

**Fall 2004 Genomics Exam #1**  
**Genomic Medicine and Sequencing Tools**

There is no time limit on this test, though I have tried to design one that you should be able to complete within 6 hours, except for typing and web searches. There are three pages for this test, including this cover sheet. You are not allowed discuss the test with anyone until all exams are turned in at 11:30 am on Friday October 1. **EXAMS ARE DUE AT CLASS TIME ON FRIDAY OCTOBER 1.** You may use a calculator, a ruler, your notes, the book and the internet. This is a challenging test, so do NOT put it off too long. You may take it in as many blocks of time as you need to.

The **answers to the questions must be typed within this Word file.** If you do not write your answers in the appropriate location, I may not find them. You will need to capture screen images as a part of your answers which you may do without seeking permission since your test answers will not be in the public domain. Paste the images within your Word file at the appropriate places. Print one hard copy (B&W or color, either is fine) to turn in no later than Friday at 11:30 am in class. In addition, please email me a copy of your Word file, also due by 11:30 am.

**-3 pts if you do not follow this direction.**

**Please do not write or type your name on any page other than this cover page.**

Staple all your pages (INCLUDING THE TEST PAGES) together when finished with the exam.

Name (please print):

Write out the full pledge and sign:

How long did this exam take you to complete (excluding typing)?

**30 Points**

1) Start with this partial sequence:

MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKT

- a) From what protein is this sequence?
- b) With what disease or diseases is/are this protein associated?
- c) Show me a picture of its biochemical pathway.
- d) Describe the cellular “function” of this protein? Provide the URL(s) for your source(s).
- e) Are there any alternative spliced forms of this protein? Support your answer with data.
- f) Based on your answer to part b, what can you deduce about this protein and its cellular roles?
- g) On what chromosomes are the human, mouse and rat orthologs? Support your answer with a single image.
- h) What is the Rat Accession Number for the mRNA/cDNA? What is the human accession number?
- i) What differences are there between human and rat orthologs at the amino acid level?
- j) Look at amino acid 53 in rat v. human, and you will see they are different (from h above). The rat amino acid 53, when found in humans, is associated with one of the diseases from your answer in part b of this question. Using the physical properties of the two amino acids being compared, explain why this difference in the protein could have a functional consequence. Use screen shots to support your answer.
- k) How can wt rats have amino acid #53 their way but if we have it we have a disease?

**20 Points**

- 2) Use the ECR Browser (see accompanying paper) to answer the following questions. You will need access to the paper to help you navigate.
- a) Look at chromosome 21 in this region: 23674384-23707716. Take a screen shot and then interpret what you see.
  - b) Go to 26024345-26079899 on chromosome 21. Interpret what you see.
  - c) Comment on the degree of conservation you found in these two regions and hypothesize on the significance of this conservation (with attention paid to coding vs. non-coding).

**20 Points**

- 3) There is another mystery yet to be solved: human protein called TAF1L.
- a) What is its function?
  - b) Where is it expressed?
  - c) Where is this gene located in the genome?
  - d) What have we learned in class that connects parts b and c from above?
  - e) Is there a functional mouse ortholog? Support your answer with data.
  - f) Any known diseases associated with this locus?

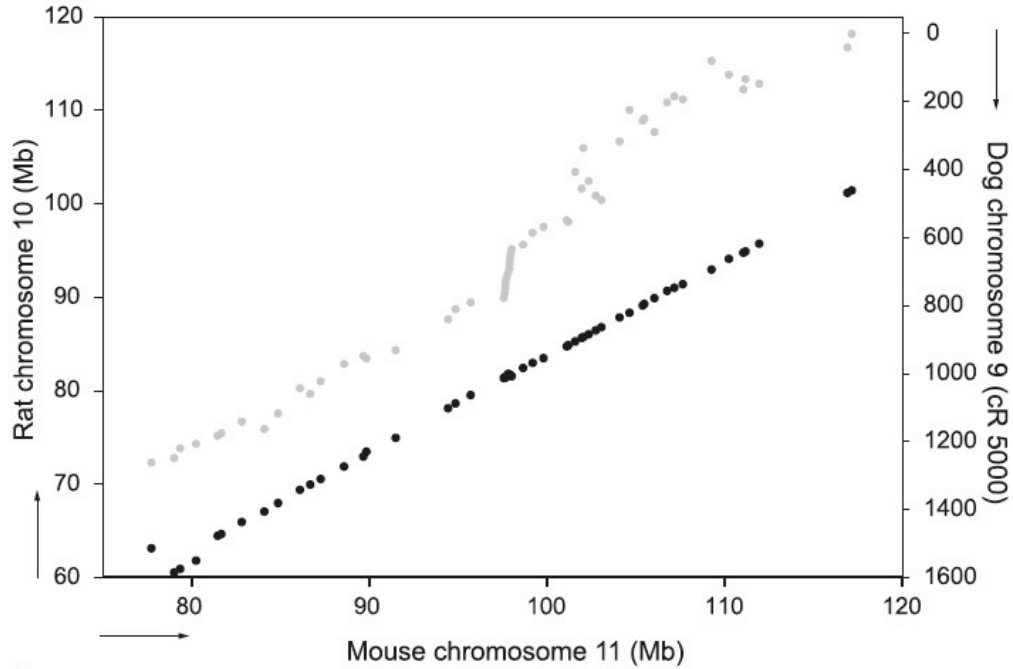
**15 Points**

- 4) a) Interpret figure 1A and B as fully as you can. Do not use information from part b to augment your interpretation for part a.
- b) Open the Exam\_PDF\_2.pdf file. Explain figure 3. Figure 2 is intended to help you if you want to see details. You will notice that some of the text has “accidentally” been lost. Do not try to

track down this paper. You have all the information you need in the figures and the figure legends.

c) Having seen part b, What can you add to your answer for part a?

A.



B.

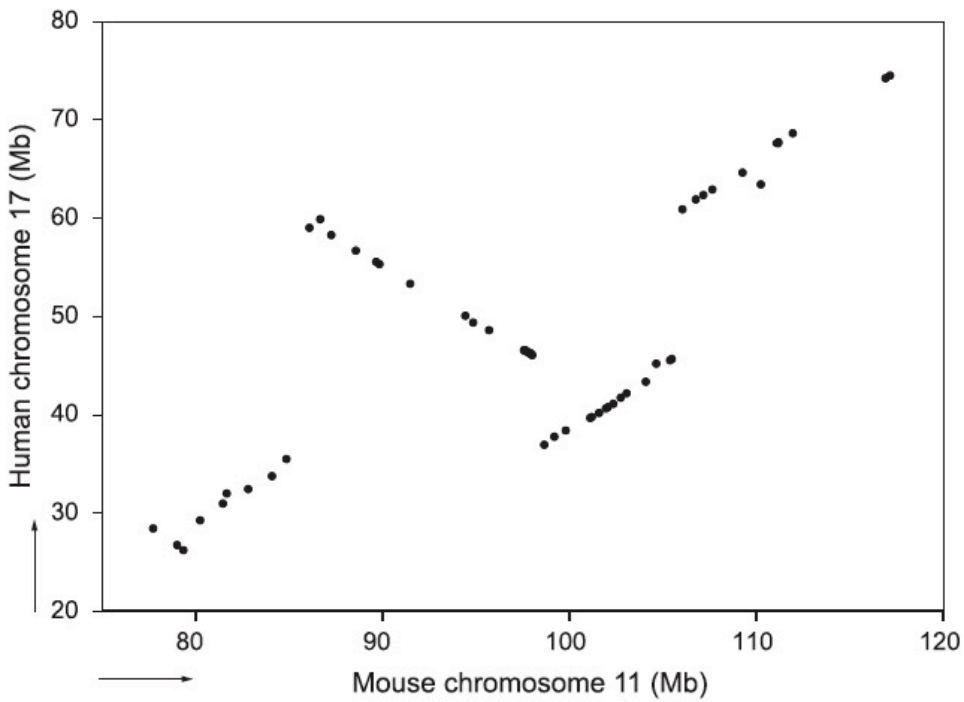


Fig. 1. Scatterplot comparison of colinearity of gene positions. Arrows indicate chromosomal orientation from centromere to telomere. (A) Black-filled circles: murine

gene positions in Mb, on the x axis, versus rat gene positions in Mb, left y axis. Gray-filled circles: murine gene positions in Mb, on the x axis, versus dog gene positions in cR5000, right y axis. (B) Murine gene positions in Mb, on the x axis, versus human gene positions in Mb, y axis.

### 15 Points

5)

a) Describe as fully as you can this protein:

```
MTLTTKLSALAIAGIMAVIGAPMVTQSAMASGRAPAPDAATTQPKLVTGDITSTDQSGTHLFFG
KNIVRNAKGAIMKVDRTWPAAVPAPLPDVRADSSSTRMMLLGPVVDLAVNEHPAGVFYRIPALATA
SNGDLLASYDLRPGSAADAPNPNSIVQRRSRDNGRTRGPQTVIHAGTLGRRKVGYSYSDPSYLVDP
ATGHILNFHVKS YDRGFATSEVGTDPDDRHLVLAHVSTSTDNGHTWTYRDITREITPDPTTRTR
FVASGQGIALLHGHAGGLIAQMTVRNSVGGQQAQSIYSDDHGITWHAGNPVGRMMDENKVVELS
DGTLMNSRDAARSGRRKVAYSHDGGTLWGPVKLVDDLIDPTNNAQIIRAYPNARAGSAKARIL
LFTNARNATERVNGTLSVSCDDGRTWVSHQTYMPGEVGYTTAAVQSDGALGVLWERDGI RYSTI
PMGWLNSVCPVAPSGRPTSSEPTSGTSLPLTATPSGSLHGGASSRPTSLPHTGD
```

*Be sure to include its function(s), species of origin, and any other aspects you can discover. Support your data with URLs or other citations.*

\*\*There is a problem with this sequence

```
GAGTTCTGGTTCGCTGTCATCAAAGTCGTCGCGATTCTTGCGATGATCGTGCTGGGTGTCCTTA
TCATTGCAACTGGCCTGGGTGGTGGCCCTCCGACCGGGATAGGTAACCTGTGGCGACACGGAGG
ATTCTTTCCAACCGGCATCAGCGGGATGCTGTGCGGTTTTGTGCTGGTGATGTTTCAGCTTTGGA
GGGTTCGAGCTCATCGGGATTACGGCAGGGGAGGCTGACGATCCGCGTCGGTCTATTCCGCGAG
CGATCAATCAAGTCGTGTATCGGATCCTCATTTTTCTACATCGGTGCAATTTTCGGTCATGTTGTG
TCTTTTTCCATGGAACCAGATCGGCAAGGCAGGCAGCCCCTTCGTGACGATCTTCGACAAAATC
GGAGTCGCAGGTGCGGCGAATATCCTCAATGTTGTGGTGCTTACCCTTCCATGTTCGGCCTACA
ACTCGGGCCTATACTCCAACGGGCGGATGCTTTACAGCTTGGCCGCTCAGCACAACGCTCCCGG
GATCTTCTGGAAGACGAATCGGCTGGGGGCGCCGTGGGTGGGAGTGCTCGCCTCCTCGGTGGTG
ACGGCAACGGCGGTGCTGCTGACGTACTTGATTCCTGGAAAGGTGTTTTGTACATCATCTCGA
TCGCCTTGATCTCTGGGGTCATCAATTGGACGATGATCATCATCACCACCTAAAGTTTTCGGCG
AAGGATCGGTCTGAAGGTGTCGCAGCGTTGGAATTTCCGATGCCGGTAATCCCCTCACCAGT
TACGTGGTGTGGTTTTTCTGGCGCTCGTGGTGGTCATCATGGCGATGATGCCGAGCTACCGAG
TGGCACTCGTTGTTGGTCCCCTGCTGGTGGCGTTGCTGTGGGTGGGTTATGACGTGTCCTGCCT
GGTGCGACGCCGTCATGCCTGA
```

b) Is this coding sequence? Support your answer with data. Use screen shots if they help you document your case.

c) What is the problem with this sequence and hypothesize how the problem happened?