

Research article

Open Access

Analysis of 16S rRNA gene sequences and circulating cell-free DNA from plasma of chronic fatigue syndrome and non-fatigued subjects

Suzanne D Vernon*¹, Sanjay K Shukla², Jennifer Conradt², Elizabeth R Unger¹ and William C Reeves¹

Address: ¹Division of Viral and Rickettsial Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA 30333, USA and ²Clinical Research Center, Marshfield Medical Research Foundation, 1000 North Oak Avenue, Marshfield WI 54449-5790, USA

Email: Suzanne D Vernon* - svernon@cdc.gov; Sanjay K Shukla - Shukla.Sanjay@marshfieldclinic.org;

Jennifer Conradt - Conradt.Jennifer@marshfieldclinic.org; Elizabeth R Unger - eru0@cdc.gov; William C Reeves - wcr1@cdc.gov

* Corresponding author

Published: 23 December 2002

Received: 25 October 2002

BMC Microbiology 2002, 2:39

Accepted: 23 December 2002

This article is available from: <http://www.biomedcentral.com/1471-2180/2/39>

© 2002 Vernon et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: The association of an infectious agent with chronic fatigue syndrome (CFS) has been difficult and is further complicated by the lack of a known lesion or diseased tissue. Cell-free plasma DNA could serve as a sentinel of infection and disease occurring throughout the body. This type of systemic sample coupled with broad-range amplification of bacterial sequences was used to determine whether a bacterial pathogen was associated with CFS. Plasma DNA from 34 CFS and 55 non-fatigued subjects was assessed to determine plasma DNA concentration and the presence of bacterial 16S ribosomal DNA (rDNA) sequences.

Results: DNA was isolated from 81 (91%) of 89 plasma samples. The 55 non-fatigued subjects had higher plasma DNA concentrations than those with CFS (average 151 versus 91 ng) and more CFS subjects (6/34, 18%) had no detectable plasma DNA than non-fatigued subjects (2/55, 4%), but these differences were not significant. Bacterial sequences were detected in 23 (26%) of 89. Only 4 (14%) CFS subjects had 16S rDNA sequences amplified from plasma compared with 17 (32%) of the non-fatigued ($P = 0.03$). All but 1 of the 23 16S rDNA amplicon-positive subjects had five or more unique sequences present.

Conclusions: CFS subjects had slightly lower concentrations or no detectable plasma DNA than non-fatigued subjects. There was a diverse array of 16S rDNA sequences in plasma DNA from both CFS and non-fatigued subjects. There were no unique, previously uncharacterized or predominant 16S rDNA sequences in either CFS or non-fatigued subjects.

Background

Chronic fatigue syndrome (CFS) is a complex illness defined by unexplained disabling fatigue and a combination of non-specific accompanying symptoms [1]. There are no consistent anatomic lesions or clinical chemistry abnormalities. While no known infectious agents or immuno-

logic perturbations have been consistently associated with CFS [2–7], the illness has many features suggestive of an infectious disease. Fatigue, muscle and joint pain, sore throat, and swollen glands are all common symptoms shared by infection and CFS. In addition, many people with CFS describe the onset of their illness as sudden or



Figure 1

Titration of *E. coli* template DNA for 16S rDNA PCR assay, using 515F and RD1 primers. Lanes: **L**, 100-bp DNA ladder; **1**, 14 ng; **2**, 1.4 ng; **3**, 0.14 ng; **4**, 1.4×10^{-2} ng; **5**, 1.4×10^{-3} ng; **6**, 1.4×10^{-4} ng; **7**, 1.4×10^{-5} ng; **8**, 1.4×10^{-6} ng; **9**, 1.4×10^{-7} ng; **10**, 1.4×10^{-8} ng; **11**, 1.4×10^{-9} ng; **12**, Negative control (sterile distilled water); **13**, Positive control (purified *E. coli* DNA).

"flu-like", anecdotally suggesting a possible infectious etiology [8]. However, identification of an infectious agent specifically associated with CFS has eluded conventional laboratory analysis. Extensive seroepidemiologic surveys to detect antibody responses to numerous known viral, bacterial, and rickettsial agents have failed to show a difference between CFS cases and normal controls [9,10].

Searching for known or novel infectious agents in persons with CFS is complicated by the lack of a known lesion or diseased tissue to sample. Recently, circulating cell-free DNA from plasma and serum [11] has been shown to contain sequences of tumor [12], viral [13], and bacterial origin [14]. This plasma DNA therefore serves as a sentinel of occult disease occurring in diverse sites throughout the body. We used this plasma DNA to search for previously uncharacterized as well as known bacterial pathogens. To do so, we used broad range PCR of the bacterial 16S ribosomal RNA gene (rRNA) [15]. This broad range amplification scheme has been successful in detecting and characterizing bacterial pathogens from several disease states and in several types of clinical specimens [16].

We determined the level of circulating plasma DNA in CFS subjects as one possible indicator of increased cellular turnover or chronic inflammation. We also amplified and sequenced the 16S rRNA gene to search for known and or

previously uncharacterized bacterial agents in cell-free circulating DNA to determine if a bacterial pathogen was associated with CFS.

Results

Evaluation of methods

We first determined whether the bacterial 16S rDNA sequence could be amplified if present in cell-free plasma DNA. To do this, we spiked 100 ng of plasma DNA with 14 ng to 1.4×10^{-9} ng of purified *Escherichia coli* template DNA. As shown in Figure 1, as little as 1.4×10^{-3} ng of *E. coli* DNA could be detected after amplification of the 16S rRNA gene by using the 515F and RD1 primers. This amount is the equivalent of 1.9×10^2 *E. coli* genomes per 100 ng cell-free plasma DNA. All amplified products were sequenced and confirmed to be the same *E. coli* strain (data not shown).

Characterization of plasma DNA

The amount of cell-free DNA ranged from 0 to 1320 ng per ml of plasma (average 128 ng DNA), with 91% (81/89) having detectable levels. CFS subjects tended to have less plasma DNA than non-fatigued subjects (average 91 versus 151 ng), but this difference was not significant. Six (18%) of 34 plasma samples from CFS subjects had no detectable cell-free DNA, whereas only 2 (4%) of 55 non-fatigued subjects had no detectable cell-free DNA in their plasma ($P = 0.08$). No differences in plasma DNA concentration were noted between subjects when grouped by sex, age, CFS onset type, or duration of illness.

All 89 samples, whether plasma DNA was isolated or not, were evaluated for the presence of bacterial 16S rDNA sequences. Overall, 23 (26%) subjects had 16S rDNA sequences amplified and characterized. The average number of distinct 16S rDNA sequences in the 23 subjects was 9 (range 3–14). The CFS subjects had on average 11 distinct 16S rDNA sequences and the non-fatigued subjects had an average of 9 distinct 16S rDNA sequences. The plasma DNA of 4 (14%) of 28 CFS subjects had 16S rDNA sequences compared with 17 (32%) of 53 of non-fatigued subjects ($P = 0.03$). No differences were noted in detection of bacterial 16S rDNA sequences between subjects when grouped by sex, age, CFS onset type, or duration of illness. There was no correlation between the amount of plasma DNA and the ability to amplify 16S rDNA sequences (Figure 2) since these sequences were detected in samples from CFS and non-fatigued subjects with plasma DNA concentrations that ranged from 24 to 294 ng/ml.

All 16S rDNA-amplified products were sequenced to identify the prokaryotic origin. Each of the sequences was either identical or highly similar (97% or higher) to prokaryotic sequences in GenBank (data not shown). To determine whether a particular bacterial sequence was

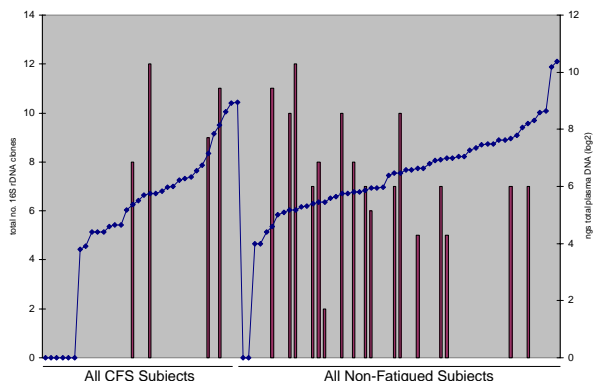


Figure 2

Graphical representation of the number of 16S rDNA sequences in each subject (primary y axis and represented as bars) in relation to the subjects' plasma DNA concentration (secondary y axis and represented as points on the line). All 34 CFS subjects are represented on the left side of the graph and all 55 non-fatigued subjects are shown on the right side. Each group was sorted from lowest to highest plasma DNA concentration to illustrate the lack of correlation between DNA concentration and 16S rDNA sequences.

found in CFS cases versus non-fatigued controls, a cluster analysis was performed. There were no 16S rDNA sequences that were unique or predominant in either the CFS or the non-fatigued group, as indicated by the random distribution and lack of clustering of CFS or non-fatigued subjects (Figure 3). There was also no indication that bacteria known to cause prolonged fatiguing illness (e.g., *Coxiella* sp. or *Borellia* sp.) were more prevalent in CFS subjects.

Discussion

Since CFS has no known anatomic lesion, we decided to examine the levels of cell-free plasma DNA as a systemic indicator of disease. Plasma DNA was isolated from most of the CFS and non-fatigued subjects. We detected a higher concentration of plasma DNA in the non-fatigued subjects than in the CFS subjects and there were fewer non-fatigued subjects who were plasma DNA negative than CFS subjects, however, these differences were not significant. The physiologic significance and the source of cell-free DNA in the plasma are not fully appreciated, but we suspect that it may result from cellular degradation. To date, plasma DNA has been used as a relatively noninvasive sample to detect ongoing pathogenic events, such as minimal residual disease or cancer [10]. Our data show

no significant difference in the level of plasma cell-free DNA in CFS versus non-fatigued subjects, indicating that there may be no unusual cell turnover in this population of CFS subjects. It is also possible that cell-free plasma DNA concentration is not a sensitive indicator for increased cellular turnover or chronic inflammation.

Despite an exhaustive search for known pathogens by conventional laboratory methods, no single pathogen has been consistently identified as a causal agent of CFS. Almost every known viral and bacterial agent that can cause fatiguing illness has been tested for in CFS subjects, and there has been no difference in the prevalence of these agents between CFS and healthy subjects [9,17]. One explanation is that the pathogen associated with CFS is novel or previously uncharacterized. To search for prokaryotic agents that might be specifically associated with CFS, we used consensus PCR primers to the conserved 16S rRNA subunit to detect and characterize these sequences. Of the 89 subjects, 4 CFS subjects and 17 non-fatigued subjects had 16S rDNA sequences amplified. This difference in the presence or absence of the 16S rDNA amplified product between CFS and non-fatigued was not related to differences in plasma DNA concentration. There were no unique or previously uncharacterized prokaryotic sequences identified in either the CFS or non-fatigued group. Rather, a diverse array of known prokaryotic sequences was found circulating in the plasma.

While it is unlikely that the 16S rDNA sequences characterized here are due to experimental or environmental contamination, we cannot exclude the possibility that the vacutainer tubes used for blood collection were a source of bacterial DNA detected. However, the vacutainer tubes were not likely a significant source since all plasma samples, whether DNA was present or not, were subjected to amplification for the 16S ribosomal subunit. All of the plasma DNA-negative samples were negative for the 16S rDNA-amplified product. In addition, water controls taken through the entire extraction process were consistently negative. Finally, only 23 of the 89 plasma samples were positive for 16S rDNA sequences. If our results are a reflection of the occurrence of 16S rDNA sequences in healthy subjects, it is plausible to hypothesize that the presence, rather than absence, of these sequences reflects the normal physiologic state and symbiotic relationship between humans and microbes. This is not the first illustration of the apparent symbiotic relationship that exists between humans and bacteria. The assessment of blood from healthy subjects by amplification of the 16S rRNA gene revealed numerous bacterial sequences that were not found in reagent controls [18]. Weber et al [19] have identified a number of microbial and viral transcripts from human cDNA libraries by computational subtraction method. Not surprisingly, the human body has been referred to as

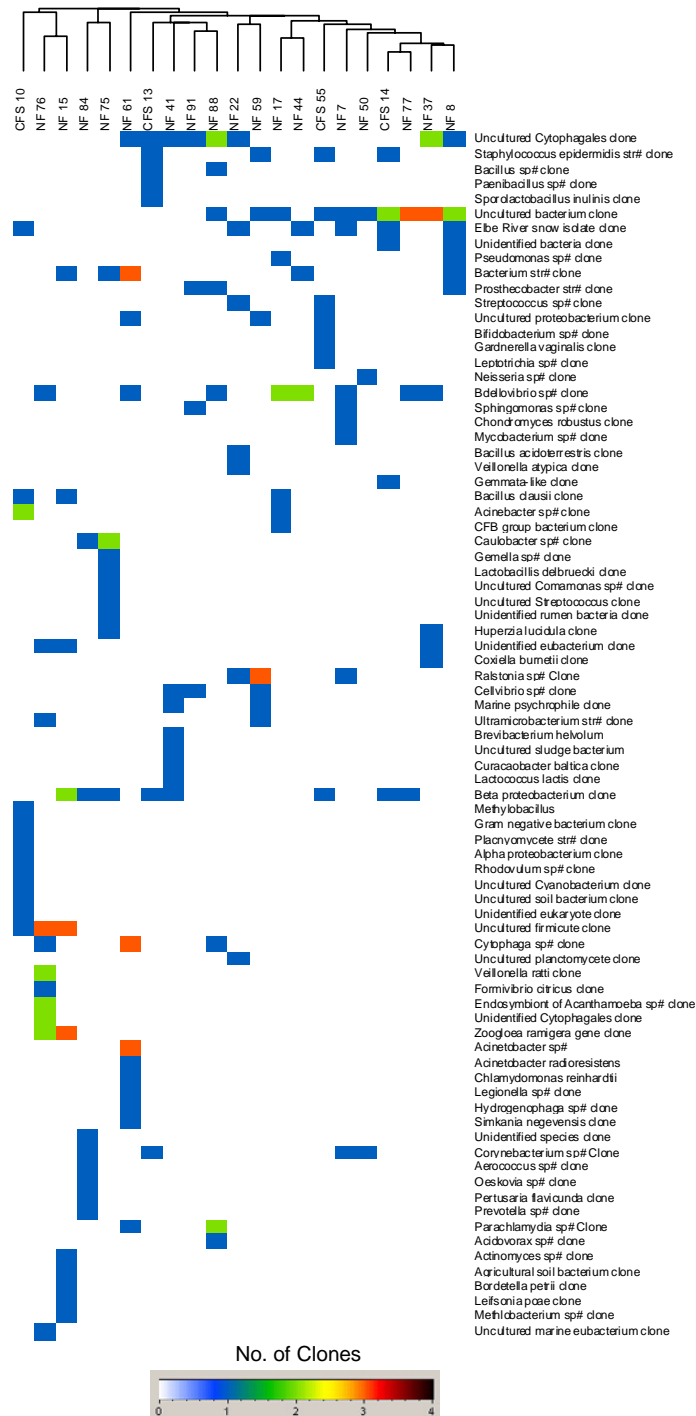


Figure 3
 Cluster analysis of the 23 subjects positive for 16S rDNA sequences from the 300 bp cloned insert to determine whether a particular bacterial sequence was found in CFS cases versus non-fatigued controls. The subjects' classification as non-fatigued (NF) or CFS is shown at the top of the columns. The identification of the 16S rDNA sequence is shown at the right. The colored block indicates the presence and number of clones of that particular bacterial sequence; white is negative, blue is one clone, green is two clones, red is three clones, and black is four clones.

"microbial observatory" [20]. Further analysis of the microbial flora that exists at various sites within the body in both healthy and diseased persons should further our understanding of the interrelationships between microbes and their human hosts.

The experimental design used for this study has some limitations for addressing our hypothesis that a novel pathogen is associated with CFS. The plasma DNA sample may not be the ideal sample for detecting prokaryotic sequences. Granulocyte cell subsets may be more appropriate since this peripheral blood cell fraction contains neutrophils and other scavenger cells important for viral and bacterial clearance. These peripheral blood samples may not all have been collected and processed optimally for preservation of DNA in plasma but all samples from all subjects were processed similarly. Finally, the CFS subjects were years past the onset of illness and may have cleared the agent that provided the trigger for illness.

Conclusions

DNA isolated from the plasma can be used to investigate the association of pathogens with occult disease. Those CFS subjects with plasma DNA had slightly lower concentrations than the non-fatigued subjects and there were more CFS subjects with no detectable DNA in plasma. There was a diverse array of 16S rDNA sequences in plasma DNA from both CFS and non-fatigued subjects. Future assessment of 16S rDNA sequences in peripheral blood will focus on the granulocyte cell subset.

Subjects and methods

Case and control subjects

As part of a longitudinal population-based study of CFS in Wichita, Kansas [21], peripheral blood specimens were collected during clinical evaluation of fatigued subjects identified as having potential CFS ("CFS-like") and a random selection of non-fatigued subjects. Among persons clinically evaluated at baseline, samples from 34 subjects who met the 1994 CFS case definition [1] and 55 non-fatigued subjects with sufficient plasma stored were selected for analysis.

Blood was collected in sodium citrate vacutainer tubes and shipped by overnight courier to the Centers for Disease Control and Prevention (CDC) Serum Bank Facility in Lawrenceville, Georgia. The uncoagulated blood was diluted 1:2 with physiologic saline and separated on Ficoll to collect plasma and mononuclear cells. The diluted plasma was stored in 1-ml aliquots at -70°C until needed.

Plasma DNA isolation

A 1-ml aliquot of plasma was concentrated to approximately 250 µl in a Centricon Centrifugal Filter Device YM-100 (Millipore Corporation, Bedford MA) and DNA was

extracted by using the QIAmp DNA Mini Kit (Qiagen Inc., Valencia, CA), according to manufacturers' instructions. A DyNA Quant 200 Fluorometer (Amersham Biosciences, Inc., Piscataway, NJ) was used to determine DNA concentration.

Polymerase chain reaction (PCR)

A 50-µl amplification reaction consisted of 5 µl of 10 × PCR buffer (100 mM Tris HCl, pH 8.3; 500 mM KCl), 2 mM of MgCl₂, 0.2 mM of dATP, dCTP, dGTP, dTTP, 2.5 U of *Taq* polymerase, 10 pmol each of forward and reverse primers, and 5 µl of the template plasma DNA. Water samples that were taken through the plasma DNA concentration and extraction process were included as samples to identify background bacterial sequences present in reagents and supplies. Two sets of 16S rDNA primers 515F/RD1 and 515F/806R [16,17] were used in separate amplifications. These primers yield amplification products of ~1045 and ~300 bp, respectively. The 515F/806R primers were included to amplify smaller templates and low copy number targets. The PCR was performed in a PE 9700 thermocycler with an initial incubation at 94°C for 4-min, followed by 35 cycles of 94°C for 30 sec, 55°C for 30 sec, and 72°C for 90 sec. A final step consisted of a 5-min extension at 72°C. The amplified products were resolved in 1.5% Nusieve agarose gel and photographed under UV light by using a GelDoc 2000 imaging system (Bio-Rad Laboratories).

16S rDNA sequence determination

The 1045- and 300-bp amplified products were purified by gel exclusion chromatography to remove unincorporated nucleotides and enzymes, and then either sequenced directly or after cloning into a vector. The direct sequencing was done with Cy5-labeled, nested sequencing primers 806R, and 515F in an ALFexpress sequencer (Amersham Pharmacia). All products amplified with the 515F/806R primer set were cloned into a PGEM-T easy vector (Promega Corp., Madison, WI) or into a TOPO TA cloning vector for sequencing (Invitrogen Corp., Carlsbad, CA) by following the manufacturer's protocol. Ninety six clones from each reaction amplified with the 515F/806R primers were selected for further investigation. Each clone was grown overnight in 2.0 ml of LB broth and a plasmid miniprep was prepared. Unique clones were determined by PCR-RFLP of the amplified inserts. The inserts in the plasmid were amplified in a 30 µl PCR reaction using T3 and T7 vector based primers followed by double digestion of the products with *MspI* and *HinPI*. The clones containing unique inserts were identified on the basis of unique restriction pattern and were selected for sequencing. Typically, ten to twenty unique restriction patterns were identified resulting in the sequencing of at least 10 clones from each sample. The plasmid inserts were sequenced in both directions using vector-based sequencing

primers T3, T7, or SP6 in an ABI 377 sequencer (Perkin Elmer Corp., Norwalk CT).

Data analysis

We used the chi square to identify differences in plasma DNA concentrations between the CFS and non-fatigued subjects. Subjects were stratified on the basis of sex, age (<45 years vs >45 years), CFS onset type (sudden vs gradual), and duration of illness (<5 years vs > 5 years) and compared by using a non-parametric Wilcoxon two-sample test. Cluster analysis was accomplished using BioNumerics (Applied Maths, Antwerp Belgium). For sequence analysis, contiguous sequences were generated using the DNASTAR program for each clone from sense and anti-sense strand sequences. The GenBank database was searched by using the BLAST tool to identify bacterial sequence similarities. A sequence similarity of 97% or greater was considered as an acceptable identification.

Author's contributions

SDV contributed to the conception and design of this study, the analysis of the results and drafted the manuscript. SKS contributed to the design of the experimental approach, implemented the experimental approach, analyzed the bacterial sequences, assisted in interpretation of the results and assisted in drafting the manuscript. JC conducted all amplification, cloning and sequence reactions and assisted in sequence analysis. ERU and WCR contributed to the conception and design of this study, interpretation of results and assisted in drafting the manuscript.

All authors read and approved the final manuscript.

Acknowledgments

This research was supported in part by extramural research funding from the Viral Exanthems and Herpesvirus Branch in the Division of Viral and Rickettsial Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention to SKS at the Marshfield Medical Research Foundation. Informed consent was obtained from all subjects, and human experimentation guidelines of the US Department of Health and Human Services were followed in the conduct of this research. The authors thank the Sedgwick County Department of Public Health, the Centers for Disease Control and Prevention CFS Research Group, the CFIDS Association of America, and the National Chronic Fatigue Syndrome and Fibromyalgia Association for their collaboration and Kurt Reed for helpful discussions.

References

- Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG and Komaroff A **The chronic fatigue syndrome: a comprehensive approach to its definition and study.** *Ann Intern Med* 1994, **121**:953-959
- Heneine W, Woods TC, Sinha SD, Khan AS, Chapman LE, Schonberger LB and Folks TM **Lack of evidence for infection with known human and animal retroviruses in patients with chronic fatigue syndrome.** *Clin Infect Dis* 1994, **18(Suppl)**:121-125
- Swanink CM, Melchers WJ, van der Meer JW, Vercoulen JH, Bleijenberg G, Fennis JF and Galama JM **Enteroviruses and the chronic fatigue syndrome.** *Clin Infect Dis* 1994, **19**:860-864
- Evengard B, Briese T, Lindh G, Lee S and Lipkin WI **Absence of evidence of Borna disease virus infection in Swedish patients with chronic fatigue syndrome.** *J Neurovirol* 1999, **5**:495-499
- Klimas NG, Salvato FR, Morgan R and Fletcher MA **Immunologic abnormalities in chronic fatigue syndrome.** *J Clin Microbiol* 1992, **28**:1404-1410
- Straus SE, Fritz S, Dale JK, Gould B and Strober W **Lymphocyte phenotype and function in the chronic fatigue syndrome.** *J Clin Immunol* 1993, **13**:30-40
- Mawle AC, Nisenbaum R, Dobbins JG, Gary HE Jr, Stewart JA, Reyes M, Steele L, Schmid DS and Reeves WC **Immune responses associated with chronic fatigue syndrome: A case-control study.** *J Infect Dis* 1997, **175**:136-141
- Taylor RR, Friedberg F and Jason LA **A clinician's guide to controversial illnesses: Chronic fatigue syndrome, fibromyalgia, and multiple chemical sensitivities.** Sarasota, Fla: Professional Resource Press 2001,
- Mawle AC, Nisenbaum R, Dobbins JG, Gary HE Jr, Stewart JA, Reyes M, Steele L, Schmid DS and Reeves WC **Seroepidemiology of chronic fatigue syndrome: A case-control study.** *Clin Infect Dis* 1995, **21**:1386-1389
- Buchwald D, Ashley RL, Pearlman T, Kith P and Komaroff AL **Viral serologies in patients with chronic fatigue and chronic fatigue syndrome.** *J Med Virol* 1996, **50**:25-30
- Jen J, Wu L and Sidransky D **An overview on the isolation and analysis of circulating tumor DNA in plasma and serum.** *Ann NY Acad Sci* 2000, **906**:8-12
- Wong IH, Johnson PJ, Lai PB, Lau WY and Lo YM **Tumor-derived epigenetic changes in the plasma and serum of liver cancer patients. Implications for cancer detection and monitoring.** *Ann N Y Acad Sci* 2000, **906**:102-105
- Dong SM, Pai SI, Rha SH, Hildesheim A, Kurman RJ, Schwartz PE, Mortel R, McGowan L, Greenberg MD and Barnes WA **Detection and quantitation of human papillomavirus DNA in the plasma of patients with cervical carcinoma.** *Cancer Epidemiol Biomarkers Prev* 2002, **11**:3-6
- Cursons RT, Jeyerajah E and Sleight JW **The use of polymerase chain reaction to detect septicemia in critically ill patients.** *Crit Care Med* 1999, **27**:860-861
- Wilson KH, Blitchington RB and Greene RC **Amplification of bacterial 16S ribosomal DNA with polymerase chain reaction.** *J Clin Microbiol* 1990, **28**:1942-1946
- Fredricks DN and Relman DA **Infectious agents and the etiology of chronic idiopathic diseases.** *Curr Clin Top Infect Dis* 1998, **18**:180-200
- Koelle DM, Barcy S, Huang ML, Ashley RL, Corey L, Zeh J, Ashton S and Buchwald D **Markers of viral infection in monozygotic twins discordant for chronic fatigue syndrome.** *Clin Infect Dis* 2002, **35**:518-525
- Nikkari S, McLaughlin JJ, Bi W, Dodge DE and Relman DA **Does blood of healthy subjects contain bacterial ribosomal DNA?** *J Clin Microbiol* 2001, **39**:1956-1959
- Weber G, Shendure J, Tanenbaum DM, Church GM and Meyerson M **Identification of foreign gene sequences by transcript filtering against human genome.** *Nature Genetics* 2002, **30**:141-142
- Relman DA **The human body as microbial observatory.** *Nature Genetics* 2002, **30**:131-133
- Reyes M, Nisenbaum R, Hoaglin DC, Unger ER, Emmons C, Randall B, Stewart JA, Abbey S, Jones JF and Gantz N **Prevalence and incidence of chronic fatigue syndrome in Wichita, Kansas.** *Arch Intern Med*

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

