The Use of Pyrosequencing for Genomic Sequence Determination

Rongahi, Mostafa. (2001). "Pyrosequencing Sheds Light on DNA Sequencing." *Genome Research*. **11**: 3-11.

Zhou, G., Tomoharu, K., et al. (2006). "Enzyme System for Improving the Detection Limit in Pyrosequencing." *Anal. Chem.* **78**(13): 4482-4489.

Model 1 How is pyrosequencing done?

1. What is the energy-producing by-product of the addition of a NTP to the DNA strand, as catalyzed by DNA polymerase?

2. Use the following figure to describe how pyrophosphate release is used to sequence DNA, paying special attention to the role that nucleotide extension plays in the process.

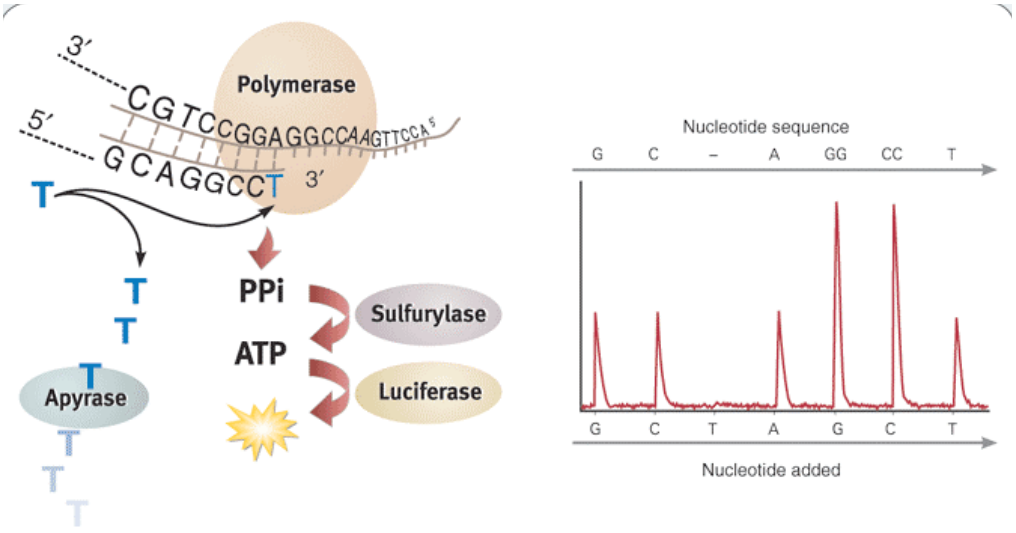
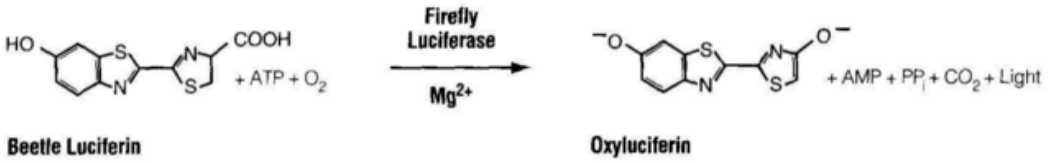


Figure 1 | The principle of Pyrosequencing and the output Pyrogram™. Double peak heights indicate incorporations of two nucleotides in a row.

3. Why does the first T base on the chart not show a peak?

4. Luciferase is an enzyme that catalyzes the conversion of luciferin to oxyluciferin and light in the presence of ATP and O_2, using the equation below.



Will the amount of light produced be doubled if the amount of ATP reacting is doubled? Explain.

MODEL 2 Interpreting a "real" pyrogram

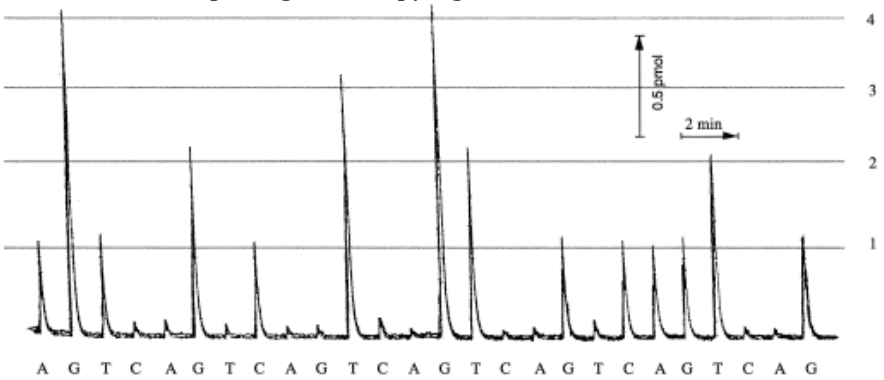
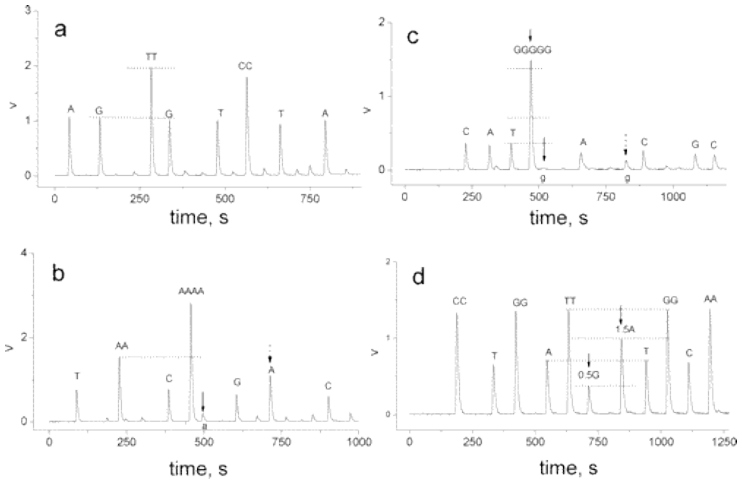


Figure 4 Pyrogram of the raw data obtained from liquid-phase pyrosequencing. Proportional signals are obtained for one, two, three, and four base incorporations. Nucleotide addition, according to the order of nucleotides, is indicated below the pyrogram and the obtained sequence is indicated above the pyrogram.

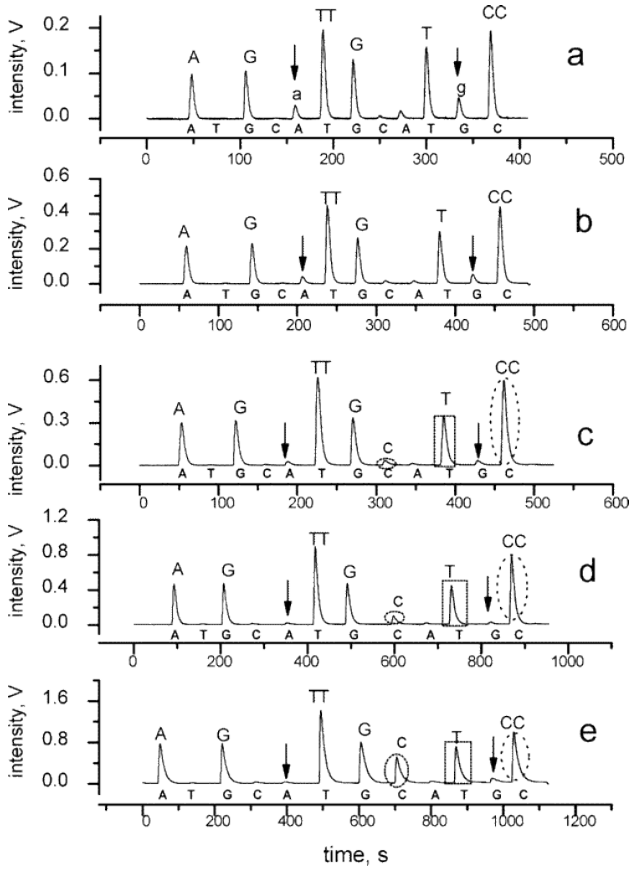
3. What do the dNTP bases under the diagram represent?

4. Why does the peak height vary among the different peaks?

5. Determine the sequence of the DNA strand from the above pyrogram.



6. For DNA sequences with nucleotide repeats, pyrogram peak height is proportional to the number of sequential bases added, up to a certain number of repeats. Given this information, use the figure above to determine the number of nucleotides at which the peak height is no longer proportional to the number of bases added.



7. Using the figure above, and given that the complementary sequence is AGTTGTCCT, identify possible plus frame and minus frame shifts in the DNA sequence given that a-e contain decreasing concentrations of Apyrase (6.7, 3.3, 2.2, 1.1 and 0.6 mU/uL). Does this figure make sense? Why or why not?

Model 3 **Designing a more efficient assay**

Table 1. Kinetic Data of Enzymes Involved in Pyrosequencing

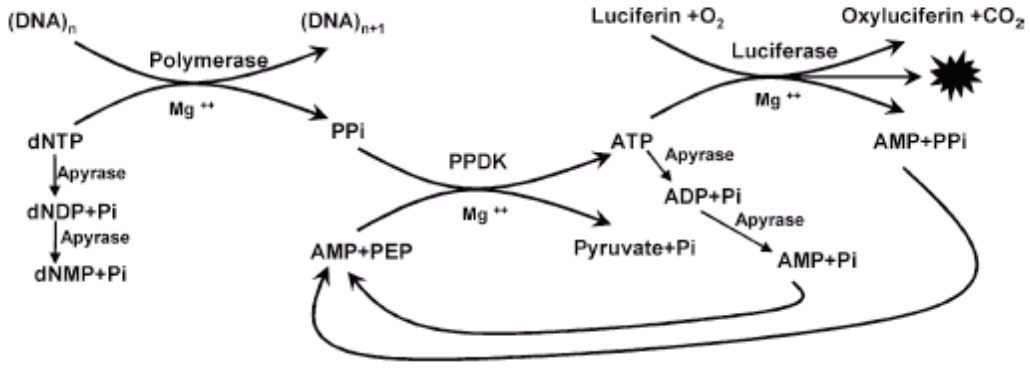
Enzyme	K_M (μM)	k_{cat} (S^{-1})
Klenow Polymerase[a]	0.18 (dTTP)	0.92
ATP sulfurylase[b]	0.56 (APS) 7.0 (PPi)	38
Firefly luciferase[c]	20 (ATP)	0.015
Apyrase[d]	120 (ATP) 260 (ADP)	500 (ATP)

[a]Van Draanen et al. 1992.
[b]Nyren and Lundin 1985.
[c]DeLuca and McElroy 1984.
[d]Traverso-Cori et al. 1965.

1. Using Table 1, write in complete sentences how each of the following concentrations of free nucleotide will react with the Pyrosequencing mixture: 0.05 uM and 300 uM. Give an estimate of the ideal concentration of free nucleotide and explain your choice.
2. Pyrosequencing is carried out by either a three-enzyme system (solid-phase) or a four-enzyme system (liquid-phase). In the three-enzyme system, the sample is washed after each successive nucleotide addition. In the four-enzyme, Apyrase is used to degrade free nucleotide. In this manner, the iterative additions of nucleotides causes an increase in sample volume of ~200 nanoliter/min. Using this value, and Table 1 (assuming that all enzymes are working optimally), calculate the volume change for sequencing a 500 bp strand.

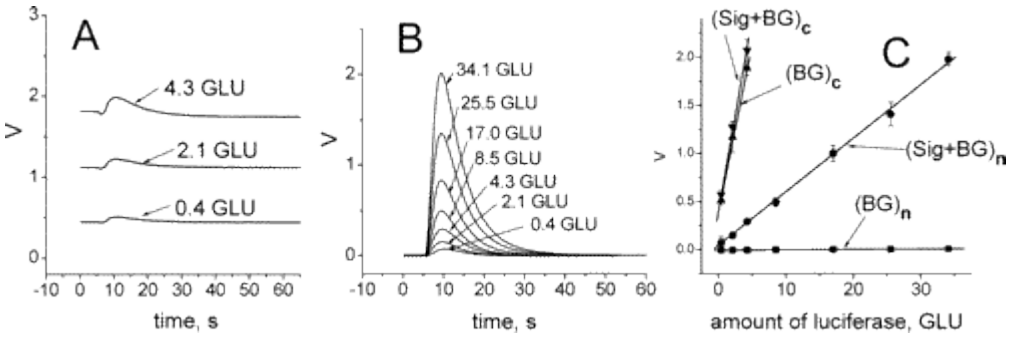
3. Given your answer to Q2 qualitatively explain what the starting volume of sample should be if using the four-enzyme system. Does this seem like an efficient genome sequencing model?

4. Given the following revised pathway (Zhou, 2006)…



Why might PPDK (pyruvate orthophosphate dikinase) be included instead of sulfurylase?

Figure 6 Zhou



5. Zhou et al were trying to find a more efficient enzyme system with which to quantify base addition. Using figure 6 C of the Zhou paper, decide whether or not Zhou found a more effective enzyme system, given that (BG)c and (Sig+BG)c are the background and signals plus backgrounds in Figure 6A and (BG)n and (Sig+BG)n are the backgrounds and signals +backgrounds in B, respectively.

Model 3 Optimization of Pyrosequencing

Figure 3 Ronaghi

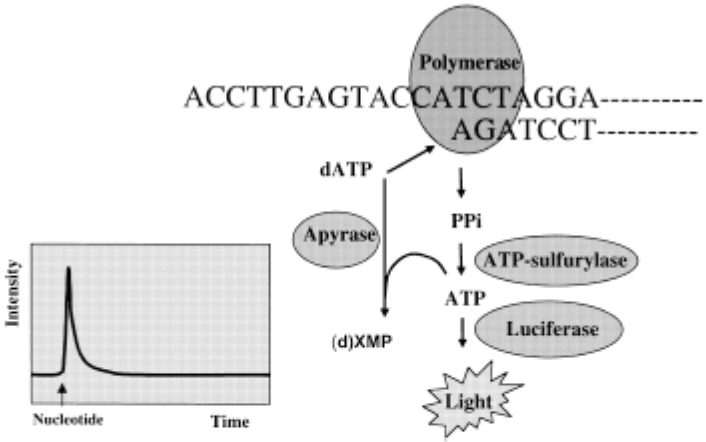


Figure 3 Schematic representation of the progress of the enzyme reaction in liquid-phase pyrosequencing. Primed DNA template and four enzymes involved in liquid-phase pyrosequencing are placed in a well of a microtiter plate. The four different nucleotides are added stepwise and incorporation is followed using the enzyme ATP sulfurylase and luciferase. The nucleotides are continuously degraded by nucleotide-degrading enzyme allowing addition of subsequent nucleotide. dXTP indicates one of the four nucleotides.

1. One of the first major improvements to pyrosequencing was the use of dATPαS instead of dATP. Using Figure 3 from Ronaghi, explain the reasoning behind this substitution.

2. Current CCD camera technology suggests that the smallest dateable reaction should contain at least 5000 DNA template molecules. Suggest another method of quantifying base addition that might require fewer templates.

3. If you were a trying to sequence the human genome using pyrosequencing would you chose to use Solid-Phase template sequencing, or Liquid-Phase sequencing? What about a small bacterial genome?

4. One of the advantages of pyrosequencing is that the addition of nucleotides is watched in real-time as bases are added to a template. Since light emission is monitored after *each* base addition, there is no need to run samples on an electrophoresis gel. What DNA/RNA structures would make gel separation problematic?

5. Considering the throughput efficiency of pyrophosphate release, is this technique better suited for de novo sequencing or re-sequencing?

6. Evaluate this POGIL using the following criteria:
Flow of questions that lead to understanding (10 points)
Questions focus on the most relevant topic (10 points)
Appropriate use of at least two figures and/or tables in question (10 points)
Challenging questions (10 points)
Depth of answers (10 points)