This section of the examination will be worth 50 points. Take your time, read the questions carefully and answer them as completely as possible. Many questions have more than one possible correct answer. Be sure to explain your logic in arriving at the answer given. Good luck!

1. (6 points) Your laboratory works on the strange organisms that live around hydrothermal vents in the deep ocean as a model system for the first multicellular organisms. Your PI has developed a new method of culturing such organisms, making it possible to grow the wormlike animals found around the vents in the laboratory. One of the first things that needs to be done is to construct the molecular tools that will be required to characterize your assigned animal, the Pompeii worm (Alvinella pompejana) which can survive an environment as hot as 80° C. The ultimate goal will be to establish an A. pompejana genome project including whole genome sequencing and mapping, an EST project and DNA microarrays.

   The first goal is to make a genomic library. What type of library will you make, i.e., which type of vector? Justify your choice. What type of equipment will be required to make your library?

   You should choose to make a BAC or PAC library. BAC is best for genome sequencing because it accepts large inserts, is stable and the vector is small, facilitating shotgun sequencing

   Not so much equipment required other than standard molecular biology laboratory equipment, electroporator and PFGE – pulsed field gel electrophoresis. PFGE is indispensable for isolation of large DNA as needs to be used for making good genomic libraries.

2. (4 points) Describe a method to make a physical map of the A. pompejana genome in order to facilitate large-scale sequencing.

   Use large insert genomic library to construct a map. Map the clones by fingerprinting, map as you go, or hybridization. Restriction mapping of the whole genome was NOT an acceptable answer.
3. (5 points) You received an *E. coli* strain with the following genotype from a neighboring laboratory for the purposes of propagating your genomic library: mcrA, Δ(mrr-hsdRMS-mcrBC), ΔlacX74, deoR, recA1, araD139, Δ(ara-leu)7697, galU, galK, endA1, nupG (in every case above, the bacteria are DEFICIENT in the indicated gene product)

a) Is this a good strain for the type of genomic library you have chosen to make, i.e., does it have the necessary genetic markers for your library to be stable and readily screened?

b) If so, what are the desirable markers that the strain has. If not, which ones are missing?

c) Would the strain be suitable if you had made a YAC library? Why?

a) suitable for PAC and BAC

b) is restriction deficient, and deoR. Some also pointed out that the strain should have lacZΔM15 for blue white selection if BACs were being used.

c) strain is not suitable for YAC library because yeast artificial chromosomes can only be propagated in YEAST

4. (5 points) A colleague has experimentally determined that the *A. pompejana* genome is 110 Mb – right between *C. elegans* (97 Mb) and *Drosophila melanogaster* (120 Mb). Describe a sequencing strategy that could allow the rapid generation of a draft genome sequence. How might you combine the mapping proposed in your answer to question 2 to facilitate the completion of the genome sequence?

Whole genome shotgun will generate a rapid draft sequence. Combining this with whole genome map made in 2 will enable closing gaps.
5. (6 points) As a side project, you decide to see if the *A. pompejana* genome contains homeobox genes. You dig into the laboratory archives and find a cDNA probe that contains the *Drosophila melanogaster* Antennapedia homeobox. What is the best way to find whether the *A. pompejana* genome contains homeobox genes? If so, how will you isolate genomic clones containing these homeobox genes? Let’s say you find 8 *A. pompejana* homeobox genes. Describe a quick way to tell whether they are located in one or more clusters as in *Drosophila* or *C. elegans? 

- Genomic southern with *A. pompejana* DNA probed with Antp homeobox to work out conditions
- Screen the genomic library you made using the Antp probe using these conditions
- Once you recover the 8 genes, start hybridizing them back to the large insert clones or to Southern of PFGE electrophoresis of 8-cutter digest of genomic DNA. Note whether more than 1 homeobox gene maps to each clone or fragment

6. (4 points) If you next wanted to isolate cDNA clones that encode the homeobox genes identified in your answer to question 5, what would be the best way to work out hybridization conditions that allow cDNA identification?

- The best approach is to do a low-stringency genomic Southern to develop hybridization conditions that maximize the signal/noise ratio and then use this to screen a cDNA library.
7. (6 points) Remember that you also need to provide material for the EST project. This means that it is time to make cDNA libraries, right? Assume that the libraries you make will be used for more than just EST sequencing. What sort of vector will you choose? Should you go to the trouble of enriching the library for full-length cDNAs? If so, how? Should the libraries be standard, normalized, or subtracted? Justify your answer. If normalized or subtracted libraries are required, describe generally how you will make them.

Plasmid vector (NOT PAC or BAC)

Yes you should enrich for full-length cDNAs since the library will be used for multiple purposes

Cap trap, oligo-capping or cap-affinity chromatography gets full-length mRNA which should yield a library enriched for full-length cDNAs

The libraries should be normalized since EST sequencing is contemplated and we don’t want to sequence the same thing many times

Make normalized libraries by making driver from the library you wish to normalize, then hybridizing it back to ss-cDNA from that library to a low Cot value (5-10). After removing hybrids, use the remaining cDNA to make the normalize library

8. (4 points) What are the major differences between normalized and subtracted cDNA libraries? If you want to use a cDNA library to isolate genes expressed specifically in the tail of *A. pompejana* compared with the head, would it be better to normalize or subtract the probe that you will use? Explain your reasoning.

Normalized libraries are depleted in abundant genes and enhanced in rare genes by self-hybridization.

Subtracted libraries are depleted in genes that are common between two sources

A subtracted probe is appropriate here since you wish to identify genes specifically expressed in the tail.
9. (6 points) The head of *A. pompejana* only encounters temperatures of about 22 ºC compared with the tail, which must tolerate temperatures as high as 80 ºC. You hypothesize that the expression of one or more genes in the tail makes the tissue tolerant to high temperatures. In answering question 8, you ended up with 10 cDNAs that are expressed in the tail but not the head. As it turns out, all of these cDNAs encode the same transcription factor. Describe how you will identify cDNAs encoding proteins that interact with this transcription factor.

Any approach that detects interacting proteins is ok (biochemical fractionation, expression cloning, IVEC, two-hybrid, phage display) provided you described how to follow up and get the cDNA itself. Approaches that detect DNA-protein binding like DNA affinity chromatography, Singh screening or yeast one-hybrid were not acceptable.

10. (4 points) Your next task is to characterize the mRNA product of the gene encoding this transcription factor. What is the best way to determine how many sizes of transcripts are expressed? How will you determine what tissues express the transcription factor and at what levels?

Only Northern blot will identify the number and sizes of transcripts.

You could use Northern blots from different tissues to determine where the transcripts are expressed. Alternatively, RNase protection, RT-PCR or in situ hybridization would yield the same information.

Quantitation requires the use of some standard, either a synthetic standard or an internal standard. Any of the RNA detection methods can be made quantitative through the use of an appropriate standard.