

Espressione eterologa......
Come e Perché esprimere la nostra proteina d'interesse in un organismo ospite

# Espressione Genica in Sistemi Eterologhi

- 1. Perché esprimere geni in sistemi eterologhi??
  - 2. Quali sistemi eterologhi utilizzare???
- 2. I sistemi eterologhi utilizzati sono *Escherichia coli*, *Saccharomyces cerevisiae*, cellule di insetto, cellule vegetali e cellule di mammifero in coltura. L'espressione in *E.coli* è di gran lunga la più semplice e, forse per questo, la più utilizzata come prototipo di espressione genica in sistemi eterologhi

# Espressione in sistemi "eterologhi"

### Espressione transiente:

#### in cellule, per studio di attività, localizzazione sub-cellulare

il vettore resta nella cellula come frammento extracromosomico, senza integrarsi nel genoma cellulare; il gene verrà espresso per breve tempo, di solito non più di 72 ore. E' un processo più rapido, facile ed economico ma ovviamente permette studi limitati nel tempo.

# Espressione stabile:

# In organismi, per purificazione, studi in vitro, produzione anticorpi

il DNA esogeno si integra stabilmente nel genoma: il gene sarà espresso «sempre» e sarà trasmesso anche alla progenie. La produzione di una linea/organismo stabilmente trasformata richiede un tempo più lungo dell'espressione transiente

Expression system	Most common application	Advantages	Challenges
Mammalian	Functional assays     Structural analysis     Antibody production     Expression of complex proteins     Protein interactions     Virus production	Highest-level protein processing     Can produce proteins either transiently, or by stable expression     Robust optimized transient systems for rapid, ultrahigh-yield protein production	Gram-per-liter yields only possible in suspension cultures     More demanding culture conditions
Insect	Functional assays     Structural analysis     Expression of intracellular proteins     Expression of protein complexes     Virus production	Similar to mammalian protein processing     Can be used in static or suspension culture	More demanding culture conditions than prokaryotic systems     Production of recombinant baculovirus vectors is time consuming
Yeast	Structural analysis     Antibody generation     Functional analysis     Protein interactions	Eukaryotic protein processing     Scalable up to fermentation (grams per liter)     Simple media requirements	Fermentation required for very high yields     Growth conditions may require optimization
Bacterial	Structural analysis     Antibody generation     Functional assays     Protein interactions	Scalable     Low cost     Simple culture conditions	Protein solubility     May require protein- specific optimization     May be difficult to express some mammalian proteins
Algal	Studying photosynthesis, plant biology, lipid metabolism     Genetic engineering     Biofuel production	Genetic modification and expression systems for photosynthetic microalgae     Superb experimental control for biofuel, nutraceuticals, and specialty chemical production     Optimized system for robust selection and expression	Nascent technology     Less developed compared to other host platforms
Cell-free	Toxic proteins Incorporation of unnatural label or amino acids Functional assays Protein interactions Translational inhibitor screening	Open system; able to add unnatural components     Fast expression     Simple format	Scaling above multimilligram quantities may not be costly

# Pichia pastoris e Kluyveromyces lactis

- ✓ P. pastoris lievito metilotrofico che utilizza metanolo come fonte di C
- ✓ K. lactis è un lievito aerobico che utilizza lattosio come fonte di C
- ✓ Organismi considerati GRAS (Generally Regarded as Safe)
- ✓ Crescita ad alta densità in tempi brevi
- ✓ Sistemi di esporto delle proteine eterologhe più efficienti
- ✓ Glicosilazione proteine simile a quella di mammifero
- ✓ Possibilità di scale up (fermentatori)
- ✓ Disponibilità di promotori inducibili e/o costitutivi
- ✓ Limitata disponibilità di ceppi e vettori





HIGH LOW

Per esprimere un gene et rologo bisogna definire due componenti: SPEED



BEVS/INSECT CELL

YEAST

BACTERIA

-un ospite per l'espressione

COST





BACTERIA

YEAST

BEVS/INSECT CELL

MAMMALIAN

TYPICAL YIELD



MAMMALIAN



BEVS/INSECT CELL



BACTERIA



YEAST

POST -TRANSLATION MODIFICATION



BACTERIA



YEAST



BEVS/INSECT CELL



MAMMALIAN

**FDA APPROVAL** 



BEVS/INSECT CELL



YEAST



BACTERIA



MAMMALIAN

#### Vantaggi

- Vasta scelta di vettori di clonaggio
- Vasta scelta di ceppi batterici
- Controllo relativamente semplice dell'espressione del gene
- Ottima resa nella produzione di proteine ricombinanti (25% del totale)
- La proteina ricombinante può essere espressa come proteina di fusione Produzione di proteine eterologhe in *E.coli* La proteina ricombinante può essere secreta nel terreno di crescita
- Basso costo di produzione
- Processi di fermentazione ben conosciuti
- Facile possibilità di "scale-up"

- La proteina ricombinante può mancare di modificazioni post-traduzionali (es. glicosilazioni)
- Batteri Gram-negativi contengono endotossine
- Presenza di proteasi
- La sovra-produzione di proteine può portare alla formazione di "corpi d'inclusione", aggregati insolubili che rendono difficoltosa la purificazione della proteina e possono ridurre la sua attività biologica

### **Promotore**

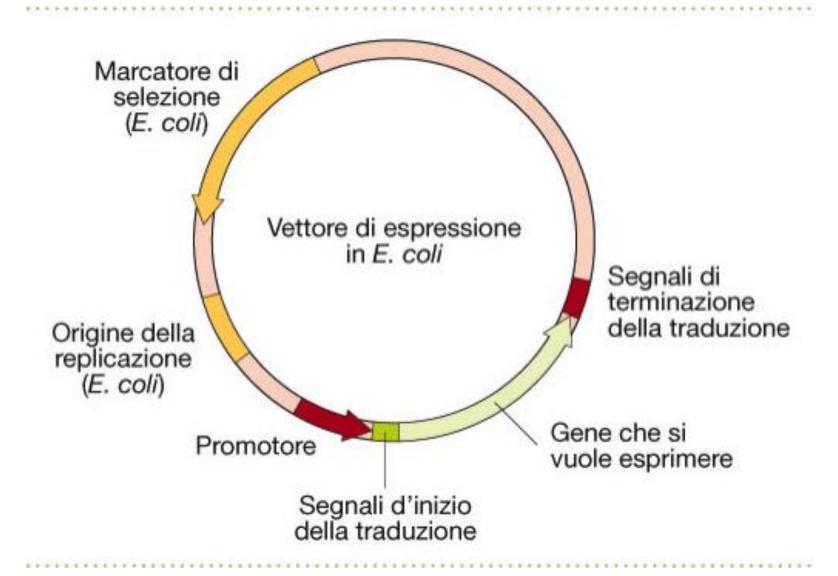
Il livello di espressione di un gene dipende in larga misura dalla forza del promotore che lo controlla determinando la frequenza con la quale l'RNA polimerasi inizia la trascrizione. Sono stati quindi isolati ed ottimizzati un certo numero di promotori forti di E.coli che sono presenti nella maggior parte dei vettori d'espressione attuali. In più, sono disponibili promotori in parte o totalmente sintetici sulla base delle sequenze consensus ottimali.

# Stabilita'

La resa di un prodotto di espressione dipende anche dalla stabilità della proteina. La stabilità delle proteine dipende dalla presenza di aminoacidi stabilizzanti all'estremità N-terminale e di aminoacidi destabilizzanti all'estremità C-terminale. Modificando la sequenza codificante una proteina, possiamo alterarne la composizione aminoacidica ed aumentarne la stabilità.

#### Uso di ceppi difettivi in proteasi

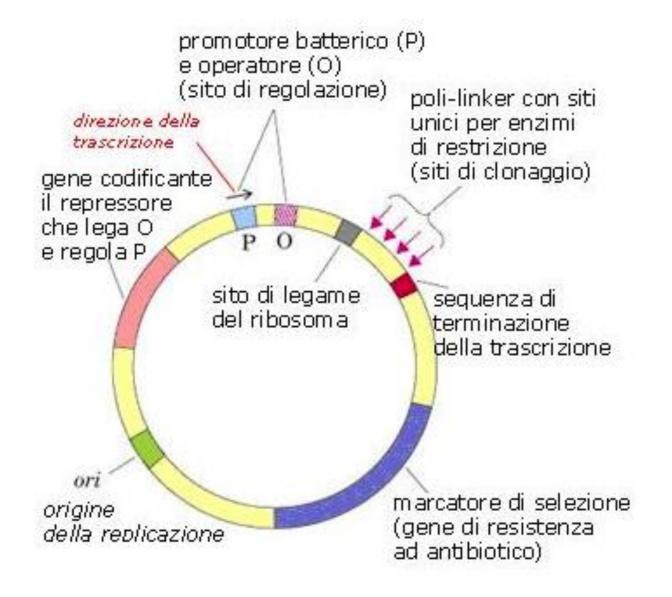
Un modo per ottimizzare la stabilità dei prodotti d'espressione, consiste nel minimizzare la degradazione proteolitica a carico delle proteine espresse



# Perchè i vettori d'espressione sono regolati

• Una proteina eterologa tende ad essere identificata come "estranea" e degradata dalla cellula che attiva specifiche proteasi. La possibilità di indurre/regolare l'espressione della proteina permette di ridurre la degradazione proteolitica aumentando la resa.

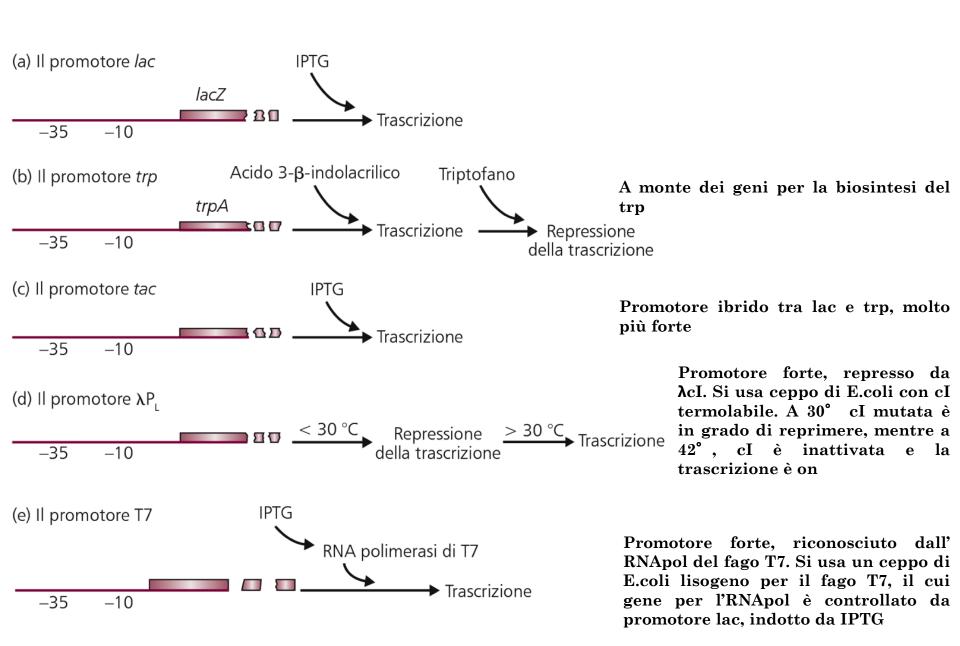
• Alcune proteine possono essere tossiche o, comunque, interferire con la crescita dell'ospite di espressione. In alcuni casi fino al 25-50% delle proteine totali sono costituite dalla proteina ricombinante, a discapito delle proteine che assicurano il normale metabolismo di E.coli. La possibilità di limitare l'espressione della proteina alla sola fase di induzione permette il normale sviluppo della cellula.



#### Assenza di lattosio RNA polimerasi DNA Promotore Z y a Repressore attivo I monomeri del repressore formano un tetramero che si lega all'operatore e blocca la trascrizione. Presenza di lattosio DNA Promotore 0 Z V a Lattosio (induttore) Il repressore inattivo non è in grado di legare il promotore Il lattosio lega e la trascrizione può avvenire. e inattiva il repressore. DNA Promotore У 0 Z a

Per indurre l'espressione di geni eterologhi clonati in vettori d'espressione regolati da questo sistema si utilizza l'induttore gratuito IPTG (isopropil-beta-tio-D-galattopiranoside)

# Principali sistemi di regolazione



I sistemi di regolazione procariotici sono sempre "leaky". Una repressione assoluta, infatti, impedirebbe al sistema di essere ri-attivato. La completa repressione dell'operone del lattosio per esempio, reprimendo completamente anche l'espressione della lattosio permeasi, impedirebbe l'assunzione di lattosio eventualmente presente nel mezzo di coltura

Analogamente, il principale problema di molti vettori d'espressione è quello di essere troppo "leaky" e di esprimere un pò (troppa) proteina anche in condizioni di non induzione

Sono stati quindi messi a punto dei sistemi di regolazione "stringenti" che minimizzano il livello di espressione basale di geni repressi.

- Aumentare il dosaggio genico dei geni di regolazione (lacI)
- Creare delle regolazioni a cascata
- Promotore dell'operone dell'arabinosio che è quantitativamente represso da quantità crescenti di arabinosio

Il promotore araBAD è regolato sia positivamente che negativamente dal prodotto del gene araC, un regolatore trascrizionale che forma un complesso con l'L-arabinosio.

#### Off:

In assenza di arabinosio, il dimero AraC contatta i siti  $O_2$  e  $I_1$  dell' operone araBAD, formando un loop sul DNA.

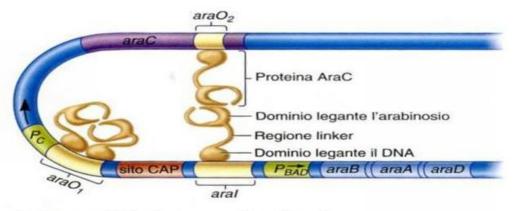
#### On:

L-arabinosio lega AraC, e cio' rilascia il loop e permette la trascrizione.

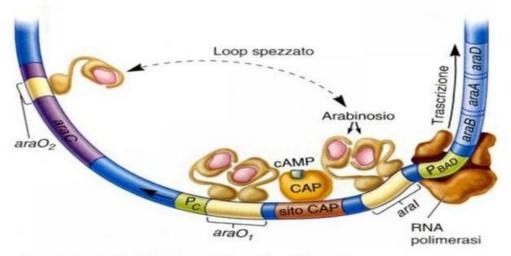
Il complesso cAMP Activator Protein (CAP)-cAMP lega il DNA e stimola il legame di AraC ad  $I_1$  e  $O_1$ .

Il livello d'espressione basale puo' quindi essere represso aggiungendo glucosio che abbassa i livelli di cAMP che a sua volta riduce il legame di CAP. Di conseguenza l'attivazione trascrizionale è ridotta.

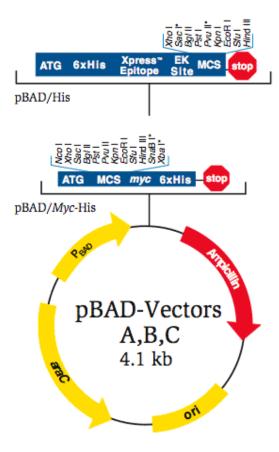
# Regolazione da Arabinosio pBAD



(a) Operone inibito in assenza di arabinosio

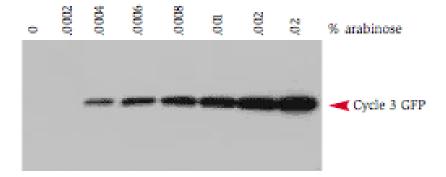


(b) Operone attivato in presenza di arabinosio

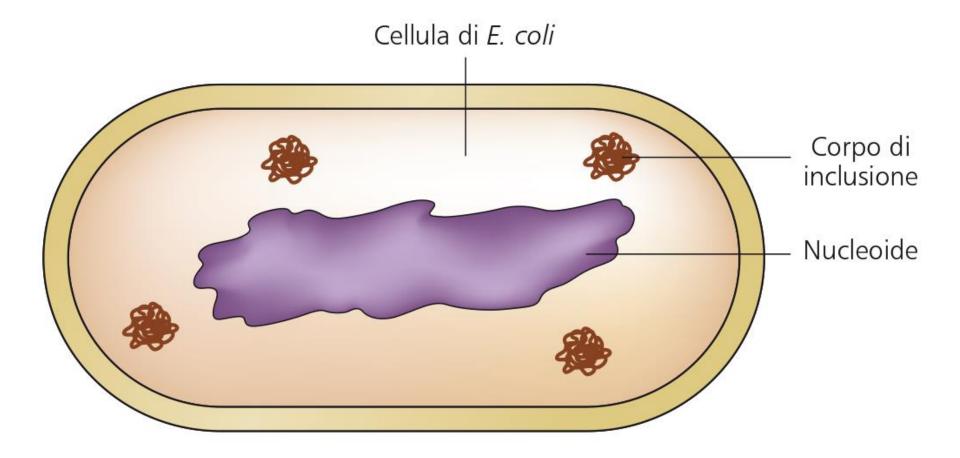


\* Frame-dependent variations

Figure 2 - Western blot of arabinose induction



# Corpi di inclusione



la proteina nativa deve essere recuperata mediante denaturazione *in vitro* e refolding

# Corpi di inclusione

Se il folding della proteina non è corretto

Se la proteina necessita di ponti disolfuro per il folding

....la proteina potrebbe essere insolubile e formare corpi di inclusione

In questo caso....

- la proteina ricombinante e' fino al 50% delle proteine totali
- e' protetta dalla degradazione
- non può essere tossica per la cellula

#### Ma non è facilmente disponibile

- isolamento dei corpi inclusi
- denaturazione in presenza di agenti denaturanti come guanidina o urea o condizioni riducenti (DTT)
- refolding della proteina mediante lenta rimozione del denaturante con dialisi,

# Esporto della Proteina

#### Vantaggi

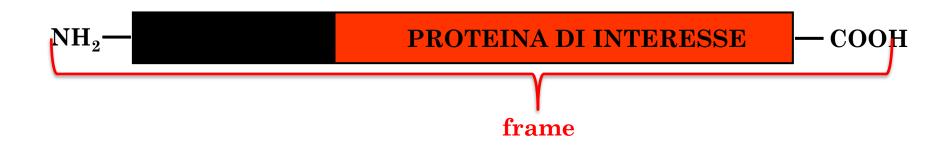
- La stabilità della proteina può aumentare (meno proteasi)
- Permette l'accumulo di proteine tossiche
- La proteina può essere isolata più facilmente

#### Svantaggi

· Resa ridotta

#### Come?

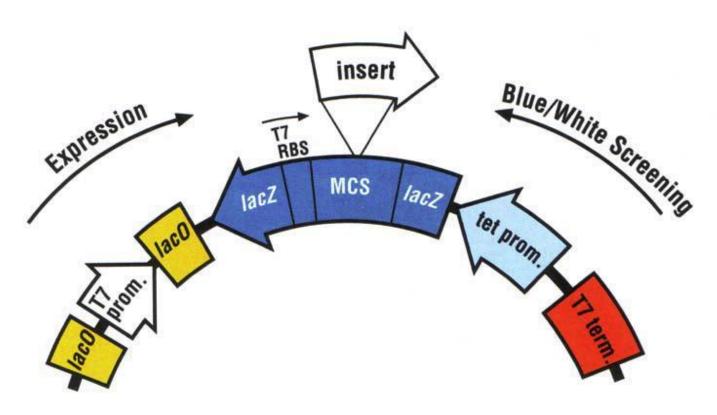
• Peptide segnale, aggiunto all'estremità NH<sub>2</sub> della proteina



# Espressione di proteine native

Un vettore d'espressione può essere usato per produrre una proteina nativa o una proteina chimerica (fusa ad un tag)

La produzione di una proteina nativa può essere necessario per produzione di anticorpo, o studi funzionali. Potrà essere purificata solo dall'SDS-PAGE

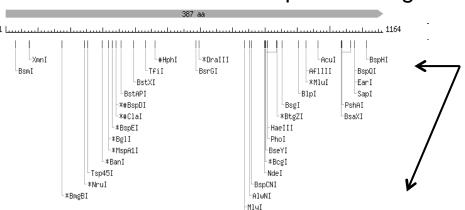


#### Sito di clonaggio multiplo del vettore di espressione

Bamhi Ecori Saci Sali Hindili Noti Xhoi
ATCGGAATTAATTC**GGATCCGAATTCGAGCTC**CGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCA

#### Sequenza del gene

Analisi di restrizione della sequenza del gene



PleI

>qi|49175990:3153377-3154540 Escherichia coli str. K-12 substr. MG1655

CACGCGCAAACACTGGCTATCGTCCTGCACTGTGGAATGAAAAACGCGATACCAAGCGCGCTAAGC
TGCTGCAATATGCTGAACGCGTCTGGAACATCACTGAAGGTTCCGATGATGAGCGTATTGACGCCGCGAT
TGCCGCAACCCGCAATTTCTTTGAGCAATTAGGCGTGCCGACCCACCTCTCCGACTACGGTCTGGACGGC
AGCTCCATCCCGGCTTTGCTGAAAAAACTGGAAGAGCACGGCATGACCCAACTGGGCGAAAATCATGACA

Enzimi che NON tagliano la sequenza del gene

chromosome, complete genome

AatII Acc65I AccI AclI AfeI AflII AhdI AleI ApaI ApaLI AscI AseI AsiSI AvaI AvrII BaeGI BaeI BamHI BanII BbsI BbvCI BciVI BclI BfaI BfuAI BglII BmtI BpmI Bpu10I BpuEI BsaAI BsaBI BsaI BseRI BsiEI BsiWI BsmAI BsmBI BsmFI BsoBI BspMI BsrBI BsrDI BssSI BstBI BstEII BstNI BstYI BstZ17I Bsu36I BtsI CspCI DraI DrdI EaeI EagI EciI Eco53kI EcoNI Eco0109I EcoRI EcoRV FseI FspI HincII HindIII HpaI Hpy99I KasI KpnI MfeI MscI NaeI NarI NcoI NgoMIV NheI NmeAIII NotI NsiI NspI PacI PaeR7I PciI PflFI PflMI PmeI PmII PpuMI PsiI PspGI PspOMI PspXI PstI PvuI PvuII RsrII SacI SacII SalI SbfI ScaI SexAI SfcI SfiI SfoI SgrAI SmaI SmlI SnaBI SpeI SphI SspI StuI StyI SwaI TliI TspMI Tthl111 XbaI XcmI XhoI XmaI ZraI



Sequenza degli oligonucleotidi per il clonaggio del gene per PCR

Forward 5'-CAGGGATCCATGAACAACTTTAATCTGCAC-3'
Reverse 5'-CAGAAGCTTTTAGCGGGCGCTTCGTATATAC-3'

# TAG per affinità

Uno delle caratteristiche più importanti dei vettori d'espressione consiste nella presenza dei TAG che permettono di purificare con facilità la proteina

- Un TAG codifica per un peptide caratterizzato da un'alta affinità a qualche ligando e che può essere revertita in condizioni fisiologiche
- Il TAG può essere aggiunto all' N-ter o al C-ter della proteina (folding della proteina)

#### **TAG**

Calmodulin binding peptide (CBP)

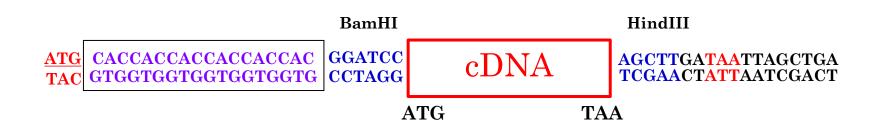
- ·6xHis
- Proteina A (IgG binding domain)
- •Glutatione S-transferasi (GST)
- Strep tag (Streptavidin binding tag)
  - ·Flag tag
  - ·Myc tag
  - ·HemAgglutinin (HA) tag

# Espressione di proteine di fusione

Spesso si preferisce esprimere proteine di fusione, per la loro maggiore stabilità, gli alti livelli di espressione e la relativa facilità con cui si purificano. Con queste proteine chimeriche si pone

il problema dello schema di lettura (frame)

		BamHI	HindIII
AGAACC <mark>ATG</mark>	CACCACCACCACCAC	GGATCCGTCGA	AAGCTTGATAATTAGCTGA
TCTT GGTAC	GTGGTGGTGGTGGTG	<b>CCTAGGCAGC</b>	<b>TTCGAACTATT</b> AATCGACT



### Tag ad istidina

Esistono numerosi vettori commerciali che includono il Tag ad istidina per esprimere fusioni N-teminali o C-terminali. Di solito nel caso delle fusioni N-terminali è presente un sito di riconoscimento per una proteasi interposto tra la sequenza che codifica il Tag ed il gene

PIS lac O-lac O-RBS ATG-6xHis-FXa Recognition Site MCS Stop Codons

Purificazione della proteina chimerica su colonna di Ni-NTA

$$C = O$$

$$C + CH_2$$

$$NH$$

$$O = C$$

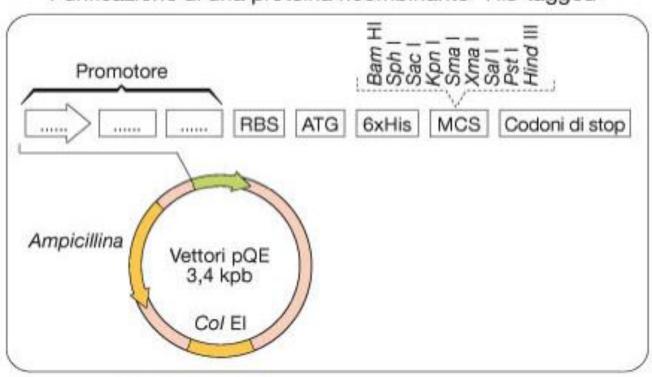
$$CH + CH_2$$

$$NH$$

$$CH_2$$

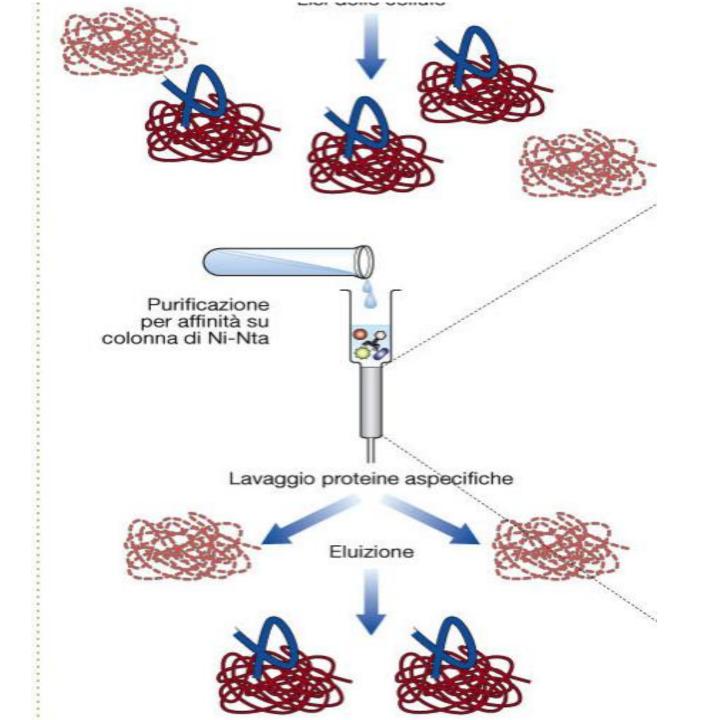
Figure 5. Interaction between neighboring residues in the 6xHis tag and Ni-NTA matrix.

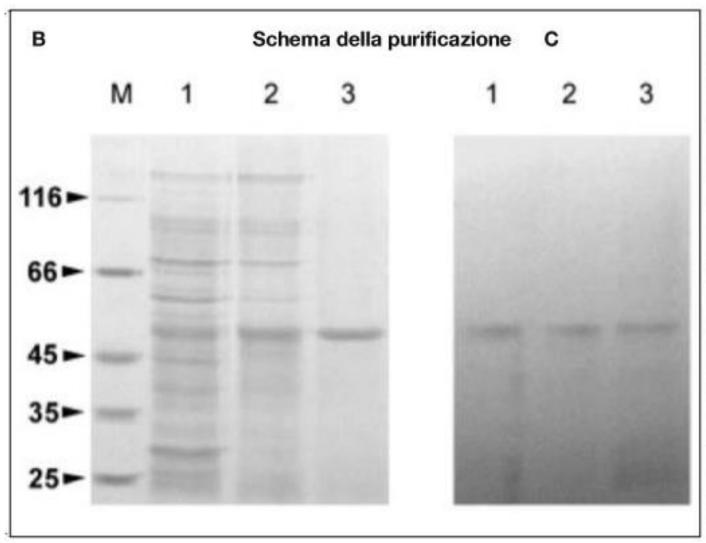
# Purificazione di una proteina ricombinante "His-tagged"



Raccolta delle cellule di E. coli dopo induzione del gene ricombinante

Lisi delle cellule

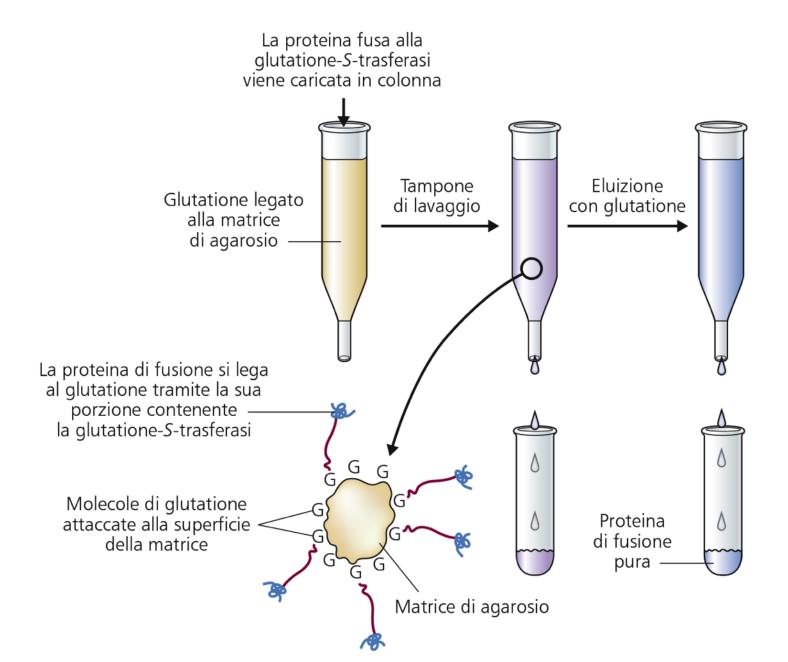




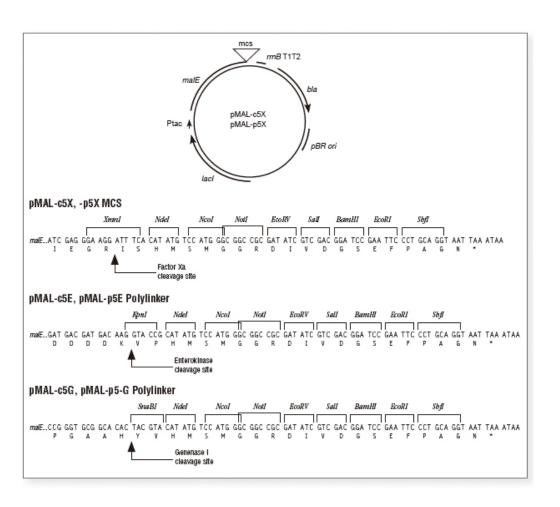
Tot Wash Eluiz.

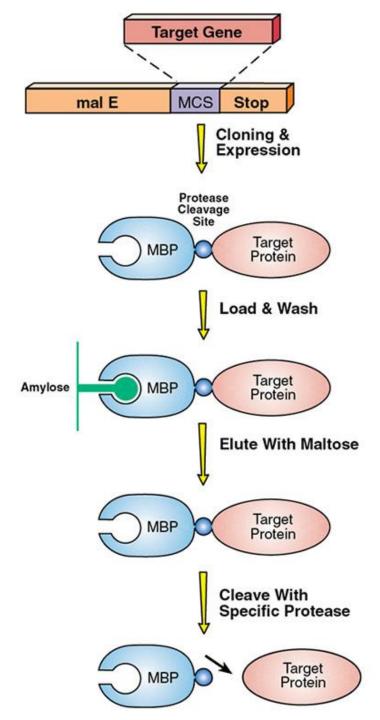
Western Ab anti His

# Tag GST



# Produzione di proteine con tag MBP



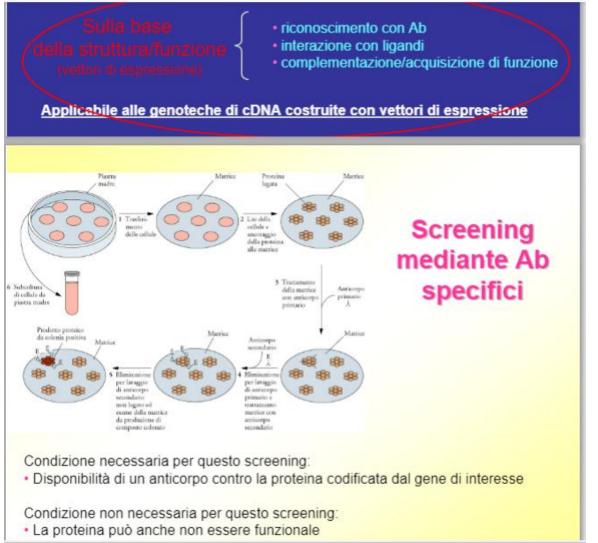


# Screening immunologico di una libreria di espressione

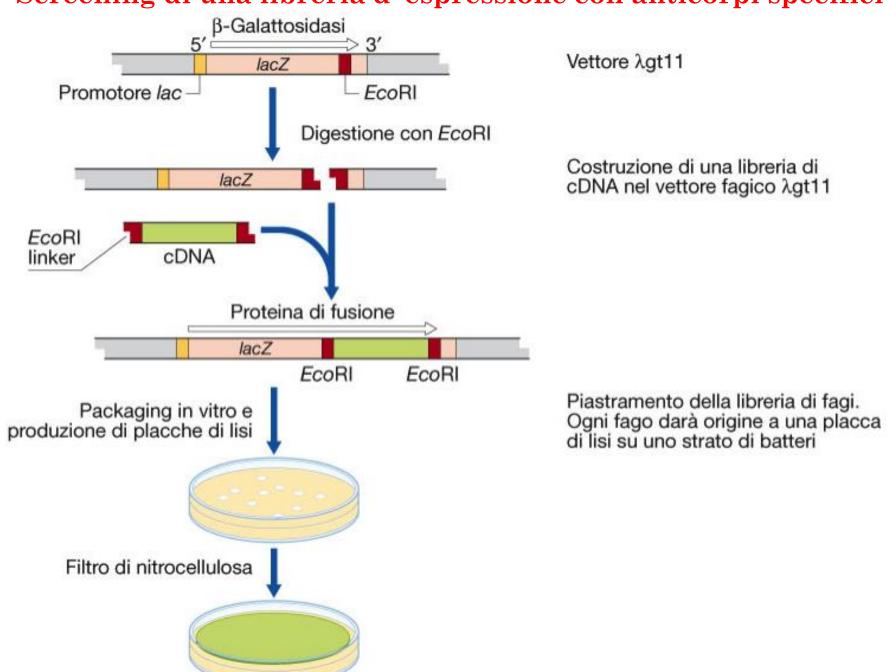
Una libreria di espressione risponde a tutti i requisiti di una libreria genica con un'unica differenza: è costruita in modo da avere promotori forti funzionali in

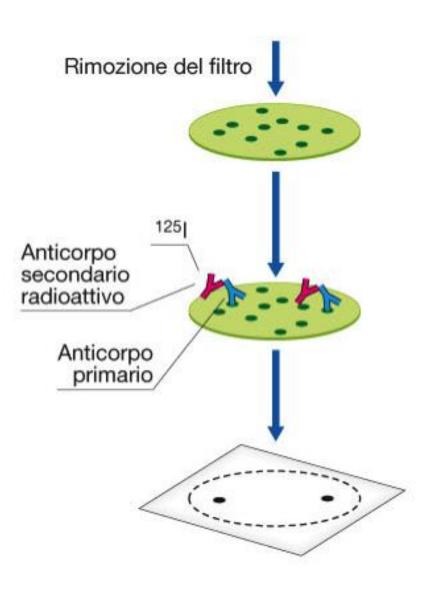
E.coli (es. il promotore lac) capaci di indurre l'espressione in E.coli dei cDNA

clonati.



# Screening di una libreria d'espressione con anticorpi specifici





Le proteine ricombinanti prodotte dai fagi ricombinanti rimangono attaccate al filtro di nitrocellulosa

Il filtro è inoculato con un anticorpo primario prodotto contro la proteina codificata del gene d'interesse. Il complesso antigene-anticorpo formatosi in corrispondenza di specifiche placche di lisi è a sua volta riconosciuto da un anticorpo radioattivo secondario 1251 che si lega all'anticorpo primario

I complessi antigene-anticorpo radioattivi danno origine a macchie scure se il filtro è esposto su una lastra per raggi X. I segnali ottenuti permettono di identificare le placche originarie che producono la proteina di fusione d'interesse

# E se non riusciamo ad esprimere la nostra proteina......

# CELL-FREE EXPRESSION SYSTEMS

In vitro transcription + in vitro translation

Lisato reticolociti coniglio

Lisato Germe Grano (Wheat Germ)

Lisato E. coli

+ ATP

**tRNAs** 

amino acidi (marcati),

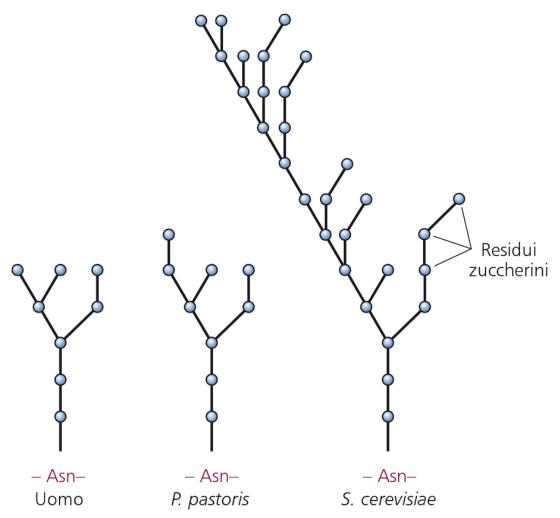
**Linearized Vector** 

# Step I: Prepare of Reaction Mixture 4 Template DNA E. coli extract Master mix T7 RNA polymerase, Amino acids, rNTPs, Ribosome, tRNAs, Energy sources, salts enzymes etc. 30°C incubation ₽ Step II: Protein Expression (in reaction tube) RNA Template DNA Transcription Polymerase Ribosome 🔩 Amino Acids **Translation Translational** Factors **Energy Source**

Spesso le proteine eucariotiche sono sottoposte a modificazioni posttraduzionali (glicosilazione), necessarie per la funzionalità.

La glicosilazione è estremamente rara nei batteri, quindi....

# Espressione in eucarioti



# Espressione in Saccharomyces cerevisiae

#### Pro

- · Produzione grandi quantità di proteina
- Disponibilità di diversi promotori inducibili (GAL, PHO5, CUP1)
- Organismo "sicuro" anche per la produzione di proteine utilizzate per principi farmaceutici

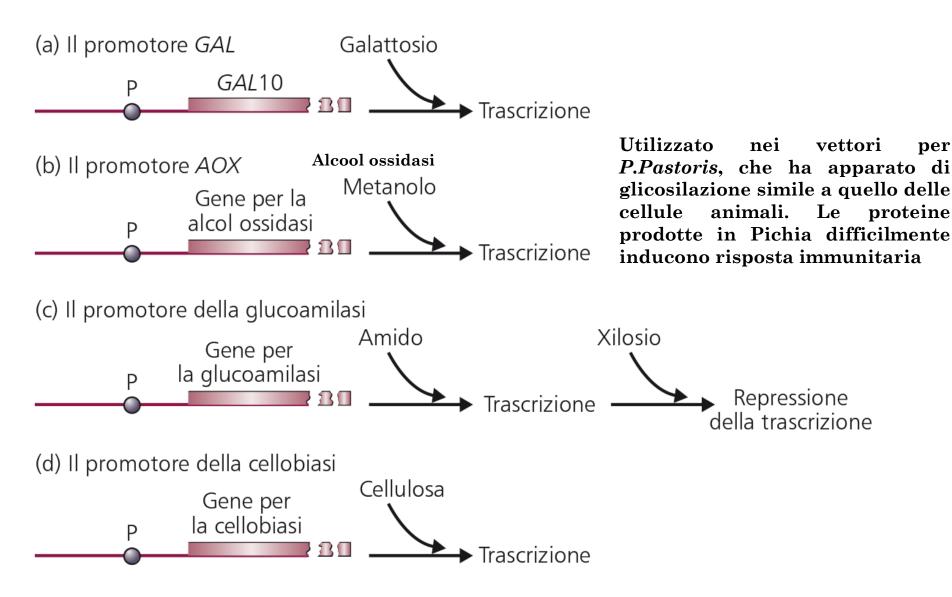
#### Contro

- Iperglicosilazione
- Privo di sistema efficiente di secrezione
- Bias nell'utilizzo dei codoni

Fattore prepro-α (peptide segnale)

cDNA clonato

Seq. di riconoscimento per una endoproteasi di lievito



Aspergillus nidulans (promotore della glucoamilasi) e *Trichoderma reseei* (promotore della cellobiasi) hanno buona attività di glicosilazione e capacità di secernere le proteine.

# Controllo dell' espressione di proteine in eucarioti Sistema Tet-on/Tet-off

Sistema che deriva da Operone Tet di E.Coli
(in presenza di tetraciclina, si attivano geni per la resistenza alla Tc)
Sistema a 3 elementi: Tet-Repressor: Repressore, Tc/Dox: Tetraciclina,
TRE: Elemento cis di risposta alla Tet (dominio tetO)

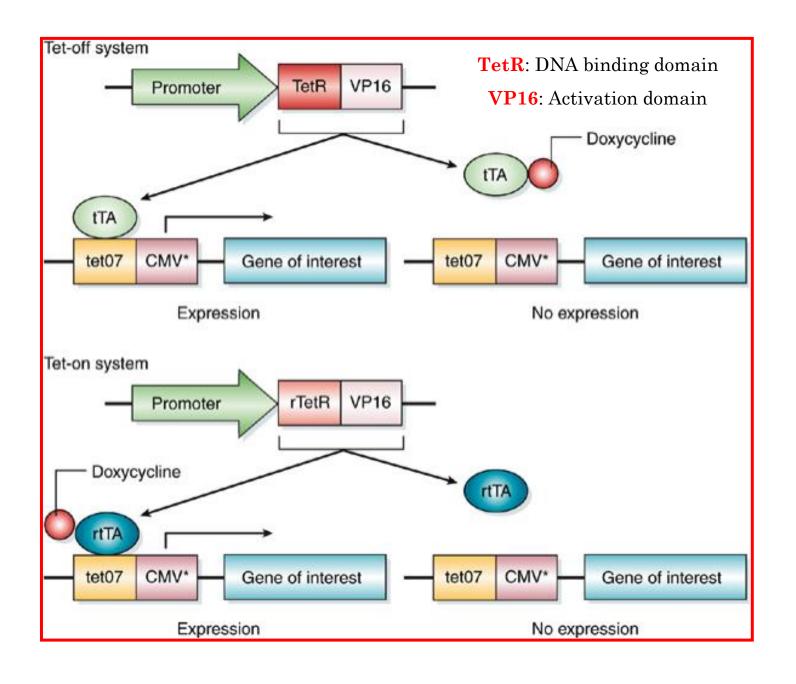
#### Tetracycline/Doxycycline-regulated gene expression

#### Tet-off: la tetraciclina spegne

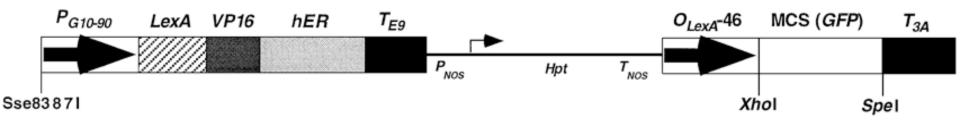
The chimeric tetracycline transactivator protein (tTA) consists of the TetR domain that recognizes tetO, fused to the VP16 transactivating domain of herpes simplex virus. tTA binds to seven tetO domains (tetO7) connected to a minimally active CMV promoter (CMV\*), resulting in gene activation. Dox binds tTA and prevents gene activation.

#### Tet-on: la tetraciclina accende

TetR is mutated to rTetR, which must have doxycycline present in order to bind to tetO7 and activate gene expression.



# Controllo dell' espressione di proteine in eucarioti Sistema a 2 componenti dell' human Estrogen Receptor (hER)



Il sistema XVE consiste di tre unità trascrizionali:

- 1. promotore sintetico costitutivo PG10-90 che controlla l'espressione della proteina chimerica (LexA BD-VP16-hER)
- 2. promotore costitutivo::Hpt (Hygromicina resistenza)
- 3. otto copie dell'operatore di LexA (OLexA) e un MCS per il gene target

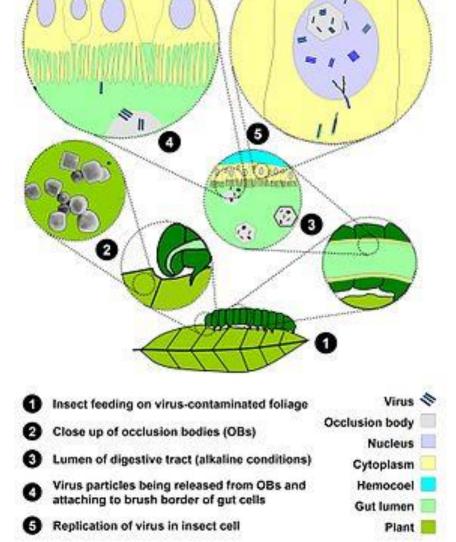
La proteina chimerica è prodotta in modo costitutivo, ma rimane sequestrata nel citoplasma in assenza dell'induttore.

In presenza di \( \beta\)-estradiolo migra nel nucleo e attiva la trascrizione del gene d'interesse, in modo proporzionale alla dose di \( \beta\)-estradiolo

Clonando il gene d'interesse nel MCS è possibile attivarne la trascrizione e controllarne il livello variando la concentrazione di \( \mathcal{B}\)-estradiolo nel mezzo di coltura.

# Espressione in Baculovirus

Famiglia di virus dsDNA infettanti artropodi molto utilizzati in ricerca, grazie alla elevata specificità



- 1. l'artropode si nutre di un substrato infetto, ad es. una foglia con associati i corpi di occlusione (OB), che permettono la conservazione del capside e del suo contenuto in un ambiente esterno. Gli OB sono formati da poliedrina, proteina degradata solo all'interno del tubo digerente dell'artropode, garantendo le massime chance di sopravvivenza al virus.
- 2. Il virus comincia a infettare le cellule.
- 3. Il virus, all'interno della cellula, trascrive il suo DNA circolare, il quale verrà poi tradotto dalla cellula ospite e darà inizio, successivamente, alla sintesi degli enzimi virali e alla riproduzione del virus.
- **4.** Sono prodotte quindi altre unità per la propagazione nell'ospite.
- 5. Replicazione del virus nella cellula dell'insetto

Le cellule di insetti offrono alti livelli di espressione proteica con modifiche post-traduzionali simili alle cellule di mammifero, facile scale-up e crescita in liquido ad alta densità per l'espressione su larga scala. I sistemi di espressione del baculovirus sono sistemi potenti e versatili per produrre alti livelli di espressione di proteine ricombinanti nelle cellule di insetti.

# Baculovirus Multicapsid nucleopolyhedrovirus

