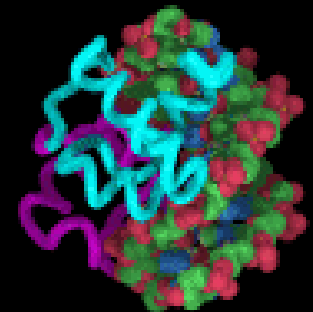


Primer design using genetic algorithm

DIEGO POGGIOLI

Soft Computing

Bologna, 18 Maggio 2007



PCR

Polymerase Chain Reaction

Tecnica di biologia molecolare, fu ideata nel 1983 da Kary B. Mullis, che consente la moltiplicazione (*amplificazione*) di frammenti di acidi nucleici dei quali si conoscano le sequenze nucleotidiche iniziali e terminali.

L'amplificazione mediante PCR consente di ottenere in vitro molto rapidamente la quantità di materiale genetico necessaria per le successive applicazioni.

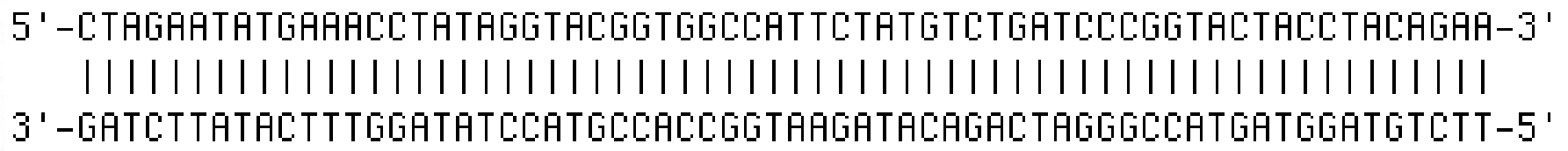
- DNA da amplificare;
- dNTP;
- Primer ;
- elementi di supporto (ad es. ioni magnesio), necessari per costituire l'ambiente adatto alla reazione;
- DNA polimerasi.





Choosing primers

Because DNA polymerase can add subunits (nucleotides) only to the 3' end of the primer, the primer has to be situated "upstream" – i.e., more 5' than – the sequence to be copied:





Cycle 1

Step 1. Denature the template DNA

Raise temperature to 92-94°C

5'-CTAGAATATGAAACCTATAGGTACGGTGGCCATTCTATGTCTGATCCCGGTACTACCTACAGAA-3'
|||||
3'-GATCTTATACTTTGGATATCCATGCCACCGGTAGATACAGACTAGGGCCATGATGGATGTCTT-5'

Cycle 1

Step 2. Anneal the primers

Drop the temperature to ~65°C



Cycle 1

Step 3. Primer extension

Allow DNA polymerase to copy the template -- 72°C

5'-CTAGAATATGAAACCTATAGGTACGGTGGCCATTCTATGTCTGATCCCGGTACTACCTACAGAA-3'

|||||
3'-GGCCATGATGG-5'

DNA polymerase

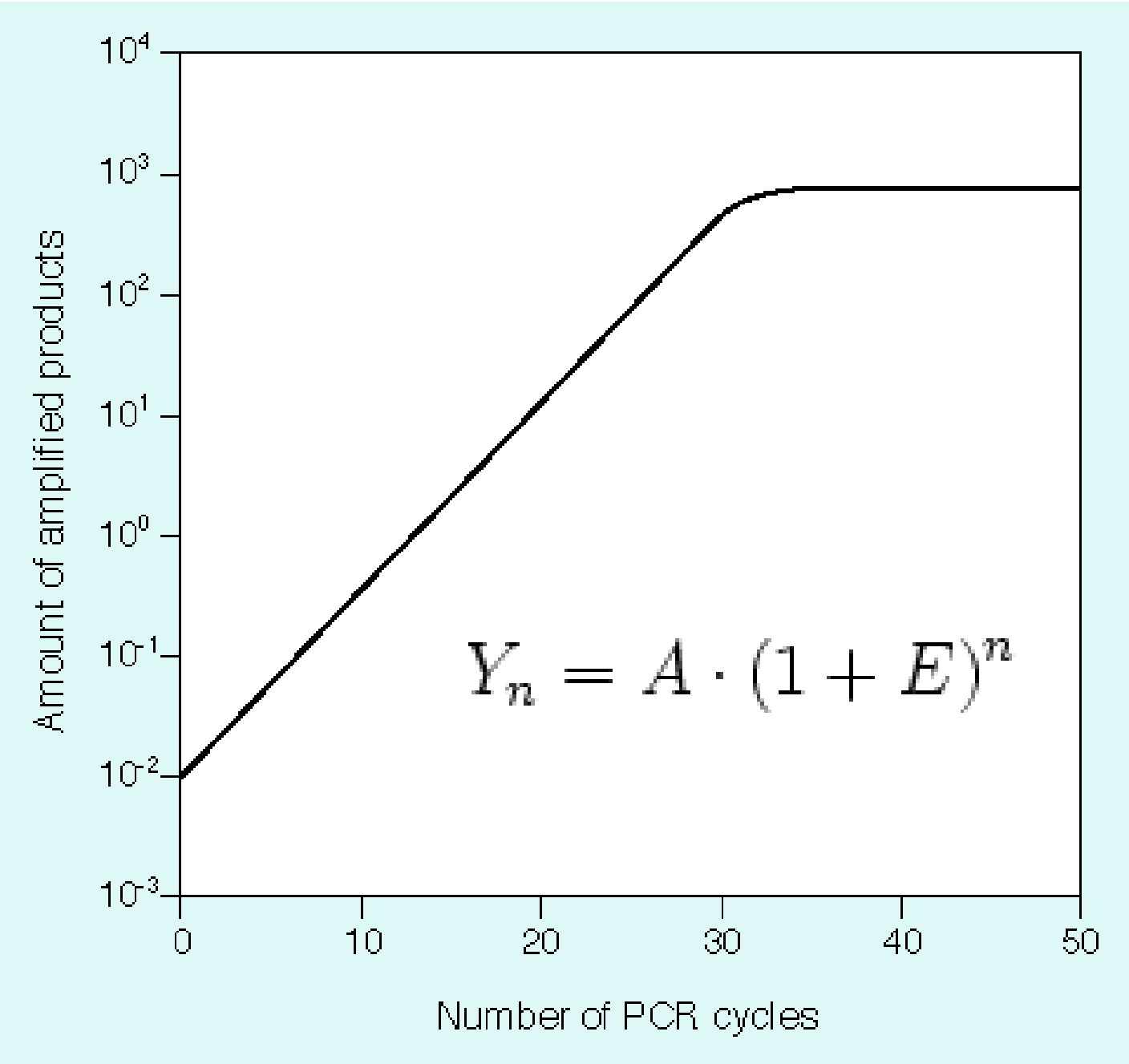
5'-ATGAAACCTATAG-3'

|||||

3'-GATCTTATACTTTGGATATCCATGCCACCGGTAGATACAGACTAGGGCCATGATGGATGTCTT-5'



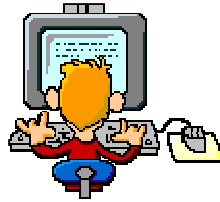
PCR Demo



parametri

- Lunghezza (16 – 28 bp; $\Delta < 3$)
- Temperatura di Melting ($\Delta < 5$)
- Composizione in basi ($40\% < GC < 60\%$)
- 3'-terminale (G/C)
- Ripetizioni e auto-complementarietà di sequenza
- Complementarità tra coppie di primer

Disegno dei PRIMER



'GeneFisher'
Meyer *et al.*
(1995)

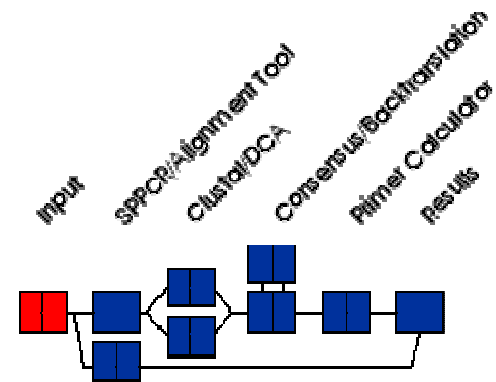
Consensus- DEgenerate Hybrid
Oligonucleotide Primer (CODEHOP)
Rose *et al.* (1998)

'Primer Premier'
Singh *et al.* (1998)

Dynamic programming to design primers.
Kämpke *et al.* (2001)

'GeneFisher' Meyer *et al.* (1995)

Interactive Primer Design



- Allineamento della sequenza target di DNA con altre sequenze note di DNA, con la stessa funzione
- Disegno del primer che può essere in accordo con la sequenza incognita.

Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences

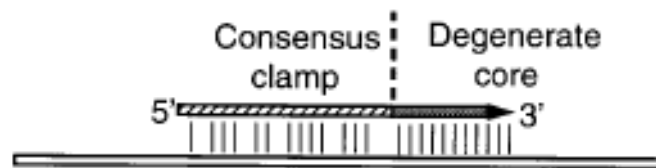
Timothy M. Rose, Emily R. Schultz, Jorja G. Henikoff¹, Shmuel Pietrokovski¹,
Claire M. McCallum¹ and Steven Henikoff^{1,2,*}

Department of Pathobiology, School of Public Health and Community Medicine, University of Washington, Seattle, WA 98195, USA and ¹Fred Hutchinson Cancer Research Center and ²Howard Hughes Medical Institute, 1100 Fairview Avenue N, Seattle, WA 98109-1024, USA

Received December 22, 1997; Revised and Accepted February 12, 1998

CODEHOP

Primer-to-template annealing (1/degeneracy):



Primer-to-product annealing (all primers):



	Y	G	D	T	D	S	V	F	M	S	C	R
hHSV1	TAC	GGG	GAC	ACG	GAC	TCC	ATA	TTT	GTG	CTG	TGC	CGC
hVZV	TAT	GGA	GAT	ACG	GAT	TCT	GTG	TTT	ATC	CGA	TTC	AAG
eHV1	TAC	GGA	GAC	ACC	GAC	TCC	GTG	TTT	ATC	AAG	TTT	GTG
hHV6	TAT	GGT	GAT	ACG	GAT	AGC	ATC	TTT	ATG	TCT	GTC	AGA
hHV7	TAT	GGT	GAT	ACT	GAC	AGT	CTT	TTT	GTT	ACT	TTC	AAA
hCMV	TAC	GGG	GAC	ACG	GAC	TCC	ATA	TTT	GTG	CTG	TGC	CGC
gpCMV	TAC	GGG	GAC	ACG	GAC	AGC	ATC	TTT	GTC	ATA	TGC	GGC
mCMV	TAC	GGC	GAC	ACC	GAC	AGT	GTG	TTT	GCG	GCT	TTC	TAC
HVS	TAT	GGA	GAC	ACA	GAC	TCT	CTA	TTT	GTA	GAA	TGT	GTT
hEBV	TAC	GGG	GAC	ACG	GAC	TCG	CTG	TTT	ATC	GAG	TGC	CGG
iHV1	TAT	GGG	GAT	ACG	GAT	AGT	ACG	ATG	CTG	TAC	CAC	CCA

5' TAY GGN GAY ACN GAC TCC GTG TTT GTC GCA TGC CG 3'
GDTD1B 3' ATR CCN CTR TGN CTG AGG CAC AAA CAG CGT ACG GC 5'
* * * *

- Non considera i siti di restrizione

Primer Premier: Program for Design of Degenerate Primers from a Protein Sequence

Vinay K. Singh, A.K. Mangalam¹, S. Dwivedi¹ and S. Naik¹

Premier Biosoft International, Palo Alto, CA, USA and ¹Sanjay Gandhi Post Graduate Institute of Medical Sciences, Lucknow, India

BioTechniques 24:318-319 (February 1998)

- Fornisce un grafico degli enzimi di restrizione sulla sequenza di DNA mediante la ricerca di siti di restrizione sulle sequenze di DNA allineate
- In accordo con il grafico viene disegnato il primer che contiene all'interno il sito di restrizione
- In alcuni casi non riesce a trovare la soluzione perchè le proprietà non possono essere soddisfatte

Efficient primer design algorithms

Thomas Kämpke¹, Markus Kieninger² and Michael Mecklenburg³

¹Forschungsinstitut für anwendungsorientierte Wissensverarbeitung FAW, Helmholtzstr. 16, 89081 Ulm, Germany, ²jSoft GmbH, Bahnhofstr. 7, 73447 Oberkochen, Germany and ³INTERACTIVA Bioteknik AB & BT Biomedical Technology, IDEON Research Park, Ole Römersväg 12, S 223 70 Lund, Sweden

Received on March 8, 2000; revised on September 14, 2000; accepted on October 19, 2000

<http://doprimer.interactiva.de>

→ Programmazione dinamica

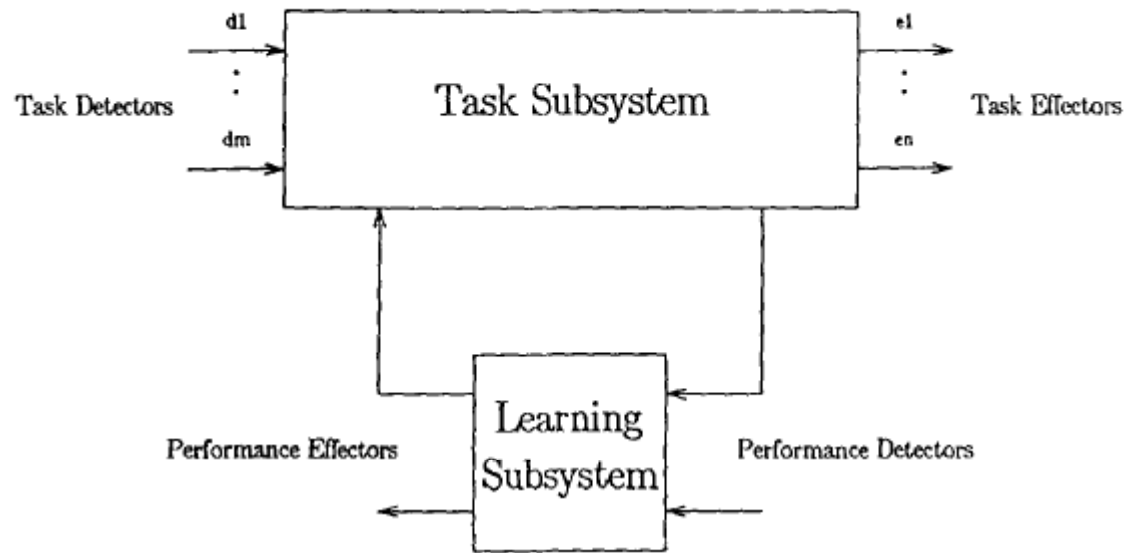
- Primer multipli per sequeze target multiple
- Alcuni criteri importanti non considerati (es. G/C al 3' end)
- Rimosse alcune potenziali soluzioni che richiederebbero lunghi tempi di calcolo



GA Algoritmi Genetici

- generazione, in maniera casuale, una **popolazione iniziale**;
- creazione di una sequenza di nuove popolazioni, o **generazioni**. In ciascuna iterazione, gli individui della popolazione corrente sono usati per creare la generazione successiva, e a questo scopo si compiono degli ulteriori passi:
 - ciascun membro della popolazione corrente è valutato calcolandone il rispettivo **valore di fitness** (idoneità);
 - si determina un opportuno ordinamento di tali individui sulla base dei valori di fitness;
 - gli individui più promettenti sono selezionati come genitori;
 - a partire da tali individui si genera un pari numero di individui della generazione successiva, e ciò può avvenire secondo due modalità distinte, vale a dire effettuando cambiamenti casuali su un singolo genitore (**mutazione**) oppure combinando opportunamente le caratteristiche di una coppia di genitori (**incrocio**);
 - gli individui così generati vanno a sostituire i genitori consentendo la formazione della generazione successiva;
- infine, l'algoritmo s'interrompe quando uno dei criteri d'arresto è soddisfatto (numero massimo di iterazioni, controllo del miglioramento della funzione di fitness).

DARWIN, CHARLES, 1809-1882



- Modifica del set di parametri
- Modifica della composizione dei dati
- Modifica del codice

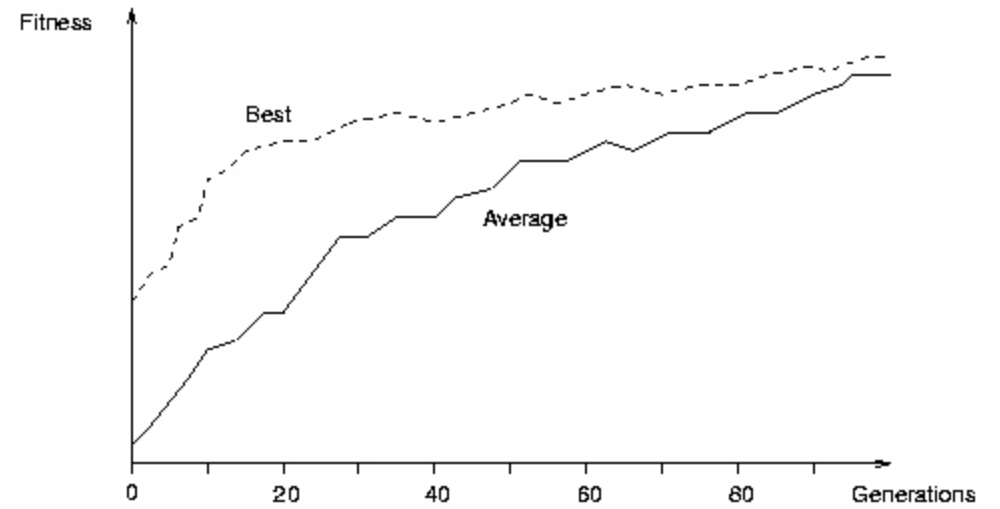
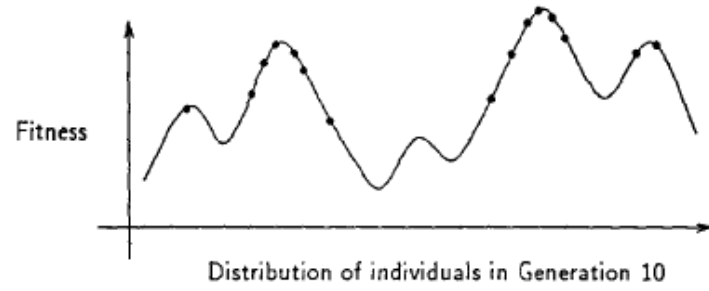
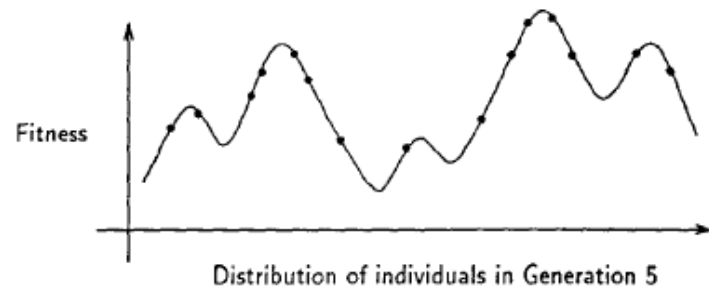
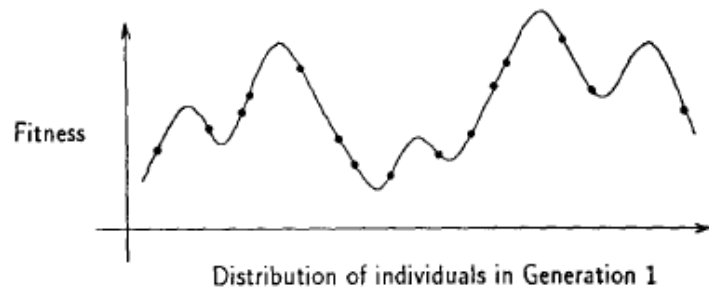


Figure 5: A Typical GA Run

Gli Schemata e il Teorema dello Schema

Uno *schema* è un modello di valori del gene che possono essere rappresentati (nella codifica binaria) da una stringa di caratteri dell'alfabeto $\{0,1,\#\}$.

“1010”: 10##, #0#0, ##10 e ##1# ...

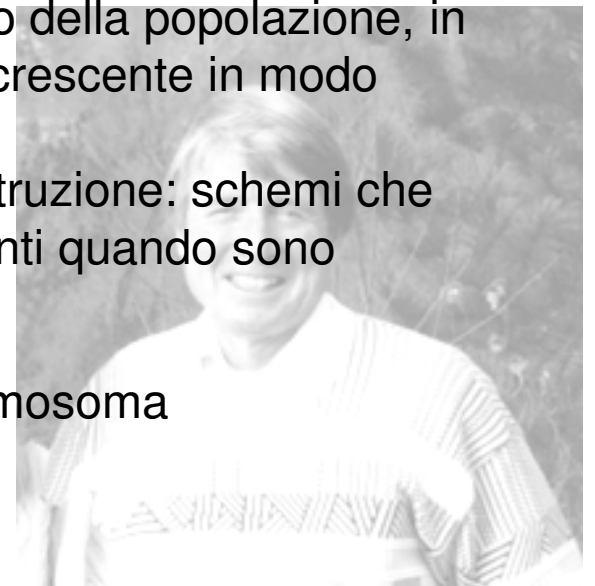
→ Agli individui della popolazione viene data la possibilità di riprodursi; il numero di opportunità che ogni individuo riceve è in proporzione al suo fitness, quindi i migliori individui contribuiscono maggiormente ai geni della generazione successiva

→ Alto valore di fitness = buoni schemi

→ Assegnare probabilità di riproduzione in numero sempre maggiore agli individui che hanno il fitness più elevato rispetto al resto della popolazione, in modo che i buoni schemi abbiano un numero di prove crescente in modo esponenziale nelle generazioni successive

→ Ipotesi **Building Block**: trovare buoni blocchi di costruzione: schemi che lavorano bene insieme e tendono a portare miglioramenti quando sono incorporati nello stesso individuo

1. i geni correlati siano vicini all'interno del cromosoma
2. ci sia poca interazione tra i geni



ESPLORAZIONE E SFRUTTAMENTO

1. la popolazione è infinita
2. la funzione fitness riflette accuratamente l'utilità della soluzione
3. i geni in un cromosoma non interagiscono significativamente

Problemi di Fitness Range

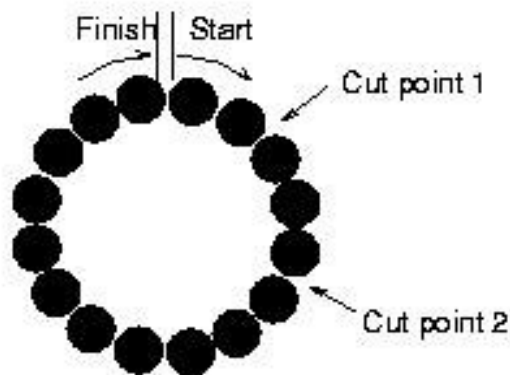
Convergenza Prematura

Fine lenta

Tecniche di Selezione dei Genitori → "roulette wheel selection"
→ fitness rimappato implicitamente
→ fitness rimappato esplicitamente

Generation gaps e steady-state replacement

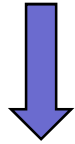
Tecniche di Crossover → 2-point Crossover
→ Crossover Uniforme



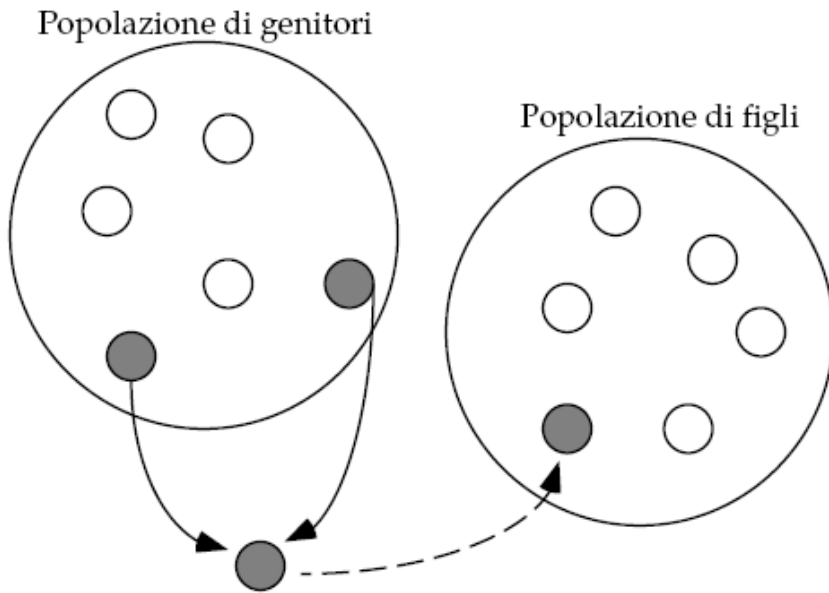
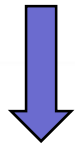
Crossover Mask	1	0	0	1	0	1	1	1	0	0
Parent 1	1	0	1	0	0	0	1	1	1	0
	↓			↓		↓	↓	↓		
Offspring 1	1	1	0	0	0	0	1	1	1	1
		↑	↑		↑				↑	↑
Parent 2	0	1	0	1	0	1	0	0	1	1

Modello di sovrapposizione di popolazione

Modello di generazione discreto



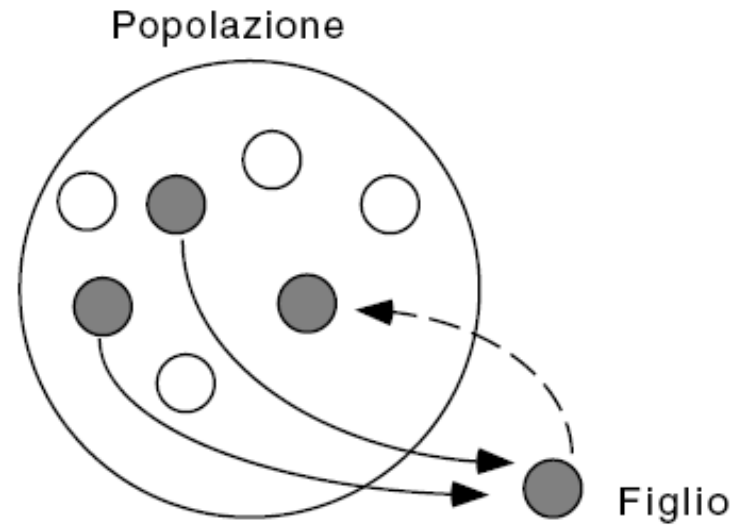
Rimpiazzamento generazionale

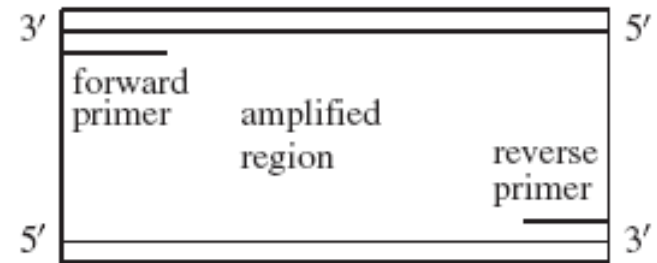
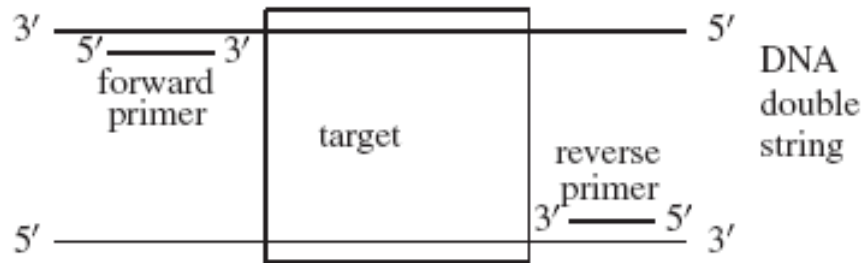


Modello di generazione continuo



Rimpiazzamento a stato continuo (modello statico di popolazione)





Le proprietà per il disegno dei primer sono trattate nella funzione di fitness, le soluzioni che non obbediscono a queste regole sono eliminate dalla competizione.

Siti di restrizione

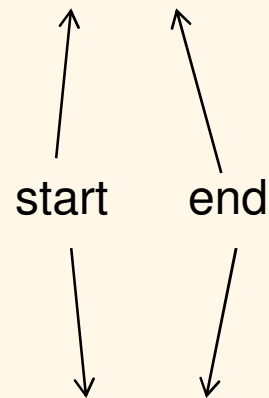
Specificità

Definizione dell' algoritmo proposto

$G_D = \text{AATCGACCAT} \dots$

$\overline{G_D} = \text{TTAGCTGGTA} \dots$

$B_f = \{b_i \mid i \text{ è l'indice di } G_D \text{ tra } F_s \text{ e } F_e\}$



$B_r = \{b_i \mid i \text{ è l'indice di } \overline{G_D} \text{ tra } R_s \text{ e } R_e\}$

Posizioni attuali dei primer

$$P_t = (F_s, F_e, R_s, R_e)$$

Posizioni relative dei primer

$$P_t' = (F_s, \alpha, \beta, y)$$

$$\alpha = (F_e - F_s)$$

$$\beta = (R_s - F_e)$$

$$y = (R_e - R_s)$$

$$(145, 164, 989, 1011)$$

$$(145, 19, 85, 22)$$

$$|P1| = \#G + \#C + \#A + \#T$$

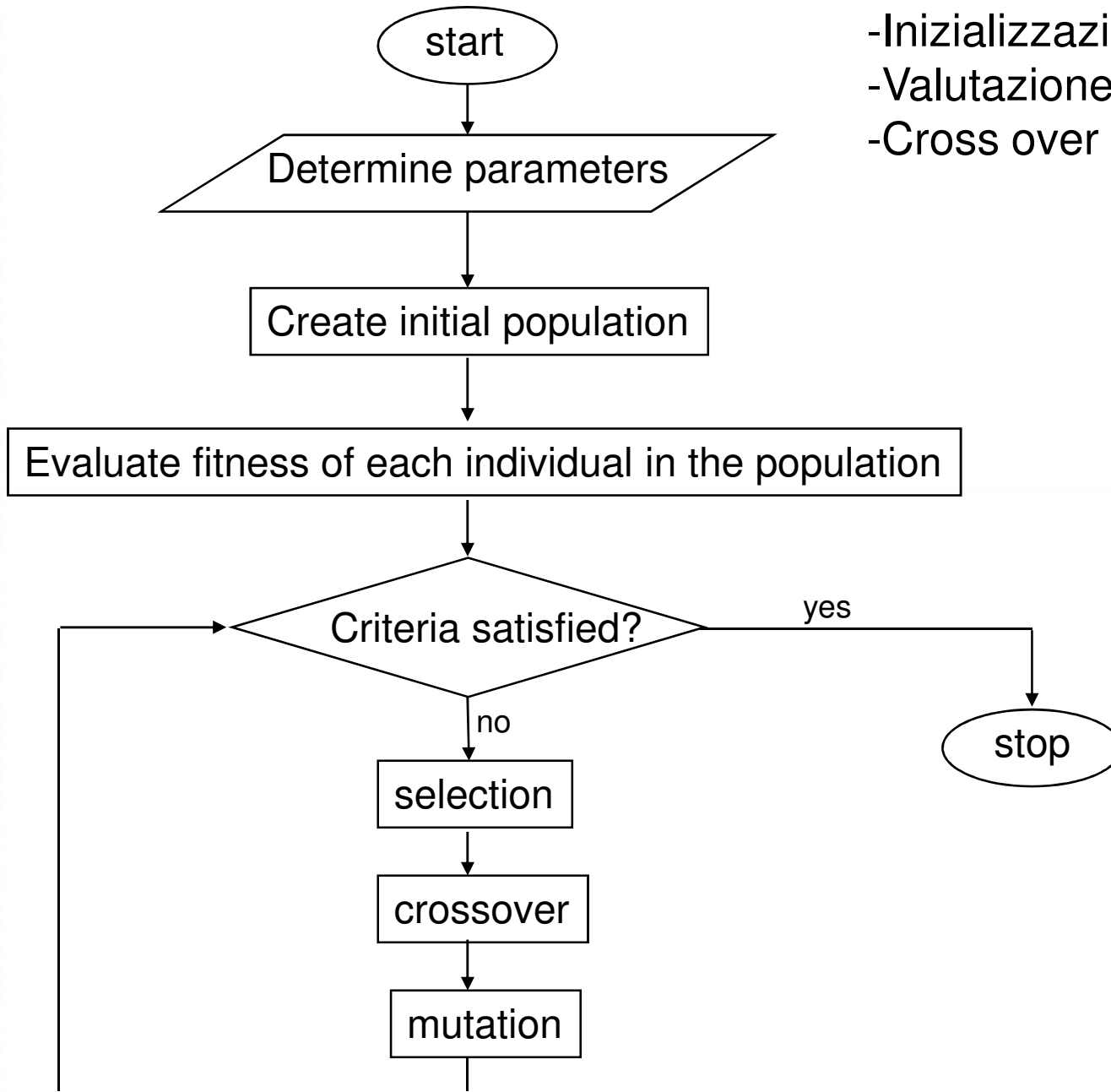
$$Tm(P1) = (\#G + \#C) * 4 + (\#A + \#T) * 2$$

$$GC(P1) = \frac{\#G + \#C}{|DA|} \times 100\%$$

$$\text{Specificità} \rightarrow \text{Uni}(P1) = \begin{cases} 0, & \text{se } P1 \text{ appare in } Gd \text{ una volta} \\ 1, & \text{se } P1 \text{ appare in } Gd \text{ più di una volta} \end{cases}$$

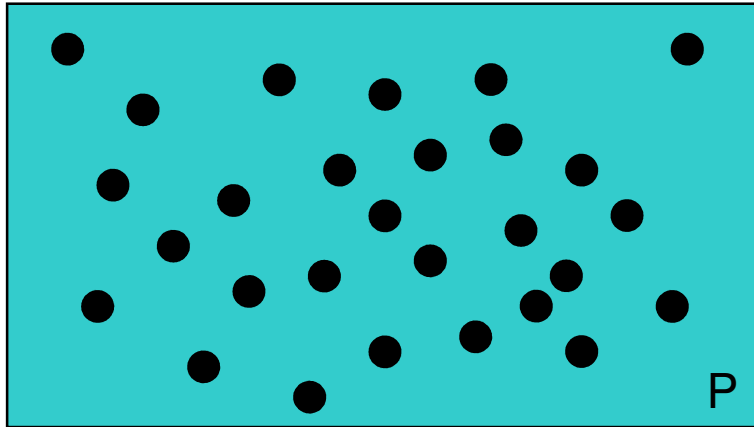
$$\text{Terminazione } 3' \rightarrow \text{Term}(P1) = \begin{cases} 0, & \text{se il } 3' \text{ finale è } G, C \text{ o } GC \\ 1, & \text{altrimenti} \end{cases}$$

Algoritmo proposto per il disegno di primer



- Inizializzazione del processo
- Valutazione
- Cross over e Mutazione

-Inizializzazione del processo



500 ● = Pt

Pt = (Fs₁₁, Fe₁₂, Rs₁₃, Re₁₄)
 (Fs₂₁, Fe₂₂, Rs₂₃, Re₂₄)
 (Fs₃₁, Fe₃₂, Rs₃₃, Re₃₄)
 (Fs₄₁, Fe₄₂, Rs₄₃, Re₄₄)

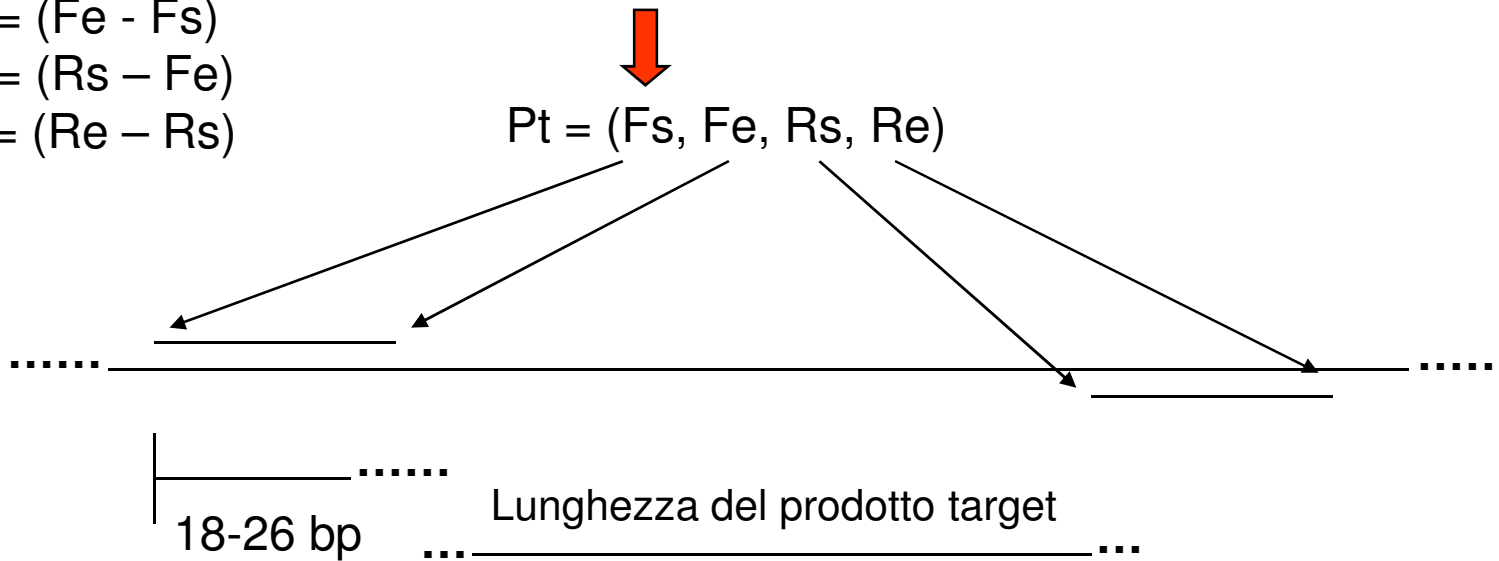
Pt' = (Fs, α, β, y)

α = (Fe - Fs)

β = (Rs - Fe)

y = (Re - Rs)

Pt = (Fs, Fe, Rs, Re)



-Valutazione

$$\text{Lunghezza} \rightarrow \text{Leng(Pt)} = \begin{cases} 0, & \text{se } 18 \leq |Bf|, |Br| \leq 26 \\ 1, & \text{altrimenti} \end{cases}$$

$$\text{Differenza tra lunghezze} \rightarrow \text{Lengd(Pt)} = \begin{cases} 0, & \text{se } \text{ABS}(|Bf| - |Br|) \leq 3 \\ 1, & \text{altrimenti} \end{cases}$$

$$\text{Differenza } T_m \rightarrow \text{Tmd(Pt)} = \begin{cases} 0, & \text{se } \text{ABS}[T_m(Bf) - T_m(Br)] \leq 5 \\ 1, & \text{altrimenti} \end{cases}$$

$$\text{GCp(Pt)} = \begin{cases} 0, & \text{se } 40\% \leq \text{GC}(Bf), \text{GC}(Br) \leq 60\% \\ 1, & \text{altrimenti} \end{cases}$$

-Valutazione

$$\text{Specificità} \rightarrow \text{Uni}(P1) = \begin{cases} 0, & \text{se } P1 \text{ appare in } Gd \text{ una volta} \\ 1, & \text{se } P1 \text{ appare in } Gd \text{ più di una volta} \end{cases}$$

$$\text{Uni}(Bf) = 0$$

$$\text{Uni}(Br) = 0$$

$$\text{Uni}(Pt) = \text{Uni}(Bf) + \text{Uni}(Br)$$

$$\text{Terminazione } 3' \rightarrow \text{Term}(P1) = \begin{cases} 0, & \text{se il } 3' \text{ finale è } G, C \text{ o } GC \\ 1, & \text{altrimenti} \end{cases}$$

$$\text{Term}(Bf) = 0$$

$$\text{Term}(Br) = 0$$

$$\text{Term}(Pt) = \text{Term}(Bf) + \text{Term}(Br)$$

-Valutazione

Sono consentite sequenze ripetute non invertite e sequenze auto-complementari >3bp

$$Sc(Pt) = \begin{cases} 0, & \text{se non c'è auto-complementarietà di Bf e Br} \\ 1, & \text{altrimenti} \end{cases}$$

$$PC(P) = \begin{cases} 0, & \text{se non c'è complementarietà a coppie tra Bf e Br} \\ 1, & \text{altrimenti} \end{cases}$$

$$\text{Terminazione 3'} \rightarrow \text{Term}(P1) = \begin{cases} 0, & \text{se il 3' finale è G, C o GC} \\ 1, & \text{altrimenti} \end{cases}$$

...XXXX[GC]

~~...XXXX[GC]{0,3-n}~~

-Valutazione

Siti di restrizione

Restriction enzyme	Sequence
<i>Apa</i> I	GGGCCC
<i>Avr</i> II	CCTAGG
<i>Bam</i> HI	GGATCC
<i>Bgl</i> II	AGATCT
<i>Dra</i> I	TTTAAA

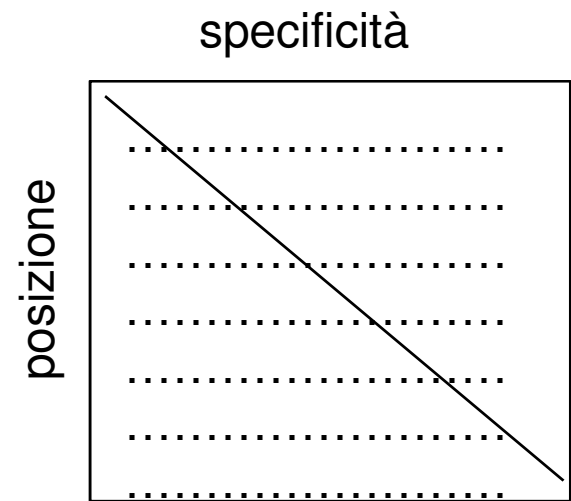
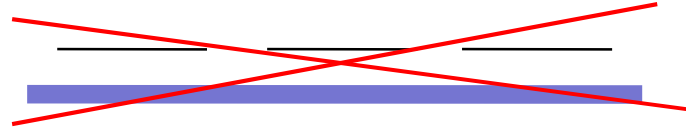
$$Rt(Pt) = \begin{cases} 0, & \text{se c'è un sito di restrizione in Bf o Br} \\ 1, & \text{altrimenti} \end{cases}$$

- Pattern match
- $[(Le-3) \leq |Pm| \leq Le]$

Lunghezza del sito di restrizione

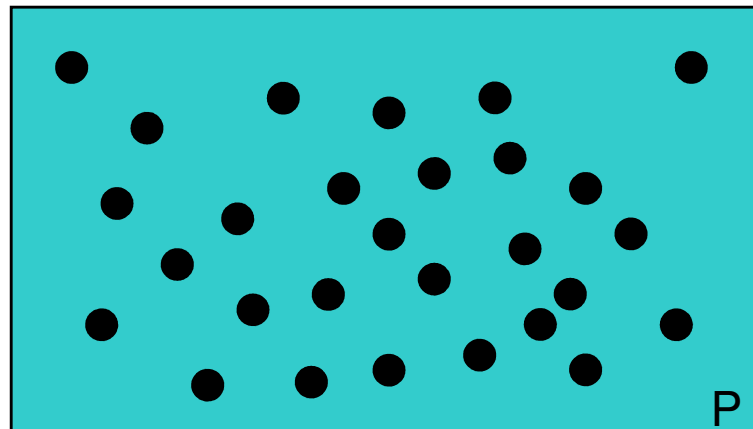
-Valutazione

Specificità



-Valutazione

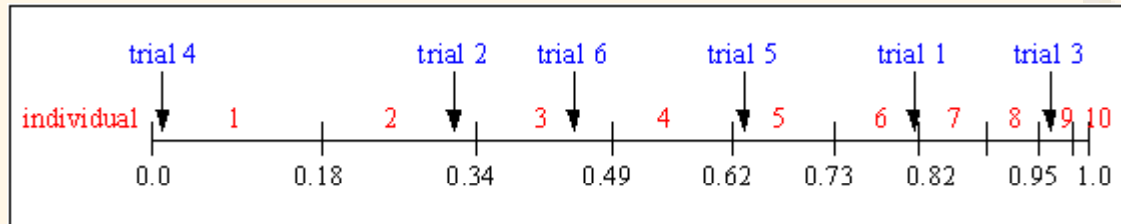
$$\text{Fitness}(Pt) = \text{leng}(Pt) + 3 * \text{lengd}(Pt) + \text{Tmd}(Pt) + 3 * \text{GCp}(Pt) + 3 * (\text{Term}(Bf) + \text{Term}(Br)) + 50 * \text{Uni}(Pt) + 10 * \text{Sc}(Pt) + 10 * \text{PC}(Pt) + \text{Rt}(Pt)$$



-Selezione

Roulette Wheel method

(campionamento stocastico con sostituzione)

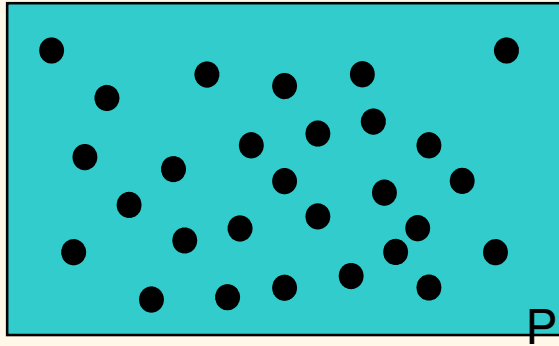


Dopo la selezione la “mating population” consiste degli individui:
1, 2, 3, 5, 6, 9.

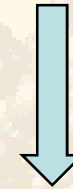
Number of individual	1	2	3	4	5	6	7	8	9	10	11
fitness value	2.0	1.8	1.6	1.4	1.2	1.0	0.8	0.6	0.4	0.2	0.0
selection probability	0.18	0.16	0.15	0.13	0.11	0.09	0.07	0.06	0.03	0.02	0.0

**Individui con un più alto peso (inverso del valore di fitness)
hanno quindi maggior probabilità di essere selezionati**

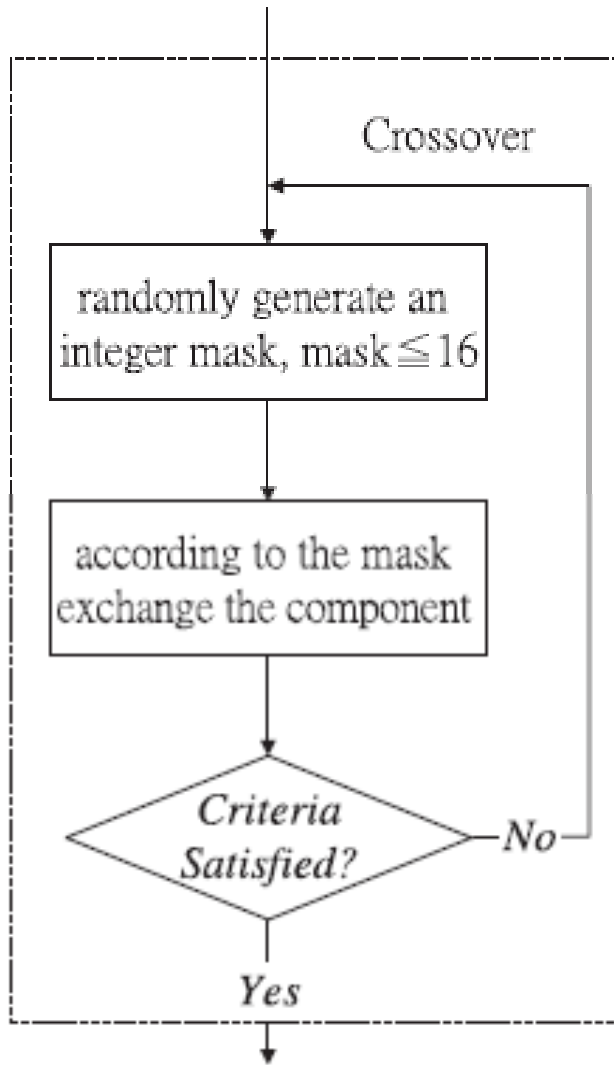
-Selezione



$$\text{Fitness}(Pt) = \text{leng}(Pt) + 3 * \text{lengd}(Pt) + \text{Tmd}(Pt) + 3 * \text{GCp}(Pt) + 3 * (\text{Term}(Bf) + \text{Term}(Br)) + 50 * \text{Uni}(Pt) + 10 * \text{Sc}(Pt) + 10 * \text{PC}(Pt) + \text{Rt}(Pt)$$



-Cross over



$R < 16$

Forma binaria di R per decidere quali componenti di X e Y dovranno essere cambiati

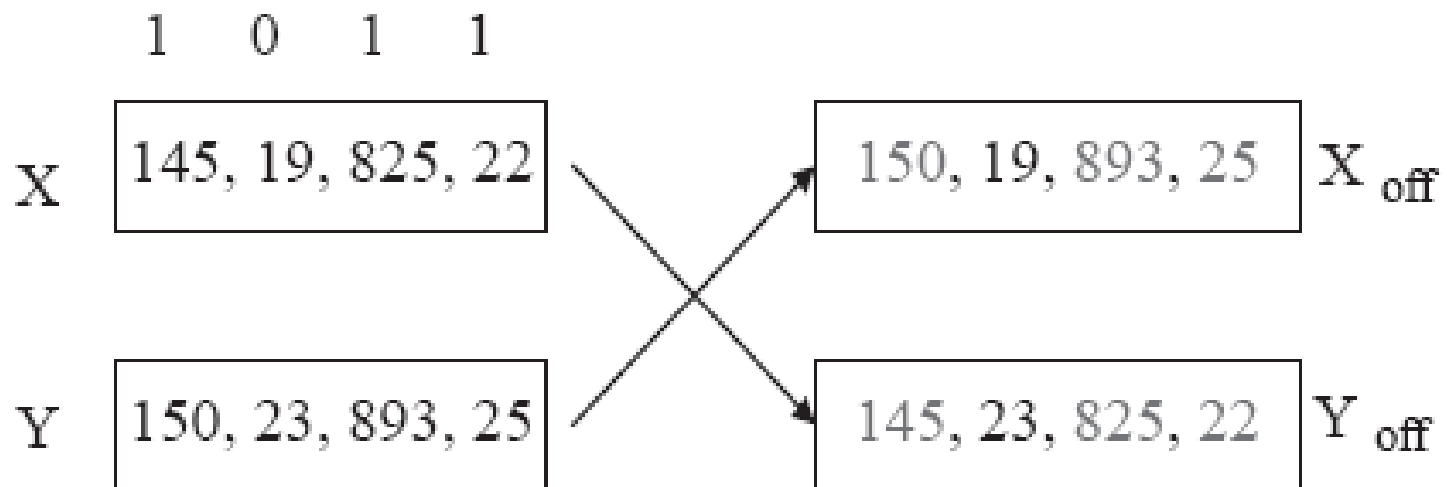
Es:

$R = 11$

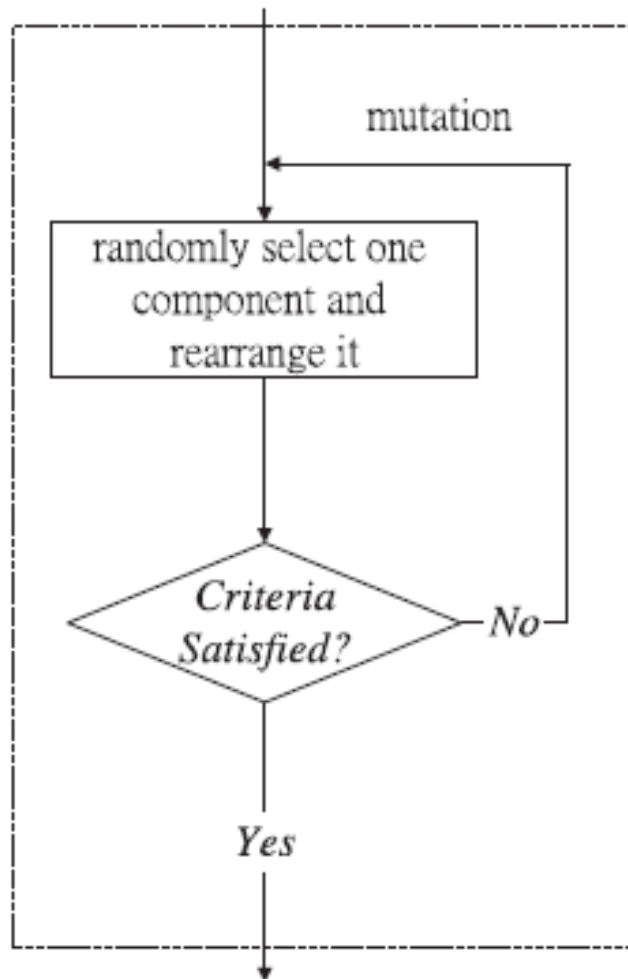
Forma binaria = 1011

Controllo delle violazioni delle restrizioni

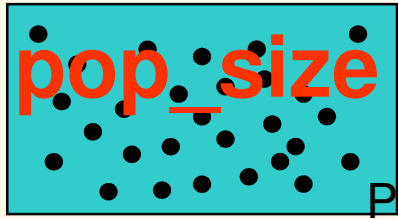
-Cross over e Mutazione



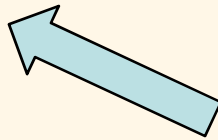
- Mutazione



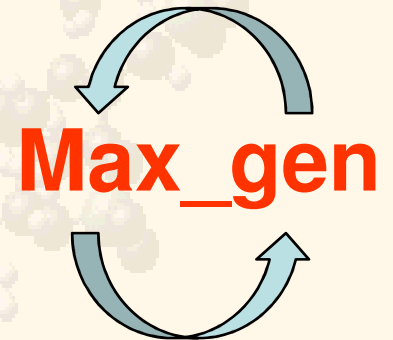
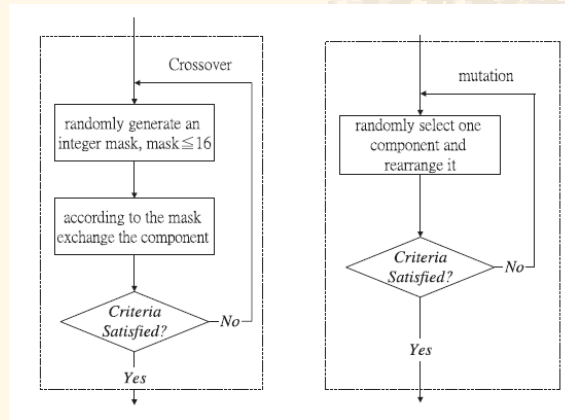
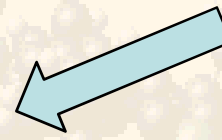
-parametri GA



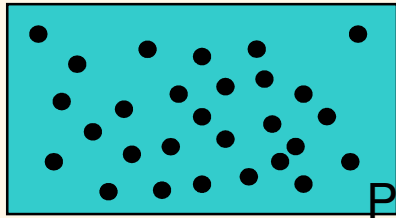
$$\text{Fitness}(Pt) = \text{leng}(Pt) + 3 * \text{lengd}(Pt) + \text{Tmd}(Pt) + 3 * \text{GCp}(Pt) + 3 * (\text{Term}(Bf) + \text{Term}(Br)) + 50 * \text{Uni}(Pt) + 10 * \text{Sc}(Pt) + 10 * \text{PC}(Pt) + \text{Rt}(Pt)$$



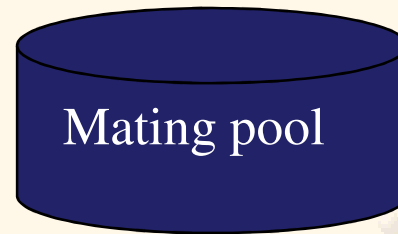
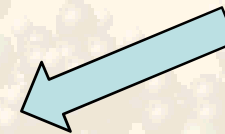
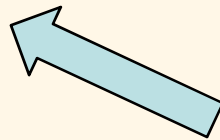
pop_size > M



-parametri GA

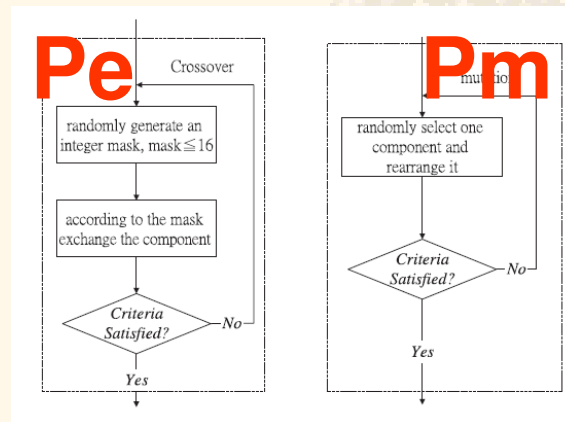


$$\text{Fitness}(Pt) = \text{leng}(Pt) + 3 * \text{lengd}(Pt) + \text{Tmd}(Pt) + 3 * \text{GCp}(Pt) + 3 * (\text{Term}(Bf) + \text{Term}(Br)) + 50 * \text{Uni}(Pt) + 10 * \text{Sc}(Pt) + 10 * \text{PC}(Pt) + \text{Rt}(Pt)$$



Se un numero random compreso tra 0 e 1 è minore di P_e , il C.O. viene eseguito.

$\uparrow P_e \rightarrow$ processo più veloce, soluzione meno precisa



Se un numero random compreso tra 0 e 1 è minore di P_m , la mutazione viene eseguita

$\uparrow P_m \rightarrow$ processo più lento, soluzione più precisa

-Test

1) *Pseudomonas mendocina* PHA synthase 1 (phaC1), PHA depolymerase (phaZ) and PHA synthase 2 (phaC2) genes

≠ lunghezza

2) Complete cds and *Homo sapiens* CDK2-associated protein 1 (CKD2AP1) coding DNAs (CDs) (Gen-Bank acc. no. NM_004642; 523..870).

Algoritmo proposto

Vs.

Prime3

GeneFisher

SGD

	The proposed algorithm	Primer 3	GeneFisher	SGD
Forward (F) primer (5'→3')	TTTCATCCTGGTAACTCTG	TGCCACTGCTGATCTTCAAC	CTCGAACTGAAGAACGTCA	AACCATTTCGTATTCCGCA
Reverse (R) primer (5'→3')	ATCCGTCTAGAGACTTTCAT	CTGGATTCTTCAGGCTCTGG	AGCCGATTTGTAGCAGGA	TTTGGGCATTCATGAAGG
Position (F)	1814–1832	1968–1987	231–249	1890–1907
Position (R)	2905–2924	3064–3083	1413–1430	2987–3005
Product size (bp)	1112	1116	1200	1116
Primer length (F/R) (mer)	19/20	20/20	19/18	18/18
Melting temperature (F/R) (°C)	54/56	60/62	56/54	51/51
Temperature difference (°C)	2	2	2	0
GC content (F) (%)	42	50	47	44
GC content (R) (%)	40	55	50	44
3'-Terminus (F)	G	C	A	A
3'-Terminus (R)	T	G	A	G
Self-complementarity	No	No	Yes	No
Pair-complementarity	No	No	No	No
Specificity	Yes	No	No	No
Restriction site	'GGTACC' (<i>Bam</i> HI)/ 'CCTAGG' (<i>Avr</i> II)	N/A	N/A	'AAATTT' (<i>Dra</i> I)/ 'GGGCC' (<i>Apa</i> I)

2° esperimento

	The proposed algorithm	Primer 3	GeneFisher	SGD
Product size (bp)	303	267	252	348
Primer length (mer)	18/18	18/18	16/17	21/18
GC content (forward) (%)	39	61	63	47
GC content (reverse) (%)	50	56	53	50
Melting temperature (F) (°C)	50	58	54	53
Melting temperature (R) (°C)	54	56	52	56
Temperature difference in pair (°C)	4	2	2	3
Specificity	Yes	No	No	No
3'-Terminus (F)	C	A	A	C
3'-Terminus (R)	T	T	A	T
Self-complementarity	No	No	Yes	Yes
Pair-complementarity	No	No	No	Yes

Siti di restrizione



Specificità

– Test bagnato

Algoritmo proposto



CDK2AP1CDs



forward: 5-ATGTCTTACAAACCGAAC-3; (1-18)

reverse: 5-CAGTCCTCTAGCGTGAAT-3; (285-303)

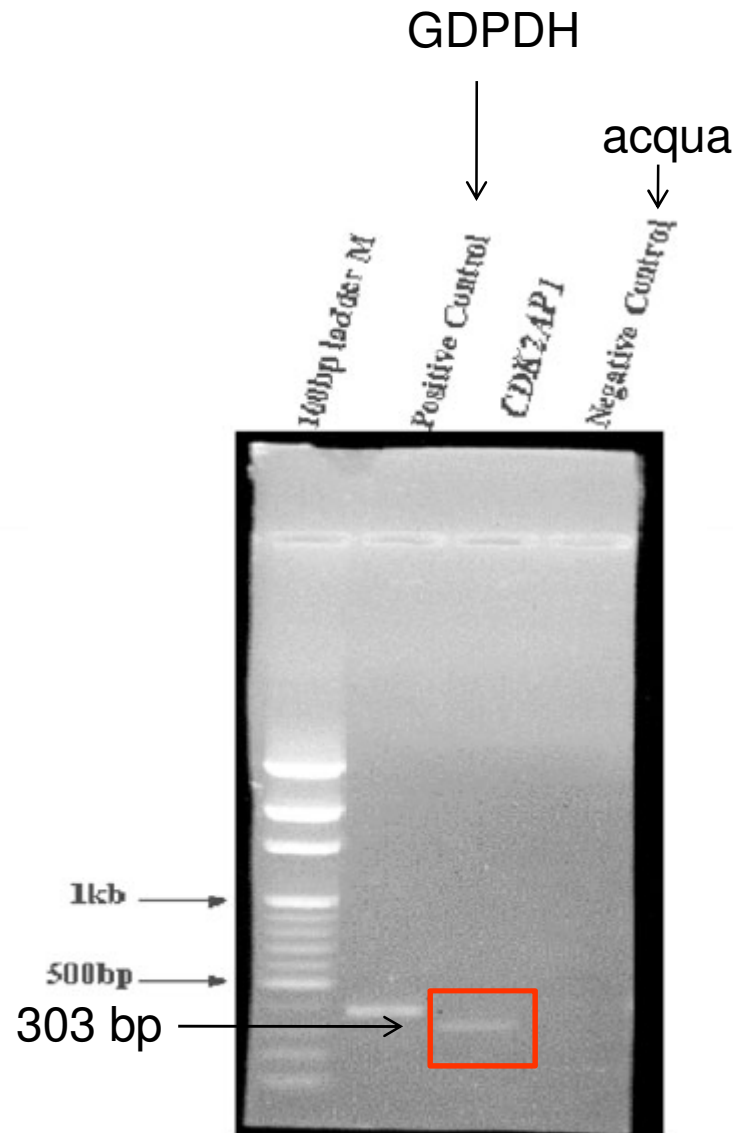


Condizioni sperimentali: 15µl reagente: 50 ng HeLa cell cDNA, 10× reaction buffer 1.5 µl, dNTP (1.25 mM) 1.8 µl, forward e reverse primers (10 mM) 0.5 µl each, *Taq polymerase* (5 U/µl) 0.05 µl e *deionizzata*, acqua distillata.

Programma PCR: 95°C per 5 min; 95°C per 45 s, 55°C per 45 s, 72°C per 45 s per 35 cicli 72°C per 10 min e 4°C per 10 min.

- **Test bagnato**

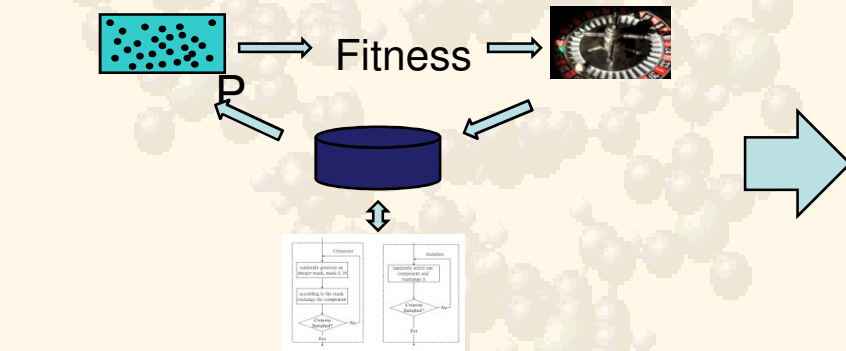
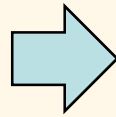
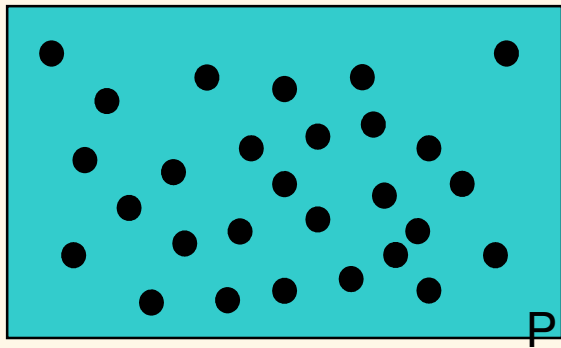
Elettroforesi su gel di agarosio 1% e 0,1% di etidio bromuro.



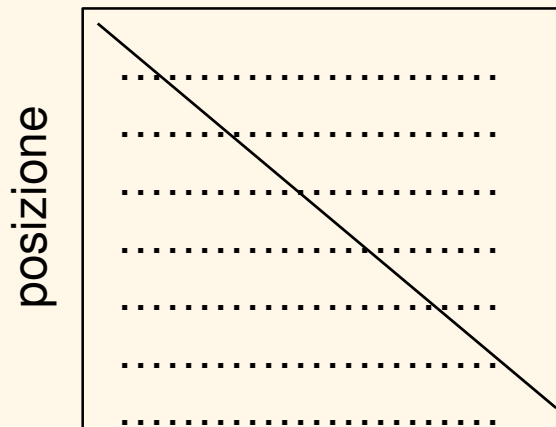
– **Conclusioni**

- Qualità → buona riuscita dell'esperimento
- Siti di restrizione
- Ricerca della soluzione migliore indipendentemente dalla lunghezza della sequenza
- Specificità ← per essere soddisfatta l'algoritmo può "allentare" alcune restrizioni come il 3' terminale e il contenuto in GC

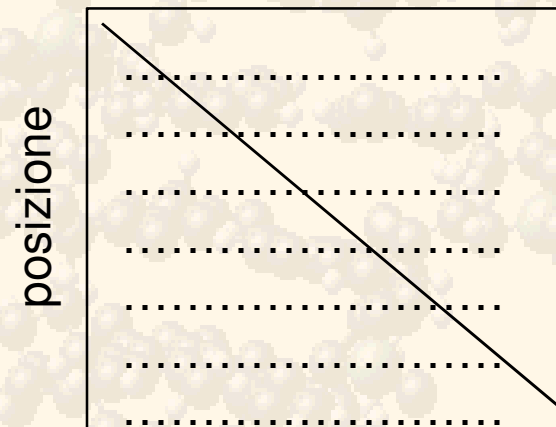
-Conclusioni



specificità



$|Pm|$ lunghezza pattern match

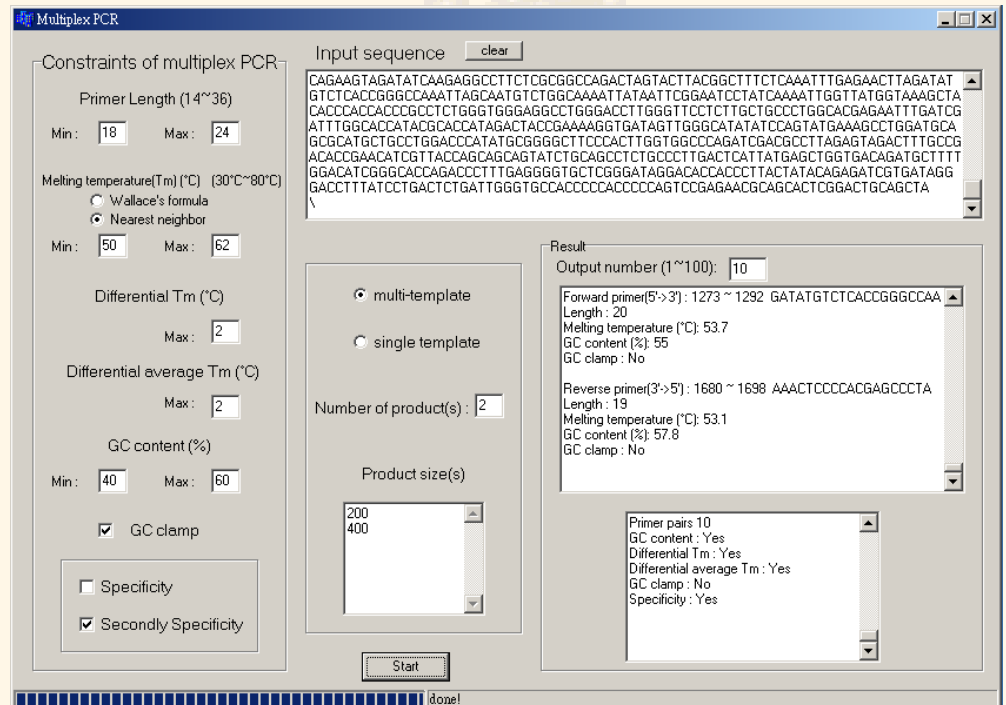


MultiPrimer: a Software for Multiple Primer Design for Gene Family

- Ricerca un set di coppie di primer per un set di sequenze di famiglie geniche o targets multipili per una singola sequenza di cDNA.
- Si può inserire una lunga sequenza (ex: take RBM15), su cui il programma ricerca diverse coppie di primer in differenti regioni con differenti lunghezze (per esperimenti di PCR in tubi differenti).
- Inserendo diverse sequenze (ex: take ANXA1, ANXA3, ANXA5), l'algoritmo ricerca una coppia di primer per ogni sequenza (per esperimenti di PCR in un unico tubo e non causa mis-priming).

Parametri di default:

- self annealing
- self-end annealing
- pair annealing and pair-end
- the population size = 1000
- the generation = 100
- the crossover rate = 80%
- mutation rate = started from 1%, and grows up to as high as 90% as generation increases (For each generation the mutation grows 1%).



Parameter Area

Multiplex PCR

Constraints of multiplex PCR

Primer Length (14~36)
Min: Max:

Melting temperature(Tm) (°C) (30°C~80°C)
 Wallace's formula
 Nearest neighbor
Min: Max:

Differential Tm (°C)
Max:

Differential average Tm (°C)
Max:

GC content (%)
Min: Max:

GC clamp

Specificity

Secondly Specificity

Input sequence

```
CAGAAGTAGATATCAAGAGGCCTTCTCGCGGCCAGACTAGTACTACGGCTTCTCAAATTTGAGAAGTACTAGATAT  
GTCTCACCAGGGCCAAATTAGCAATGTCTGGCAAATATAATTCGGAATCCTATCAAATTTGGTTATGGTAAAGCTA  
CACCACCCACCCGCTCTGGGTGGGACCCCTGGCACTCTGCTGCCCTGGCAGAGAAATTTGATCG  
ATTTGGCACCATACGCACCATAGACTATATATCCAGTATGAAAGCCTGGATGCA  
GCGCATGCTGCCTGGACCCATATGCGAGATCGACGCCTTAGAGTAGACTTTGCCG  
ACACCGAACATCGTTACCAGCAGCAGTATCTGCAGCCTCTGCCCTTGACTCATTATGAGCTGGTGCAGATGCTTTT  
GGACATCGGGCACCAGACCCTTTGAGGGGTGCTCGGGATAGGACACCACCTTACTATACAGAGATCGTGATAGG  
GACCTTTATCCTGACTCTGATTGGGTGCCACCCCAACCCCACTCCGAGAACGCAGCACTCGGACTGCAGCTA  
\
```

Input Area

multi-template
 single template

Number of product(s):

Product size(s)

Result

Output number (1~100):

Forward primer(5'→3'): 1273 ~ 1292 GATATGTCTCACCAGGGCCAA
Length: 20
Melting temperature (°C): 53.7
GC content (%): 55
GC clamp: No

Reverse primer(3'→5'): 1680 ~ 1698 AAACCTCCCCACGAGCCCTA
Length: 19
Melting temperature (°C): 53.1
GC content (%): 57.8
GC clamp: No

Output Area

Primer pairs 10
GC content: Yes
Differential Tm: Yes
Differential average Tm: Yes
GC clamp: No
Specificity: Yes

done!

