMOKIKEKLPGGHRQEEQYQSQTTTTTTGGGAGYG	217
	*****.*****	

HCM	ETHEKKGMMKIKEKLPGHH	260
Durango	ETHEKKGMMKIKEKLPGHH	247
Colorado	ETHEKKGMMKIKEKLPGHH	237
	*****	

I genotipi provenienti da Durango e da Colorado presentano due delezioni a carico di regioni codificanti per due domini conservati, rispettivamente un dominio con sequenza TDEYGNP e un dominio ricco in glicine (G).

Stiamo valutando se i genotipi selvatici che hanno questi alleli del gene modificati (mutanti) siano più o meno resistenti allo stress idrico. Dal punto di vista del miglioramento genetico del girasole, tali mutanti, se presenteranno una maggiore resistenza al deficit idrico, potrebbero essere utilizzati in incroci con linee di girasole coltivato per ottenere girasole coltivato resistente al deficit idrico, che potrebbe essere coltivato in ambienti aridi.