UPTEC X 01 021
APR 2001

ISSN 1401-2138

BJÖRN LÖNNQVIST

A study of protein-protein interactions in *Staphylococcus aureus*

Master's degree project

Molecular Biotechnology Programme
Uppsala University School of Engineering

UPTEC X 01 021	**Date of issue 2001-04**

Author

Björn Lönnqvist

Title (English)

A study of protein-protein interactions in *Staphylococcus aureus*

Title (Swedish)

Abstract

A bacterial two-hybrid system using *E. coli* as the host organism was developed and used for studying protein-protein interactions in the gram positive pathogen *Staphylococcus aureus*.

Keywords

Protein, interactions, Staphylococcus aureus, two-hybrid,

Supervisors

Dr. Fredrik Åslund

Examiner

Prof. Staffan Arvidsson

Project name	Sponsors
Language **English**	Security
ISSN 1401-2138	Classification
Supplementary bibliographical information	Pages **12**

Biology Education Centre Biomedical Center Husargatan 3 Uppsala
Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 555217

A study of protein-protein interactions in *Staphylococcus aureus*

Björn Lönnqvist

Sammanfattning

Staphylococcus aureus är en av våra vanligaste bakterier t ex har 1/3 av befolkningen bakterien i näsan. Det är också en farlig opportunist som kan orsaka alvarliga sjukdomar. Fler och fler multiresistenta stammar (d v s bakterier som vi inte kan döda med antibiotika) upptäcks varje år och det är därför viktigt att studera bakterien för att hitta andra sätt att angripa den. I det här arbetet använder jag en metod för att studera proteiners bindningar till varandra. Det är en relativt ny metod och jag har haft lite problem med att få det att fungera så storskaligt som jag hade hoppats.

Examensarbete 20 poäng, Molekylär bioteknikprogrammet
Uppsala Universitet, april 2001

Introduction

There are many available methods to study protein-protein interactions, the most simple and powerful being the two-hybrid method. The method is an *in vivo* assay that derives from the idea that protein-protein interactions can be studied directly inside a cell instead of using conventional biochemical methods (often involving expression and purification of the proteins of interest). Since each cell in a library contains different sets of possibly interacting proteins (hereafter referred to as the "bait" and "prey" proteins), millions of pair wise interactions can be studied at once on relatively few agar plates. This provides the means for a brute force screening for protein-protein interactions of a complete genome (ref. 1).

Two-hybrid systems using yeast as the host organism has been around since the beginning of the nineties and has proved to be an invaluable tool to find protein-protein interactions. The map of interactions of all yeast proteins was published last year (ref. 2) and showed that 1000 proteins out of a total of about 6000 interacted with each other. This map should be considered as hypothesis generating and it has provided many important clues to function of unclassified proteins and functional linkages between cellular processes.

The yeast system has several drawbacks: *i*) it relies on importation of the "bait" and "prey" fusion proteins into the nucleus *ii*) the transformation capacity of yeast is considerably lower than that of *E. coli* and *iii*) the selection procedure is time consuming and inefficient.

Several of these disadvantages can be overcome by using a bacterial two-hybrid system, for instance by using *E. coli* as the host organism for the system. *E. coli* grows fast, has high transformation efficiency and the "bait" and "prey" fusion proteins do not need to be exported from the cytoplasm. This means that the whole screening procedure can be performed in a week.

This project aims to use and further develop a bacterial two-hybrid system in *E. coli* to screen for interacting proteins in *S. aureus*. The *E. Coli* system was developed by Simon Dove and co-workers (refs 4,6).

How a two hybrid system works

A two-hybrid system is a method designed to find interacting proteins. As the name indicates it's based on two hybrid proteins expressed in a reporter system (ref. 3).

A two-hybrid system is actually based on three different components:

1. A vector that directs the expression of one of the proteins of interest fused to a DNA binding domain, this protein is often referred to as the "bait".
2. Another vector that directs the expression of the second protein of interest fused to a transcription activation domain, this protein is referred to as the "prey".
3. One or more reporter genes places downstream of the DNA-binding site(s) recognized by the DNA-binding domain fused to the "bait".

All of these components are present in the same cell and the general idea is that when an interaction occurs between the prey and bait protein the reporter gene is expressed.

Our system

The *E. coli* two-hybrid system used in this study takes advantage of the domain structure of the α-subunit of RNA polymerase (RNAP) and the transcriptional initiation ability of the bacteriophage λ cI protein (Fig. 1).

The RNAP holoenzyme consists of the enzymatic core (subunits α_2, β and β´) and a σ-factor that directs binding to a specific class of promoters.

Many transcriptional activators in *E. coli* bind to specific sites located upstream of the transcription start point. Furthermore most of the activators that have been examined so far appear to interact with either the α-subunit (in particular the C-terminal domain (CTD)) or the σ^{70}-subunit (the most commonly used σ-factor) (ref. 5).

The transcriptional activator used in our system is the bacteriophage λ cI protein (λcI). It is a two-domain protein that binds DNA at a specific sequence called a λ-operator site. At the λ-operator site λcI binds as a dimer, and pairs of dimers bind cooperatively to adjacent sites.

The λcI-CTD mediates both dimer formation and the dimer-dimer interaction that results in cooperativity (Fig. 1A).

The λcI-NTDs primary function is to mediate binding to the DNA but it can also interact with the σ-subunit. This interaction can activate transcription but only when the λcI protein is bound at a λ-operator site centered 42 bp upstream of the transcription start point (Fig. 1B)

What Dove and co-workers did was to construct a chimera of the RNAP α-subunit. They replaced the α-subunits CTD with the CTD from λcI. The chimera would then display a dimeric target that could be contacted by an appropriately positioned λcI-dimer (Fig. 1C). They showed that this chimera increased the transcription of the reporter gene (ref. 4).

Since this result suggested that any protein-protein interaction that created a link between the α-subunit and the λcI protein could initiate transcription they realized that this could be turned into a detection system for protein-protein interactions.

They constructed a system where the prey protein is fused to the α-NTD replacing the α-CTD and the bait protein is fused to the C-terminus of the full-length λcI. They then created a reporter strain that contained a test promoter that bear the λ operator centered 62 bp upstream of the two reporter genes, β-lactamase and lacZ.

If the bait and prey proteins interact it will stabilize the RNA polymerase and activate transcription of the reporter genes (Fig 1D).

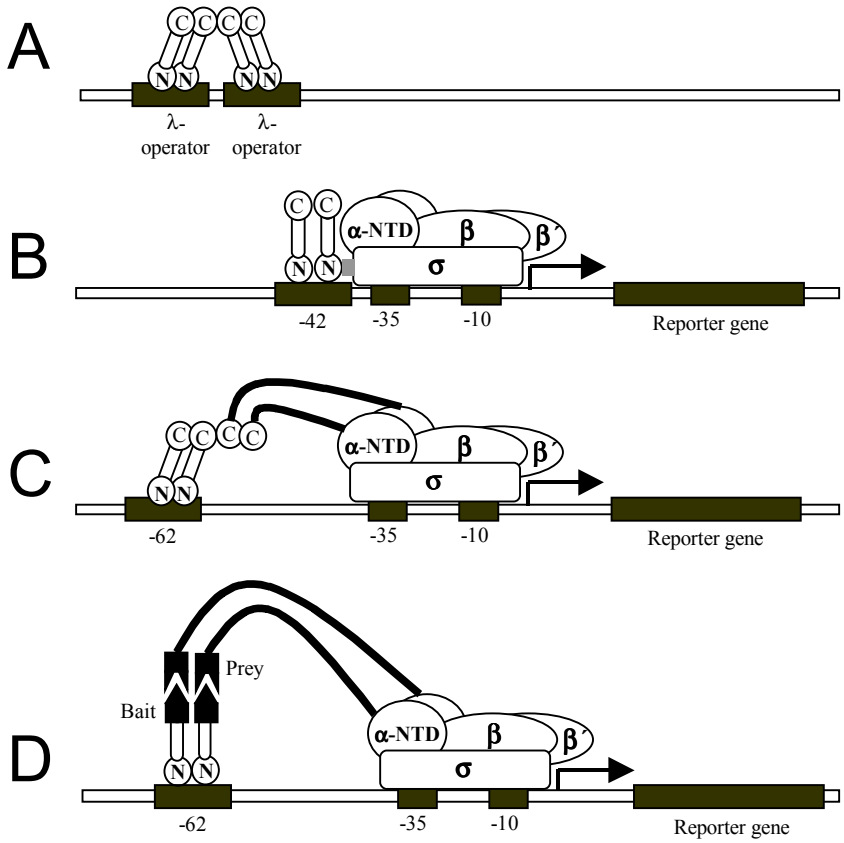

Figure 1

A The NTD of λcI binds to the DNA at a λ-operator site and the CTD mediate dimer formation and dimer-dimer interactions. **B** When the λ-operator is centered at -42 the λcI NTD can interact with the σ-subunit and initiate transcription. **C** If the λ-operator is centered at –62 the NTD can not interact with the σ-subunit but when the CTD of the α-subunit is replaced by the CTD of λcI the dimer-dimer interaction is enough to stabilize the RNAP and initiate transcription. **D** The two hybrid system in its final form. The CTD of the α-subunit is replaced by the prey protein and the bait protein is fused to the C-terminus of λcI. Interaction between the bait and prey proteins stabilizes the RNAP and initiate transcription of the reporter genes.

Materials and methods

Media, growth conditions and genetic techniques

E. coli cells were grown in LB for culturing, in SOB for preparing electrocompetent cells and on NA-plates. *S. aureus* was grown in TSB and on NA-plates. Where needed, antibiotics were used at the following concentrations: chloramphenicol (25 μg/ml), kanamycin (50 μg/ml), tetracycline (10 μg/ml) and carbenicillin (0.5–1.5 mg/ml). For detection of β-galactosidase activity cells were grown on NA-plates containing (40 μg/ml) of X-gal. IPTG was routinely used at a concentration of 50μM.

Restriction enzymes used were from Life Technologies or from Boehringer Mannheim and used according to manufacturers directions. The PCR was performed using Dynazyme (finnzymes) and standard touchdown protocol.

The test plates for the screening contained chloramphenicol, tetracycline, carbenicillin, X-gal and IPTG at the above concentrations. The carbenicillin concentration was 0.5 mg/ml for the first screen but was increased to 1.5 mg/ml for the second and the library screen.

Transformations of plasmids into *E. coli* were performed using electroporation. Competent cells were prepared as follows: An o/n culture were diluted 1/100 and grown until OD_{600}=0.5. Cells were harvested by centrifugation and washed 3 times with dH_2O and a final time with 10% glycerol. The cells were then resuspended in an appropriate volume of 10% glycerol, quick frozen and stored at –70°.

The electroporations were made using the following settings: 200Ω, 25μF and 2.3 kV/cm (Gene pulser, Biorad). The cells were taken up in 1 ml of SOB and incubated for 1h at 37° before being spread out onto NA plates with the appropriate antibiotics.

Plasmid preparations were made with Quiagen mini spin plasmid kit and performed according to manufacturers directions. Gel extractions were performed with Pharmacia Biotech gel extraction kit.

Plasmids and strains

The bait plasmid pλcI is a derivative of pACYC184 that confers chloramphenicol resistance and directs transcription of λcI under the control of the lacUV5 promoter, which can be induced by IPTG. At the 3' end of the λcI gene there is a short alanine linker and after that several restriction sites to facilitate the fusion construction. We have cloned the bait proteins into the BglII site, which gives a three residue long alanine linker

The prey plasmid pBRStar is a derivative of pBR322 that confers tetracycline resistance and directs transcription of the first 248 residues of the α subunit of the RNA polymerase. The expression is (as for pλcI) regulated by the lacUV5

promoter and is therefore IPTG inducible. At the 3' end there is a three residue alanine linker and two restriction sites, NotI and BamHI, that can be used to insert the prey protein.

The *E. coli* strain used for the screening was USF´3, which is lacZ$^-$ to enable selection for β-galactosidase. The F' episome, that confers kanamycin resistance, contains lacI^q to repress the lacUV5 promoter. The reporter genes lacZ and β-lactamase are controlled by the test promoter which is a lac promoter derivate that contains the λ operator centered 62 bp upstream of the reporter gene.

First screen

As prey proteins for the first screen 11 individually cloned *S. aureus* genes were used. Each gene was cloned into the pBRStar plasmid between the NotI and the BamHI sites. The plasmids were electroporated into the reporter strain and grown up individually over night. The over night cultures were then pooled together and competent cells containing a mix of the eleven plasmids were prepared.

The bait plasmid contained a random library of *S. aureus* genomic DNA. Genomic DNA was prepared from the S. aureus strain 8325-4. Several different digestions were made with the restriction enzyme Sau3A and the digested DNA purified either on agarose gels or with isopropanol precipitation. Sau3A cuts approximately once in every 400 bases, which gives fragments of varying lengths. The bait vector was digested with BglII. Sau3A and BglII have compatible ends so the digested genomic DNA could be ligated into the bait vector at that site.

The bait plasmid exists in three different versions, pλcI-31, pλcI-32 and pλcI-33. One, two or three bases have been inserted between the end of the λcI protein and BglII site to create plasmids which transcribes the bait proteins in all three reading frames.

The bait plasmid library was electroporated into the reporter strain containing the pool of the eleven prey proteins. The cells were then spread out on test plates. After o/n incubation at 37° positive colonies (i.e. blue colonies) were restreaked on test plates and a plasmid prep was made containing both the prey and the bait plasmid.

The bait plasmid from the positive colonies were then electroporated into a reporter strain containing a prey plasmid without any inserted prey protein i.e. expressing only the NTD of the α subunit, to confirm that an interaction between the bait and prey proteins were necessary to induce transcription of the reporter gene.

The cells were again grown on test plates and those samples that gave rise to blue colonies were discarded as false positives.

A bait plasmid without any insert i.e. only expressing the λcI gene were electroporated into the reporter strain containing the pool of prey plasmids to rule out any false positives arising from the interaction between λcI and the prey protein.

The positive colonies that passed the "false positive test" were selected for sequencing. A PCR was performed on the bait and prey plasmids to amplify the insert and the products were sequenced from the 5' end and size determinated on agarose gels.

Creating library plasmids

Three modified versions of the prey plasmid were needed to enable screening of two genomic libraries against each other. The three different bait plasmids (pλcI-31, pλcI-32 and pλcI-33) were used as templates for a PCR reaction resulting in a 150bp fragment containing the BglII site and the different linker that creates the three different reading frames. The PCR fragment were digested with NotI and BamHI and ligated into pBRStar to create pBstar31, pBstar32 and pBstar33. As for pλcI these three plasmids facilitate insertion of a genomic fragment in a BglII site and the only difference between the three variants is that they express the inserted fragment in different reading frames.

A new library was created in the prey plasmids with the same procedure as for the bait plasmid.

Library screen

The prey library was electroporated into the reporter strain and competent cells prepared as above. The bait library was then electroporated into the reporter strain containing the prey library and the culture selected on test plates. Positive colonies were picked and restreaked on test plates before plasmid preparation. The positive colonies were then subjected to two different false positive tests as described above for the first screen.

After discarding the false positives the unknown inserts of both the prey and bait plasmids were sequenced.

Results and discussion

First Screen

The 11 prey proteins (Table 1) chosen to be tested against the genomic bait library were selected either because they had an important role in *S. aureus* or because they were of unknown function and finding possible interactors might give a clue about their function.

Gene	*SarA*	*AgrA*	*411-4*	*546-2*	*DHS16512*	*312-2*	*PdHC*	*RpoBup*	*325-2*	*GidBH*	*168292*
Size (bp)	342	717	1698	936	720	1075	1407	609	543	715	740

Table 1 The eleven prey proteins used in the first four screens

The first screen where the three different bait plasmid libraries (pλcI-31, pλcI-32 and pλcI-33) were transformed into the reporter strain containing a pool of the 11 prey plasmids gave 40 positive colonies (Table 2). 17 colonies were restreaked on test plates and plasmid preps for those were made containing a mix of both prey and bait plasmid. 4 colonies were discarded due to various problems during this step but 13 were subjected to the false positive test described above. 6 turned out to be false positives. 2 positive colonies and 2 false positive were selected for sequencing and several different sequencing templates were tested to improve the result. Unfortunately none of the sequenced inserts were a functional *S. aureus* protein in the correct reading frame.

Second screen

The second screen was performed in parallel with the first and under the same conditions. The false positive frequency was much lower but none of the sequenced inserts were of any real significance.

Third and fourth screen

For the third and fourth screens the carbenicillin concentration on the test plates were increased to 1.5 mg/ml. Increasing the carbenicillin concentration had proved to decrease the number of positive colonies (data not shown) and was therefore believed to subject the protein-protein interactions to a stronger selection pressure.

After two electroporations of the bait libraries into the reporter strain containing the 11 prey plasmids 18 positive colonies had been collected and restreaked on test plates. When subjected to the false positive test 13 passed completely 3 were definitely false positives and 2 were suspected false positives. The 13 positive colonies were sequenced but only 3 translated to any part of a functional protein.

Screen	*Chloramphenicol # colonies*	*Test plate # colonies*	*Ratio positive/colonies*	*Positive colonies picked*	*False positives*
1st	$1 \bullet 10^6$	40	10^{-5}	13	6
2nd	$5 \bullet 10^6$	600	10^{-4}	12	0
3rd and 4th	-	-	-	18	5

Table 2 The results from the first four screens

Library screen

The genomic libraries in both the bait and the prey plasmids were prepared as described above. Three different Sau3A digests of genomic DNA from *S. aureus* strain 8325-4 were made. Two partial digestions with 1/100 diluted enzyme and a digestion time of 30 min and 1 hour respectively One complete digest where undiluted enzyme was allowed to digest the genomic DNA o/n. The digested DNA was isopropanol precipitated and ligated into the three different kinds of bait and prey plasmids.

The partial and complete digestions were meant to create fragments of varying lengths and give a high complexity to the library. The complexity of the libraries was determined when the ligations were electroporated into *E. Coli.* All the three libraries gave rise to at least 10^6 colonies which was regarded as high enough complexity to represent the 2500 genes of *S. aureus.*

The prey library was electroporated into the reporter strain and competent cells prepared. The bait library was then transformed into the cells and spread out onto test plates containing 1.5 mg/ml carbenicillin. Six colonies were picked and restreaked on test plates for further studies. Plasmids were prepared from the colonies and both the prey and bait plasmids were subjected to a false positive test with an empty bait/prey plasmid as partner. Five colonies passed both tests but the sequence information was as disappointing as for the previous screens.

Improvements and further studies

So far the two-hybrid system hasn't given us much information but there are possible improvements of the system to be made:

1. The libraries have high complexity but since half of the inserts will be in the wrong direction and there is a slight bias towards smaller fragments there might not be any interacting proteins in the libraries. Creating more libraries by different methods might solve this problem. For example the vector can be changed as to facilitate inserts from a sonicated library or the Sau3A digestions might be separated by size on agarose gels.
2. The selection system might be improved. The increased carbenicillin concentration improved the selection but the colonies were often not visible until after two days of incubation. This should be compared with the growth of a colony containing plasmids encoding fusion proteins that are known to interact, which give rise to large colonies in one day. So perhaps none of the positive colonies picked were in fact true interactions.

At first we will try to use the system on a smaller scale by cloning individual genes into the bait and prey vectors. We will use three different mammalian proteins, Socs 1-3 and their suspected interactors elongin b and c. If the first results are positive we will also try mutated forms of the above proteins. Perhaps the system can be improved when less complex prey and bait libraries are used.

Abbreviations

CGR	Center for genomic research
CTD	carboxy-terminal domain
DNA	deoxyribonucleic acid
E. coli	*Escherichia Coli*
IPTG	isopropyl-β-D-thiogalactopyranoside
LB	Luria-Bertani Broth
MTC	Molecular and tumor biological center
NA	Nutrient agar
NTD	amino-terminal domain
o/n	over night
OD	optical density
PCR	polymeric chain reaction
RNA	ribonucleic acid
RNAP	RNA polymerase
S. aureus	*Staphylococcus aureus*
TSB	Tryptic Soy Broth
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

Acknowledgments

First I would like to express my deepest gratitude towards my supervisor Dr. Fredrik Åslund at CGR, Karolinska Institute, Stockholm. He has encourage and inspirited me when the results were disappointing and he has always believed in me and allowed me to try out my new (and sometimes stupid) ideas.

I would also like to thank Agneta Wahlqvist, Lena Norenius and all the other members in Staffan Arvidssons group at MTC, Karolinska Institute, Stockholm for assisting me in my laboratory work and always answering my questions with a smile. Without them I would never have been able to complete my project and the warm and friendly environment has made the months just fly by.

Finally I would like to thank professor Staffan Arvidsson, my examiner, for taking me into his group without hesitation and for all the great discussions about my project and science in general.

References

1 Hu, J. C., Kornacker, M. G. & Hochschild, A. (2000) *Methods* **20**, 80-94

2 Uetz *et al*. (2000) *Nature* **403**, 623-627

3 Colas, P. & Brent, R. (1998) *Trends Biotechnol.* **16**, 355-363

4 Dove, S. L., Joung, J. K. & Hochschild, A. (1997) *Nature (London)* **386**, 627-630

5 Hochschild, A. & Dove, S. L. (1998) *Cell* **92**, 597-600

6 Dove, S. L. & Hochschild, A. (1998) *Genes Dev.* **12**, 745-754

Appendix A

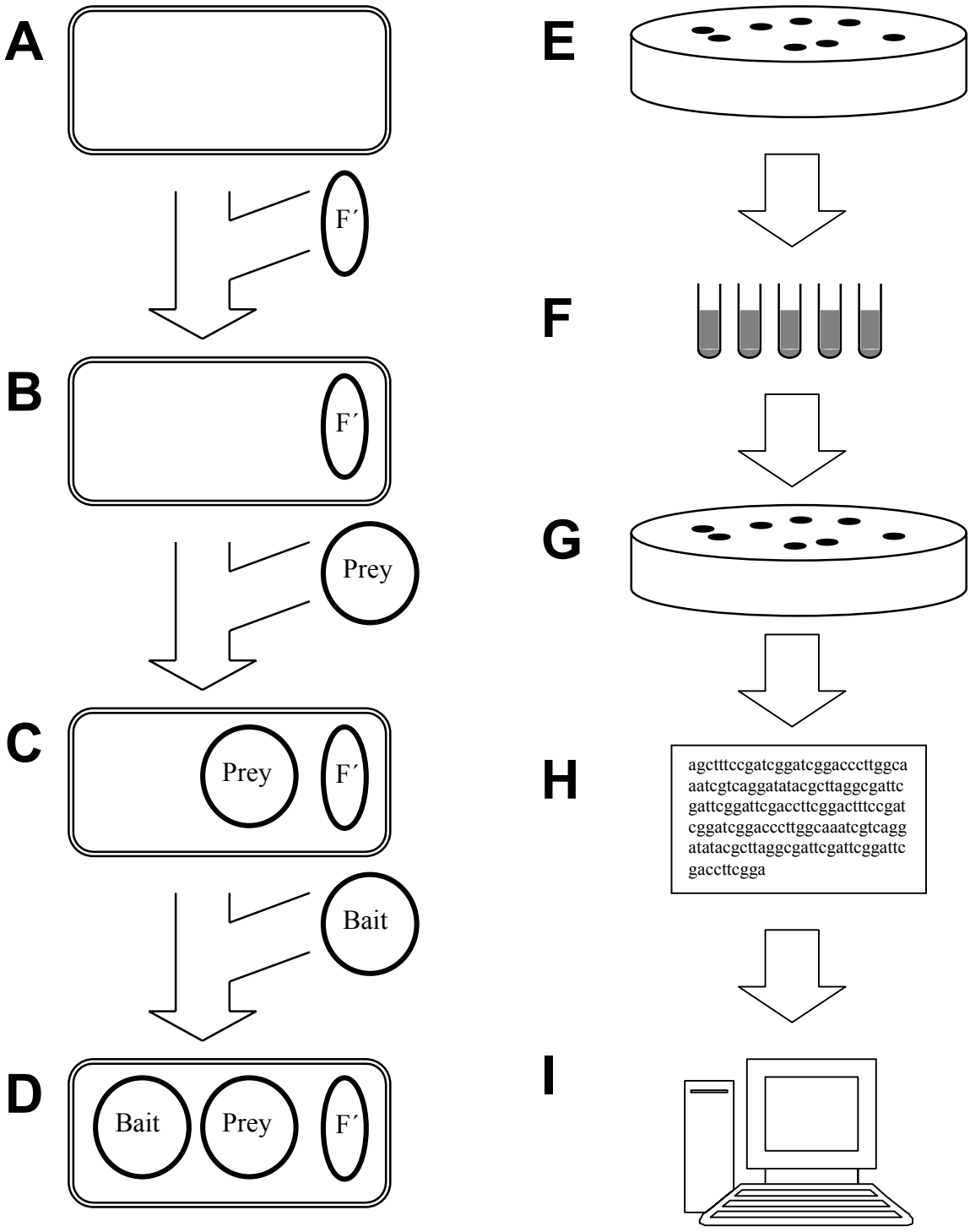

Figure 2

A,B The F´ episome containing the reporter genes and the lacI^q is mated into the E.Coli strain. **C** Electrocompetent cells are prepared and the prey plasmid is transformed into the reporter strain. **D** Electrocompetent cells from the reporter strain containing the prey plasmid are prepared and the bait plasmid or the bait plasmid library are transformed into that cell line. **E** Cells are selected on test plates and **F** plasmids are prepared from positive colonies. **G** The plasmids are screened for false positives and clones that pass the test are **H** sequenced. **I** The sequences are analyzed by various bioinformatic methods.