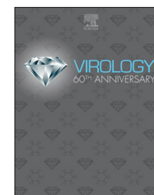
Virology 520 (2018) 21–29

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/virology



Viremia preceding multiple sclerosis: Two nested case-control studies

Emilie Hultin[a], Laila Sara Arroyo Mühr[a], Zurab Bzhalava[a], Maria Hortlund[a], Camilla Lagheden[a], Peter Sundström[b], Joakim Dillner[a,*]

[a] Department of Laboratory Medicine, Karolinska Institutet, Huddinge SE-141 86, Sweden
[b] Department of Pharmacology and Clinical Neuroscience, Umeå University, Umeå SE-901 87, Sweden

ARTICLE INFO

Keywords:
Multiple sclerosis
Anelloviridae
Metagenomic sequencing
Serum

ABSTRACT

Infections have been suggested to be involved in Multiple Sclerosis (MS). We used metagenomic sequencing to detect both known and yet unknown microorganisms in 2 nested case control studies of MS. Two different cohorts were followed for MS using registry linkages. Serum samples taken before diagnosis as well as samples from matched control subjects were selected.

In cohort1 with 75 cases and 75 controls, most viral reads were *Anelloviridae*-related and > 95% detected among the cases. Among samples taken up to 2 years before MS diagnosis, Anellovirus species TTMV1, TTMV6 and TTV27 were significantly more common among cases. In cohort2, 93 cases and 93 controls were tested under the pre-specified hypothesis that the same association would be found. Although most viral reads were again related to *Anelloviridae*, no significant case-control differences were seen. We conclude that the *Anelloviridae*-MS association may be due to multiple hypothesis testing, but other explanations are possible.

1. Introduction

Multiple Sclerosis (MS) is an autoimmune disease with inflammation of the central nervous system, with the most common onset in young adults. Although many factors, including genetic and environmental/lifestyle factors, have been implicated in either triggering MS or modulating the subsequent disease course, results vary substantially between studies (McKay et al., 2016).

Genetic factors (both HLA and non-HLA genes) have been linked to MS susceptibility (Sawcer et al., 2011; Beecham et al., 2013), but cannot fully explain the initiation of MS. Environmental factors as tobacco smoking (Salzer et al., 2014; Hedstrom et al., 2013, 2011), lack of sun exposure and low levels of vitamin D (Salzer et al., 2013, 2012), have all been suggested to be associated with MS, with varying degrees of evidence to back those claims.

Viruses have for a long time been of interest as potential environmental risk factors for MS, under the hypothesis that virus-triggered autoimmunity would cause inflammation, demyelination and nerve damage in the central nervous system. Several infectious agents have been investigated, for instance, human cytomegalovirus (CMV) (Vanheusden et al., 2015; Sundqvist et al., 2014), Epstein–Barr virus (EBV, HHV-4) (Thacker et al., 2006; Handel et al., 2010), human herpes virus 6 (HHV-6) (Leibovitch and Jacobson, 2014; Broccolo et al., 2013; Pietilainen-Nicklen et al., 2014), human parvovirus B19 (Nakashima et al., 1999), Chlamydia pneumoniae (Bagos et al., 2006; Munger et al., 2003; Krone et al., 2008) and Anelloviruses (Sospedra et al., 2005). Furthermore, epidemiological data suggest that parasite infections may be also associated with a lower risk of autoimmune diseases such as MS (Correale and Farez, 2011).

Progress in identifying infectious agents as risk factors has been slow as most studies have typically only focused on one candidate infection at a time and/or used study designs fraught with epidemiological biases. By performing an unbiased detection (not based on predetermined DNA sequences) using massively parallel sequencing, it is possible nowadays to identify all microorganisms (known and yet unknown) that are present in biological specimens from any patient.

Our aim was to analyze MS-patients with donated serum samples preceding the MS diagnosis by performing unbiased metagenomic sequencing using a strong epidemiological study design: two case-control studies (one exploratory and one validation study) nested in prospectively followed cohorts.

2. Materials and methods

2.1. Patients

The first cohort consisted of i) a maternity cohort and ii) a population-based cohort, both enrolled in Northern Sweden. The Northern

* Corresponding author.
E-mail address: joakim.dillner@ki.se (J. Dillner).

https://doi.org/10.1016/j.virol.2018.04.006
Received 8 November 2017; Received in revised form 26 March 2018; Accepted 10 April 2018
0042-6822/ © 2018 Elsevier Inc. All rights reserved.

Sweden Maternity Cohort (NSMC) contained 124,000 serum samples collected since 1975 from pregnant women tested for rubella immunity at the first antenatal visit after confirmed pregnancy (usually around week 12 of gestation), stored at −20 °C (Salzer et al., 2012). The Northern Sweden Health and Disease Study cohort (NSHDS) contained 167,500 blood samples collected since 1985 in population-based health programs, stored at −70 °C (Salzer et al., 2012). Two subcohorts of the NSHDS were used in this study: The Västerbotten Intervention Program (n = 113,000 samples) (a population-based health-promoting program which started in 1985) and the Västerbotten Mammography Screening Program (N = 54,500 samples), which biannually invites all women aged 50–60 and resident in Västerbotten county. The joint cohort contained a total of 291,500 samples from 164,000 unique individuals, when this study was performed. All samples were collected in the same way, except that the samples in NSMC that were collected between 1975 and 1987 were heat-treated before storage. The second cohort was the Southern Sweden Microbiology Cohort, which enrolled all subjects with samples submitted for microbiological serology in Southern Sweden. This includes also the samples submitted from pregnant women for testing of rubella immunity (same principle as the NSMC, see above).

Registry linkages were performed to identify samples taken from subjects who later became MS-cases, but had no symptom suggestive of MS noted in medical records prior to sample collection. Two controls matched for sex, cohort, sampling date and age (decreasing priority, best match applied) were selected for each case (Salzer et al., 2012). For this study, only one control per case was used, for reasons of high cost of metagenomics analyses.

In the first cohort, 75 MS-patients (65 females, 10 males) who had donated serum samples preceding the MS-diagnosis with 0–6 years were identified together with 2 matched healthy controls for each case (150 controls). The serum samples were collected from 1976 to 2005, the birth years ranged from 1936 to 1983 and the ages ranged from 16 to 59 years. Cases and controls were dichotomized in 2 data sets taken shortly or a long time before the MS diagnosis. For 48 cases (37 from maternity cohort and 11 from population cohort) samples were taken within 0–4 years before diagnosis and this was, together with 48 matched controls (the control with the lowest sample ID number for each case) considered to be the set with highest risk in the first cohort. A second data set included 27 cases of serum samples (21 from maternity cohort and 6 from the population cohort) taken 5–6 years preceding diagnosis, together with 27 matched controls and 2 negative controls (lab-grade water, SigmaAldrich, US).

In the second cohort, registry linkages identified 93 MS-patients (84 female, 9 male) with donated serum samples stored before MS-diagnosis, together with 93 matched controls. Samples were collected from 1985 to 2007, the birth years ranged from 1953 to 1991 and ages from 12 to 38. Two data sets were analyzed, one including 46 patients (12 from the maternity cohort and 34 from other microbiological testing) with samples stored between 0 and 20 years before diagnosis, together with 46 matched controls and 2 negative controls (PBS-buffer, SigmaAldrich, US). The other data set from this cohort included 47 serum samples (11 from maternity cohort and 36 from other microbiological testing) taken 0–4 years before MS-diagnosis, 47 matched controls and 2 negative controls (PBS-buffer and lab-grade water, SigmaAldrich, US).

The study was approved by the Ethical Review Committee of Umeå, Sweden (Dnr 2011-198-31M). All methods were carried out in accordance with the approved guidelines.

2.2. Sample preparation

50 μl of each serum sample was first diluted in PBS-buffer (SigmaAldrich, US) 1:2. To reduce cellular debris, the samples were centrifuged at 21,000g for 15 min at room temperature. The serum supernatants were used for DNA extraction with the MagNA Pure nucleic acid isolation kit using MagNA Pure LC instrument (Roche, Switzerland), where the DNA was eluted in 50 μl of elution buffer.

The extracted DNA from the serum samples together with the negative controls (lab-grade water and PBS-buffer (SigmaAldrich, US)) were subjected to whole genome amplification (WGA) using Illustra™ Ready-To-Go GenomiPhi™ HY DNA Amplification Kit (GE Healthcare, United Kingdom). WGA was performed according to the manufacturer´s instructions with some modifications. 5 μl of DNA template was mixed with 25 μl of denaturation buffer, diluted to 50 μl with lab-grade water (SigmaAldrich, US), incubated at 95 °C for 3 min and then cooled on ice for 3 min. The Ready-To-Go GenomiPhi HY cake was reconstituted with the cooled and denatured DNA. The samples were incubated at 30 °C for 7 h, inactivated at 65 °C for 10 min and cooled to 4 °C. The amplified products were diluted 1:2 in lab-grade water (SigmaAldrich, US). The WGA products were quantified using QuantiFluor-ST (Promega, US), a fluorometric assay quantifying dsDNA, according to manufacturer's user guide. The concentrations of the WGA products were between 0.6 and 8 μg/μl. The DNA fragments size distributions were checked using the Bioanalyzer DNA 12,000 chip (Agilent Technologies, US).

2.3. Metagenomic sequencing

DNA libraries were prepared from 50 ng of WGA DNA using the Nextera DNA Sample Preparation kit according to the user guide revision B (Illumina, US). Dual indexing was used during the adapter amplification step to facilitate multiplex sequencing of up to 96 samples. All individual libraries were validated, normalized to 4 nM and combined in 4 different pools, with each case together with its matched control.

Prior to sequencing, the library pools were denatured and diluted. The first pool, containing 96 individual indexed libraries (48 cases and 48 controls, first cohort), was sequenced using HiSeq. 2000 (Illumina, US) at 8pM DNA solution spiked with PhiX control at 1% according to Denaturing and Diluting Libraries for the HiSeq and GAIIx Rev. A (Illumina). The library pool was sequenced in one flow cell by paired-end 101 + 101 cycles on the HiSeq. 2000 using TruSeq PE Cluster v3 and TruSeq SBS v3 reagent kits according to cBot User Guide Rev L, Sequencing Dual-Indexed Libraries on the HiSeq System User Guide Rev B and HiSeq. 2000 User Guide Rev T (Illumina).

The following 3 library pools (second data set from the first cohort and both data sets from the second cohort) were sequenced on the NextSeq. 500 (Illumina, US) at 1.8 pM DNA solution spiked with 1% of PhiX control. Each of the library pools was sequenced by paired-end 151 + 151 cycles using NextSeq. 500 High Output reagent kit (Illumina, US). The sequencing preparations were made according to the user guides Denaturing and Diluting Libraries for the NextSeq. 500 revision A and NextSeq. 500 kit Reference Guide revision F.

2.4. Bioinformatics analysis

All analyses were performed using in-house R (www.R-project.org), python (www.python.org) and bash (www.gnu.org/software/bash/) scripts that run on a high performance (40 core, 2TB RAM) Linux server. Sequences were obtained from the HiSeq and NextSeq platforms (Illumina, US) and the indices, part of the Illumina adapters, were used to assign the sequence reads to the originating samples.

Bioinformatics workflow was performed as described (Bzhalava et al., 2013). Shortly, the bioinformatic analysis started with quality checking where sequences were trimmed according to their Phred quality scores (Ewing and Green, 1998). Quality checked reads were then screened against the human reference genome hg19, as well as bacterial, phage and vector sequences downloaded from GenBank using BWA-MEM (Li and Durbin, 2010; Peng et al., 2012). The sequencing reads with > 95% identity over 75% of their length to human, bacterial, phage and vector DNA were removed from further analysis. The rest of the sequences were normalized (http://ged.msu.edy/papers/

Download English Version:

https://daneshyari.com/en/article/8751407

Download Persian Version:

https://daneshyari.com/article/8751407

Daneshyari.com