Open Access

Black Box Chimera Check (B2C2): a Windows-Based Software for Batch Depletion of Chimeras from Bacterial 16S rRNA Gene Datasets

Viktoria Gontcharova[1,2], Eunseog Youn[4], Randall D. Wolcott[1,2,3], Emily B. Hollister[5], Terry J. Gentry[5] and Scot E. Dowd[1,2,*]

[1]*Research and Testing Laboratory of the South Plains, Lubbock, TX, USA 79407*

[2]*Medical Biofilm Research Institute, Lubbock, TX, USA 79407*

[3]*Southwest Regional Wound Care Clinic, Lubbock, TX, USA 79410*

[4]*Computer Science Department, Texas Tech University, Lubbock, TX, USA 79409-3104*

[5]*Department of Soil and Crop Sciences, Texas A&M University, College Station, TX, USA 77843-2474*

Abstract: The existing chimera detection programs are not specifically designed for "next generation" sequence data. Technologies like Roche 454 FLX and Titanium have been adapted over the past years especially with the introduction of bacterial tag-encoded FLX/Titanium amplicon pyrosequencing methodologies to produce over one million 250-600 bp 16S rRNA gene reads that need to be depleted of chimeras prior to downstream analysis. Meeting the needs of basic scientists who are venturing into high-throughput microbial diversity studies such as those based upon pyrosequencing and specifically providing a solution for Windows users, the B2C2 software is designed to be able to accept files containing large multi-FASTA formatted sequences and screen for possible chimeras in a high throughput fashion. The graphical user interface (GUI) is also able to batch process multiple files. When compared to popular chimera screening software the B2C2 performed as well or better while dramatically decreasing the amount of time required generating and screening results. Even average computer users are able to interact with the Windows .Net GUI-based application and define the stringency to which the analysis should be done. B2C2 may be downloaded from http://www.researchandtesting.com/B2C2.

Keywords: Chimera, chimera detection, high throughput, pyrosequencing, Windows, next generation.

INTRODUCTION

PCR techniques, although commonly used and arguably the most powerful and accurate means of characterizing bacterial diversity also generate chimeric sequences [1]. A chimera is formed during PCR amplification if sequence synthesis which first starts at one template, is interrupted for any reason, and continues synthesis along another template. The focus of chimera detection is now on development of high throughput methods which still maintain a degree of efficiency and accuracy while allowing for flexible user input. Next generation pyrosequencing technologies, like the Roche 454, generate over 1 million high quality reads per run, between 250-600 basepairs (bp) on average using PCR principles. This can complicate chimera detection even further by amplifying the errors occurring in the process.

Pyrosequencing is a rapid sequencing process following the synthesis principle that has been made feasible by approaches like 454 Sequencing, a system for massively-parallel pyrosequencing [2]. Amplicon based 454 sequencing techniques promise the potential for previously unrealized resolution into highly diverse environments. Bacterial tag-encoded FLX Amplicon Pyrosequencing (bTEFAP) is a newly described method for evaluating microbial diversity within complex environments [3-5], based upon the 16S rRNA gene or any other gene. One of the necessary improvements in the analysis of data generated by this powerful technique and other approaches to analysis of pyrosequencing results [6, 7] is the ability to detect and eliminate chimeras.

Chimeras pose serious problems in community analysis by artificially increasing diversity. The current and most powerful chimera detection algorithms require computation of multiple sequence alignment and distance matrices, which is computationally and time intensive [8-11]. In addition, many of the existing algorithms are optimal for detection of chimeras in full length sequences (1300+ bp) [8, 9]. We introduce a standalone Windows. Net based application which can be used for high throughput, high volume screening and batch processing of 16S rRNA gene amplicon libraries. By utilizing a GUI format in the Windows environment, we provide an application which can be utilized by a wide variety of microbial ecologists who may not have expertise in programming or the Linux/Unix based environments, which are more common to bioinformaticians.

*Address correspondence to this author at the Research and Testing Laboratory of the South Plains, Lubbock, TX, USA 79407; Tel: 806-771-1134; Fax: 806-771-1168; E-mail:sdowd@pathogenresearch.org

1874-2858/10 **2010 Bentham Open**

MATERIALS AND METHODOLOGY

With a demand for a method to screen for chimeras we built a GUI-based high throughput implementation. The software is designed for a non-bioinformatician scientist and could be used for variable stringencies of chimera checking. The software is built within the. NET Windows platform which allows scientists to batch process sequence files from a wide range of projects, including short read data.

Characteristics of the B2C2 Software

Multiple files within a folder may be selected for analysis during the same process. The method extracts 100 bp long contiguous start and end regions of each FASTA format input sequence. These regions are then aligned against a custom BLAST database containing 7,199 high quality sequences consisting of a representative, verified set of bacteria with full taxonomic information (at all 7 major levels). Each species of bacteria was only included once to minimize search space. When sequences are aligned against this database, taxonomic information of each of the regions is evaluated and the percentage of matching taxonomic levels is calculated from species back to the kingdom level. The higher the taxonomic identity between both regions, the less likely is the sequence to be chimeric. In other words, if the start and end regions align to the same species, the identity is 100%, or 7 out of 7 possible levels. This evaluation considers all levels of taxonomy since sequences belonging to the same species belong to the same genus, family, order, class, phylum and kingdom. Similarly, if the start and end regions align to sequences from the same genus, but differing species, there are 6 of 7 matching levels- including kingdom, phylum, class, order, family and genus.

The stringency of the analysis is user defined, thus the accuracy is able to be customized for the strictness desired and the task at hand. The database against which the input sequences are aligned has been designed to contain all the necessary information for this type of analysis based on the seven major taxonomic levels: kingdom, phylum, class, order, family, genus, and species. Each sequence in the database contains this information and during the analysis, the number of matching taxa are assessed. For each region of each sequence, top 8 best results from a BLAST, specifically blastn, alignment are considered. The best match from the start and end regions are used in the calculation of the result.

B2C2 Output

The output is organized into four FASTA files based on user defined parameters (for definite, possible, non-chimeric and unclassified sequences). The sequence names in the files are followed by a number of matching phylogenetic levels between the start and end regions of the sequence in question, neither the potential parents nor the break-points of a potential chimera are reported. These files are designated with the file name followed by "DefiniteChimera", "PossibleChimera" and "NotChimera". In addition, a fourth file designated with "NotClassified" is created to contain sequences for which a problem occurred, such as no hits were found for one or both end regions. Because as noted for all other chimera detection algorithms, there is no guaranteed way to computationally determine if the sequences are chimeric, the files are provided for the user to determine the next appropriate step. However, the application also contains the option for the user to remove chimeric and possibly chimeric sequences from the submitted files to generate a "Final" FASTA file. This file can be used for downstream analysis applications.

Testing B2C2

To ensure that this application performed as well or better than current chimera check implementations we did a comparative study. A test set of 300 chimeric sequences was generated from randomly selected bacterial sequences. Sequences of 600 bp were formed from two parental strands snipped in the 200 to 400 range and re-concatenated to a different sequence in the 200-600 range. This process was applied to simulate 100 chimeric sequences with parents varying at the genus or species level (close relationship), 100 chimeric sequences varying on order or family levels, and 100 distantly related chimeric sequences varying at the phylum or class levels. The 600 bp sequences emulated the short reads resulting from "next generation" technologies. A fourth file, consisting of verified, non-chimeric, parent sequences was also compiled to test the ability of B2C2 to correctly classify non-chimeras. These "parents" are existing sequences in commonly used databases such as NCBI or RDP, not only present to demonstrate the ability of the engines to correctly identify non-chimeras, but also demonstrate how the engines work on potentially randomly extracted sequences that are often referenced and used in research.

Testing was performed not only on B2C2, but also on the Greengenes server, using the "Chimera check with Bellerophon (version 3)" [8, 9] and Chimera_Check from RDPII [11] for performance comparison. Other chimera checking packages are available, such as the Pintail and Mallard programs from the Cardiff School of Biosciences [12]. These applications were not selected as comparison tools for the performance of B2C2 due to preliminary pre-processing steps. These steps are not easily amenable for next-generation scale data. Sequences have to be pre-aligned by ClustalW, which not only adds a step in the processing, but increases processing time. The authors of these programs also suggest not including more than 1000 reliable (non-chimeric) sequences in the analysis due to speed concerns, encouraging the files to not exceed 500 sequences and excluding distantly related sequences. Furthermore, the user must supply the engine with not only the query sequences, but the subjects against which to align as well. This is a potentially time consuming, intricate step that not every user is willing or able to perform. Because next generation products may result in over a million reads, the suggested pre-screening processes and establishing an essential reliable set would greatly increase the processing time of the data. Mallard, Pintail and other existing chimera checking tools, although reliable and highly functional, are not the optimal tools for rapid and efficient screening of high volume datasets such as those resulting from 454 FLX and Titanium technologies (Roche Applied Science).

RESULTS

The application performed well on a test set of 300 short FASTA formatted sequences, similar to a set resulting from next-generation processing. The default parameters for B2C2 were used. These parameters were established to allow for

flexibility in the decision making process of the algorithm. Performance results are shown in Table **1a** and summarized in **1b**. Non-chimeric, essentially randomly chosen existing sequences from commonly used databases such as NCBI were also tested (last column in the table under heading "Non Chimeras"). Results demonstrate the ability of B2C2 to properly identify non-chimeric sequences. Data is organized in a table formatted to contain the prediction in the first column and the subset of test cases in the first row. The results indicate that from the distant chimeras, B2C2 accurately categorized 98% of the sequence as chimeric. Another 1% of the sequences were organized into the "Possibly Chimeric" category, generally intended for further consideration by the user. The last chimeric sequence was wrongfully characterized as chimeric. Although the performance of B2C2, like most other chimera detection algorithms, appears to become more ambiguous with the increased homology of the two ends of the simulated chimeric sequences, the engine still performed well, according to Tables **1a** and **1b**. From Table **1b**, it is easily interpreted that not only are chimeras accurately declared, but non chimeras are also detected as well, indicating a low rate of false positives.

The same 300 chimeric sequences and 100 non-chimeric sequences were submitted online to the Chimera_Check from RDPII engine. The results returned consisted of images needing to be manually, subjectively, individually and visually analyzed. Based on slopes and the designated "breaking point" in the sequences, the queries were organized into three categories of "Chimeras", "Maybe Chimeras", and "Not Chimeras". If the image consisted of a result where a clear positive slope continued until the breaking point, followed by a negative slope, the sequence was tallied into the "Chimeras" category. If the positive slope was followed by a slope of 0 and then a negative slope, the sequence was organized into the "Maybe Chimeras" section. Sequences were organized into non-chimeras if the engine returned the statement of "There was no way of breaking your sequence in two chimeric halves, so their combined match with the database was better than that of the full length sequence" or if the slope was very small or the break appeared to be close to the start or end of the sequence.

Chimera_Check from RDPII performed well on the test set and the results can be viewed in Tables **2a** and **2b**. The results indicate RDPII to be an excellent chimera detection tool for distant, medium and close chimeras. Although according to the resulting statistics, RDPII outperformed B2C2, it must be emphasized that inspection of the output is subjective and B2C2 did not asses the data with the strictest parameters possible (chimeras from sequences differing at the species level were deemed non-chimeric). These parameters may be adjusted by the user for desired stringency. Furthermore, RDPII performed worse than B2C2 on the accuracy of non-chimeric sequence detection. This tool identified only 54% of the sequence correctly, while B2C2 correctly identified 89%.

The next comparison tool, Greengenes' "Chimera check with Bellerophon (version 3)" was not able to properly evaluate the test set; the sequences came back with a "not tested" categorization. Greengenes' algorithm is specialized to evaluate full length bacterial sequences, thus not appropriate for analysis of next generation data, which contains short reads averaging 400 bp.

The platform was further tested with a new set of 300 full length sequences. The full length sequences were generated from known bacterial parents to simulate full length chimeric strands. Full length parents were used as non-chimeric sequences. Default parameters were used for the analysis. The results were scored again into the same categories: "Chimeras", "Maybe Chimeras" and "Not Chimeras". The additional "Problems" category was also used during this analysis for sequences not returning a result. From Greengenes' results,

Table 1a. Results of test set Implementation on B2C2. The Software Classified Chimeras from Three sets, Distant Chimeras, Medium Chimeras, and Close Chimeras, into Four Resultant files: "NotChimera", "PossibleChimera", "DefiniteChimera", and a "NotClassified" File

Category	Distant Chimeras	Medium Chimeras	Close Chimeras	Not Chimeras
Problems	0	0	0	0
Not Chimeras	1	11	35	89
Maybe Chimeras	1	11	4	3
Chimeras	98	78	61	8
Totals	100	100	100	100

Table 1b. Summary of Results of Test Set Implementation on B2C2. The Software Classified Chimeric and Non-Chimeric Sequences. Three Possible Classification Categories Resulted: Chimera, Possible Chimera and Non-Chimera

		Actual	
		Chimeras	**Non-Chimeras**
Predicted	**Chimeras**	237	8
	Possible Chimeras	16	3
	Non-Chimeras	47	89

"putative chimera" sequences were scored as "Chimeras", "Appears to be clean" and sequences matching the Core Set were scored as "Not Chimera", "A sub-threshold chimera" were counted as "Maybe Chimeras", and "not tested" sequences were put into the "Problems" category. Greengenes' performance noticeably increased by lengthening the short reads to their full length and the results are provided in Tables **3a** and **3b**. Greengenes appears to take the more conservative approach in chimeric detection, a very small margin of chimeric sequences was actually declared to be chimeric. However, 62% of non chimeric sequences were deemed as such. This percentage is still not as good at the performance seen with B2C2, but it is improved from the 54% seen with the RDPII chimera checker.

Table 2a. Results of Test Set Implementation on Chimera_Check for RDPII. The Software Classified Chimeras from Three Sets, Distant Chimeras, Medium Chimeras, and Close Chimeras, into Four Resultant Files: Not Chimeras, Maybe Chimeras, Definite Chimeras, and a Problem File

Category	Distant Chimeras	Medium Chimeras	Close Chimeras	Not Chimeras
Problems	0	0	0	0
Not Chimeras	0	2	5	54
Maybe Chimeras	0	2	20	20
Chimeras	100	96	75	26
Totals	100	100	100	100

Table 2b. Summary of Results of Test Set Implementation on Chimera_Check for RDPII. The Software Classified Chimeric and Non-Chimeric Sequences. Three Possible Classification Categories Resulted: Chimera, Possible Chimera and Non-Chimera

		Actual	
		Chimeras	**Non-Chimeras**
Predicted	**Chimeras**	271	26
	Possible Chimeras	22	20
	Non-Chimeras	7	54

Table 3a. Results of Second Test Set (Composed of Longer Sequences) Implementation on the Greengenes Server. The Software Classified Chimeras from Three Sets, Distant Chimeras, Medium Chimeras, and Close Chimeras, into Four Resultant Files: Not Chimeras, Maybe Chimeras, Definite Chimeras, and a Problem File

Category	Distant Chimeras	Medium Chimeras	Close Chimeras	Not Chimeras
Problems	12	10	5	0
Not Chimeras	69	67	82	62
Maybe Chimeras	14	14	7	22
Chimeras	5	9	6	16
Totals	100	100	100	100

Table 3b. Summary of Results of Second Test Set (Composed of Longer Sequences) Implementation on the Greengenes Server. The Software Classified Chimeric and Non-Chimeric Sequences. Three Possible Classification Categories Resulted: Chimera, Possible Chimera and Non-Chimera

		Actual	
		Chimeras	**Non-Chimeras**
Predicted	**Chimeras**	20	16
	Possible Chimeras	35	22
	Non-Chimeras	218	62

The same long sequences were also analyzed with the B2C2 software. Our goal for the software was to develop a fast, reliable and easy to use process for our short read, high volume output. We were not able to find a suitable automated method in the past, thus the only solution was to create our own. However, the software is not limited to next generation output, and B2C2 can be used on all FASTA formatted bacterial 16S rRNA gene sequences. The software performed well with the new full length test set (averaging 1300 bp) and the default parameters. The results are provided in Tables **4a** and **4b**. The results demonstrate significantly better performance when compared with Greenegenes. Close chimeras showed to be the greatest of difficulty, although still produced good results with 71% of the sequences being declared chimeric or possibly chimeric compared to 13% as shown by Greenegenes. A greater percentage of non chimeric sequences was also correctly identified by B2C2 than Greengenes, 88% as opposed to 62%.

DISCUSSION

All tests of B2C2 were performed using default conditions to simulate an easy and comfortable situation for a scientist assessing data. Identity of less than 60% in the start and end regions of the query sequences were organized into the "DefiniteChimera" file, identities greater than 60% and less than 85% were distributed into the "PossibleChimera" file and those sequences with an identity of greater than 85% were put into the "NotChimera" file. These percentages allow for reliable results without removing all room for error. The upper bound of 85% allows for sequences where the ends don't differ or only differ at the species level to be deemed non-chimeric.

Chimera detection is not an exact science, thus we felt it was not appropriate to default a parameter to the highest strictness possible. Due to this criteria, chimeric sequences where the start and end regions are from the same genus will be deemed non chimeric. However, as previously mentioned, the user is able to adjust setting according to his/her specifications and raising the upper bound to guarantee the start and end regions aligning to the same species.

We feel it is important mentioning that it was difficult to establish a precise scoring mechanism for RDPII, due to a needed visual judgement call on the user's part, thus making a non uniform evaluation based on the resulting graphs. During this analysis, we gave RDPII a favorable judgement when results were unclear. Furthermore, because manual inspection of each of the submitted sequences has to be performed, the software is not an appropriate tool for analysis of high throughput data containing millions of individual reads, particularly because the user will then need to return to the original data to find and remove the chimeric sequences. It would take a considerable amount of time to not only submit sequences and wait for the results, but to individually assess the results and decide the fate of each sequence from a graph.

The previously existing software was able to assess the datasets tailored to the engines, however, both Greengenes' and RDPII's algorithms have the drawback of slow turnaround times (even on the small test sets of 400 sequences) – taking up to several hours per run. Greengenes' limitation on the number of input sequences per analysis run and the manual analysis necessary for Chimera_Check from RDPII results make the engines inconvenient for high volume bacterial sequences analysis. Also, the overall performance of B2C2 is competitive with both engines used for comparison. Both Chimera_Check and Greengenes did not perform as well as B2C2 in classifying non-chimeric sequences, and the

Table 4a. Results of Test Set Implementation on B2C2 Using Full Length Bacterial Sequences. The Software Classified Chimeras from Three Sets, Distant Chimeras, Medium Chimeras, and Close Chimeras, into Four Resultant Files: NotChimera, Possible Chimera, DefiniteChimera, and a NotClassified File

Category	Distant Chimeras	Medium Chimeras	Close Chimeras	Not Chimeras
Problems	0	0	0	0
Not Chimeras	0	9	29	88
Maybe Chimeras	1	17	7	7
Chimeras	99	74	64	5
Totals	100	100	100	100

Table 4b. Summary of Results of Test Set Implementation on B2C2 Using Full Length Bacterial Sequences. The Software Classified Chimeric and Non-Chimeric Sequences. Three Possible Classification Categories Resulted: Chimera, Possible Chimera and Non-Chimera

		Actual	
		Chimeras	**Non-Chimeras**
Predicted	**Chimeras**	237	5
	Possible Chimeras	25	7
	Non-Chimeras	38	88

accuracy on chimeric sequences is close with RDPII and noticeably better than Greengenes.

B2C2 has the ability to perform well in a speedy fashion implementing a methodology that has not been made publically available. The process follows a logical and simplified method to attain results comparable to those from the existing popular methods. The advantage of the new application is not only in performance, but also processing time, ease of use and less limitations on file and sequence size.

CONCLUSIONS

B2C2 is the next step in addressing the need for a useful and efficient tool applicable for chimera detection in next generation 16S rRNA-based output such as that generated by the bTEFAP method. The software currently available for traditional chimera checking, although precise, accurate, and very useful for appropriate data, are not an efficient solution for "next generation" output. B2C2 is a standalone, user-friendly application, which can be downloaded from http://www.researchandtesting.com/B2C2 and run by anyone for bacterial 16S rRNA gene sequence data. Currently, this software is used in our lab to deplete chimeras from pyrosequencing results as a part of a processing pipeline [13-16]. B2C2 can be incorporated into a pipeline as a preprocessing step from which chimera depleted files may be further analyzed in an appropriate fashion. As with all chimera detection algorithms the precision of the results tends to decline as the parents of chimeric sequences get phylogenetically more closely related. To increase the stringency, even to 100%, the user can adjust the settings and determine the optimal set of parameters for their specific data and need. This type of platform is very useful, in accuracy, flexibility and simplicity, for labs and individuals conducting research on 16S rRNA gene sequences and dealing with large numbers of reads. B2C2 is an efficient, self-explanatory application capable of handling chimera evaluation on an unprecedented scale.

ACKNOWLEDGEMENTS

Funding: This work was supported by internal development funds of Medical Biofilm Research Institute and the Research and Testing Laboratory of the South Plains.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publishers Web site along with the published article.

REFERENCES

[1] Shuldiner AR, Nirula A, Roth J. Hybrid DNA artifact from PCR of closely related target sequences. Nucleic Acids Res 1989; 17: 4409.

[2] Rothberg JM, Leamon JH. The development and impact of 454 sequencing. Nat Biotechnol 2008; 26: 1117-24.

[3] Dowd SE, Wolcott RD, Sun Y, McKeehan T, Smith E, Rhoads D. Polymicrobial nature of chronic diabetic foot ulcer biofilm infections determined using bacterial tag encoded FLX amplicon pyrosequencing (bTEFAP). PLoS One 2008; 3: e3326.

[4] Dowd SE, Sun Y, Wolcott RD, Domingo A, Carroll JA. Bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP) for microbiome studies: bacterial diversity in the ileum of newly weaned Salmonella-infected pigs. Foodborne Pathog Dis 2008; 5 :459-72.

[5] Dowd SE, Callaway TR, Wolcott RD, *et al.* Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). BMC Microbiol 2008; 8: 125.

[6] Dethlefsen L, Huse S, Sogin ML, Relman DA. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. PLoS Biol 2008; 6: e280.

[7] Sogin ML, Morrison HG, Huber JA, *et al.* Microbial diversity in the deep sea and the underexplored "rare biosphere". Proc Natl Acad Sci USA 2006;103:12115-20.

[8] DeSantis TZ, Hugenholtz P, Larsen N, *et al*. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 2006; 72: 5069-72.

[9] DeSantis TZ, Jr., Hugenholtz P, Keller K, *et al.* NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. Nucleic Acids Res 2006; 34: W394-W9.

[10] Huber T, Faulkner G, Hugenholtz P. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. Bioinformatics 2004; 20: 2317-9.

[11] Cole JR, Chai B, Marsh TL, *et al.* The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. Nucleic Acids Res 2003; 31: 442-3.

[12] Ashelford KE, Chuzhanova NA, Fry JC, Jones AJ, Weightman AJ. At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. Appl Environ Microbiol 2005; 71: 7724-36.

[13] Wolcott RD, Gontcharova V, Sun Y, Zischakau A, Dowd SE. Bacterial diversity in surgical site infections: not just aerobic cocci any more. J Wound Care 2009; 18: 317-23.

[14] Wolcott RD, Gontcharova V, Sun Y, Dowd SE. Evaluation of the bacterial diversity among and within individual venous leg ulcers using bacterial tag-encoded FLX and titanium amplicon pyrosequencing and metagenomic approaches. BMC Microbiol 2009; 9: 226.

[15] Pitta DW, Pinchak WE, Dowd SE, *et al*. Rumen bacterial diversity dynamics associated with changing from bermudagrass hay to grazed winter wheat diets. Microb Ecol 2010; 59(3): 511-22..

[16] Suchodolski JS, Dowd SE, Westermarck E, *et al.* The effect of the macrolide antibiotic tylosin on microbial diversity in the canine small intestine as demonstrated by massive parallel 16S rRNA gene sequencing. BMC Microbiol 2009; 9: 210.

Received: June 18, 2010 Revised: June 25, 2010 Accepted: June 30, 2010

© Gontcharova *et al.*; Licensee *Bentham Open.*

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.