A method for DNA virus detection and quantification during pregnancy based on noninvasive prenatal testing whole genome sequencing results

V. CHESNAIS[1], A. OTT[1], E. CHAPLAIS[1], S. GABILLARD[1], C. VAULOUP-FELLOUS[2], A. BENACHI[3], J.-M. COSTA[4], and E. GINOUX[1]

1 Life&Soft, 8b avenue Descartes, 92350, Plessis-Robinson, France; 2 Hôpital Paul Brousse, 12 avenue Paul Vaillant Couturier, 94800, Villejuif, France; 3 Hôpital Antoine-Béclère, 157 rue de la Porte de Trivaux, 92140, Clamart, France; 4 Laboratoire CERBA, 11 rue de l'Équerre, 95310, Saint-Ouen-l'Aumône, France

Context:

With 10–100 billion fragments per milliliter of plasma, circulating cell-free DNA is an information-rich window into human physiology, with rapidly expanding applications in genetic prenatal diagnosis. The whole genome sequencing (WGS) of cell-free plasma DNA is classically used to diagnosis fetal aneuploidy during pregnancy. Because WGS has also become a standard tool in pathogen discovery in biological samples [1], the purpose of this study is to propose a new method to detect and quantify circulating viral DNA during pregnancy using the same sequencing results than noninvasive prenatal testing whole genome sequencing data.

HCMV infection and impact on offspring

At-risk subjects

Immunonegative women (40% to 60%)

Infection risk during the pregnancy

1% to 4% of women acquire HCMV during their pregnancy

It includes reinfections and reactivations

Vertical transmission (mother→fetus)

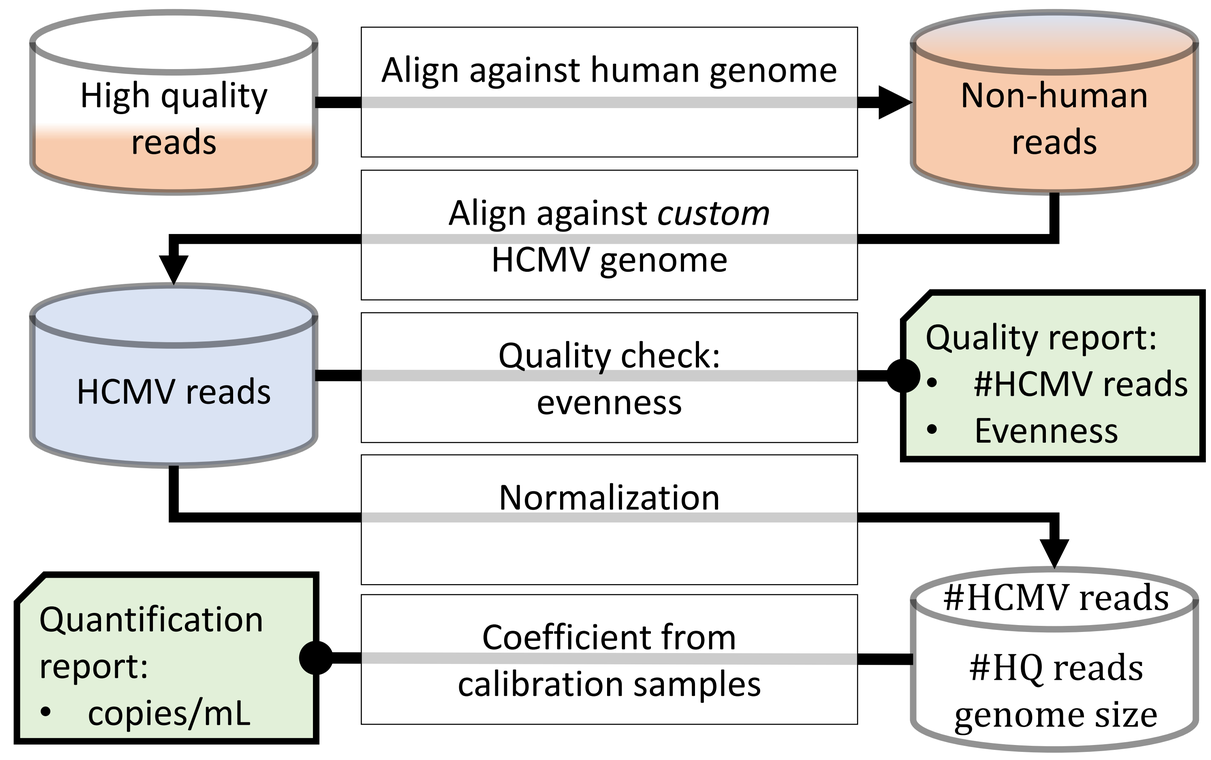
40%-50% risk of transmission to the fetus.

Consequences of congenital HCMV infection

- 12% of newborns have symptoms; 40%-58% have permanent sequelae.
- 13% of remaining develop permanent sequalae in the first year.

Global prevalence of congenital CMV infection has been estimated to be 0.7% [2]

Analysis workflow



HCMV detection in blood samples

Calibration samples

Aim: convert a normalized number of reads into a **clinical meaningful** viral concentration

To mimic infected maternal plasma samples, we mixed DNA plasma sample of non infected women and DNA of HCMV infected samples to make HCMV range samples of known viral load. Sequencing of these samples are performed several times in duplicate and in distinct series with different dilutions.

Cohort patients

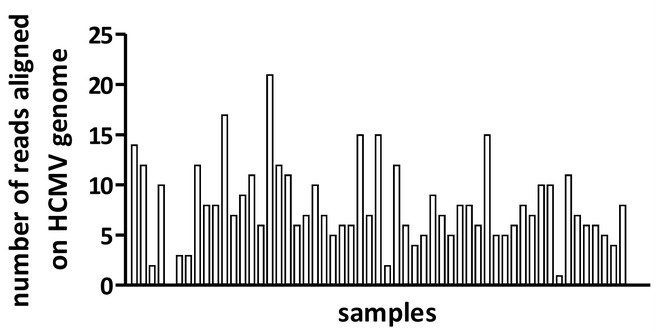
We then studied a cohort of 538 plasma samples from pregnant women to identify HCMV viremia positive samples. The immune status was known for 357 of them (66%) including 204 positive (38%).

Once an individual has been infected once over her life time, she remains positive to immune tests.

Custom HCMV genome

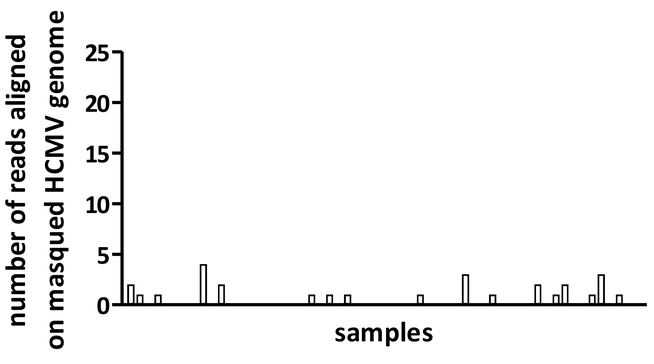
Non-specific read alignment.

At first we aligned reads on HCMV genome. Even negative controls obtained about 10 aligned reads.



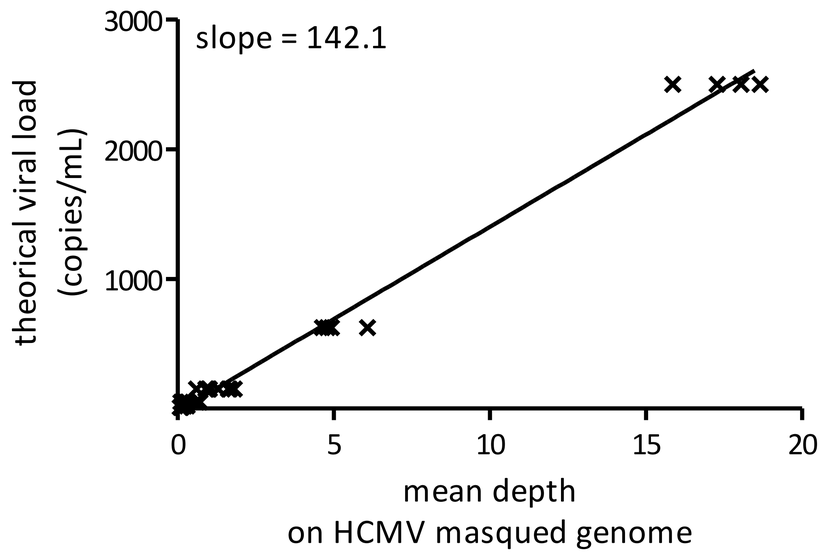
Masked genome

We removed some regions of the HCMV genome. These low complexity regions tend to recruit unspecific reads. Masking only 4% of the genome dramatically increase the specificity.



Results

Concentration of HCMV is linearly correlated with the normalized number of reads.



Bootstrap was performed to obtain a 95% IC on slope → copies/mL

Cohort patients

For 2 samples a depth superior to the previously determined threshold was observed. One patient was immunopositive but the other one had an unknown immune status.

Please note that an infected individual remains immunopositive his whole life. A positive immune status is not a proof of a ongoing infection.

Go further: application to HBV

HBV challenge

The hepatitis B virus (HBV) genome is smaller than HCMV (~80 folds). The availability of the method to identify such infection is limited by the sequencing deepness.

Encouraging results

Only 369 patients had an HBV serology available, with 3 positives one. Two patients with positive serology were detected by our analysis and 1 were missed. For the third positive patient detected by our method, the serology was not available.

Please note that for HBV immunopositivity is a sign of acute infection.

References

[1] Q. Wang, P. Jia, Z. Zhao, VirusFinder: software for efficient and accurate detection of viruses and their integration sites in host genomes through next generation sequencing data, PloS One. 8 (2013) e64465. doi:10.1371/journal.pone.0064465.

[2] S.C. Dollard, S.D. Grosse, D.S. Ross, New estimates of the prevalence of neurological and sensory sequelae and mortality associated with congenital cytomegalovirus infection, Rev. Med. Virol. 17 (2007) 355–363. doi:10.1002/rmv.544.