Bio Sci 203 Lecture 25 - Genomic libraries, etc

- Bruce Blumberg (blumberg@uci.edu)
  - office – 2113E McGaugh Hall
  - 824-8573
  - lab x46873, x43116
  - office hours Wednesday 11-12.


- Link is also on main class web site

- Goals:
  - impart an appreciation of how to approach gene isolation and characterization
    - how to do recombinant DNA manipulations
    - I may begin each lecture with discussion of an important technique
  - present a practical introduction to techniques
    - library construction and use
    - gene identification
    - functional analysis
  - point out some of the pitfalls of various methods and why certain methods are not appropriate to answer particular questions
- Please feel free to ask me questions at any time
Introduction - Sanity checks in molecular biology

• Sanity checks are one of the most important and overlooked aspects of molecular biology
  – are the data reasonable on their face?
  – Troubleshooting failed experiments is an important skill but prevention is better yet
  – small time savings from shortcuts can waste weeks or months

• example 1
  – tube labeled as 23.7 mg/ml plasmid DNA obtained from another laboratory (friend in a world class lab)
  – diluted 1:237 to get a nominal 0.1 mg/ml solution
  – DNA was used directly in a transfection experiment which didn’t work
  – experiment repeated many times over several months and didn’t work a single time

• example 2
  – did a plasmid maxiprep, resuspended DNA in buffer and checked absorbance using built in program in spectrophotometer
  – obtained readings as follows:
    • OD 260 1.45
    • OD 280 1.4
    • OD 320 1.5
Sanity checks in molecular biology (contd)

• example 3
  – plasmid maxiprep yielded readings as follows
    • OD 260  3
    • OD 280  3
    • OD 320  0.01
  – DNA wouldn’t digest or label

• example 4
  – plasmid prep yielded the following spectrophotometer readings
    • OD 260  3
    • OD 280  2.3
    • OD 320  0.02
  – question was whether the DNA was any good since the ratio was only 1.3

• example 5
  – I diluted some DNA and read the following:
    • OD 260  0.013
    • OD 280  0.007
    • OD 320  0.000
  – ratio is 1.85 -> clean DNA, right?
Sanity checks in molecular biology (contd)

- example 6
  - transformed 1 pg of DNA into standard laboratory grade competent E. coli then plated 10% and 90% of the transformation onto LB-Amp plates
  - 10% plate gave too many colonies to count (>5000)
  - 90% plate gave a lawn of bacteria
  - picked colonies from 10% plate -> didn’t grow on fresh LB-Amp plates
  - diagnosis? Awesome transformation, right?
  - 5000 colonies * 10 = 5 x 10^4/pg or 5 x 10^10 cfu/ug
    - reasonable expectation for transformation?
      - Lab grade cells ~10^7 cfu/ug sc DNA
      - ultracompetent cells ~10^9 cfu/ug DNA
      - electroporation ~10^9 cfu/ug
  - common flaw?
    - Incorrect assumptions
      - instrument readings are always reliable
        - every machine has a linear range
        - no measuring device is accurate at the limits of its range
        - common contaminants and their effects
      - believed data that should have been suspect on their face
        - concentrations
        - number of colonies
Rules for success in the lab

• Trust no one - assume nothing
  – never believe a concentration unless you measured it
  – Never rely on the identity of a plasmid unless you sequenced it
  – never rely on stock reagents for important experiments
    • especially if you didn’t make it

• do the controls
  – a negative result without a positive control is meaningless
  – a positive result without a negative control is meaningless
  – first assumption for a failed experiment should be that you made a mistake
    • controls play a critical role in troubleshooting
  – adequate controls prevent needless repetition
  – if you don’t have time or materials to do the controls you shouldn’t do the experiment

• time is money
  – your wasted time is painful for you but costly for us
    • wasted time -> more teaching
    • waste time -> scooped by competition
Basic techniques and rules of thumb

• How does one precipitate nucleic acids?
  – What are important considerations?
    • What else is in the solution?
      – Proteins?
      – Small fragments
      – oligonucleotides
    • What will the NA be used for next?
      – Cloning
      – transfection
      – in vitro transcription/translation
      – enzymatic labeling
  • what is the concentration now and what is desirable concentration?
How to precipitate DNA

• Examples of how to ppt DNA?

• Does the method used make any difference in the final product?
Basic techniques and rules of thumb (contd)

- Precipitation by PEG
  - polyethylene glycol (6000-8000 mw) selectively precipitates macromolecules by size
  - ds-DNA is quantitatively precipitated by >6.7% PEG
  - size selectivity requires >2 fold difference
  - 10% PEG, 2.5 M NaCl is good choice to get rid of small fragments and oligos
  - centrifugation temperature is critical
    - rt >>>>> 4% WHY?
  - no real difference after incubation - quantity related

FIGURE 1. PEG precipitation of small DNA fragments. Samples were electrophoresed in 2% agarose/TAE buffer with 1 μg/ml ethidium bromide at 102 V (6 V/cm) for 35 min with a cooling fan. Lane S, 250 ng of the 50 bp DNA Ladder. Final PEG concentrations were 10%, 8.3%, 6.7%, 5%, 3.3%, and 1.7% in lanes 1–6, respectively. Panel A. Samples were centrifuged immediately after mixing with PEG. Panel B. Samples were incubated for 10 min with PEG prior to centrifugation.
Basic techniques and rules of thumb (contd)

• PEG precipitation (contd)
  – advantages
    • very clean DNA
    • quantitative precipitation
    • some ability to size fractionate
  – disadvantages
    • a bit time consuming
    • pellets are absolutely clear - good technique is required
    • occasionally difficult to resuspend large amounts of NA after PEG ppt.
  – common applications
    • cleaning up sequencing reactions
    • cleaning up templates to be used for sequencing
    • preparation of phage and phage DNA
Basic techniques and rules of thumb (contd)

- Salt precipitation
  - LiCl
    - RNA can be selectively precipitated from DNA and proteins by 2.5 M LiCl
    - this also eliminates nucleotides
    - incubate on ice > 30’ then spin at rt
    - great purification method if the RNA you will make is to be used for in situ hybridization or microinjection
  - NH$_4$-acetate
    - proteins can be selectively precipitated by making sample 2.5 M in NH$_4$-acetate and incubating for 30’
    - after removal of precipitated proteins by centrifugation, the nucleic acids can be recovered by adding 2-2.5 volumes of ethanol
- advantages
  - can achieve selective precipitation of NA
  - best choice to get rid of nucleotides and cap analog from in vitro transcription reactions
- disadvantages
  - some size selection - small RNAs not precipitated very well
Basic techniques and rules of thumb (contd)

- Alcohol precipitation
  - what type of alcohol is best, ethanol or 2-propanol?
  - Volume
    - 2-2.5 volumes of etoh
    - 1 volume of 2-propanol
  - cleanliness required
    - 2-propanol does not ppt proteins well
    - less volatile than ethanol
  - what salts are used
    - NaCl 0.3 M
      - Cl⁻ ions inhibit in vitro TNT reactions
    - LiCl 0.8 M
      - Li⁺ inhibits reverse transcriptase
    - Na-acetate 0.3 M
      - good general purpose salt
    - NH₄-acetate 2.5 M
      - gives cleanest DNA, no protein, small fragments or nucleotides
      - NH₄ inhibits PNK and TdT
  - Effect of order of addition
    - DNA as a crystallization reaction
    - cleanest DNA comes from slow precipitation, i.e., add ethanol first, mix well then add salt.
    - Try this with plasmid DNA to see the difference
Basic techniques and rules of thumb (contd)

- Factors affecting recovery of DNA (Crouse and Amorese (1987) Focus 9:2 p3-16)
  - centrifugation
    - temperature 22° C > 4° C  WHY?
    - time 30’ > 15’
    - volume small > large

What is wrong with data in table below?

Table 1. Effect of time and temperature on ethanol precipitation with ammonium acetate.

<table>
<thead>
<tr>
<th>DNA Concentration</th>
<th>70° C</th>
<th>-20° C</th>
<th>0° C</th>
<th>22° C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0'</td>
<td>10'</td>
<td>30'</td>
<td>overnight</td>
</tr>
<tr>
<td>5 µg/ml</td>
<td>85</td>
<td>80</td>
<td>91</td>
<td>91</td>
</tr>
<tr>
<td>0.5 µg/ml</td>
<td>62</td>
<td>46</td>
<td>52</td>
<td>50</td>
</tr>
<tr>
<td>0.05 µg/ml</td>
<td>29</td>
<td>29</td>
<td>30</td>
<td>32</td>
</tr>
<tr>
<td>0.005 µg/ml</td>
<td>25</td>
<td>27</td>
<td>38</td>
<td>33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DNA Concentration</th>
<th>0'</th>
<th>10'</th>
<th>30'</th>
<th>overnight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0'</td>
<td>10'</td>
<td>30'</td>
<td>overnight</td>
</tr>
<tr>
<td>5 µg/ml</td>
<td>88</td>
<td>94</td>
<td>94</td>
<td>96</td>
</tr>
<tr>
<td>0.5 µg/ml</td>
<td>86</td>
<td>64</td>
<td>65</td>
<td>92</td>
</tr>
<tr>
<td>0.05 µg/ml</td>
<td>63</td>
<td>40</td>
<td>36</td>
<td>87</td>
</tr>
<tr>
<td>0.005 µg/ml</td>
<td>37</td>
<td>33</td>
<td>39</td>
<td>85</td>
</tr>
</tbody>
</table>

Note: Data shown in bold type had ≥80% recovery.

- 5 ug/ml
- 0.5 ug/ml
- 0.05 ug/ml
- 0.005 ug/ml

Figure 1. Effect of incubation temperature on ethanol precipitation with ammonium acetate. All solutions were incubated overnight at the designated temperature and centrifuged for 15 min at 22°C. DNA concentrations were 5 µg/ml (×), 0.5 µg/ml (■), 0.05 µg/ml (▲), and 0.005 µg/ml (●).
Basic techniques and rules of thumb (contd)

- Rule of thumb for nucleic acid precipitation
  - for amounts > 5 ug/ml
    - calculate amounts of solution required
      - example
        » 100 µl of DNA solution
        » 10 µl of Na-acetate
        » 220 µl of ethanol
    - add 2 volumes of room temp ethanol
    - then add 1/10 volume of Na-acetate
    - mix well and centrifuge for at least 15’ at rt
      - incubation at rt or on ice is optional but may improve recovery
    - rinse pellet with 70% ethanol to remove salts, then dry briefly.
  - for amounts < 5 ug/ml or unknown quantities
    - calculate amounts
    - add 2 volumes of room temp ethanol
    - add 1/10 volume of Na-acetate
    - mix well and incubate on ice for at least 30’ or preferably several hours to overnight
    - spin 30’ at rt
    - rinse pellet with 70% ethanol and dry briefly
  - for RNA
    - incubation should be on ice to minimize the activity of any RNases present.
Basic techniques and rules of thumb (contd)

• Alcohol precipitation (contd)
  – advantages
    • most versatile choice for concentrating NA
    • can concentrate DNA from arbitrary volume quickly
  – disadvantages
    • occasional contamination of 100% alcohol with heavy metal ions can lead to random breakage
      – less frequent with glass bottles
    • not as easy to remove small fragments and nucleotides as with salt or PEG precipitation
What are genomic libraries?


• What types of libraries are useful?

• Why do we care which one will be used?
Vectors for constructing genomic libraries

- bacteriophage lambda
  - 15-20 kb inserts
  - infection and phage recovery
- cosmids
  - up to 40 kb inserts
  - infection and plasmid purification
- Fosmid
  - Up to 40 kb inserts
  - Infection and plasmid purification (high copy)
- bacteriophage P1
  - up to 95 kb inserts
  - infection and plasmid purification
- PAC (P1 artificial chromosome)
  - no theoretical limit to insert size (~150-300 kb)
  - transformation and plasmid purification
- BAC (bacterial artificial chromosome)
  - no theoretical limit to insert size (150-300 kb)
  - transformation and plasmid purification
- YAC (yeast artificial chromosome)
  - no limit to insert size practical maximum is ~1 mb
  - transformation and plasmid purification

Genomic libraries (contd)

• What do we commonly use genomic libraries for?

• Considerations before making a genomic library?
Genomic libraries (contd)

• What do we commonly use genomic libraries for?
  – Genome sequencing
  – gene cloning prior to targeted disruption or promoter analysis
  – positional cloning
    • genetic mapping
      – Radiation hybrid
      – STS (sequence tagged sites)
    • chromosome walking
    • gene identification from large insert clones
    • disease locus isolation and characterization

• Considerations before making a genomic library
  – what will you use it for, i.e., what size inserts are required?
    • Walking to a clone
    • isolation of genes for knockouts
  – Are high quality validated libraries available?
    • Caveat emptor
      – Drosophila ~50% of clones are not traceable to original plates
      – Research Genetics Xenopus tropicalis BAC library is really Xenopus laevis
        » diploid vs tetraploid
    • apply stringent standards, your time is valuable
Genomic libraries (contd.)

• Considerations before making a genomic library (contd)
  – availability of equipment?
    • PFGE
    • laboratory automation
    • if not available locally, it may be better to use a commercial library when available

• Goals for a genomic library
  – Faithful representation of genome
    • clonability and stability of fragments essential
    • >5 fold coverage is desirable (i.e., base library should have a complexity of five times the estimated genome size to have a 95% probability of identifying a clone.
  – easy to screen
    • plaques much easier to deal with colonies UNLESS you are dealing with libraries spotted in high density on filter supports
  – easy to produce quantities of DNA for further analysis
Construction of a genomic library

- Prepare HMW DNA
  - bacteriophage λ or cosmids
    - partial digest with frequent (4) cutter followed by sucrose gradient fractionation or gel electrophoresis
      - Sau3A (^GATC) most frequently used, compatible with BamHI (G^GATCC)
    - why can’t we use rare cutters?
  - Ligate to phage or cosmid arms then package in vitro
    - Stratagene >>> better than competition
    - Vectors that accept larger inserts
      - prepare DNA by enzyme digestion in agarose blocks
        - why?
  - Partial digest with frequent cutter
  - Separate size range of interest by PFGE (pulsed field gel electrophoresis)
  - ligate to vector and transform by electroporation

- What is the potential flaw for all these methods?
Construction of a genomic library (contd)

• What is the potential flaw for all these methods?
  – Unequal representation of restriction sites, even 4 cutters in genome
  – large regions may exist devoid of any restriction sites
    • tend not to be in genes

• Solution?
  – Shear DNA or cut with several 4 cutters, then methylate and attach linkers for cloning
  – benefits
    • should get accurate representation of genome
    • can select restriction sites for particular vector (i.e., not limited to BamHI)
  – pitfalls
    • quality of methylases
    • more steps
    • potential for artefactual ligation of fragments
      – molar excess of linkers