

Produzione di proteine ricombinanti

Applicazioni principali delle proteine ricombinanti

- STUDIO DELLA RELAZIONE TRA STRUTTURA E ATTIVITA'
- .-PROTEINE DI INTERESSE TERAPEUTICO (anticorpi, insulina, ormone crescita).
- PROTEINE DI INTERESSE COMMERCIALE (enzimi).
-

Le proteine ricombinanti possono essere espresse come

- *Wild type* equivalenti alle proteine naturali
- Mutanti varianti nelle quali sono state inseriti dei cambiamenti nella sequenza mediante tecniche di mutagenesi

Espressione di una proteina ricombinante

1. Clonaggio del gene o del cDNA (screening di librerie di cDNA, RT-PCR, RACE, TAIL PCR)
2. Inserimento nel vettore di espressione (promotore inducibile o costitutivo)
3. Trasformazione dell'ospite (shock termico, elettroporazione, liposomi)
4. Selezione dei ricombinanti (resistenza ad antibiotici, vantaggio selettivo)
5. Espressione e purificazione della proteina (intracellulare o secreta)

Strategia di clonaggio di sequenze codificanti note

- Analisi di restrizione del cDNA/gene che deve essere clonato
- Scelta e del vettore ed analisi del suo MCS
- Progettazione oligonucleotidi contenenti estensioni con i siti di restrizione scelti
- Amplificazione per PCR o RT pCR del DNA da clonare
- Digestione del frammento amplificato e del vettore con gli enzimi di restrizione
- Purificazione dei frammenti
- Ligazione e trasformazione

Strategia di clonaggio di sequenze codificanti ignote

- Screening di librerie genomiche o di cDNA
- RACE Rapid Amplification of cDNA ends
- TAIL PCR

Requisito necessario

Conoscenza di alcune porzioni della sequenza amminoacidica per sintetizzare oligonucleotidi da utilizzare nelle reazioni di PCR

Progettazione di oligonucleotidi a partire da una sequenza amminoacidica

Si sintetizzano delle miscele di oligonucleotidi in modo tale che sia presente almeno un oligo che possa appaiarsi allo stampo

Il codice genetico é degenerato

		Seconda base del codone					
		T	C	A	G		
Prima base del codone	T	TTT → Phe F	TCT → Ser S	TAT → Tyr Y	TGT → Cys C	Terza base del codone	T
		TTC → Phe F	TCC → Ser S	TAC → Tyr Y	TGC → Cys C		C
		TTA → Leu L	TCA → Ser S	TAA → STOP	TGA → STOP		A
		TTG → Leu L	TCG → Ser S	TAG → STOP	TGG → Trp W		G
	C	CTT → Leu L	CCT → Pro P	CAT → His H	CGT → Arg R		T
		CTC → Leu L	CCC → Pro P	CAC → His H	CGC → Arg R		C
		CTA → Leu L	CCA → Pro P	CAA → Gln Q	CGA → Arg R		A
		CTG → Leu L	CCG → Pro P	CAG → Gln Q	CGG → Arg R		G
	A	ATT → Ile I	ACT → Thr T	AAT → Asn N	AGT → Ser S		T
		ATC → Ile I	ACC → Thr T	AAC → Asn N	AGC → Ser S		C
		ATA → Ile I	ACA → Thr T	AAA → Lys K	AGA → Arg R		A
		ATG → Met M	ACG → Thr T	AAG → Lys K	AGG → Arg R		G
	G	GTT → Val V	GCT → Ala A	GAT → Asp D	GGT → Gly G		T
		GTC → Val V	GCC → Ala A	GAC → Asp D	GGC → Gly G		C
		GTA → Val V	GCA → Ala A	GAA → Glu E	GGA → Gly G		A
		GTG → Val V	GCG → Ala A	GAG → Glu E	GGG → Gly G		G

COME PROGETTARE UN OLIGONUCLEOTIDE A PARTIRE DA UNA SEQUENZA AMINOACIDICA

- Scegliere la sequenza aminoacidica a minore degenerazione
- Escludere la terza base dell'ultimo codone dal conteggio
- Considerare il «Codon usage»
- Inserire la deossiinosina

Met-Leu-Pro-Gln-Lys-Trp-Asp-Gly-Ser-Met-Asn-Gly-Arg

1 6 4 2 2 1 2 4 6 1 2 4 6

Pro-Gln-Lys-Trp-Asp-Gly

4 2 2 1 2 4

CCN-CAA/G-AAA/G-TGG-GAT/C-GGN 18 basi degenerazione 128

17 basi degenerazione 32

Met-Leu-Pro-Gln-Lys-Trp

1 6 4 2 2 1

ATG-TTA/G-CCN-CAA/G-AAA/G-TGG 18 basi degenerazione 32 (96)

Frequenza di codoni usati in *Saccharomyces cerevisiae*

Saccharomyces cerevisiae [gbpln]: 11937 CDS's (5695112 codons)

fields: [triplet] [amino acid] [fraction] [frequency: per thousand] ([number])

UUU F 0.59 26.1 (148716)	UCU S 0.26 23.5 (134028)	UAU Y 0.56 18.9 (107653)	UGU C 0.63 8.0 (45462)
UUC F 0.41 18.2 (103686)	UCC S 0.16 14.2 (80799)	UAC Y 0.44 14.6 (83397)	UGC C 0.37 4.7 (26534)
UUA L 0.28 26.7 (151850)	UCA S 0.21 18.8 (106932)	UAA * 0.47 1.0 (5692)	UGA * 0.31 0.7 (3804)
UUG L 0.28 27.0 (153619)	UCG S 0.10 8.5 (48582)	UAG * 0.22 0.5 (2682)	UGG W 1.00 10.3 (58581)
CUU L 0.13 12.1 (69114)	CCU P 0.31 13.6 (77417)	CAU H 0.64 13.7 (78282)	CGU R 0.15 6.5 (36777)
CUC L 0.06 5.4 (30538)	CCC P 0.15 6.8 (38522)	CAC H 0.36 7.7 (44112)	CGC R 0.06 2.6 (14638)
CUA L 0.14 13.3 (75972)	CCA P 0.42 18.2 (103680)	CAA Q 0.69 27.5 (156586)	CGA R 0.07 3.0 (17099)
CUG L 0.11 10.4 (59114)	CCG P 0.12 5.3 (29915)	CAG Q 0.31 12.1 (69006)	CGG R 0.04 1.7 (9848)
AUU I 0.47 30.4 (173137)	ACU T 0.35 20.3 (115551)	AAU N 0.59 36.3 (206516)	AGU S 0.16 14.2 (80728)
AUC I 0.26 17.0 (97055)	ACC T 0.21 12.6 (71608)	AAC N 0.41 24.9 (141737)	AGC S 0.11 9.7 (54958)
AUA I 0.27 17.8 (101628)	ACA T 0.30 17.8 (101336)	AAA K 0.58 42.2 (240515)	AGA R 0.48 21.3 (121292)
AUG M 1.00 20.9 (119008)	ACG T 0.14 7.9 (45120)	AAG K 0.42 30.7 (174597)	AGG R 0.21 9.2 (52441)
GUU V 0.39 22.0 (125255)	GCU A 0.38 21.1 (120208)	GAU D 0.65 37.8 (215249)	GGU G 0.47 24.0 (136434)
GUC V 0.21 11.6 (65876)	GCC A 0.22 12.5 (71233)	GAC D 0.35 20.2 (115047)	GGC G 0.19 9.7 (55103)
GUA V 0.21 11.8 (67408)	GCA A 0.29 16.1 (91967)	GAA E 0.71 45.8 (261115)	GGA G 0.22 10.9 (62176)
GUG V 0.19 10.6 (60537)	GCG A 0.11 6.1 (34942)	GAG E 0.29 19.1 (108666)	GGG G 0.12 6.0 (34032)

Coding GC 39.64% 1st letter GC 44.53% 2nd letter GC 36.55% 3rd letter GC 37.85%

Genetic code 1: Standard

COME PROGETTARE UN OLIGONUCLEOTIDE A PARTIRE DA UNA SEQUENZA AMINOACIDICA

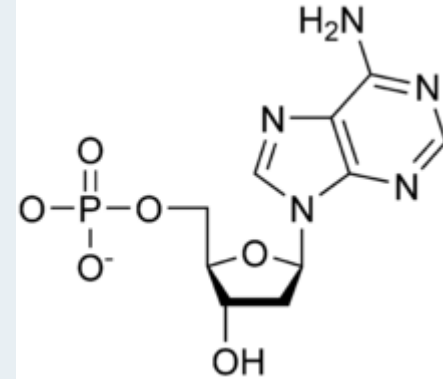
DEOSSIIINOSINA

I → **A**

I → **C**

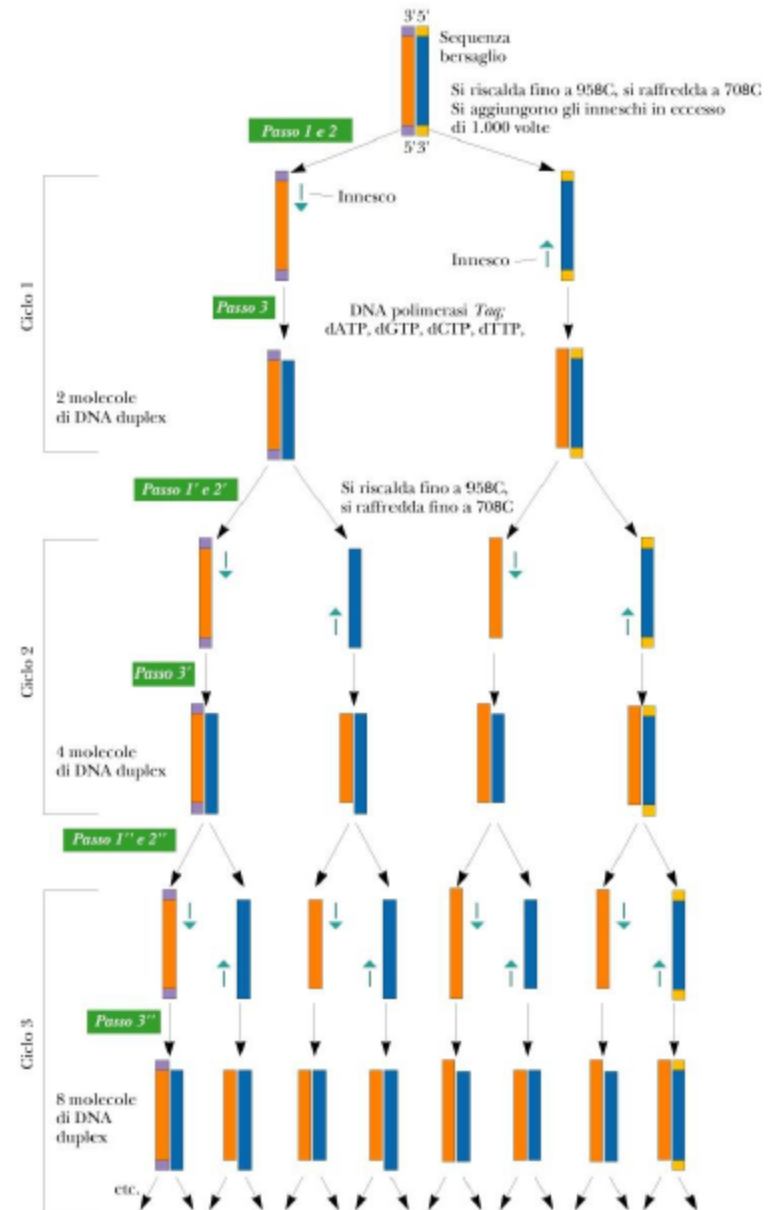
I → **G**

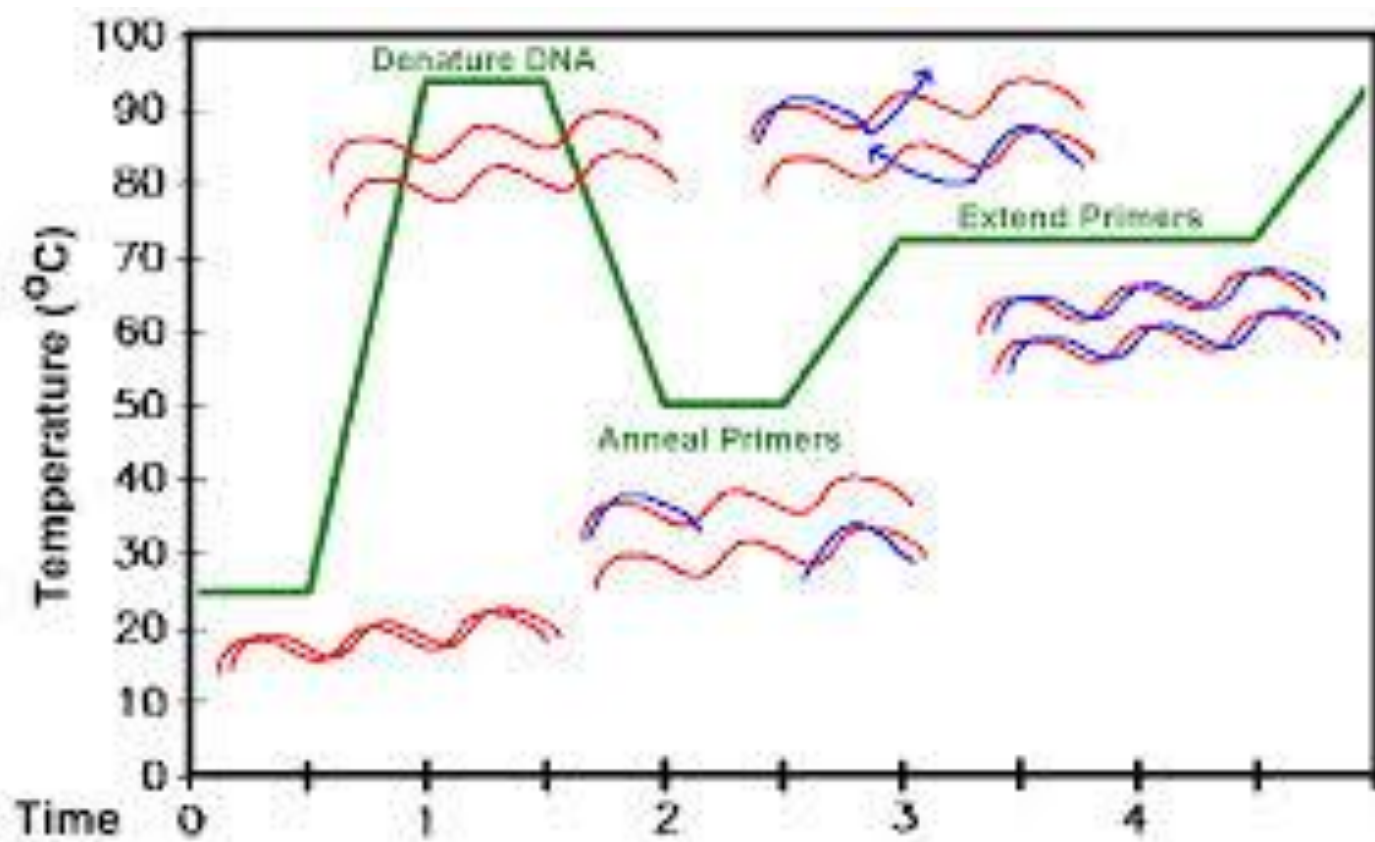
I → **T**



PCR (Polymerase Chain Reaction)

Figura 9.20 La reazione a catena della polimerasi (PCR). Gli oligonucleotidi complementari ad una data sequenza di DNA ne innescano la sintesi. La DNA polimerasi Taq termostabile sopravvive a diversi cicli di riscaldamento. Teoricamente la quantità di sequenza specificata dai primer è raddoppiata ad ogni ciclo.



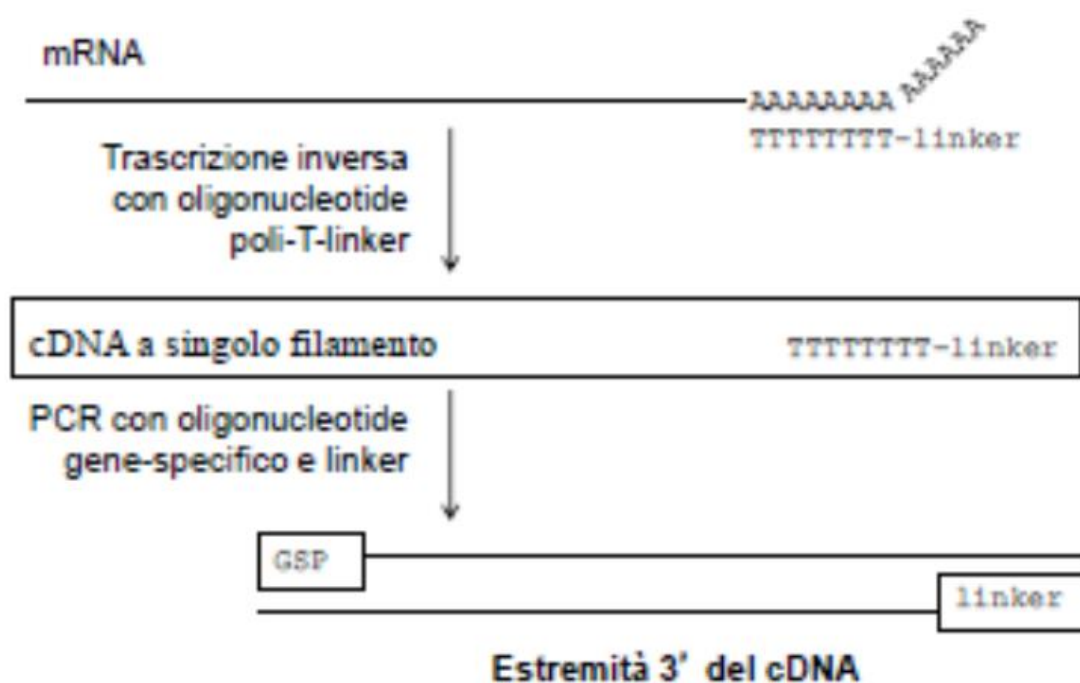


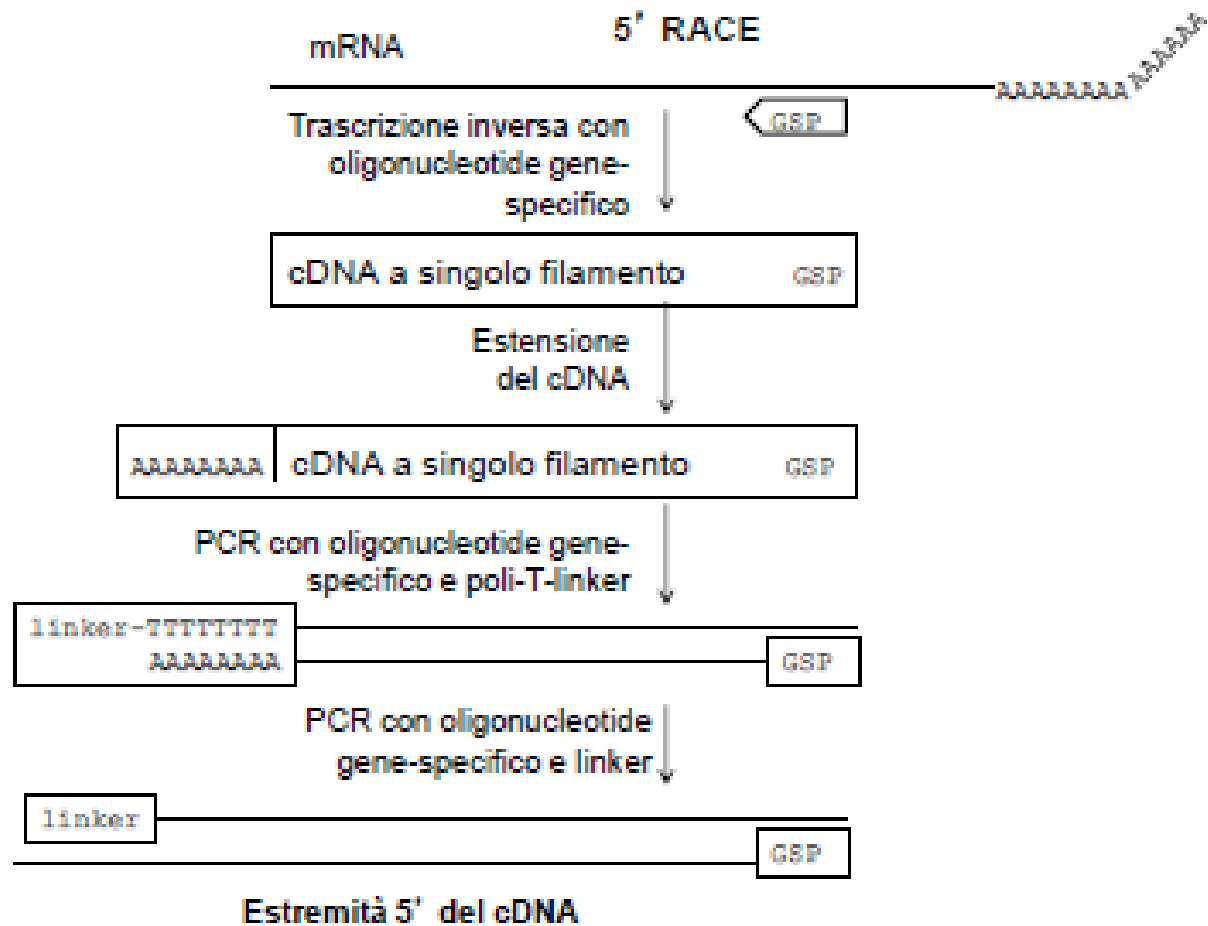
APPLCAZIONI DELLA PCR

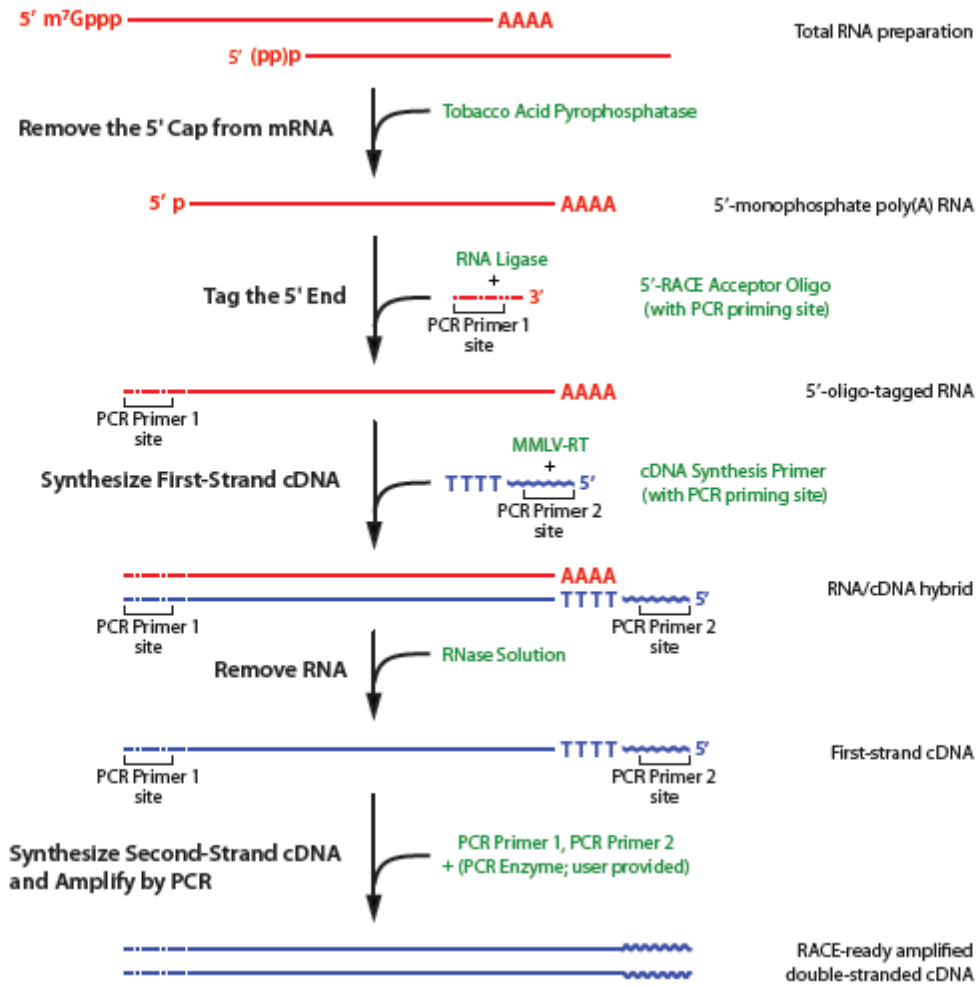
RACE Rapid Amplification of cDNA ends

l'mRNA viene trascritto in cDNA dalla trascrittasi inversa utilizzando un innesco generico (poli-T+adattatore, 3' RACE) o gene-specifico (5' RACE). Il cDNA viene poi amplificato con l'oligo adattatore e un oligo gene-specifico.

3' RACE





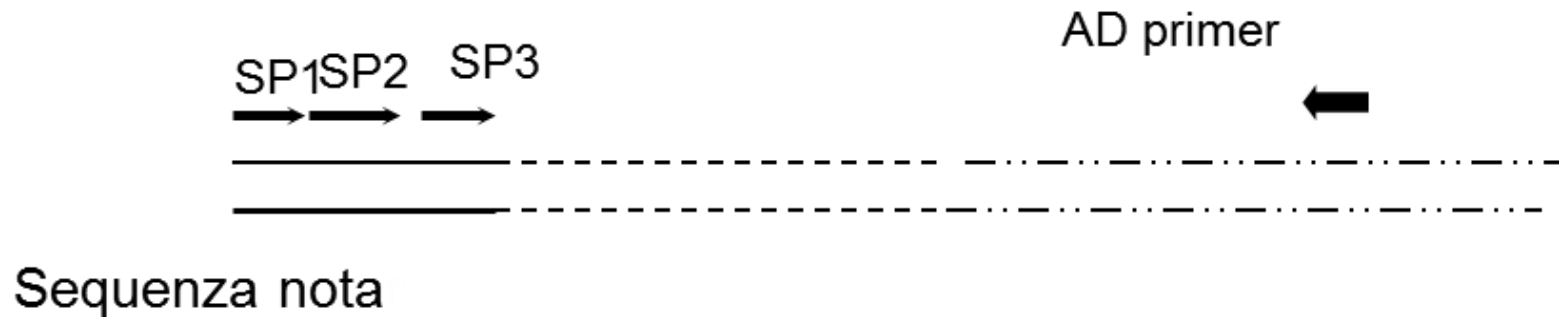


TAIL PCR (THERMAL ASYMMETRIC INTERLACED PCR)

TAIL-PCR è una tecnica che permette di amplificare frammenti di DNA adiacenti a sequenze note

- consiste in tre reazioni consecutive di PCR in cui vengono utilizzati tre primer specifici consecutivi SP1 SP2 e SP3 e un primer arbitrario AD degenerato avente una T_m più bassa

le amplificazioni specifiche e aspecifiche sono controllate termicamente



Applicazioni

- Isolamento regioni codificanti presenti al 5'
- Isolamento di promotori
- Amplificazione di sequenze poco rappresentate da genomi complessi

TAIL-PCR

Progettazione primers

Oligonucleotidi specifici (SP1-SP2-SP3):

- *Nested*
- Complementari alla sequenza codificante
- $T_m = 58-63^\circ\text{C}$

Oligonucleotidi degenerati :

- Corti
- T_m più bassa di quella degli oligo specifici ($T_m = 44-45^\circ\text{C}$)

Temperatura di annealing

variabile

Ordine dei cicli

si alternano cicli ad alta e bassa stringenza

-

-

Temperatura di *annealing*

Cicli ad alta stringenza (thermal asymmetric)

Annealing temperature = 63°C

Cicli a ridotta stringenza (thermal symmetric)

Annealing temperature = 44°C

Cicli a bassa stringenza

Annealing temperature = 30°C

**Cycling Conditions Used for TAIL-PCR
on the GeneAmp System 9600**

Reaction	File no.	Cycle no.	Thermal condition
Primary	1	1	92°C (2 min), 95°C (1 min)
	2	5	94°C (15 s), 63°C (1 min), 72°C (2 min)
	3	1	94°C (15 s), 30 °C (3 min), ramping to 72°C over 3 min, 72°C (2 min)
	4	10	94°C (5 s), 44°C (1 min), 72°C (2 min)
	5	12 ^a	94°C (5 s), 63°C (1 min), 72°C (2 min) 94°C (5 s), 63°C (1 min), 72°C (2 min) 94°C (5 s), 44°C (1 min), 72°C (2 min)
	6	1	72°C (5 min)
Secondary	7	10 ^a	94°C (5 s), 63°C (1 min), 72°C (2 min) 94°C (5 s), 63°C (1 min), 72°C (2 min) 94°C (5 s), 44°C (1 min), 72°C (2 min)
	6	1	72°C (5 min)
Tertiary	8	20	94°C (10 s), 44°C (1 min), 72°C (2 min)
	6	1	72°C (5 min)

Note. The program files in each reaction were linked automatically.

^a These are nine-segment super cycles each consisting of two high-stringency and one reduced-stringency cycle (see Fig. 1).

I PCR primers
SP1 e AD

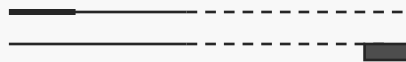
- 5 cicli ad alta stringenza Tm 63
- 1 ciclo a bassa stringenza Tm variabile (ramping)
- 10 cicli a ridotta stringenza Tm 44

12 TAIL-cycling
(super cicli)

1 ciclo a ridotta stringenza
(thermal symmetric)

2 cicli ad alta stringenza
(thermal asymmetric)

prodotto specifico
(type I)



Product yield
High or middle
(detectable or undetectable)

prodotto non specifico
(type II)



High
(detectable)

Prodotto non specifico
(type III)



Low
(undetectable)

Passaggi	Cicli	Programmi	Materiali
Prima	1 5 1 15	94°C, 5 min 94°C, 30 s; 70 °C, 30 s; 72°C, 2 min 94°C, 30 s; 30 °C, 3 min; 72 °C, 0.3°C/s; 72 °C, 2 min 94°C, 30 s; 70 °C, 30 s; 72 °C, 2 min 94°C, 30 s; 70 °C, 30 s; 72 °C, 2 min 94°C, 30 s; 50 °C , 1 min;72 °C, 2 min	5 µM del primer GCADXa 0.15 µM sp1 0.2mM di dNTPs
Seconda	1 15 1	94°C, 5 min 94°C, 30 s; 70 °C, 30 s; 72 °C, 2 min 94°C, 30 s; 70 °C, 30 s; 72 °C, 2 min 94°C, 30 s; 50 °C ,1 min; 72 °C,2min 72°C, 7 min	5 µM del primer GCADXa 0.2 µM sp2 0.2mM di dNTPs Templato (prodotto della PCR primaria diluito 1000 volte)
Terza		Il programma è lo stesso della PCR secondaria	5 µM del primer GCADXa 0.2 µM sp3 0.2mM di dNTPs Templato (prodotto della PCR secondaria diluito 1000 volte)

II PCR

PCR con SP2 e AD (10 super cicli)

templato è il prodotto della prima reazione di PCR diluito 1000-volte



Prodotto specifico

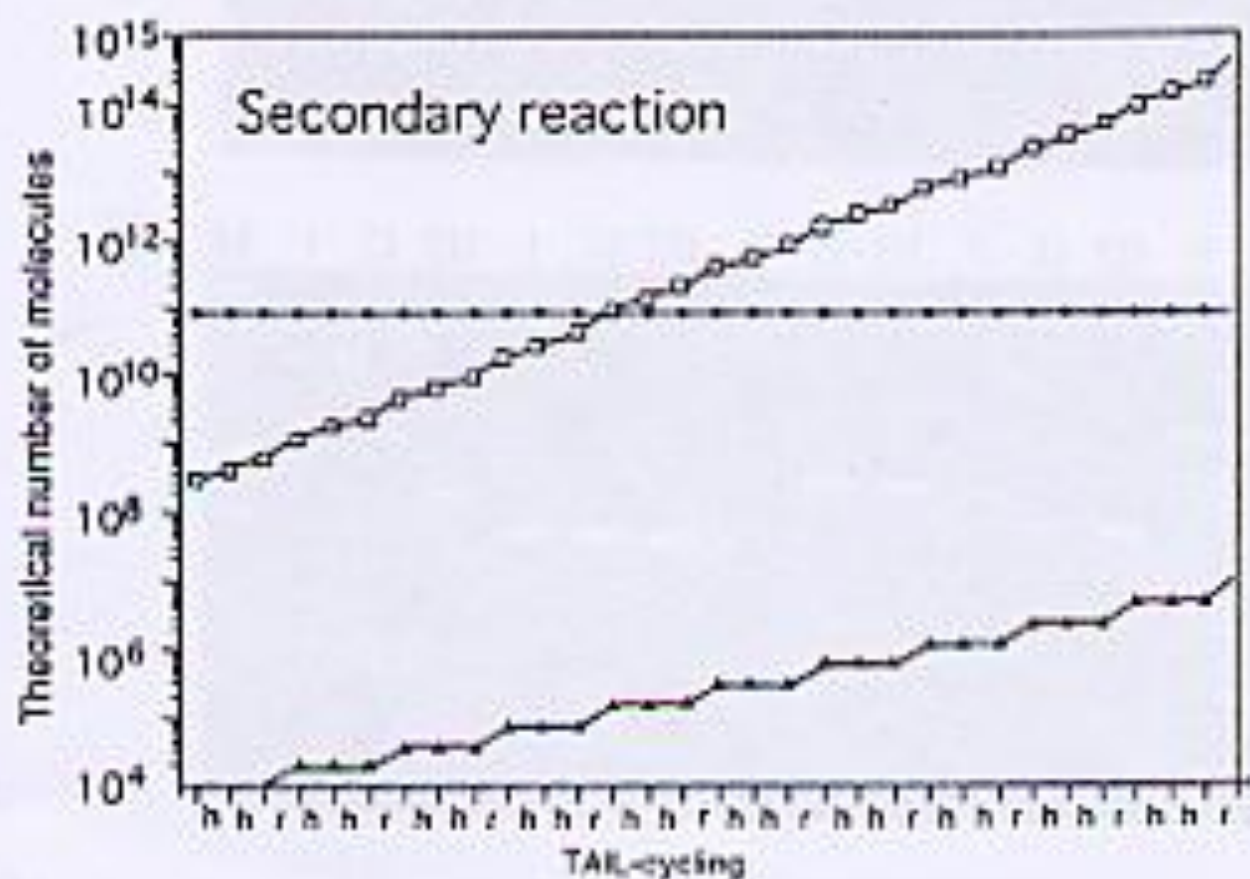


Prodotto nonspecifico (type III)



Prodotti : Alta concentrazione

Bassa concentrazione)



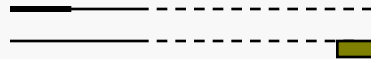
III PCR

PCR con SP3 E AD (20 cycles)

- templatò è il prodotto della seconda reazione di PCR diluito 1000-volte



Prodotti specifici

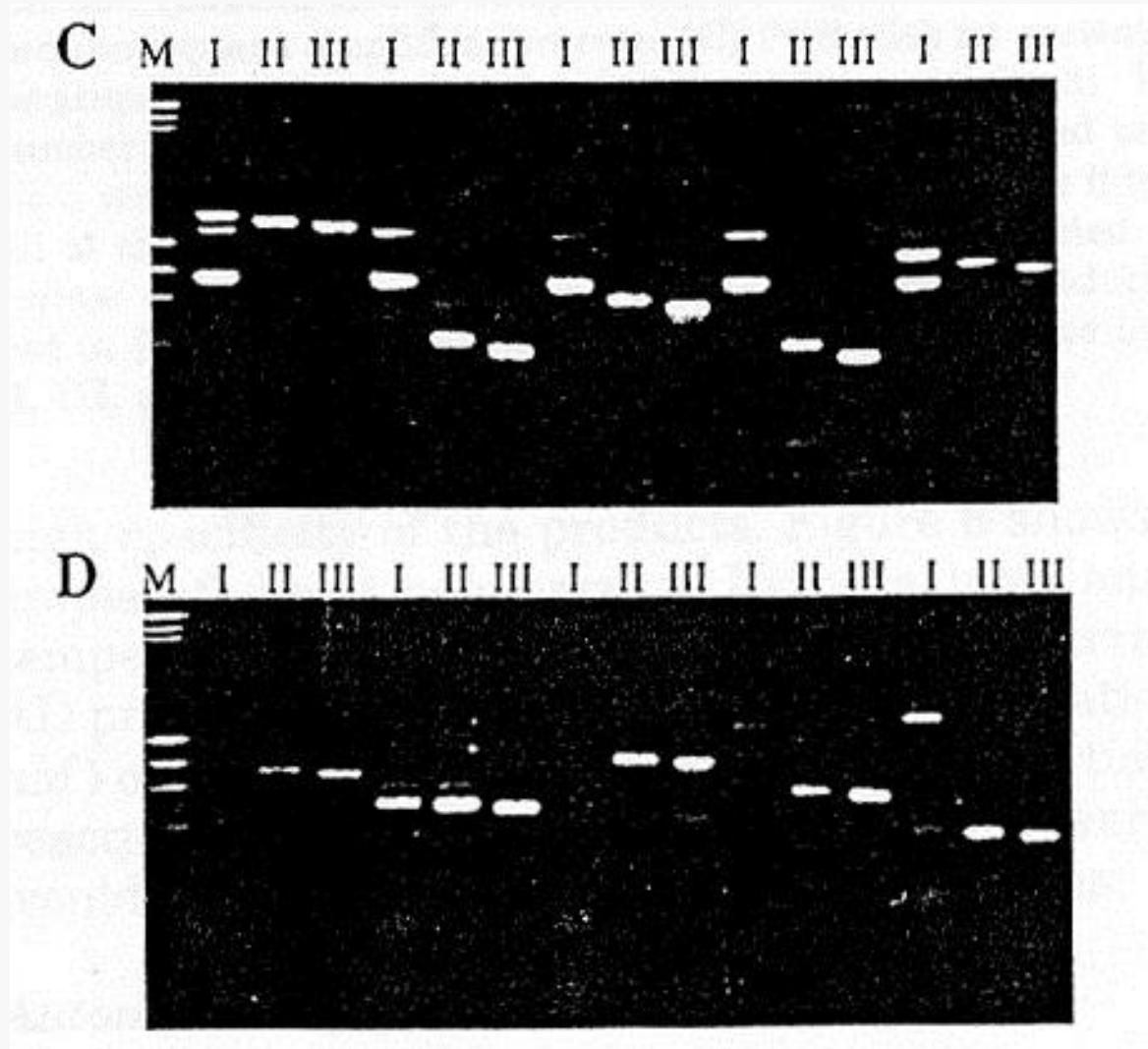


Gel di Agarosio

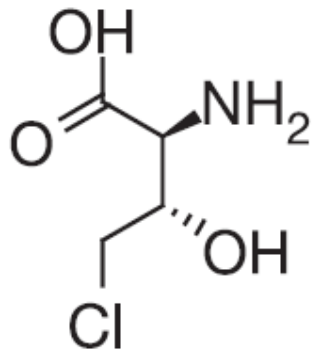


Sequenziamento

YAC clones



Identificazione del cluster biosintetico della 4-clorotreonina in *Streptomyces* sp. OH- 5093



4-Cl-L-Thr

Attività antifungina

Identificazione dell'alogenasi in *Streptomyces* sp. OH-5093

1. Allineamento proteine ortologhe a SyrB2

NYDRHLD: 5'- AAC TAC GAC CG(CG) CAC CT(GC) GA-3'

```

5748800 1 MSKEFLTAEQRAAFKMGGLGPPDAYSPEEMRETKWRTRLDLDRSAAAYQDLDAISGGTNI...NYDRHLD DFLASHTCRPEICDQVESLGGNVLQWTEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
66045846199.412e-75 1 MSKEFLTAEQRAAFKMGGLGPPDAYSPEEMRETKWRTRLDLDRSAAAYQDLDAISGGTNI...NYDRHLD DFLASHTCRPEICDQVESLGGNVLQWTEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
72610045155.41e-69 1 HQNFETPEEIAAFRAQGGGPPFLYSPSEAEERYKTRALDFDRE...YAYVDLPTTSP...NYDRHLD NLGSEIGRQEIIVQALNSHLLGGDVLQWRSEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
67661461162.713e-68 1 .STRFASQQLAIFERMGGLGPPFLYSDQMEANLQALRPLDNTKHAIVSQERSISGVINL...NYDRHLD DFLASHTCRPEICDQVESLGGDVLQWRSEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
67760175152.811e-67 1 MS.QFFLTPDEIAAFRECGGLGPPFLYSEDEAEERYQIRALDFDRE...YAYVLELPPNSP...NYDRHLD NLGSEHMRAEIVHKLNAHLLGGDVLQWRSEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
37575141155.313e-65 1 . . . . . FSTKQELQFPRDGGGLGPPFLYSEEMKQMHNTIRALDFWTRHAP...YDAPLQVAT...NYDRHLD DFLTSVFRKEIVHVRVSHLLGGDVLQWRSEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
28871835155.311e-64 1 NTO.FSTKQELQFPRDGGGLGPPFLYSEEMKQMHNTIRALDFWTRHAP...YDAPLQVAT...NYDRHLD DFLTSVFRKEIVHVRVSHLLGGDVLQWRSEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
67544685154.412e-64 1 . . . . . TQFETPEEIAAFRAQGGGPPFLYSPSEAEERYKTRALDFDRE...YAYVDLPTTSP...NYDRHLD NLGSEIGRQEIIVQALNSHLLGGDVLQWRSEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
48766802143.810e-64 1 . . . . . YLSVEEQFPRDGGGLGPPFLPPEEIEELWQETIRMLDDETAFFPKSM...NYDRHLD QLQLDIISHPKVAHRVSSLLGGNVLQWRTEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
48766806143.812e-63 1 . . . . . YLSVEEQFPRDGGGLGPPFLPPEEIEELWQETIRMLDDETAFFPKSM...NYDRHLD QLQLDIISHPKVAHRVSSLLGGNVLQWRTEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
23452292142.910e-63 1 . . . . . YLSVEEQFPRDGGGLGPPFLPPEEIEELWQETIRMLDDETAFFPKSM...NYDRHLD QLQLDIISHPKVAHRVSSLLGGNVLQWRTEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
49659917143.614e-62 1 . . . . . YLSVEEQFPRDGGGLGPPFLPPEEIEELWQETIRMLDDETAFFPKSM...NYDRHLD QLQLDIISHPKVAHRVSSLLGGNVLQWRTEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
49659919142.514e-62 1 . . . . . YLSVEEQFPRDGGGLGPPFLPPEEIEELWQETIRMLDDETAFFPKSM...NYDRHLD QLQLDIISHPKVAHRVSSLLGGNVLQWRTEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
23452293143.114e-62 1 . . . . . RYLSVEEQFPRDGGGLGPPFLPPEEIEELWQETIRMLDDETAFFPKSM...NYDRHLD QLQLDIISHPKVAHRVSSLLGGNVLQWRTEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
49659921142.713e-61 1 . . . . . RYLSVEEQFPRDGGGLGPPFLPPEEIEELWQETIRMLDDETAFFPKSM...NYDRHLD QLQLDIISHPKVAHRVSSLLGGNVLQWRTEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
48766808143.611e-60 1 . . . . . RYLSVEEQFPRDGGGLGPPFLPPEEIEELWQETIRMLDDETAFFPKSM...NYDRHLD QLQLDIISHPKVAHRVSSLLGGNVLQWRTEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
48766804143.615e-60 1 . . . . . RYLSVEEQFPRDGGGLGPPFLPPEEIEELWQETIRMLDDETAFFPKSM...NYDRHLD QLQLDIISHPKVAHRVSSLLGGNVLQWRTEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
24985426149.716e-60 1 NTDLHLDTEQVAFKMGGLGPPFRYSQSEIEDIKERRARLDTFHSQSVNPGKASTKDFP...NYDRHLD DFLASHTCRPEICDQVESLGGDVLQWRSEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
67755172142.819e-58 1 MQKRFPLSEAEKFFPEMGGLGPPFLYSEEMTRVYKGVEMLDLTAKAAP...PNSK...NYDRHLD LSLQIISHPKVAHRVSSLLGGDVLQWRSEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF
67763421142.812e-57 1 MQKRFPLSEAEKFFPEMGGLGPPFLYSEEMTRVYKGVEMLDLTAKAAP...PNSK...NYDRHLD LSLQIISHPKVAHRVSSLLGGDVLQWRSEFFPKYPCRGTDWQDQDTPHAGSGKQIWPENESFGGITITWTAFTDANIANGLQF

```

```

5748800 161 IPGQMSMNYDDEKRMITPDANNVVDGVRGFGCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
66045846199.412e-75 1 IPGQMSMNYDDEKRMITPDANNVVDGVRGFGCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
72610045155.41e-69 1 VFGHEEMFYDSSKMAVDASRIINQBDGDKRFGCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
67661461162.713e-68 160 IPGHHTWYDFDHRNIFPEPKIKKLTGGKRGFYCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
67760175152.811e-67 156 IPGHEEMFYDSSKMAVDASRIINQBDGDKRFGCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
37575141155.313e-65 153 IPGHSHQMNIDTRGHTFPVYTNINNVYQGYRFGCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
28871835155.311e-64 156 IPGHSHQMNIDTRGHTFPVYTNINNVYQGYRFGCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
67544685154.412e-64 155 IPGHEEMFYDSSKMAVDASRIINQBDGDKRFGCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
48766802143.810e-64 149 IPGHHTWYDFDHRNIFPEPKIKKLTGGKRGFYCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
48766806143.812e-63 149 IPGHHTWYDFDHRNIFPEPKIKKLTGGKRGFYCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
23452292142.910e-63 149 IPGHHTWYDFDHRNIFPEPKIKKLTGGKRGFYCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
49659917143.614e-62 149 IPGHHTWYDFDHRNIFPEPKIKKLTGGKRGFYCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
49659919142.514e-62 149 IPGHHTWYDFDHRNIFPEPKIKKLTGGKRGFYCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
23452293143.114e-62 150 IPGHHTWYDFDHRNIFPEPKIKKLTGGKRGFYCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
48766808143.611e-60 150 IPGHHTWYDFDHRNIFPEPKIKKLTGGKRGFYCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
48766804143.615e-60 150 IPGHHTWYDFDHRNIFPEPKIKKLTGGKRGFYCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
24985426149.716e-60 159 IPGGHARTWYDFDHRNIFPEPKIKKLTGGKRGFYCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
67755172142.819e-58 153 IPGGHARTWYDFDHRNIFPEPKIKKLTGGKRGFYCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA
67763421142.812e-57 153 IPGGHARTWYDFDHRNIFPEPKIKKLTGGKRGFYCYDYQLQDNDNPPDASAVYQMKAGQFIFASHTINASHSPHSGESQEMRNGASRVPVSTFHWVPPDSHIEEYCARISLBYKGA

```

2. Scelta delle regioni altamente conservate per la costruzione di oligonucleotidi degenerati

GF/Vf/YGYDY : 3'-CC(CG) (CA)A(CG) A(TA)G CC(CG) ATG CTG ATG-5'

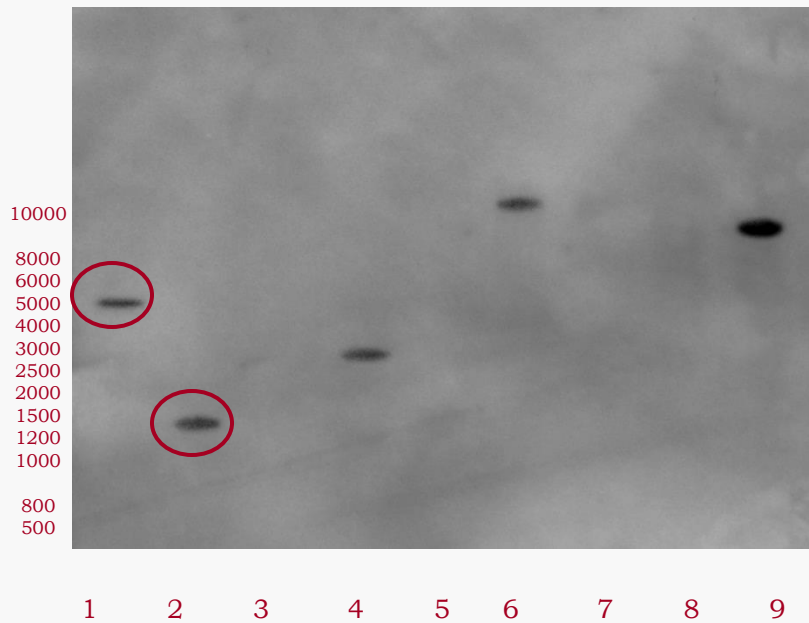
c) G₁₉₄-Y₂₀₀ : 3'-CC(CG) (CA)A(CG) A(TA)G CC(CG) ATG CTG ATG-5'

5748808	161	IPGTQNSMNYDE	TKRMTYEPDANNSVVKDGVRRG	FFGYDYRQLQIDENWKPDEASAVPMQMKAGQPIIFWSTLMHHSYFPH
66045846 99.4 2e-75	161	IPGTQNSMNYDE	TKRMTYEPDANNSVVKDGVRRG	FFGYDYRQLQIDKDWKPEASAVPMQMKAGQPIIFWSTLMHHSYFPH
72610045 55.4 1e-69	156	VPCTHEEMFYDES	SRKNQFDPNINALEKDGIKRGG	FFGYDYRSLQKQDPDWKPEDESKAVSIVMKPGQFVIFWSTLMHHSYFPH
67661461 62.7 3e-68	160	IPCTHRTMNYDES	SKVMEYREDAINNREKQGVRRG	FFGYDYRQLQKQDPNWSPEDESAKSMVMRQGPPIIFWSTLMHHSYFPH
67760715 52.8 1e-67	156	MPCTHEEMFYDES	SKKMTFDPNINALEKDGIKRGG	FFGYDYRSLQKQDPDWKPEDESNVSIKMKPGQFVIFWSTLMHHSYFPH
37575141 55.3 3e-65	153	IPGSHKQMNIDE	TRGMTFDPVTNNNVKGYARG	FNGYDYTLQKQDKSNKPEDEASAVPIVMKAGQFVIFWSTLMHHSYFPH
28871835 55.3 1e-64	156	IPGSHKQMNIDE	TRGMTFDPVTNNNVKGYPRG	FNGYDYTLQKQDKSNKPEDEASAVPIVMKAGQFVIFWSTLMHHSYFPH
67544685 54.4 2e-64	155	IPCTHEEMFYDES	SKMAYDASKINTQEKDGIKRG	FFGYDYRSLQKQDPNWKPEDESAVSIITMQAGECVIFWSTLMHHSYFPH
48766802 43.8 8e-64	149	IPGTHNTWYFDE	HRIEPEPEKINKKLTGGKKTG	VYGYDYKLLKLDPNWEPDESAVHLEMQPGQFILFTSRCMHGSEPN
48766806 43.8 2e-63	149	IPGTHNTWYFDE	HRIEPEPEKINKKLTGGKKTG	VYGYDYKLLKLDPNWEPDESAVHLEMQPGQFILFTSRCMHGSEPN
23452292 42.9 8e-63	149	MPGTHNTWYFDE	HRIEPEPKINQKLTGGKKTG	VYGYDYKLLKLDPNWEPDESAVHLEMQPGQFILFTSRCMHGSEPN
49659917 43.6 4e-62	149	IPGTHNTWYFDE	HRIEPEPEKINKKLTGGKKTG	VYGYDYKLLKLDPNWEPDESAVHLEMQPGQFILFTSRCMHGSEPN
49659919 42.5 4e-62	149	MPGTHNTWYFDE	HRIEPEPKINQKLTGGKKTG	VYGYDYKLLKLDPNWEPDESAVHLEMQPGQFILFTSRCMHGSEPN
23452293 43.1 4e-62	150	MPGTHNTWYFDE	HRIKFEPDINQKLTGKQKSG	FYGYDYKLLKLDPNWEPDESAVHLEMQPGQFILFTSRCMHGSEPN
49659921 42.7 3e-61	150	MPGTHNTWYFDE	HRIKFEPDINQKLTGKQKSG	FYGYDYKLLKLDPNWEPDESAVHLEMQPGQFILFTSRCMHGSEPN
48766808 43.6 1e-60	150	MPGTHNTWYFDE	HRIKFEPEKINQKLTGKQKSG	FYGYDYKLLKLDPNWEPDESAVHLEMQPGQFILFTSRCMHGSEPN
48766804 43.6 5e-60	150	MPGTHNTWYFDE	HRIKFEPEKINQKLTGKQKSG	FYGYDYKLLKLDPNWEPDESAVHLEMQPGQFILFTSRCMHGSEPN
24985462 49.7 6e-60	159	IPCTQHARFYDES	KAAAYSQDKHGLV.KNGVAR	GFYGYDYRDLQVDPDWKPEDESKAIYFTMKKGEPIIFWSTLLHGSYFPH
67755172 42.8 9e-58	153	MPGSHRTWFFDE	KRNIPPEPENNRVLDDGQKSG	FYGYDYKLLKLDPNWEPDEARAVHMEVVKPGQFPIFTSRCMHGSEPN
67763421 42.8 2e-57	153	MPGSHRTWFFDE	KRNIPPEPENNRVLDDGQKSG	FYGYDYKLLKLDPNWEPDEARAVHMEVVKPGQFPIFTSRCMHGSEPN

3. Amplificazione per PCR e sequenziamento

```
1  N Y D R H L D V D L L S E H I M R E E I  
1  AACTACGACCGCCACCTCGACGTCGACCTGCTCTCGAGCACATCATGAGGGAGGAGATC  
  
21  V D R V V S I L G P D V L C W R T E F F  
61  GTCGACCGGTCGTCTCCATCCTGGGGCCCGACGTCCTGTGCTGGCGCACGGAGTTC TTC  
  
41  P K Y P G D E G T D W H Q A A T F A N A  
121  CCCAAGTACCCGGGTGACGAGGGGACC GACTGGCACCAGGCGGCGACGTTCCCAACGCC  
  
61  S K K P Q I M W P S E E G R P A F G G T  
181  AGCAAGAAGCCGAGATCATGTGGCCGTCCGAGGAGGGCCGGCCCGCTTTCGGCGGCACC  
  
81  I T V W T A F T D S T L E N G C L Q L M  
241  ATCACCGTGTGGACCGCTTTCACCGACTCGACGCTGGAGAACGGCTGCCTCCAGCTGATG  
  
101  P G T H K H M N Y D E S L A M D Y D P N  
301  CCCGGCACCCACAAGCACATGAACTACGACGAGTCGCTGGCCATGGACTACGACCCCAAC  
  
121  A I N S R E K D G V K R G F F G Y D Y  
361  GCCATCAACAGCCGTGAGAAGGACGGCGTCAAGCGGGCTTCTTCGGCTACGACTAC
```

Analisi di *Southern blotting*

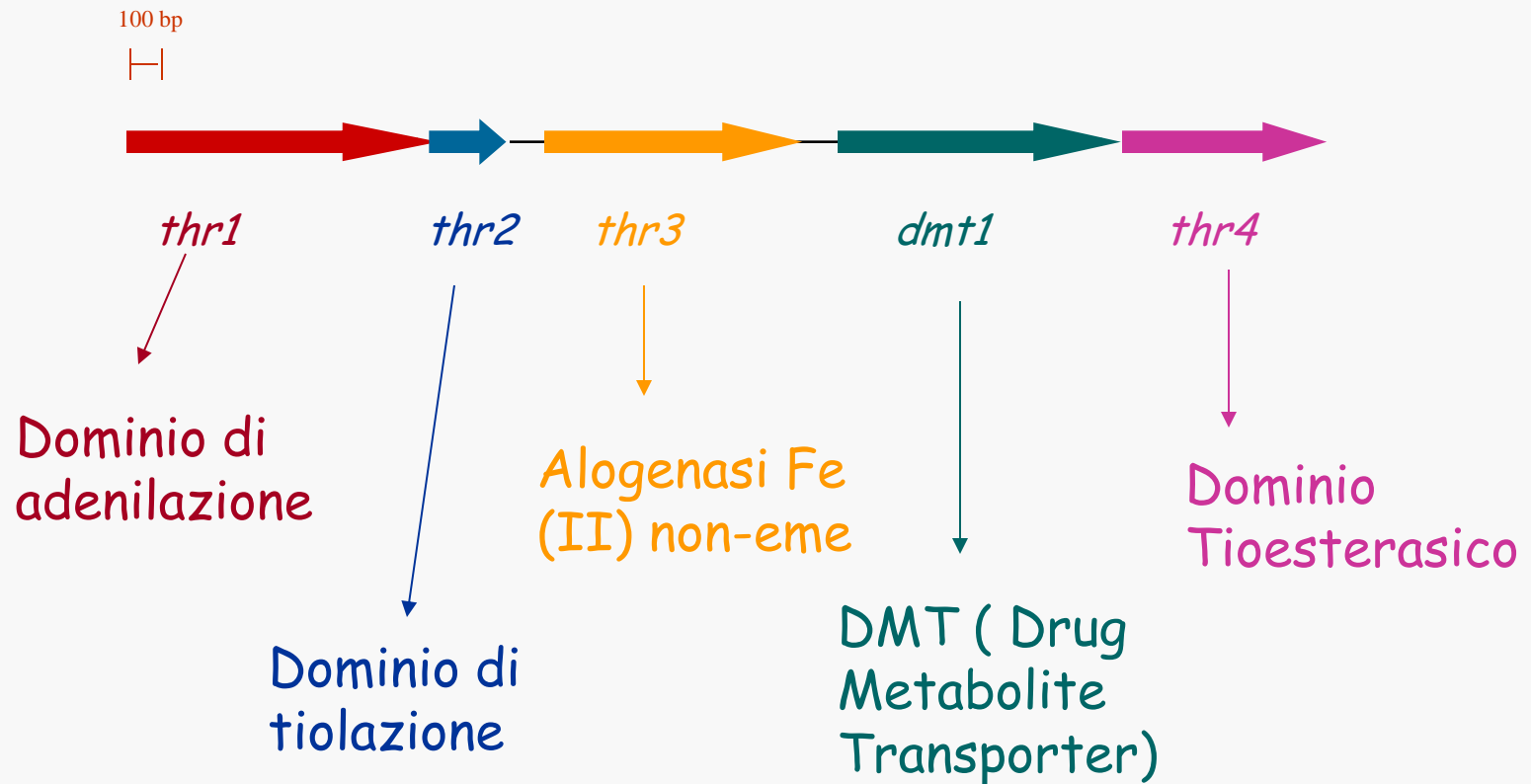


Southern blotting effettuato sul DNA genomico di *Streptomyces* sp. OH-5093 digerito con BamHI (1), PstI (2), RsaI (3), SacII (4), SmaI (5), EcoRI (6), XhoI (7), SphI (8), BfaI (9). La sonda utilizzata è JV2 marcata con ^{32}P .

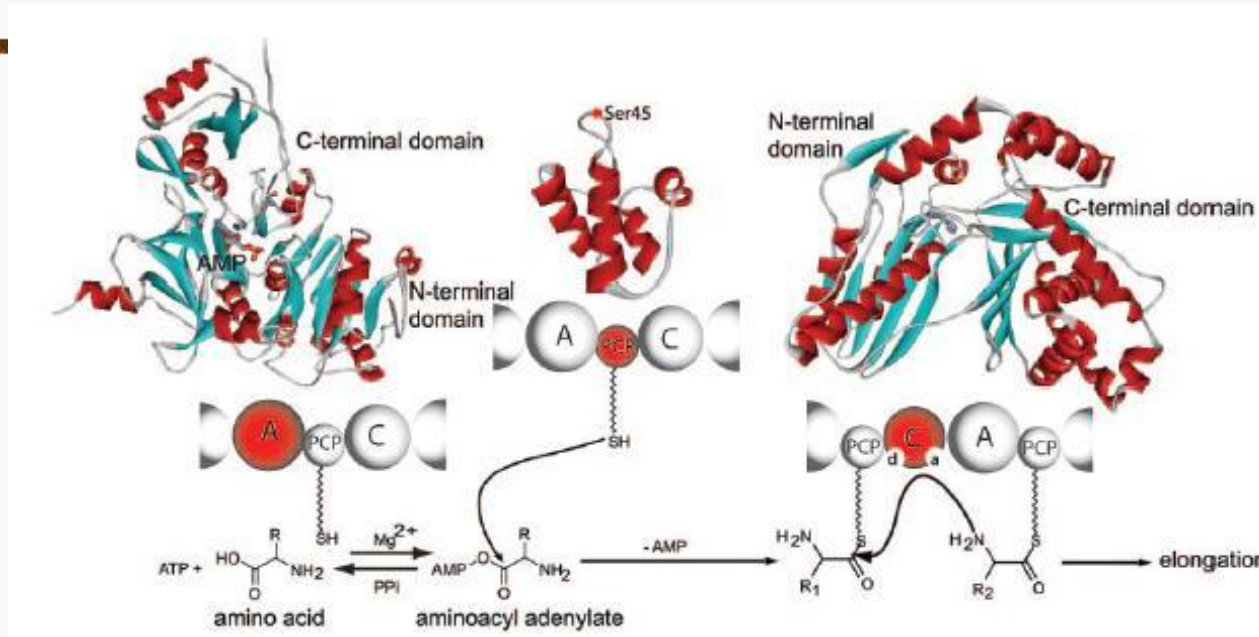
➤ L'alogenasi è presente in singola copia

- Clonaggio dei frammenti in pBlueScript
- Screening su colonia con sonda marcata radioattivamente

Organizzazione del cluster *thr*



MECCANISMO DELLA SINTESI PEPTIDICA NON RIBOSOMIALE



- **A** seleziona e attiva l'amminoacido;
- **PCP** lega l'amminoaciladenilato alla 4'-fosfopanteteina, gruppo prostetico del dominio stesso;
- **C** catalizza la formazione del legame peptidico.

A3

1 A Y I I Y T S G T T G R P K G V P V R H
 1 GCCTACATCATCTACACCTCCGGCACACGGCCGCCCAAGGGTCCCGCTCCGGCAC

21 G N V T A L F E A C S R L F S F S A D D
 61 GGGAAATGTACCCGCTCTCTTCGAGGCTGTCTCCGCTGTTCTCCTTCCTCCGGCCGACGAC

41 R W L L F H S M A F D F S V W E I W G A
 121 CGCTGGCTGCTCTTCCACTCCATGGCTTCGACTTCTCGGTCTGGGAGATCTGGGTGGC

61 L S T G A E L V V L P Y W T A R T P V E
 181 CTGAGCACCGGCGCGGAACTGGTCGTCTGCTGCTACTGGACAGCCCGACCCCGCTGAG

81 T A R V V R D R G I T V L N Q T P T A F
 241 ACCGCGGAGTCTACGCGACCGGGGATCAGCGTGTGAACAGACGCGGACCGCCTC

101 G A L T T A V L G E G I D L P E L R Y V
 301 GGCGCGCTGACGACAGCGTCTCGGGCGAGGCATCGACTCGCGGAACTGCGCTACGTC

121 V F G G E K L T P A V V R P W A K R F G
 361 GTCTTCGGCGGCAGAAAGCTCACGCCCGCGTGGTACGCCCTGGGCCAAGCGTTCGGT

A5

141 L D R P H L I N M Y G I T E T T V H A T
 421 CTGGACCGGCGCACCTGATCAACATGTACGGCATCACCGAGACCAGGTGCACGCCAC

161 F H R L T E D D L A A E D S V I G R P L
 481 TTCACCGGCTGACCGAGGACGACCTGGCAGCGGAGACTCGGTCATCGGCGGCGCGCTG

181 P G F T H R I V T E D G R D A A T G E P
 541 CCGCGCTTACCCACCGGATCGTCACCGAGGACGGGCGGACCGCGGACCGGCGGAGCGC

A6

201 G E L W L A G P Q V S E G Y L N R P E L
 601 GGTGAACTGTGGCTGGCCGGTCCACAGGTGAGCGAGGGTATCTCAACCGGCCGGAACTC

221 T A E R F T T G P A R D G G A P P P R Y
 661 ACCGCGAAGCGTTCACCCAGGTCCCGCGGGACGGCGGGGCGCGCCACCGCGCTAC

A8

241 Y H S G D L V S R R A G G D L V Y Q G R
 721 TACCACTCGGCGACCTGTCAGCCGGAGCGGGGGCGACCTGGTCTACCAAGGCCGA

261 A D L Q V K L R G H R I E L S D V E A A
 781 GCCGACCTGCAGGTGAAGCTCCGAGGCCATCGCATCGAACTCAGCGATGTGAGGCCGCG

281 V R T H P A V V D A V V W V H E F A P G
 841 GTGGCACGACCCCGGTTGGTCGACGCGGTTGGTGGTGCACGAGTTCGACCGGGC

301 D S R L V C A Y T A Q A S E E G T G P D
 901 GACAGCCGCTGTGTGGCCTACACAGCACAGGCTCCGAGGAGGACCGGACCCGAC

321 A R A L R A H V K T V L P S Y M Q P S Q
 961 GCCGTGCGCTGGCGCCCATGTGAAGACAGTCTGCGCTGTACATGACGCCCTCGCAG

A10

341 Y L A L P E L P R T I N G K A D R A S V
 1021 TATCTGGCGTGGCCGAGCTGCGCGCACGATCAACGGAAAGCCGACCGGGCTCCGTG

361 A R A F D E R R -
 1081 GCCCGAGCCTTCGACGAAAGGAGATGA

Thr1

Dominio di adenilazione

Presenza dei motivi conservati

Specificità per il riconoscimento della treonina

Identificazione regione 5' del gene thr1
mediante TAIL PCR


Thr1

Dominio di adenilazione

 Regioni del promotore

 Possibili codoni d'inizio

 Shine - Dalgarno

 Motivi conservati nei domini di adenilazione

 Primer utilizzati nella TAIL-PCR

```
1  R L G R M R R Q R R P A A I - A A W P L
3  CGACTGGGACGCATGAGGCGACAGCGGCGACCGGCCATTGAGCGGCCTGGCCCCTT

21  L L S G P F L V N S Q R I H R L R F A V
63  TTGCTAAGTGGGCCGTTTCTCGTCAACTCCCAGCGGATTACAGACTCCGATTGCTGTCT

41  K T - V Q P E C P N P - R S H V T K I T
123 AAGACATAGGTTTCCAGCCAGAGTGTCCAAACCCTTGACGATCACATGTAACTAAGATCACA

61  F I Y G G V R T D I D R V T M - C D F G
183 TTTATCTACGGCGGGTCCGGACAGACATCGATCGTGTAACGATGTAGTGCCTGACTTCGGG

81  P R W T G G F G H L H N G A I V S E K K
243 CCACGTTGGACTGGTGGATTCCGGTCATCTACACAATGGAGCCATTGTGTCCGAAAAGAAG

101  D Q Y P S P G Y S E F G A T L V D L F S
303 GACCAGTACCCCTCGCCAGGATATTCGGAATTCGGCGCGACACTGGTTGACTTGTCTTCC

121  R A A M E M P D R T A L H I D D E K I S
363 AGGGCAGCCATGGAAATGCCCGACCGCACGGCGCTCCACATCGACGACGAGAAGATCTCG
      A1
141  Y G L I H S W A E G L A D L L H D A G V
423 TACGGCCTCCTCCATTCTGGGCGAGGGCCTCGCGGATCTTCTCCACGACGCCGGGGTA

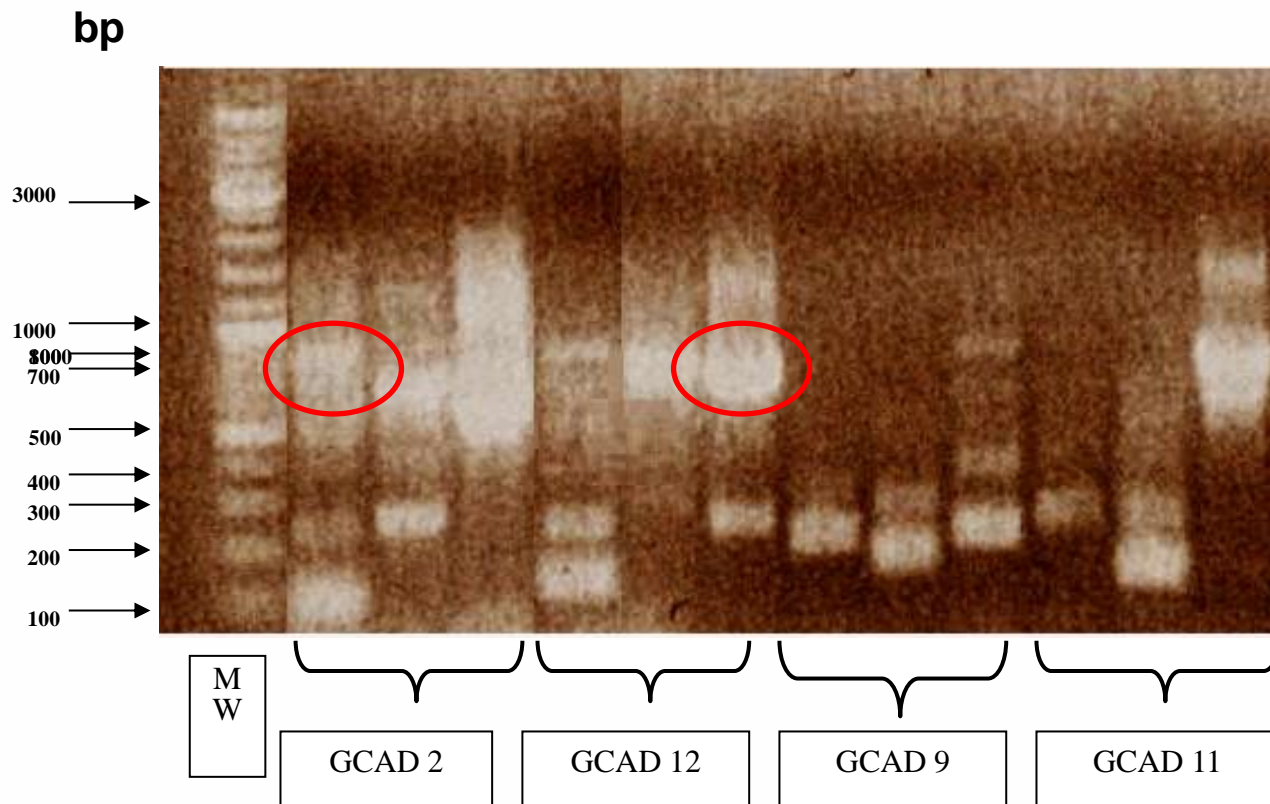
161  R K G D R V A L R M P P G A N A I A A M
483 CGGAAGGGTGATCGTGTGCGACTGCGTATGCCGCGGGGCGGAACGCGATCGCCGCCATG
      A2
181  L G I L R A G A A Y V P L D I R N P P A
543 CTGGGAATCCTCCGGGCCGGCGCCGCTATGTTCCGCTCGACATTGCAACCCACCCGCT

201  R N A F I V T D S Q V V A L V G D P G D
603 CGCAACGCCTTCATCGTCCACGACAGCCAGGTGGTCCGCGTGGTGGGAGATCCCGGTGAT

221  I P E Y T G P L V T E E N V A A L R D R
663 ATTCCCGAATACAGGGGCCGCTCGTCCACGAGGAGAACGTGGCAGCACTGCGGGACCGT

241  E A I K D P G P E R P G P Q D V A Y I I
723 GAGGCGATCAAGGATCCCGGGCCCGAGCGGCCCGCAGGACGTGCGCTACATCATC
      A3
261  Y T S G T T G R P K G V P V R H G N V T
783 TACACCTCCGGCACACGGGCCGCCCAAAGGGTCCCCGTCGGGCACGGGAATGTCACC
      sp3
281  A L F E A C S R L F S F S A D D R W L L
843 GCTCTCTTCGAGGCCTGCTCCCGCTGTTCTCCTTCTCCGCGGACGACCGCTGGCTGCTC
      A4 sp2
301  F H S M A F D F S V W E I W G A L S T G
903 TTCCACTCCATGGCCTTCGACTTCTCGGTCTGGGAGATCTGGGGTGCCTGAGCACCGGC
      sp1
321  A E L V V L P Y W T A R T P V E T A R V
963 GCGGAACTGGTCGTCCCTGCCGTACTGGACAGCCCCGACCCCGTCGAGACCGCCCGATC
```

IDENTIFICAZIONE DELLA REGIONE 5'- DEL GENE *thr1*




Thr1

Dominio di adenilazione

 Regioni del promotore

 Possibili codoni d'inizio

 Shine - Dalgarno

 Motivi conservati nei domini di adenilazione

 Primer utilizzati nella TAIL-PCR

```
1 R L G R M R R Q R R P A A I - A A W P L
3 CGACTGGGACGCATGAGGCGACAGCGGCGACCGGCCATTGAGCGGCCTGGCCCCTT

21 L L S G P F L V N S Q R I H R L R F A V
63 TTGCTAAGTGGGCCGTTTCTCGTCAACTCCCAGCGGATTACAGACTCCGATTGCTGTG

41 K T - V Q P E C P N P - R S H V T K I T
123 AAGACATAGGTTTCCAGCCAGAGTGTCCAAACCCTTGACCGATCACATGTTAACTAAGATCACA

61 F I Y G G V R T D I D R V T M - C D F G
183 TTTATCTACGGCGGGGTCCGGACAGACATCGATCGTGTAAACGATGTAGTGGCGACTTCGGG

81 P R W T G G F G H L H N G A I V S E K K
243 CCACGTTGGACTGGTGGATTCCGGTCATCTACACAATGGAGCCATTGTGTCGGAAGAAG

101 D Q Y P S P G Y S E F G A T L V D L F S
303 GACCAGTACCCCTCGCCAGGATATTCGGAATTCGGCGCGACACTGGTTGACTTGTTCCTCC

121 R A A M E M P D R T A L H I D D E K I S
363 AGGGCAGCCATGGAAATGCCCGACCGCACGGCGCTCCACATCGACGACGAGAAGATCTCG
A1
141 Y G L I H S W A E G L A D L L H D A G V
423 TACGGCCTCCTCCATTCTGGGCGGAGGGCCTCGCGGATCTTCTCCACGACGCCGGGGTA

161 R K G D R V A L R M P P G A N A I A A M
483 CGGAAGGGTGATCGTGTGCGACTGCGTATGCCGCGGGGCGGAACGCGATCGCCGCCATG
A2
181 L G I L R A G A A Y V P L D I R N P P A
543 CTGGGAATCCTCCGGGCCGGCGCCGCTATGTTCCGCTCGACATTGCAACCCACCCGCT

201 R N A F I V T D S Q V V A L V G D P G D
603 CGCAACGCCTTCATCGTCCACGACAGCCAGGTGGTCCGCGTGGTGGGAGATCCCGGTGAT

221 I P E Y T G P L V T E E N V A A L R D R
663 ATTCCCGAATACAGGGGCCGCTCGTCCACGAGGAGAACGTGGCAGCACTGCGGGACCGT

241 E A I K D P G P E R P G P Q D V A Y I I
723 GAGGCGATCAAGGATCCCGGGCCCGAGCGGCCCGCAGGACGTGCGCTACATCATC
A3
261 Y T S G T T G R P K G V P V R H G N V T
783 TACACCTCCGGCACACGGGCCGCCCAAAGGGTCCCCGTCGGGCACGGGAATGTCACC
sp3
281 A L F E A C S R L F S F S A D D R W L L
843 GCTCTCTTCGAGGCCTGCTCCCGCTGTTCTCCTTCTCCGCGGACGACCGCTGGCTGCTC
A4 sp2
301 F H S M A F D F S V W E I W G A L S T G
903 TTCCACTCCATGGCCTTCGACTTCTCGGTCTGGGAGATCTGGGGTGGCGTGGACCCGGC
sp1
321 A E L V V L P Y W T A R T P V E T A R V
963 GCGGAAGTGGTCTGCTGCGTACTGGACAGCCCGCACCCCGTCCGAGACCGCCCGATC
```

STRATEGIE DI MUTAGENESI

Mutagenesi sito-specifica

Per valutare il ruolo strutturale/funzionale di specifici aminoacidi, si producono poche varianti sulla base di ipotesi predefinite

Mutagenesi casuale (evoluzione guidata in vitro)

Per modificare le proprietà di una proteina (termostabilità, specificità di substrato), si produce un gran numero di varianti che devono essere analizzate per identificare quelle migliori

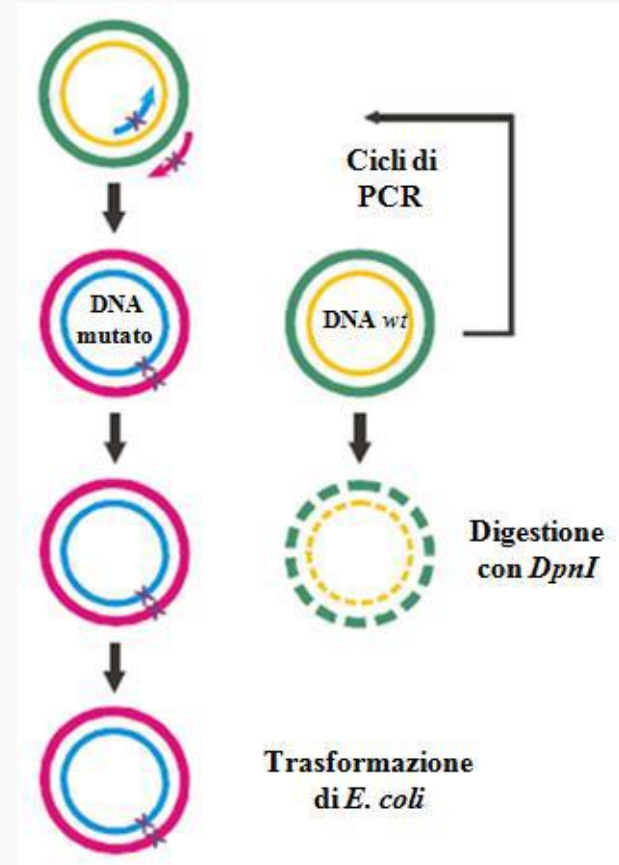
PCR error-prone

DNA shuffling

saturation

Mutagenesi sito-specifica QuikChange

1. PCR di amplificazione dell'intero plasmide con due oligo complementari che portano la mutazione
2. Digestione con DpnI per degradare il template wild type (metilato)
3. Trasformazione di *E. coli*
4. Verifica mediante sequenziamento del DNA



Mutagenesi sito-specifica: introduzione simultanea di mutazioni multiple (overlap extension PCR)

- PCR per amplificare diversi frammenti contenenti le mutazioni A B C parzialmente sovrapposti
- I frammenti amplificati vengono mescolati e sottoposti a due cicli di denaturazione ed annealing
- PCR con oligonucleotidi esterni per permettere di completare l'estensione dei filamenti mancanti

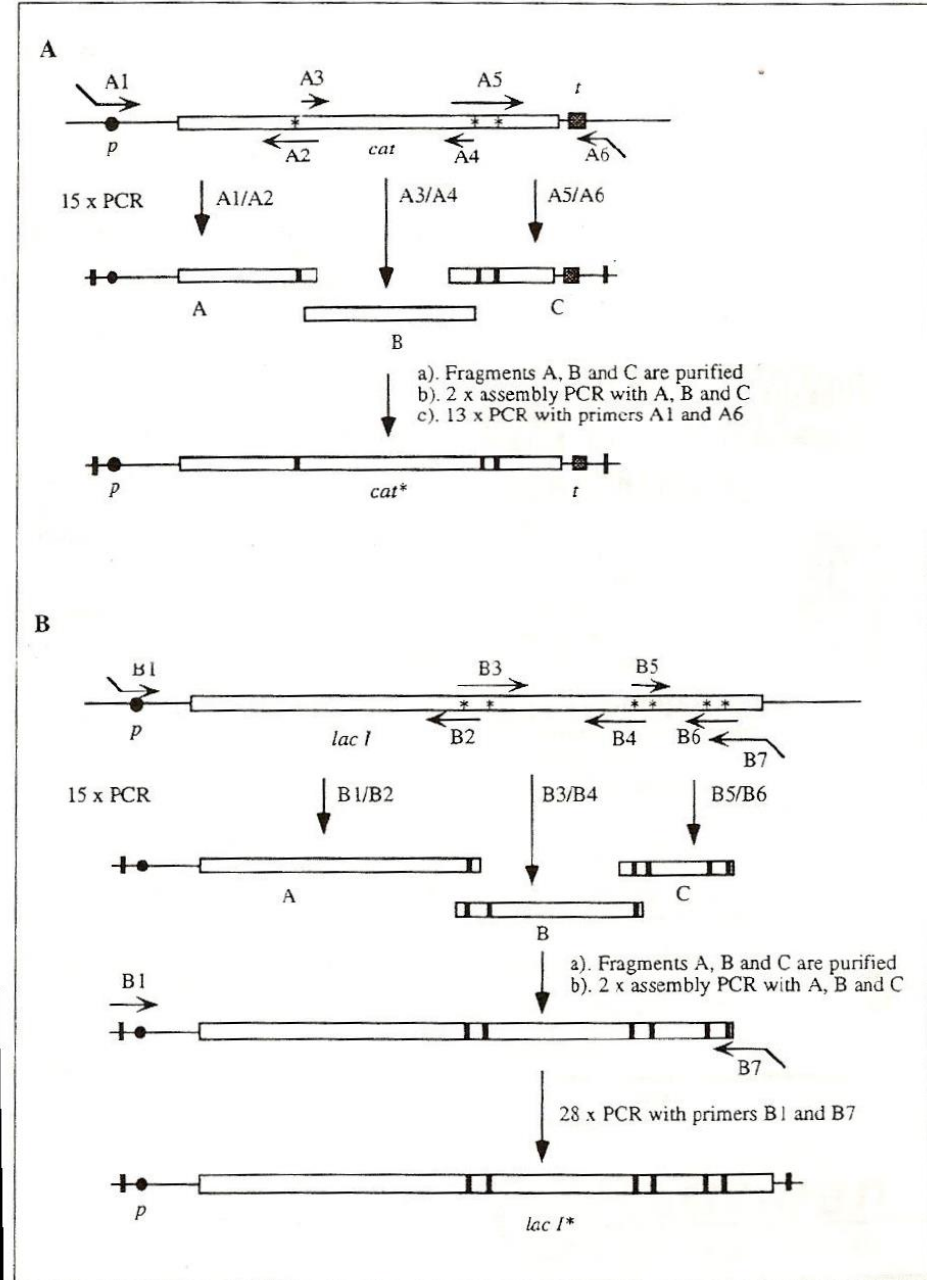


Figure 1. Outline of the simultaneous multisite mutagenesis procedure. The *cat* (A) and *lacI* (B) gene segments were amplified, purified and assembled to produce complete genes with introduced mutations. *P*, promoter; *t*, terminator; *, restriction site that is to be changed; and solid bar, introduced mutation.

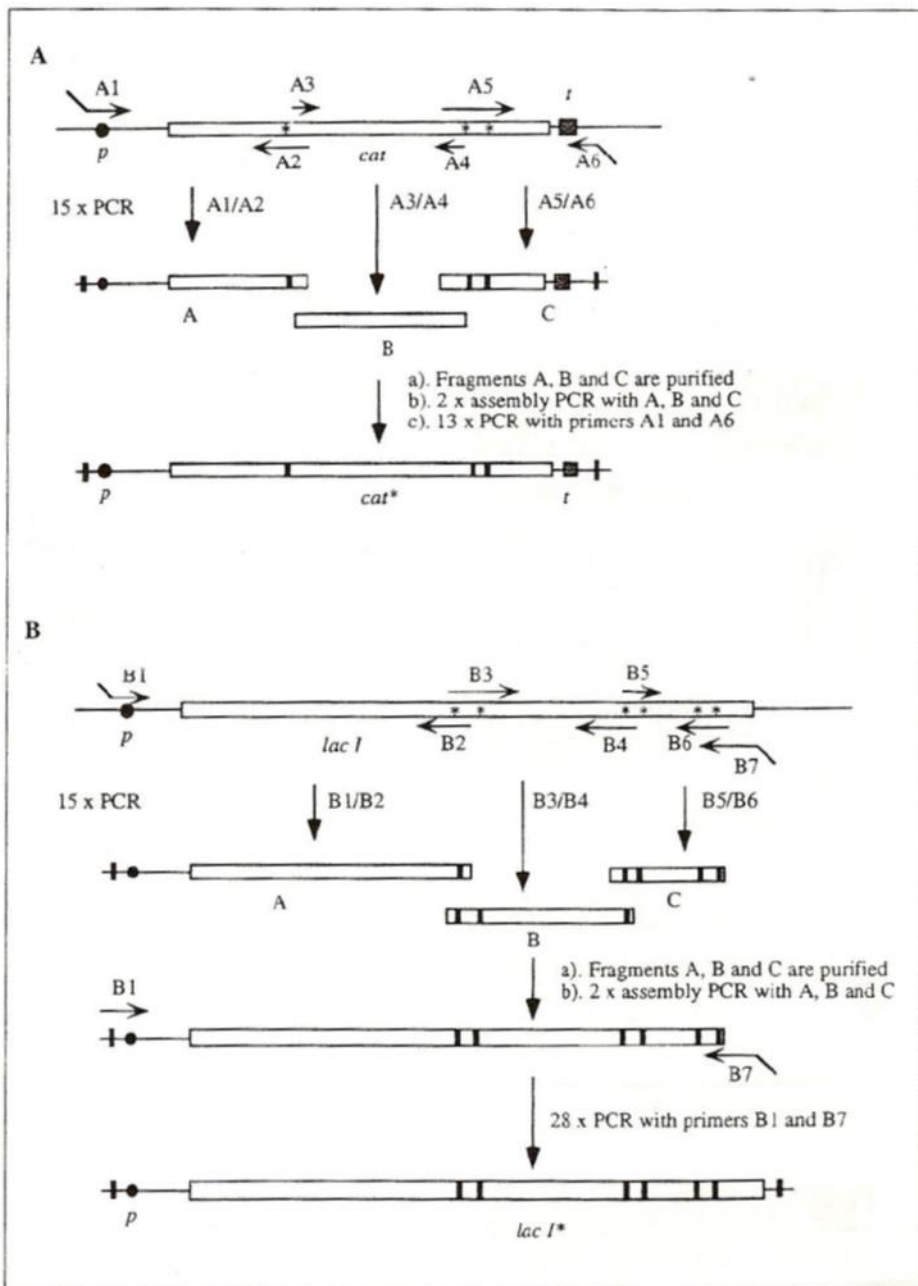


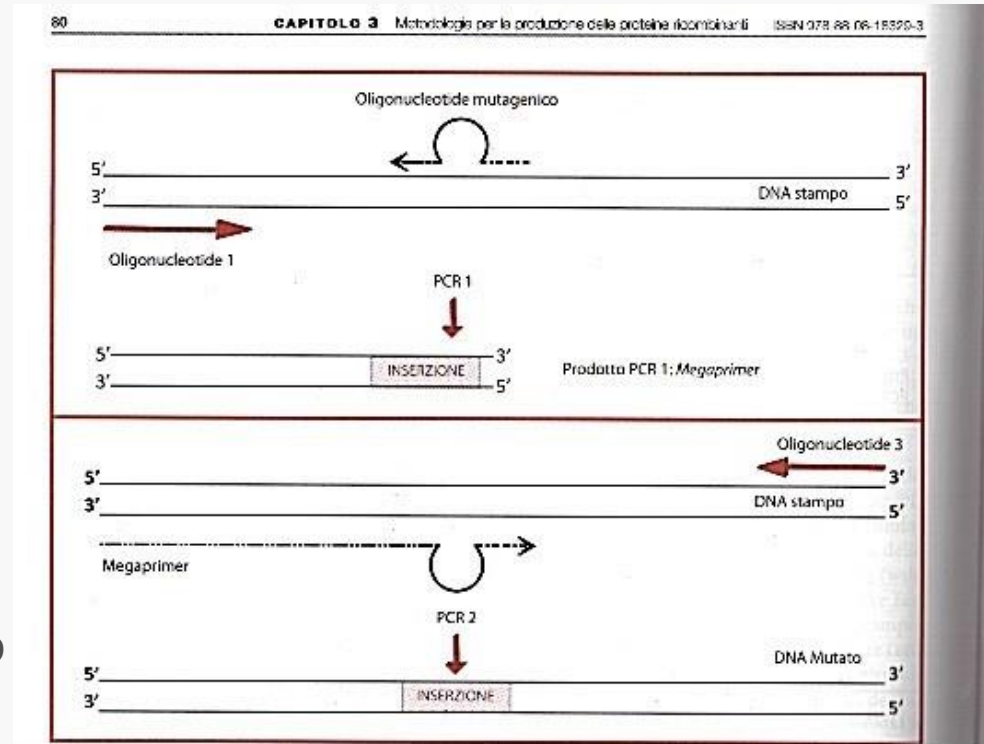
Figure 1. Outline of the simultaneous multisite mutagenesis procedure. The *cat* (A) and *lacI* (B) gene segments were amplified, purified and assembled to produce complete genes with introduced mutations.

Mutagenesi sito-specifica: introduzione simultanea di mutazioni multiple (megaprimer)

Si eseguono due PCR consecutive con
due diverse coppie di oligonucleotidi

IPCR :oligo esterno e oligo contenente
l'inserzione fiancheggiata da almeno 10
basi per l'appaiamento allo stampo a
monte e a valle

IIPCR : megaprimer (frammento ottenuto
dalla IPCR) e oligo 3 esterno



Mutagenesi casuale

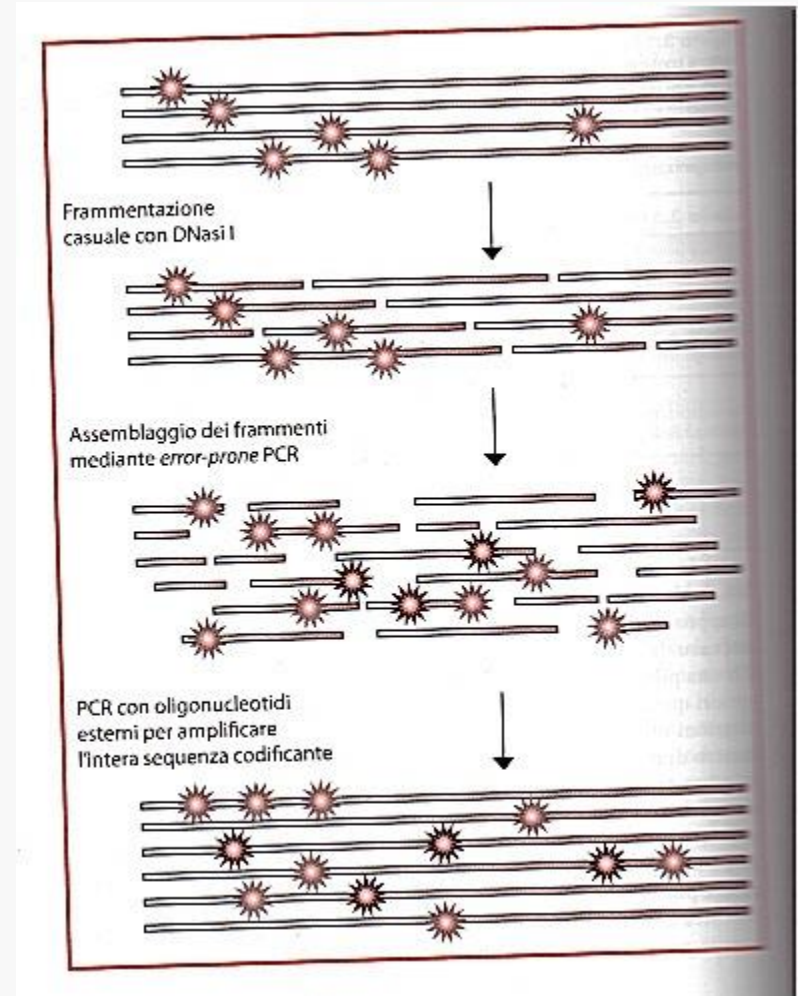
Evoluzione guidata in vitro: error-prone PCR

- PCR in condizioni tali da aumentare la frequenza di errore della Taq polimerasi:
 - aggiungendo Mn^{2+} (riduce la specificità di appaiamento delle basi),
 - sbilanciando le concentrazioni dei desossiribonucleotidi (se uno dei dNTP diventa limitante, potrebbe essere incorporato al posto un altro dNTP),
 - aumentando la concentrazione di Mg^{2+} (stabilizza l'appaiamento di basi non complementari),
 - aumentando la concentrazione di Taq polimerasi (aumenta la probabilità di allungamento di estremità che derivano da inneschi non corretti)

Mutagenesi casuale Evoluzione guidata in vitro: DNA shuffling

Il DNA proveniente da geni omologhi
) o da esperimenti di mutagenesi casuale
viene

- Digerito con DNAsi o enzimi di restrizione
- Riassembleto in una PCR senza oligo
- Amplificato con oligo esterni per ricostituire un'intera sequenza codificante ibrida



SCREENING

PER IDENTIFICARE I MUTANTI CON LE PROPRIETA' DESIDERATE