1.) Please provide a brief description of the rationale/motivation for pursuing this prediction and the importance/benefit of experimentally validating this gene's function.

The YP_780411 gene, RPE_1481, is a putative restriction enzyme. The domains responsible for recognition and modification are very similar between this gene and members of the LlaGI and HauI families, but the two families have adopted quite different modes of DNA cleavage. It is hoped that once the recognition specificity is known for a number of members of either family, the particular amino acids that determine given base pair recognition can be identified bioinformatically and altered experimentally, leading to the rational design of restriction enzymes of altered specificity (1). When blasted against genes of known function, the closest match is to HP0910, with 24% sequence identity, and an E_val of only 0.18. This marginal match to HP0910, a confirmed methyltransferase, suggests part of this protein's function. Although the cluster CLSK958780 has only 5 members, the proposed target is highly homologous to a large number of proteins, with BLAST e-values of less than 1e-50 to more than 100 proteins from three phyla, most in currently incompletely sequenced organisms. Thus, information gained about this protein is likely to inform the function of many potential restriction enzymes.

1. Morgan, R.D. and Luyten, Y.A. (2009) Rational engineering of type II restriction endonuclease DNA binding and cleavage specificity. *Nucleic Acids Res*, **37**, 5222-5233.

2.) Please describe what is experimentally known about this gene. Results from a thorough Pubmed search should be described, and is an essential element of a successful application.

RPE_1481 is a fusion of an endonuclease domain and a DNA methyltransferase domain that share a single DNA recognition domain that provides DNA specificity for both activities. It is highly similar to the c-terminal half of the recently characterized restriction enzyme LlaGI (2), from which the annotation of "DNA/RNA helicase" derives; however, this gene lacks the helicase domain found in LlaGI. This gene is highly similar throughout its sequence to a closely related family of enzymes, including DrdV and HauI, which cut specifically at a distance of 11/9 from the adenine that they modify.

2. Smith, R.M., Josephsen, J. and Szczelkun, M.D. (2009) The single polypeptide restriction-modification enzyme LlaGI is a self-contained molecular motor that translocates DNA loops. *Nucleic Acids Res*, **37**, 7219-7230.

3.) Please provide a brief description of proposed experimental procedures. Details should include method of protein purification, the proposed assay, any specific reagents that may exist in the lab that will facilitate experimentation, and an estimate of the time required.

The gene will be cloned and expressed in *E. coli*. Following cellular disruption, the protein will be purified using a Heparin HiTrap column, with elution by a KCl gradient.

Endonuclease activity will be assayed by incubating various amounts of enzyme in reaction buffer containing 1 µg substrate DNA per 50 µl for one hour at 37°C. Reactions will be terminated by addition of stop solution, and reaction products will be analyzed by electrophoresis.

3. Morgan, R.D., Dwinell, E.A., Bhatia, T.K., Lang, E.M. and Luyten, Y.A. (2009) The MmeI family: type II restriction-modification enzymes that employ single-strand modification for host protection. *Nucleic Acids Res*, **37**, 5208-5221.

4.) Please provide a justification of the budget. The most successful bids will be for costs up to $10,000, and the justification may be brief. If unusual circumstances require funds in excess of this, please specifiy in detail the reasons.

The requested funds will pay for supplies associated with cloning, expression and biochemical characterization of the gene. No salary support requested. Experimentation will be performed by a rotating graduate student under the supervision of a lab postdoc.

5.) Please provide up to three Pubmed IDs which demonstrate your lab's experience with the proposed assays.

PMID: 19578066	PMID: 19567736	PMID: 10748211

6.) Please provide any additional information relevant to this proposal.