



Representational Difference Analysis and the Search for Unrecognized Pathogens

Step 1: Introduction and Experimental Design

Representational Difference Analysis (RDA)

RDA is one of a handful of molecular techniques that have recently been employed to investigate the possible pathogens involved in the development of human diseases. The technique compares DNA present in infected and non-infected tissue to reveal sequences unique to the infected tissue. RDA has been successfully used to detect previously unknown organisms associated with various disease conditions. For example, this technique was used to identify human herpes virus 8, a previously unknown virus now recognized to be a causative agent in the development of Kaposi's Sarcoma.

Non A-E Post-transfusion Hepatitis

Several human viruses have been discovered which cause chronic hepatitis and cirrhosis, including the viruses associated with hepatitis A, B, C, D, and E. Despite screening of the blood supply for these viruses, some surgical patients still develop symptoms of hepatitis-like infection after receiving a blood transfusion. A classic symptom of this post-transfusion hepatitis is an elevated level of the enzyme alanine aminotransferase (ALT). It appears that most people do not develop elevated levels of ALT until at least one week after transfusion, and the intervening period may represent a time when the unknown virus is at an extremely low concentration in the blood.

A relatively newly discovered positive, single-stranded RNA virus referred to as hepatitis G virus (HGV) has been suspected as the pathogen responsible for these unexplained infections. Nevertheless, PCR screening for HGV RNA revealed many patients with symptoms of non A-E post-transfusion hepatitis who were negative for the presence of HGV RNA. Another possible suspect is TTV (TT Virus, named after the initials of the patient it was first identified from), a DNA virus that appears to be associated with some cases of non A-E post-transfusion hepatitis. It is also possible that a new, undiscovered virus may be responsible for these unexplained cases of post-transfusion hepatitis.

In this scenario, pretend that you are a doctor at a research hospital and that you are investigating the possible causes of post-transfusion hepatitis. You will design a representational difference analysis to attempt to identify a sequence belonging to the pathogen responsible for non A-E post-transfusion hepatitis. According to records, approximately 1000 patients receive blood transfusions during surgery at your hospital each year. Approximately 5% of these patients develop elevated ALT levels and other symptoms of hepatitis despite testing negative for hepatitis A-E. For a representational difference analysis, you must obtain: 1.) tissue suspected of containing genetic material from the pathogen as well as the host (referred to as "tester") and 2.) tissue that contains all of the host genetic material found in the tester but is not likely to contain genetic material from the pathogen (referred to as "driver").

You will design a representational difference analysis (RDA) to identify the pathogen responsible for non A-E post-transfusion hepatitis at your hospital. The first step in designing your experiment will be to choose appropriate research subjects.

Step 2: Choosing the First Patient

After months of collecting blood samples, you have identified the following patients whose might be likely candidates for RDA. Assume that all of these patients tested negative for hepatitis A-E. They are as follows:

Patient A: Thirty-five year old male receiving surgery to correct renal artery stenosis. Blood samples taken at two weeks and at eight weeks post-surgery both revealed elevated levels of ALT. A portion of both samples was reserved for possible use in your study. Both of the patient's parents contributed units of their blood to be used in case the patient required transfusion. However, due to hospital error, the parents' blood was never used and is still in cold storage.

Patient B: Fifty-eight year old female receiving surgery for the removal of gall stones. A blood sample taken at three days post surgery revealed normal ALT levels. A blood sample taken at five weeks post-surgery revealed elevated levels of ALT. A portion of both samples was reserved for possible use in your study.

Patient C: Forty-one year old female receiving surgery for the removal of uterine fibroid tumors. A blood sample taken at four days post surgery revealed normal ALT levels. A blood sample taken at five weeks post-surgery revealed elevated levels of ALT. The blood sample taken at five weeks post-surgery was reserved for future use.

Patient D: Fifty-two year old female receiving liver transplant surgery. A blood sample. taken at four days post surgery revealed elevated ALT levels. A blood sample taken at five weeks post-surgery also revealed elevated levels of ALT. A portion of both samples was reserved for possible use in your study. This patient tested negative for HGV RNA at five weeks post surgery.

a.) On which of these patients' samples would you choose to perform RDA? Explain why the patient(s) is/are an appropriate choice.

b.) Why would the other patients not be appropriate choices?

c.) For the patient on whom you choose to perform the RDA, what would you identify as the "driver" and what would you identify as the "tester"?

Step 3: Analysis of sequence of difference product of tester amplicons

After subjecting the chosen patient's tester and driver DNA samples to RDA, you obtain one difference product consisting of many copies of a 500 bp fragment. You sequence this difference product and obtain the results shown on the attached copy of the sequencing gel. Write the sequence of nucleotides represented on this gel.

Step 4: Sequence identification

Compare the sequence you have identified to known sequences in the National Center for Biotechnology Information (NCBI) database using the BLAST search engine. Follow the instructions below for logging on to NCBI and comparing your sequence data.

- 1.) Open your Internet browser application (e.g. Netscape or Internet Explorer).
- 2.) Open location
<http://www.ncbi.nlm.nih.gov:80/blast/blast.cgi?Jform=0>
- 3.) Be sure that "blastn" is selected from the program menu. It is normally the default setting.
- 4.) Be sure that "nr" is selected from the database menu. It is also normally the default setting.
- 5.) In the entry box, type the following string:
>unknown pathogen sequence, transfusion patient
- 6.) Hit return to start a new line, and enter your sequence of nucleotide bases in lower case letters.
- 7.) When done typing in your sequence, click "Search". After the wait-time recommended by BLAST (usually less than one minute), click "Format results".
- 8.) Scroll down the results screen until you see the line that begins:

	Score	E
	(bits)	Value
Sequences producing significant alignments:		

9.) Interpret your results. The sequences listed first are those with the greatest homology to your unknown sequence. Click on a link (blue) to read more about the origin of the sequence that most closely matches yours. Click on a score to see where in the matching sequence, your sequence homology overlaps.

Note: If you had isolated a DNA fragment from a previously unknown pathogen, your sequence would not match any known sequence very closely but might have homologies (similarities) with other, related known organisms.

Step 5: Conclusions

- 1.) Does your sequence closely resemble any known pathogen sequences? If so, which ones?
- 2.) What does the presence of this sequence in the patient's tissue say about the potential cause of this disease?
- 3.) What experiment(s) would you propose to do next to confirm your results?