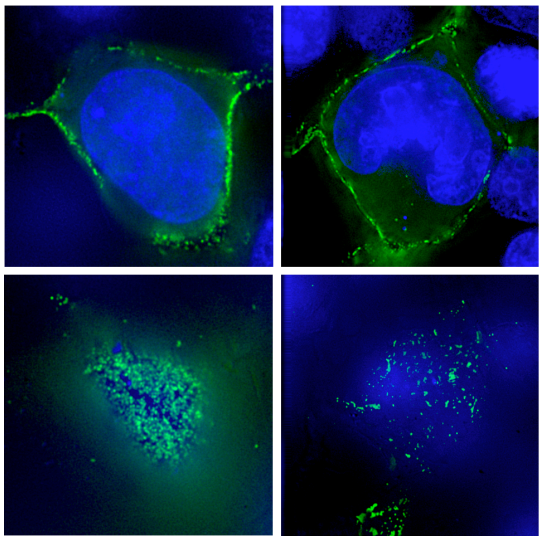

Project Proposal

- Analysis of Historical Retroviral Contributions to The Modern Human Genome

Prepared for: The Summer Research Program, Saint Mary's College of CA School of Science
Prepared by: Thomas Reynolds, Class of 2011 – Biochemistry
Faculty Mentor: Dr. Keith Garrison, Dept. of Biology

March 19, 2009

Photo Credit: Lee, Young Nam and Paul D. Beiniasz. "Reconstitution of an Infection Human Endogenous Retrovirus." PLoS Pathog. 3(1) 2007

Table of Contents

Project Proposal

Introduction

In modern biology, an organism's genome is considered its most fundamental substance–what a philosopher might call the form that shapes its physical matter. The interpretation of DNA sequences ultimately determines an organism's functions. Genomic information might even be considered the purest essence of life itself. Therefore, it might be surprising that the human genome isn't entirely human after all. The history of *Homo sapiens* is cluttered with DNA insertions, courtesy of multiple retroviral strains in the past. Studying these foreign sequences may help us understand more about important diseases such as cancer and HIV.

Target of Interest

This project will focus on a specific retrovirally-introduced sequence in the human genome, named K113. It is one of the Human Endogenous Retro-Viruses, type K (*HERV-K*), subset HML-2. *Endogenous* means that the "virus" exists as a genetic sequence in the human genome, rather than an infectious virus in nature. That does not mean it is lost to history; teams led by Lee and Dewannieaux have each cloned the resultant virus and studied its functional properties.[1,2] The *K* stands for the amino acid lysine, which corresponds to the *t*-RNA molecule which HERV-K uses to prime its synthesis.

The age of a retroviral inclusion, such as K113, is determined by the differences in its two long-terminal repeat (LTR) sections. These parts of the genome, at the beginning and end of each sequence, do not encode for anything in the virus itself; mutations may therefore accumulate in these regions over time, without affecting survival of the virus. What makes K113 of special interest is its age: roughly 200,000 to 450,000 years, judging by the complete lack of mutation in its LTRs.[3] K113 is therefore a relatively intact ancient retrovirus, preserved in the human genome like a Cretaceous-era ant trapped in amber. Nevertheless, because of its relative novelty in the human genome K113 is not universal–there are individuals who did not inherit the retroviral sequence.

---

[1] Dewannieax, Marie *et al.* " HERV-K human endogenous retroelements Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements." Genome Research. 2006

[2] Lee, Young Nam and Paul D. Beiniasz. "Reconstitution of an Infection Human Endogenous Retrovirus." PLoS Pathog. 3(1) 2007

[3] Turner, Geoffrey *et al.* "Insertional Polymorphisms of full-length endogenous retroviruses in humans." Current Biology. 11, 2001 p.p1531-1535

Experimental Overview

DNA Sample Collection

Once IRB approval is granted, the project will begin with a voluntary DNA donation campaign on the Saint Mary's campus (although alternative local sites would be desirable). Buccal (cheek) cells will be collected from the participants and stored with specialized buccal swabs (as seen in any forensic investigation or paternity suit). Donor identity will be kept strictly anonymous–only tracking numbers will be used to identify each sample, although the donor's ethnicity will be recorded for possible trend interpretation. The aim is for 100 total contributions, which will represent a sizable yet economically-feasible sample size. For statistical integrity, donors must not be symptomatic of cancer or HIV infection (see **Implications and Future Research**, p. 3). Donors will also be asked if they want to be contacted for a follow-up study. No more than five of the collected samples will be used to test and validate those procedures to follow.

DNA Extraction and PCR Amplification

Each collected sample will be laden with cell tissues, saliva, food traces, and other normal oral debris; steps will be necessary to draw out and purify the desired genetic material. After that each sample's K113 genome, if present, will be replicated by the Polymerase Chain Reaction (PCR) procedure. This will be needed for the later sequencing steps. An example protocol for the DNA Extraction and PCR Amplification experimental phase is attached to this proposal as **Appendix I**.

Gel Electrophoresis

The amplified samples above will then be analyzed by gel electrophoresis. This step will determine the presence or absence of the K113 sequence in each sample, and therefore in each donor's genome. An example protocol for Gel Electrophoresis is attached to this proposal as **Appendix II**.

Sequencing and Interpretation

Saint Mary's does not have the facilities for DNA sequencing; all the K113-positive DNA will instead be sent to the UC Berkeley DNA Sequencing Facility in Berkeley. The samples will be sequenced there for a fee, and data returned via the Internet. Once returned from UC Berkeley, the sequences will be qualitatively analyzed three ways:

1. Is K113 present in the sample at all?
2. What polymorphisms, if any, do each K113-positive sample exhibit?
3. Is there a correlation between K113, its possible polymorphic variations, and donor ethnicity?

Implications and Future Research

It is true that a K113 population study is not a novel concept. A cursory literature search reveals at least four other papers on the subject. each in different physiological conditions. K113 has been studied in individuals with breast cancer[4], testicular cancer[5], melanoma[6], and HIV[7]. It is known that the genetic abnormalities behind each condition can cause K113 gene expression to flourish.

However, to date there is little to no data on the occurrence and mutation of K113 in a large, effectively healthy sample size–that is, the general public. That is the short-term goal of this project, to generate a data set that future studies can use for baseline comparisons. This is the reason why donors must be free of either HIV infection or cancerous tissues–the presence of either condition precludes the data from accurately reflecting K113 in the general population.

A direct continuation of this project would catalogue the presence of K113 in each donor, and return to them for the collection of blood samples. As Garrison, Rakoff-Nahoum *et al* demonstrated, the human immune system can recognize HERV-K proteins and develop immune defenses against it [5, 7]. Again, however, these studies were conducted with abnormal samples. Examining the lymphocyte condition of healthy individuals, with respect to K113 presence, would generate a more detailed baseline data set for future research. A control group would come from those donors identified in this genotyping study as K113-positive or not, who would be asked if they were willing to participate in further research.

---

[4] Burmeister, Thomas *et al*. "Insertional Polymorphisms of Endogenous HERV-K113 and HERV-K115 Retroviruses in Breast Cancer Patients and Age-Matched Controls." AIDS Research and Human Retroviruses. 20, 2004. pp. 1223-1229.

[5] Rakoff-Nahoum, Seth *et al*. "Detection of T Lymphocytes Specific for Human Retrovirus K (heRV-K) in Patients with Seminoma." AIDS Research and Human Retroviruses. 22, 2006. pp. 52-56.

[6] Schiavetti, Francesca *et al.* "A Human Endogenous Retroviral Sequence Encoding an Antigen REcognized on Melanoma by Cytolytic T Lymphocytes." Cancer Research 62. 2002. pp. 5510-5516.

[7] Garrison KE, Jones RB, Meiklejohn DA, Anwar N, Ndhlovu LC, et al. (2007) T cell responses to human endogenous retroviruses in HIV-1 infection. PLoS Pathog 3(11)

Possible Timeline

Week	Task
1	Order time-sensitive reagents (*Taq* polymerase, K113 primers, etc.). Review relevant literature provided by Dr. Garrison.
2	Plan collection sites and obtain permission from relevant authorities. If possible, begin DNA Sample Collection.
3	Continue DNA Sample Collection.
4	Finish DNA Sample Collection. Necessary reagents should have arrived at this point. If so, begin DNA Extraction and PCR Amplification.
5	Continue DNA Extraction and PCR Amplification.
6	Continue DNA Extraction and PCR Amplification. Begin Gel Electrophoresis on K113-(+) samples.
7	Finish DNA Extraction and PCR Amplification. Continue Gel Electrophoresis on K113-(+) samples.
8	Finish Gel Electrophoresis. Send amplified samples to UC Berkeley DNA Sequencing Facility for Sequencing and Interpretation.
9	Perform Interpretation of Data. Analyze any apparent trends in K113 distribution and polymorphistic frequency.
10	Prepare final report, summarizing findings and suggesting causes for any trends.

N.B.: Loss or damage of samples at any point before Sequencing and Interpretation will require further sample collection. This will cause a delay in the project timeline.

Budget

Proposed Materials With Sources

Materials that can be acquired from the SMC Stock Room are not listed here.

Description	Quantity	Unit Price	Cost
AmpliTaq Gold Fast PCR Master Mix, 250 reactions (*Invitrogen*)	1	$155.00	$155.00
K113 5' Flanking Primer and Complement, .2 µM (*Operon*)	3	$20.00	$60.00
Chelex 100 Resin, 50 g (*Bio-Rad*)	1	$150.00	$150.00
Agarose, 25 g (*The Lab Depot, Inc.*)	1	$37.85	$37.85
100bp DNA Ladder (*Invitrogen*)	1	$96.00	$96.00
Cap-Shure Protective Swabs, 50 (*SecurityProUSA*)	3	$31.45	$94.35
UC Berkeley DNA Sequencing, per sample	100	$1.53	$153.00
		Total	**$746.20**

Appendix I ☣

– Example Protocol for DNA Extraction and PCR Amplification, adapted from BIO-001L Experiment #7 ("Polymorphism and DNA Identity")

1. Add 1 mL .9% saline solution to a 1.5 mL locking microfuge tube.
2. ☣ Insert cheek swab into saline tube, mix, and dispose into a biohazard receptacle.
3. Centrifuge tube(s) at an intensity of 1000*g* for 10 minutes, at room temperature.
4. ☣ Dispose of biohazardous supernatant solution, and suspend the resulting pellet in .5 mL of 10% Chelex solution.
5. Boil tubes for 10 minutes, then ice for 2 minutes.
6. Centrifuge chilled tubes again for 30 sec.
7. ☣ Transfer 200 µL of supernatant to another, clean 1.5 mL microfuge tube and put on ice. Dispose of other biohazardous products.
8. Freeze at -20˚C until ready to proceed. *End Extraction Procedure.*
9. Combine the following in a .5 mL PCR tube, cap and mix:
    a. 1 Dry PCR Bead
    b. 9 µL distilled water
    c. 5 µL K113-upstream DNA Primer Sequence
    d. 5 µL K113-downstream DNA Primer Sequence
    e. 5 µL collected genomic template from *Extraction Procedure*.
10. Program thermocycler and insert matrix of prepared PCR reaction tubes. After reaction is complete, chill tubes at 4˚C and then store at –20˚C. *End Amplification Procedure.*

Appendix II ☣

- Example Protocol for Gel Electrophoresis, adapted from BIO-001L Experiments #6 ("Analysis of DNA Using Restriction Enzymes") and #7 ("Polymorphisms and DNA Identity")

1. Prepare a submarine gel electrophoresis apparatus with a NuSieve 3% agarose gel:
    a. Add .25g anhydrous agarose to mixing bottle with 25 mL 1x TAE buffer. Mark fluid level on outside of bottle with felt-tip marker.
    b. Microwave for 30 sec., and stir until agarose is completely dissolved.
    c. Transfer bottle to 50°C water bath, and add distilled water until fluid reaches level from step a).
    d. Pour liquid gel into prepared forming tray with comb, and let set over 30 minutes.
2. Add 2.5 µL of 10x Loading Buffer to each DNA sample to be examined, and mix.
3. Using a micropipette, lead 10 µL into each well of the agarose gel.
4. Also load 5-10 µL of 100 base-pair ladder markers, for comparison.
5. Apply voltage to the gel, and let run until front line of solutions are ¾ across the length of the gel (approx. 75-90 minutes).
6. ☣ Disconnect voltage, remove gel from apparatus, and stain with ethidium bromide.
    a. Pipet 90 µL of ethidium bromide into developing tray. Add 90 mLs reverse-osmosis water.
    b. Add gel and shake for 15 minutes. Rinse twice to destain.
    c. Add 90 mLs RO water, and shake for 15 minutes. Dispose of water.
7. ☣ Photograph gel with UV camera device. Dispose of gel when done in biohazardous waste.