Developing a transgenic marker to research Huntington's disease in Drosophila melanogaster

Simon Rodney, S. Imarisio

Corresponding author: Simon Rodney (srodney@cantab.net)

Correspondence concerning this article should be addressed to Simon Rodney; Email: srodney@cantab.net

International Journal of Collaborative Research on Internal Medicine & Public Health Vol. 2 No. 6 (June 2010) Pages 199-213

ISSN 1840-4529

http://www.iomcworld.com/ijcrimph/

Paper review summary: Paper submission: February 01, 2010 Paper acceptance: May 25, 2010 Paper publication: May 31, 2010 (CiteAhead Publication)

International Journal of Collaborative Research on Internal Medicine & Public Health

Editors-in-Chief: Asst. Prof. Dr. Jaspreet S. Brar (University of Pittsburgh, USA) Forouzan Bayat Nejad (Tehran University of Medical Sciences, Iran)

Executive Editor: Mostafa Nejati (Universiti Sains Malaysia, Malaysia)

Deputy Editor: Dr. Mensura Kudumovic (University of Sarajevo, Bosnia & Herzegovina)

Associate Editors:

Dr. Monica Gaidhane (Virginia Commonwealth University, USA) Dr. Suresh Vatsyayann (FreeGP, New Zealand)

Developing a transgenic marker to research Huntington's disease in Drosophila melanogaster

Simon Rodney University of Cambridge, UK Email: srodney@cantab.net

S. Imarisio Department of Genetics University of Cambridge, UK

Abstract

Background: Huntington's disease is an autosomal dominant disease where the huntingtin protein is expanded by polyglutamines, increasing its capacity to aggregate. This results in a toxic gain of function of the protein ¹. One of the major ways soluble monomer, oligomers and aggregations can be degraded is by macroautophagy (from here after referred to as autophagy). Furthermore, defective autophagy is involved in the pathology of neurodegenerative diseases ². LC3 is a protein incorporated uniquely into the membrane of an autophagosome, which is the apparatus of autophagy, and is widely accepted as a marker of autophagy. A relatively new assay has been developed by Kimura where the different stages and activity of autophagy can be more closely studied. This system is based on the production of a fusion protein of LC3 tagged with two fluorescent markers, monomeric red fluorescent protein (mRFP) and a green fluorescent protein (EGFP) ³, and has already been used in mammalian cells.

Aims & Objectives: The aim of this project is to use this system in vivo targeting the mRFP-EGFP-rLC3 construct into the Drosophila genome by site specific integration.

Methods/Study Design: We attempted this by placing the LC3 marker into a Drosophila vector Φ C31 pUAST attB. After inserting the marker into the vector the construct was then tested using Drosophila embryonic cell cultures and analysis by fluorescence microscopy.

Results/Findings: When the construct was tested in Drosophila embryonic cells, expression of the two colours red and green were seen, showing that the construct was being expressed correctly. These were the phenotypes of the mRFP and EGFP. The intensity of these colours changed when the levels of autophagy were varied using bafilomycin and rapamycin. However, only a small qualitative difference was seen.

Conclusion: For this construct to be used as a useful marker for autophagy, a robust quantitative analysis still needs to be undertaken. Dose response curves would also produce important data in trying to ascertain what level of calibration is needed to effectively use this transgenic marker.

Keywords: Huntington's. autophagy, transgenesis

Introduction

Autophagy is an intracellular degradation process for worn out organelles or protein aggregates ⁴. It can be activated by nutrient availability, hormones and intracellular pathogens. Autophagy takes place by a phagophore (double membrane structure) surrounding a section of the cytoplasm, containing what is to be degraded, and producing an autophagosome. The autophagosome ultimately fuses with lysosomes containing hydrolytic enzymes and acids, increasing the acidity inside. This results in degradation of the proteins in question (fig 1). A protein which is unique to the phagophore and autophagosome membrane is LC3 (microtubule associated protein 1, light chain 3)⁵. LC3 binds, after post-translational modification, to the inner and outer membrane of the autophagosome making it a good candidate to be used as a marker for autophagy. The marker LC3 is tagged with a red fluorescent protein (mRFP) and a green fluorescent protein (EGFP). The mRFP is stable under acidic conditions; i.e. when a lysosome fuses with the autophagosome and releases its acidic content. Whereas the EGFP is stable at the start of the maturation process but loses its fluorescence under acidic conditions caused when the lysosome fuses with the autophagosome³. A shift in colour from green to red therefore should indicate autophagy progression. These two colours can then be compared to assess autophagy maturation and overall activity in a cell.



Figure 1: Autophagy maturation process. (a, b) Cytosolic proteins and organelles being sequestered into the autophagosome double membrane structure (c), forming a double membrane bound vesicle (d) which fuses with a lysosome containing hydrolytic enzymes resulting in formation of an autolysosome (e) causing degradation of the vesicular contents and an increase in acidity.

In neurodegnerarative diseases, such as Huntington's, neurotoxicity is brought about by expanded glutamine repeats on the huntingtin protein within neurons 1 . There is debate over whether it is the large polyglutamine aggregations which cause the toxicity or whether they are just the end product of many more soluble toxic proteins. Either way autophagy has been shown to break down both types of proteins and so has an important protective function in normal cells². Upregulation of autophagy could also be beneficial in patients with Huntington's disease ⁷. All the research done so far shows the importance of autophagy and the need for further research to thoroughly characterise whole physiologically the process both and pathophysiologically.

Kimura has shown the functioning of the protein marker mRFP-EGFP-rLC3 in mammalian cells and since there have been strong effects of rapamycin in mice Huntington's models, it would be useful to create a system in vivo to assess the different stages of the autophagy process. To do this we wanted to create a Drosophila line able to express the mRFP-EGFP-rLC3 in vivo. To achieve this we decided to use the Φ C31 based integration system. The vector of which can be seen in figure 2 below. This new system of transgenesis uses site specific recombination by the use of the bacterial and plasmid attachment sites to integrate at a specific site into the Drosophila genome ⁸ (fig.2). It has also been adapted to have an endogenously expressed integrase which increases the transgenesis efficiency.



Figure 2: Simplified cartoon demonstrating the makeup of the pUAST attB vector as well as the result of site specific recombination. Notable features are the; white mini gene used as a marker, UAS promotor to control expression only in the presence of GAL4, MCS(multiple cloning site) where our autophagy marker has been inserted, attB and attP sites where recombination takes place.

After the marker has been inserted correctly into the vector Φ C31 pUAST attB the DNA sequence can then be sent for analysis to check for mutations, correct insertion and orientation of the autophagy marker insert. Once this has proved successful the construct can be tested in Drosophila embryonic cells for green/red/yellow expression

of the construct.

Results

We wanted to clone the construct mRFP-EGFP-rLC3 into the pUAST attB vector. The construct maker was originally in a mammalian plasmid. The mRFP-EGFP-rLC3 was excised from the plasmid by double digestion using *EcoR1* and *Nhe1* and inserted into the vector which was cut open by *EcoR1*. The digested mammalian plasmid was then resolved by gel electrophoresis and the mRFP-EGFP-rLC3 was then removed and purified from the gel.

Out of 80 transformed E-Coli we found three inserts of mRFP-EGFP-rLC3 including one insertion of the correct 5' to 3' orientation. A series of DNA digestions were carried out throughout the protocol which were analysed by gel electrophoresis. The vector's original size was approximately 8.5Kb and the insert was approximately 2Kb. Therefore the expected molecular size was 10.5Kb, which can be seen in figure 3 below.

Once the bacterial E-Coli cells were transformed, plated out and lysed as outlined in the methods section below, their DNA needed to be analysed to check for the presence of an insert and afterwards the correct orientation.

The DNA content from different colonies was resolved using gel electrophoresis. The results shown (fig.3) indicate that there are two inserts present, as seen by the slower running larger DNA fragments in the lanes indicated. The difference in weight between the two bands should be approximately 2Kb which is the approximate size of the insert.



Figure 3: DNA gel electrophoresis to resolve the different sizes of the samples obtained. The bands of sample 23 and 24 are running more slowly indicating the presence of an insert.

The orientation of the insert can then be determined by using restriction enzymes to cut both the plasmid and insert at specific sites and then by careful analysis. We used Bg/II, as outlined in the methods below, which cuts once inside the vector and also once in LC3. This would result in two fragments of DNA being produced if the insert was present, including a smaller and larger fragment if the insert was correctly orientated 5' to 3' or more medium sized fragments if there was a 3' to 5' orientation as explained diagrammatically in figure 5 below:



Figure 4: Cartoon representing a DNA digestion by to determine insert orientation.

Figure 5 shows the result of our DNA digest using gel electrophoresis to separate out the fragments. It can be seen that batch 23 contains the correctly inserted fragment while 24 contains the incorrectly inserted fragment.



Figure 5: Results of a gel electrophoresis undertaken to resolve the orientation and presence of an insert when subjected to DNA digestion by BglII. Lane 1 is a DNA ladder used to calculate the absolute weights of the bands produced from the digestion. Lane 2 is a negative control of just the vector without the insert. Lanes 3, 4

and 5 are digestions all of the same bacterial colony but at increasing concentrations as explained later in the methods section. Lanes 6, 7 and 8 represent digestions at another colony at increasing levels of DNA content. Lanes 3, 4 and 5 contain an insert with the correct 5' to 3' orientation whereas lanes 6, 7 and 8 contain an insert but with the incorrect 3' to 5' orientation as explained above and analysed in the discussion later.

After we had ascertained which colony contained the correctly orientated insert we amplified and purified the corresponding DNA, to send it to sequence and eventually for microinjection. We sent the DNA off to be sequenced by making a maxi DNA prep and sending in the relevant primers as outlined in the methods. The complete result of the DNA sequence can be seen in the appendix. Figure 6 below shows a sample of the DNA sequence showing the correct annealing, alignment and no evidence of mutation. The presence and orientation of the construct can also be checked by using BLAST annealing program (www.ncbi.nlm.nih.gov/blast/) as seen in the same figure below.

We confirmed by sequence that the colony number 23 was 5'-3' orientated and there were no mutations.



Reverse DNA sequence

DNA Sequence: ...CGTCATGTCCCTGGTTTATAACTATGGTGCGAATTAATTCGTTAACAGATCTGCGGCCGCGGCGCGCGGGGGGA rLC3: ...CGTCATGTCCCTGGTTTATAACTATGGTGCG

Figure 6: Sample of the forward sequence returned from Geneservice Limited showing complete alignment to mRFP using BLAST as the alignment software. This proves that with this sample the insert has been correctly inserted and with the required 5' to 3' orientation.

The next step was to test for expression of the construct using Drosophila DMel cells, an S2 derived embryonic cell line. The cells were transfected with the pUAST attB vector and the actin-Gal4 driver, to drive the expression of the mRFP-EGFP-rLc3 construct. Before harvesting, cells were treated with different compounds to increase, or block autophagy. In particular, bafilomycin (Baf) and rapamycin (Rap) as outlined in the methods. The cells were then analysed using florescence microscopy. Figure 7 below shows the expression of the EGFP and mRFP in cells bathed in DMSO. As well as testing for general expression of the insert we tried to examine how the insert responds under different conditions. The two extra conditions we used were with bafilomycin and rapamycin. Rapamycin increases the activity of autophagy and autophagosome maturation by inhibiting mTOR. We made the prediction that due to the increased acidity, because of the increased levels of autophagy maturation, the cells should appear more red than green. We also used bafilomycin which inhibits the ATPase of the lysosome causing its acidity to decrease by increasing the diffusion of H^+ ions out of the lysosome. The cells were then analysed looking at the different fluorescence to check for expression of the vector. In parallel we checked if we were able to reproduce in Drosophila cells the data published using mammalian cells³.

The results shown in figure 7 shows that with rapamycin there is a slight increase in the amount of red fluorescence but with bafilomycin some cells showed increased red fluorescence while some showed not much change when compared to the basal levels with just DMSO. However, these results are not conclusive.

© 2010 Rodney S, Imarisio S International Journal of Collaborative Research on Internal Medicine & Public Health Vol. 2 No. 6 (June 2010) pp. 199-213



Figure 7: Photographs taken by fluorescence microscopy displaying the expression of both the mRFP, EGFP and the merging of the two together. Photographs were also taken of the cells when the conditions of autophagy have been manipulated using rapamycin 0.2μ g/ml and bafilomycin 400nM. Key: BAF = Bafilomycin, RAP = Rapamycin

Discussion

The results of the DNA digestions shown above indicated that an insert was present and correctly orientated in sample 23 (fig 4 & 5). On digestion with *BglII* the DNA will be cut just after the insert in the vector and between the EGFP and the rLC3 (fig.4). A prediction can therefore be made that there should be two bands when an insert is present in the vector. Furthermore, while in the 5'-3' orientation the bands produced should be 947 bases and 9951 base pairs in length. This can be seen on the gel with sample 23, with a very fast running band and a very slow running band. When the insert is orientated 3'-5' two fragments of 9475 bases and 1423 bases should be produced as seen by sample 24 (fig.5).

This was then proved by the DNA sequence analysis shown above (fig.6). The sequence from BLAST in figure 6 shows the sequence perfectly aligning at the beginning of mRFP and ending with rLC3. It also shows the kozak sequence just upstream of the mRFP. There is also a poly A tail just downstream of the rLC3 ensuring that the mRFP, the EGFP and the rLC3 will be translated as one unit. This is important for the functioning of the marker as it means the florescent tags will be incorporated with the rLC3 within the autophagosome membrane.

When the construct was tested in Drosophila embryonic cells, expression of the two colours red and green were seen, showing that the construct was being expressed correctly. The construct was also tested to see how it performed under the influence of bafilomycin and rapamycin. Bafilomycin is an ATPase inhibitor causing a loss of H⁺ from the lysosome and ultimately inhibiting autophagosome maturation. In contrast to this, rapamycin inhibits mTOR resulting in activation of autophagy. The application of these two drugs should therefore produce different fluorescence. Both drugs should produce more red due to increasing autophagy activity with Rap and increasing acidic dispersal with Baf. However, on analysis of the photos produced by the confocal microscope as seen in figure 7, no conclusive remarks can be made either way regarding the different levels of autophagy induced by the applied drugs. This is because of multiple reasons. Firstly, not enough magnification was used to view individual puncta, representing individual autophagocytic vesicles. Without this level of magnification the precise localisation of the marker cannot be observed resulting in more noise and less contrast. Secondly, different cells treated with bafilomycin produce conflicting results. Some were more green when compared to just DMSO while some were more red. This is probably due to not enough bafilomycin reaching all of the observed cells and can be explained by dose differences. Thirdly, there is a possibility of saturation of some of the fluorescent makers in some of the observed cells. Fourthly, the EGFP in the cells is being expressed everywhere and especially in the nucleus. This is due to the EGFP becoming incorporated into the nucleus and bringing the rest of the autophagy marker with it. This means that in order to analyse the photos produced, the noise from the nucleus needs to be ignored and discarded. In order to rectify these difficulties and produce conclusive significant results we could repeat the confocal microscopy with an increased magnification to focus in on individual puncta within each cell. Moreover it could be that the chosen cells where highly fluorescence, but actually not that appropriate to analyse the maturation of the autophagic vesicles. A higher magnification would allow for localisation of our marker signal and would be less affected by the noise produced by expression of EGFP from the nucleus. Another possibility could be to create a dose response to the different treatments. It is highly possible that the dosage of drugs used in mammalian cells could not be appropriate for Drosophila cells. A dose response curve should be

organised to try and ascertain what concentration of both rapamycin and bafilomycin are required to produce conclusive results. This could be undertaken by producing many more cell cultures each bathed in increasing concentrations of the drugs. Only once these improvements have been undertaken can the suitability of the mRFP-EGFP-rLC3 marker in Drosophila cells can be realised.

Once the suitability of this marker for autophagy is confirmed, further research can also help to decipher exactly how autophagy is involved in the pathology of Huntington's, especially concerning what type of proteins are degraded and how are they targeted.

Methods and Materials

Vector insertion

The pUAST attB plasmid was kindly supplied by Konrad Basler, while the mammalian plasmid containing the mRFP-EGFP-rLC3 marker insert was kindly supplied by Prof. Yoshimoni. The pUAST attB plasmid was linearised by using EcoRI and the mRFP-EGFP-rLC3 insert was excised by making a double digest of NHE1 and EcoRI as shown in the diagram 9 below. All the restriction enzymes and buffers were supplied by New England BioLabs (NEB) and the manufacturer's protocol was followed throughout. Unfortunately corresponding sticky ends could not be created by using restriction enzymes and so we proceeded to blunt the ends by using T4 DNA pol (NEB). 100 µM dNTPs were added and the solutions were left for 30 minutes at room temperature before adding 1 µl 10mM EDTA to inhibit the T4 DNA pol. The solutions were then heated to 70°C to denature the T4 DNA pol. Before the DNA could be ligated the mRFP-EGFP-rLC3 marker had to be separated from the rest of the digested mammalian plasmid. This was done by resolving via gel electrophoresis (8% agarose) and then extracting the relevant band using QIAGEN MiniElute Gel Extraction kit (catalogue number 28604). The manufacturer's protocol was followed. The insert and vector were then ligated using T4 DNA ligase following the manufacturer's protocol. As explained above in the results, a ratio of 3:1 insert to plasmid was used to ensure efficient ligation between the insert and vector. Before a ligation could be carried out the relative quantities of the insert and vector had to be ascertained. This was done by running a selection of both on a gel and comparing the relative intensities of the band produced. It was estimated that the band for the insert is $10ng/\mu$. This knowledge enabled us to proceed with a ligation in the ratio of 3:1 knowing that 10 µl of the insert contained 100 ng of DNA. Due to the insert being a third of the length of the vector, 100ng of each provided the 3:1 ratio we were looking for. Buffer 2 (NEB) was used for both NheI, EcoRI and T4 DNA pol while T4 DNA ligase buffer was supplied together with T4 DNA ligase by NEB.



Figure 9: Diagram indicating the enzymes used to linearise the $\Phi C31$ vector, cut out the insert from the mammalian plasmid, blunt all the ends and ligate the insert into the vector.

DNA Transformation

5 μ l of the ligation was then added to 50 μ l of super competent DH5 E-Coli and left to thaw for 30 minutes. The bacteria were then heat shocked to open up pores on the surface membrane to facilitate the transformation for ten minutes. They were then placed in ice for two minutes before being placed in nutrient broth for one hour at 37°C. The bacteria were then plated out on an agar medium infused with 100 μ g/ml ampicilin. They were then left to form colonies overnight. A colony was then taken by pipette and added to 2ml of broth which was then left to culture at 37°C.

DNA Mini Prep

Half, i.e. 1 ml, of each colony in broth was then microcentrifuged to return the bacteria, and the supernatant as discarded. The E-Coli were then lysed using the QIAprep Spin Miniprep kit from QIAGEN (catalogue number 27106) for the first 40 cultures. For the remaining 40, a mini DNA prep was undertaken using a method produced by Zhou¹⁰.

Insertion Analysis

The presence of an insertion was detected by running on an 0.8% agarose gel by gel electrophoresis. 0.8% was chosen to resolve the large fragments with enough resolution and in a small enough time frame. Vectors containing an insert were approximately 2Kb heavier and this was detected by a slower moving band on the gel. A DNA ladder was also used to allow quantification of the weights of the different bands of DNA separated by the electrophoresis. The ladder was purchased from (NEB)

The orientation was elucidated by digesting the insert using *BglII* (NEB) with buffer 3 (NEB), following the manufacturer's protocol, and analysed as explained in the results, discussion and diagram 4 above.

DNA Sequencing

The DNA was then prepared to be sent for sequencing, to Geneservice Limited, by undertaking a Maxi DNA Prep. This was undertaken by using a QIAGEN Plasmid Maxi kit (catalogue number 12662) and the protocol was followed as per the manufacturer's guidelines. Relevant primers were also sent which primed into the multiple cloning site of the vector pUAST attB a forward primer pUASTF (5'-GACTCTGATAGGGAATTG-3') and reverse primer Puastr а (5'-AATACACAAACATTATAC-3') (MWG-Biotech). The following programs were used in the analysis of the DNA sequence; BLAST (www.ncbi.nlm.nih.gov/blast/), Workbench (http://workbench.sdsc.edu/) Chromas and (http://www.technelysium.com.au/chromas.html).

Testing in Cell Culture

The pUAST attB construct containing the insert was cotransfected with actin Gal4 into Drosophila D.Mel-2 S2 embryonic cells. 9 wells were used; 3 to be used for A Western blot and 6 to be used for florescence microscopy. 3 Million Drosophila cells were in each well. Cellfectin from Invitrogen was used to help form micelles with DNA to attach to the cell membranes. These were then left to incubate at 37° C overnight. Before harvesting the cells to be used for fluorescence microscopy were treated with either Dimethyl Sulfoxide (DMSO), a final concentration of 0.2μ g/ml rapamycin, or a final concentration of 400nM bafilomycin. The cells were then rinsed with PBS and mounted using a drop of VECTASHIELD (Vector), which is a mounting medium with DAPI to stain nuclei.

Fluorescence Microscopy

The prepared slides were visualised using a camera equipped fluorescent confocal microscope (Nikon Eclipse E800) at 40x magnification. Red, green and yellow colour analysis was performed using ImageJ software (http://rsb.info.nih.gov/ij/).

References

- 1. Rubinsztein DC. The roles of intracellular protein-degradation pathways in neurodegeneration. Nature. Oct 19 2006;443(7113):780-786.
- 2. Ravikumar B, Duden R, Rubinsztein DC. Aggregate-prone proteins with polyglutamine and polyalanine expansions are degraded by autophagy. Human Molecular Genetics. 2002;11(9):1107-1117.
- 3. Kimura S, Noda T, Yoshimori T. Dissection of the autophagosome maturation process by a novel reporter protein, tandem fluorescent-tagged LC3. Autophagy. 2007;3(5):452-460.
- 4. Yorimitsu T, Klionsky DJ. Autophagy: molecular machinery for self-eating. Cell Death and Differentiation. 2005;12:1542-1552.
- 5. Kabeya Y, Mizushima N, Uero T, et al. LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing. Embo Journal. 2000;19(21):5720-5728.
- 6. Xie ZP, Klionsky DJ. Autophagosome formation: Core machinery and adaptations. Nature Cell Biology. 2007;9:1102-1109.
- 7. Rubinsztein DC, Gestwicki JE, Murphy LO, Klionsky DJ. Potential therapeutic applications of autophagy. Nat Rev Drug Discov. Apr 2007;6(4):304-312.
- 8. Groth AC, Fish M, Nusse R, Calos MP. Construction of transgenic Drosophila by using the site-specific integrase from phage phi C31. Genetics. 2004;166(4):1775-1782.
- 9. Bischof J, Maeda RK, Hediger M, Karch F, Basler K. An optimized transgenesis system for Drosophila using germ-line-specific phi C31 integrases. Proceedings of the National Academy of Sciences of the United States of America. 2007;104(9):3312-3317.
- 10. Zhou C, Yang YJ, Jong AY. MINI-PREP IN 10 MINUTES. Biotechniques. 1990;8(2):172-173.

Appendix

DNA sequenced from sample 23:

Complete sequence produced by forward primer:

Ν

Complete sequence produced by reverse primer:

TANAATACACAAACAATTAGAATCAGTAGTTTAACACATTATACACTTAAAAAATTTTATATTTACCTTAG AGCTTTAAATCTCTGTAGGTAGTTTGTCCAATTATGTCACACCACAGAAGTAAGGTTCCTTCACAAAGAT AGGGACATGACGACGTACACAACCCACACGGCAGCAGTGGGGGACTTACACAGTGTAACATAACATTGG GGGTTAGCAGGGACAGCTGCATGCAGGGAGGGTGGGTGCTGTGGGGGGGCCAGGTCGAGAGGTGGCTGTGT GTGCTGTCCCGAACGTCTCCTGGGAGGCATAGACCATGTACAGGAAGCCGTCTTCATCTCTCGCTCTC GTACACTTCAGAGATGGGTGTGGACACACTCACCATGCTGTGCCCATTCACCAGGAGGAAGAAGGCTTGG TTAGCATTGAGCTGCAGGCGCCTTCTAATTATCTTGATGAGTTCGCTCATATTCACGTGATCAGGTACAA GGAACTTGGTCTTGTCCAGGACGGGCAGCTGCTTCTCACCCCTTGTATCGCTCTATAATCACTGGGATCTT GGTGGGGTGCTGCTCCCGGATGAGCCGGACATCTTCCACTCTTTGTTCGAAGCTCCGGCGCTGTTTGAAA GGTCTTCTCGGACGGTCTAGATCTGAGTNCCGGACTTGTACAGCTCGTCCATGCCCNAGAGTGATCCCNG CGGCGGTCACGAACTNCAGCNGACCATGTGATCGCGCTTCTCGTNGGGGTCTTTGCTCNNGNGNACTGGG NNNNNTCNGNTN