Quantitative investigation of murine cytomegalovirus nucleocapsid interaction

Christopher Buser[1] **Frank Fleischer**[2]
Thomas Mertens[3] **Detlef Michel**[3] **Volker Schmidt**[2]
Paul Walther[1]

28th February 2007

[1]Electron Microscopy Facility
University Ulm, D-89069 Ulm, Germany
Phone: +49 731 50 23441
Fax: +49 731 50 23383

[2]Institute of Stochastics
University Ulm, D-89069 Ulm, Germany
Phone: +49 731 50 23531
Fax: +49 731 50 23649

[3]Institute of Virology
University Ulm, D-89081 Ulm, Germany
Phone: +49 731 50 023341
Fax: +49 731 50 023337

Corresponding author:
Frank Fleischer
Institute of Stochastics
University Ulm, D-89069 Ulm, Germany
Phone: +49 731 50 23617
Fax: +49 731 50 23649
E-mail: frank.fleischer@uni-ulm.de

Abstract

In this study we quantitatively investigate the role of the M97-protein for viral morphogenesis; in murine cytomegalovirus (MCMV) infected fibroblast cells. For this purpose a statistical analysis is performed for the spatial distribution of nuclear B-capsids (devoid of DNA, containing the scaffold) and C-capsids (filled with DNA). Cell nuclei infected with either wild type or an M97-deletion mutant were compared. Univariate and multivariate point process characteristics (like Ripley's K-function, the L-function and the nearest neighbor distance distribution function) are investigated in order to describe and quantify the effects that the deletion of M97 causes to the process of DNA-packaging into the nucleocapsids. The estimation of the function $L(r) - r$ reveals that with respect to the wild type there is an increased frequency of point pairs at a very short distance (less than approx. 100 nm) for both the B-capsids as well as for the C-capsids. For the M97-deletion mutant type this is no longer true. Here only the C-capsids show such a clustering behavior while for B-capsids it is almost non-existing. Estimations of functionals like the nearest neighbor distance distribution function confirmed these results. Thereby a quantification is provided for the effect that the deletion of M97 leads to a loss of typical nucleocapsid clustering in MCMV infected nuclei.

Keywords ELECTRON MICROSCOPY, MURINE CYTOMEGALOVIRUS, HIGH-PRESSURE FREEZING, POINT PROCESS CHARACTERISTICS, SPATIAL POINT PATTERNS, STOCHASTIC GEOMETRY

1 Introduction

Cytomegaloviruses (CMV) are DNA-viruses of the herpesviridae family. The viral genome codes for approximately 200 open reading frames (Murphy et al., 2003). In the virion, the DNA is packed in an icosahedral capsid, which is surrounded by a protein layer called the tegument, and a lipid envelope containing viral glycoproteins. The morphogenesis of herpesviruses is not completely understood. The capsid is preformed in the nucleus as a B-capsid with an internal stabilizing scaffold protein (the internal 'ring' in the B-capsid) and is subsequently loaded with the viral DNA (vDNA) yielding the C-capsid (Church and Wilson, 1997; Gao et al., 1994; Newcomb et al., 1994; Sheaffer et al., 2001). The loaded C-capsid then exits the nucleus by primary envelopment at infoldings of the inner nuclear membrane (Buser et al., 2007) and probably fusion of the primary envelope with the outer nuclear membrane, thus releasing the naked capsid to the cytoplasm. There it is tegumented and enveloped again at Golgi-derived cisternae and released by fusion of the transport vesicle with the plasma membrane (Mettenleiter, 2002).

The exact function of the M97 protein and its homolog pUL97 in human cytomegalovirus is still unclear and has been implicated in different steps of the morphogenesis (Azzeh et al., 2005; Krosky et al., 2003a and 2003b; Marschall et al., 2003 and 2005; Wagner et al., 2000; Wolf et al., 2001). In an analysis of an M97-deletion mutant by electron microscopy we found a reduction of nuclear C-capsids in favor of B-capsids and an apparent loss of the typical nucleocapsid clustering. Here we used methods from spatial statistics to quantify this nuclear phenotype. We will show that in the case of the wild type, for a distance of less than approx. 100 nm, there is an increased frequency of pairs of capsids. For the M97-deletion mutant such a behavior can only be stated for the C-capsids, while the frequency of B-capsid pairs of such a distance is diminished.

2 Materials and methods

2.1 Cells and viruses

3T3 mouse fibroblast cells were grown in Dulbecco's modified Eagle's medium (DMEM; GIBCO, Germany) supplemented with 10 % (v/v) newborn calf serum, 1 % Glutamine (w/v) and penicillin/streptomycin and passaged with trypsin-EDTA. MCMV strain (Smith strain) was used as previously described (Wagner et al., 2000), but propagated on M210B4 murine bone marrow stromal cells in minimal essential medium (MEM; GIBCO, Germany) supplemented as described for DMEM.

2.2 High pressure freezing of infected cells on sapphire discs

Sapphire discs (3 mm diameter, thickness about 50 μm; Brügger, Minusio, Switzerland distributed by Engineering Office M. Wohlwend GmbH, Sennwald, Switzerland) were cleaned in sulphuric acid, soap water and absolute ethanol and subsequently coated with 20 nm of carbon by electron beam evaporation with a BAF (Baltec, Liechtenstein). The sapphire discs were then placed in 24-well plates and the cells were allowed to grow to approximately 80 % confluency. Infection was done by inoculating with 200 μl total volume of virus suspension per well containing an MOI of 0.1-0.5 PFU/cell for 15 min at 37 $°C$. After this 1 ml medium was added per well and infection was allowed to proceed for 48 hours before high-pressure freezing. High pressure freezing was performed with the new compact high pressure freezing apparatus HPF 01 (Engineering Office M. Wohlwend GmbH, Sennwald, Switzerland). The sapphire discs were clamped between two aluminum planchettes (diameter 3 mm) (18). One planchette was flat and the other had a central cavity of 100 μm in depth and 2 mm in diameter. The cavities between the tissue and the aluminum planchettes were filled with hexadecene (28). The frozen samples were stored in liquid nitrogen until freeze-substitution.

2.3 Freeze-substitution and embedding for ultrastructural imaging

The frozen samples were carefully removed from the aluminium planchettes in liquid nitrogen and immersed in substitution medium pre-cooled to $-90\ ^{\circ}C$. The following substitution medium was used: 1.6 % (w/v) of osmium tetroxide, 0.1 % (w/v) uranyl acetate and 5 % (v/v) of water in acetone (31). We used a specially designed computer-controlled substitution apparatus (A. Ziegler and W. Fritz, University of Ulm, unpublished) to slowly warm the samples from $-90\ ^{\circ}C$ to $0\ ^{\circ}C$ over a period of $16 - 22\ h$. The samples were kept at $0\ ^{\circ}C$ and at room temperature for 30 min each, washed with acetone and embedded in a two step Epon series (Fluka, Germany) of 50 % Epon in acteone for 1 h, 100 % Epon for 6 h. The Epon was then polymerized for 48 h at $60\ ^{\circ}C$. Ultrathin sections were cut with a Leica Ultracut UCT ultra microtome using a diamond knife (Diatome, Switzerland). For ultrastructural analysis 60 nm sections were collected on bare copper grids and post-stained with uranyl acetate and lead citrate.

2.4 Transmission electron microscopy

The sections were imaged with a Zeiss EM 10 or a Philips 400 transmission electron microscope at an acceleration voltage of 80 kV. Only central cross-sectioned nuclear areas were selected for analysis in order to contain a sufficient number of capsids. The images were recorded on Agfa negative film at an identical magnification of $4,600\times$. After development, the negatives were scanned and the resulting digital images were processed with Adobe Photoshop.

2.5 Statistical analysis

For each of the sample images locations of C-capsids and B-capsids were interactively marked, as well as the borderline of the nucleus (Figure 1b). The capsid type was determined by visibility of the internal scaffold ring of B-capsids and the density of the packed DNA in C-capsids (Figure 1a). If no scaffold could be seen and the inner density was high the capsid was considered of the C-type. Capsids that could not be clearly assigned to one class (e.g. if only partially within the thin-section, arrowhead Figure 1a) or that already budded into the perinuclear space were ignored. Since A-capsids (without DNA or scaffold) were very rare (below 1 % of nucleocapsids; Buser et al., 2007) and are regarded as by-products unable to undergo further maturation, they were also not included in this analysis. The amount of ignored capsids was approximately 5 % in both wild type and delM97 samples. Later analysis was based on the marked point patterns consisting of the locations of the B-capsids and the C-capsids in sampling windows of variable but similar size and shape.

Data analysis was performed using the GeoStoch library. GeoStoch is a Java-based open-library system developed by the Institute of Applied Information Processing and the Institute of Stochastics of the University of Ulm that can be used for spatial statistical analysis of image data and spatio-geometric modelling (Mayer et al., 2004; http://www.geostoch.de).

In the following mean values of estimated point process characteristics are considered, where the means were taken for the two regarded groups of images separately. Means were taken in a pointwise sense, meaning that if $\widehat{f}_1(r), ..., \widehat{f}_n(r)$ represent the n individual estimates for a function $f(r)$ at a point-pair distance r the averaged estimate $\widehat{f}(r)$ is given as $\widehat{f}(r) = \frac{1}{n}\sum_{i=1}^{n}\widehat{f}_i(r)$. Statistical comparison of groups are based on the (exact) Wilcoxon-Mann-Whitney test.

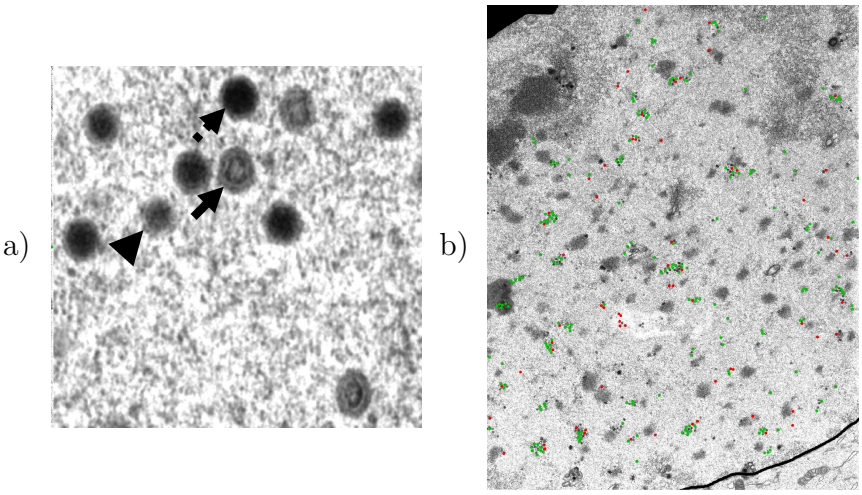


Figure 1: Data preprocessing. Panel a) illustrates the choice of capsid type while Panel b) depicts the marked capsids and borders of cell nuclei.

3 Point process characteristics and their estimators

In the following we want to introduce the (univariate and multivariate) point process characteristics and their corresponding estimators that are used in this paper. For a more detailed description of the topics mentioned the reader is refered to Baddeley et al. (2006), Diggle (2003) and Stoyan et al. (1995).

3.1 Multivariate spatial point processes

Let $X^{(1)}, ..., X^{(m)}$ be m univariate spatial point processes in $\mathbb{R}^2$, i.e., for any $i = 1, ..., m$ and for any sampling window $B \subset \mathbb{R}^2$, the random variable $X^{(i)}(B) = \#\{n : X_n^{(i)} \in B\}$ denotes the number of points $X_n^{(i)}$ of type i observed in the window B. Then $X = \sum_{i=1}^{m} X^{(i)}$ is called a multivariate point process and $X^{(i)}$ with $1 \leq i \leq m$ is called a component of X. From a different point of view, one can define X as a marked point process with mark space $M = \{1, ..., m\}$. Note that in the following we will only regard two-dimensional point processes with $m = 2$, leading to bivariate point processes and that the point processes regarded are considered to be stationary and

isotropic, meaning that their distribution is invariant with respect to translations and rotations around the origin.

3.2 Intensity

Let $\{X_n\}$ denote the points of a (stationary and isotropic) multivariate point process X in $\mathbb{R}^2$. The intensity measure $\Lambda(B)$ of X, i.e., the mean number of points in a sampling window B, is given as

$$\Lambda(B) = EX(B) = \lambda|B|, \tag{3.1}$$

where $X(B) = \#\{n : X_n \in B\}$ denotes the number of points of X in B and $|B|$ means the area of B. The constant λ is called the intensity of X. A simple estimator for λ is given by

$$\widehat{\lambda} = \frac{X(B)}{|B|}. \tag{3.2}$$

However, for the estimation of the nearest neighbor distribution, a different estimator for λ is suggested in Hanisch (1984), see also (3.11):

$$\widehat{\lambda}_H = \sum_{X_n \in X} \frac{1_{B \ominus b(o,s(X_n))}(X_n)}{|B \ominus b(o, s(X_n))|}, \tag{3.3}$$

where $1_B(x)$ is the indicator function of the set B, i.e., $1_B(x) = 1$ if $x \in B$ and $1_B(x) = 0$ if $x \notin B$. Furthermore, $s(X_n)$ denotes the distance of X_n to its nearest neighbor in $X \cap B$ and $b(x, r)$ is the ball with radius r and midpoint x (Stoyan et al., 2001). Following the recommendation in Stoyan and Stoyan (2000), λ^2 has been estimated by

$$\widehat{\lambda^2} = \frac{X(B)(X(B) - 1)}{|B|^2}. \tag{3.4}$$

Estimates for the intensity $\lambda^{(i)}$ of the component $X^{(i)}$ of X can be obtained naturally by replacing X with the relevant $X^{(i)}$ in the estimators above.

3.3 Multivariate K-function and multivariate L-function

The multivariate K-function $K_{ij}(r)$ for a multivariate point process X and its components $X^{(i)}$ and $X^{(j)}$ can be defined as the mean number of points of $X^{(j)}$ that are located in a disc around a randomly chosen point of $X^{(i)}$, divided by the intensity $\lambda^{(j)}$ of $X^{(j)}$. As a mathematical definition we get

$$K_{ij}(r) = E \sum_{X_n^{(i)} \in B} \frac{X^{(j)}(b(X_n^{(i)}, r) \setminus \{X_n^{(i)}\})}{\lambda^{(i)}\lambda^{(j)}|B|}. \tag{3.5}$$

Note that theoretically

$$K_{ij}(r) = K_{ji}(r), \tag{3.6}$$

also if $i \neq j$. Furthermore, since we want to regard only simple point processes we have that $P(X_n^{(i)} \in X^{(j)}) = 0$ for all n. In the case $i = j$ it is easy to see that $K_{ii}(r) = K^{(i)}(r)$, meaning that $K_{ii}(r)$ is the univariate K-function for the component $X^{(i)}$. Analogously to the univariate case, one can define the multivariate L-function $L_{ij}(r)$ as

$$L_{ij}(r) = \sqrt{\frac{K_{ij}(r)}{\pi}} \tag{3.7}$$

in order to get a function that is more easily interpretable. As an estimator for the function $K_{ij}(r)$ we used

$$\widehat{K}_{ij}(r) = \frac{1}{\widehat{\lambda}^{(i)}\widehat{\lambda}^{(j)}|B|} \sum_{X_n^{(i)} \in B} \sum_{X_m^{(j)} \in B} \frac{1_{b(0,r)}(|X_n^{(i)} - X_m^{(j)}|)}{\omega_B(X_n^{(i)}, X_m^{(j)})}, \tag{3.8}$$

where $|x|$ denotes the length of the vector $x \in \mathbb{R}^2$ and where $\omega_B(X_n^{(i)}, X_m^{(j)})$ is the ratio of the part of the circle around $X_n^{(i)}$ with radius $|X_m^{(j)} - X_n^{(i)}|$ that belongs to B and the total circumference of this circle. Note that in the univariate case the weights $|B|\omega_B(X_n^{(i)}, X_m^{(i)})$ can alternatively be replaced by $|B_{X_n}^{(i)} \cap B_{X_m}^{(i)}|$, where $B_{X_n}^{(i)}$ denotes

the window B shifted by the vector $X_n^{(i)}$. A natural estimator for $L_{ij}(r)$ is given by

$$\widehat{L}_{ij}(r) = \sqrt{\frac{\widehat{K}_{ij}(r)}{\pi}}. \tag{3.9}$$

Note that

$$L_{ij}(r) = r \tag{3.10}$$

if $i = j$ and the points of $X^{(i)}$ are distributed according to complete spatial randomness, or if $i \neq j$ and $X^{(i)}$ and $X^{(j)}$ are independent. Notice furthermore that a slope of $L_{ij}(r) - r$ that is positive (negative) is a sign of attraction (rejection) for point-pairs $(X_n^{(i)}, X_m^{(j)})$ at a distance r.

3.4 Multivariate nearest neighbor distance distribution

The (i to j) nearest neighbor function $D_{ij}(r)$ is the distribution function of the distance from a randomly chosen point $X_n^{(i)}$ of $X^{(i)}$ to its nearest neighbor belonging to $X^{(j)}$. Therefore $D_{ij}(r)$ can be regarded as the probability that a randomly chosen point $X_n^{(i)}$ of $X^{(i)}$ has a distance less than or equal to r to its nearest neighbor in $X^{(j)}$. For the estimation of $D_{ij}(r)$ we used a Hanisch-type estimator (Baddeley, 1998; Hanisch, 1984) given by

$$\widehat{D}_{ij}(r) = \frac{\tilde{D}_{ij}(r)}{\widehat{\lambda}_H^{(j)}}, \tag{3.11}$$

where

$$\tilde{D}_{ij}(r) = \sum_{X_n^{(i)} \in B} \frac{1_{B \ominus b(o, s_j(X_n^{(i)}))}(X_n^{(i)}) 1_{(0,r]}(s_j(X_n^{(i)}))}{|B \ominus b(o, s_j(X_n^{(i)})|}, \tag{3.12}$$

and where $s_j(X_n^{(i)})$ denotes the distance of $X_n^{(i)}$ to its nearest neighbor belonging to $X^{(j)} \cap B$. Note that in the case of $X^{(i)}$ being completely spatially random we get that

$$D_{ii}^{Poi}(r) = 1 - \exp\left(-\pi \lambda^{(i)} r^2\right), \tag{3.13}$$

where $\lambda^{(i)}$ denotes the intensity of $X^{(i)}$.

3.5 Replicated data and tests by bootstrapping

The fact that for both groups (wild type as well as mutant) several samples are available can be used in order to enhance the statistical analysis (cmp. e.g. Baddeley et al., 1993; Diggle et al., 1991, 2000; Howard et al., 1985; Stoyan and Stoyan, 2000). First of all, estimates of the (pointwise) standard error for the averaged functions can be provided by taking the sample standard deviation of the individual estimates at a specific range r and divide it by the square root of the number of samples regarded. Intervals based on such pointwise estimates for the standard error can provide hints whether two functions are significantly different or not. Furthermore, it is possible to construct parametric tests based on the standard errors and on normality assumptions. In our data situation. due to the relatively small number of replications, we instead opted for a non-parametric test by using bootstrap techniques for point process characteristics (cmp. Diggle et al., 2000; Mattfeldt et al., 2006). In particular it can be tested whether two averaged estimated L-functions $\widehat{L}_1(r)$ and $\widehat{L}_2(r)$ are significantly different with respect to a given range of r. For this purpose the statistic $D^* = \sum_{i=1}^{n} |\widehat{L}_1(r_i) - \widehat{L}_2(r_i)|$ is regarded for a specified set $\{r_i\}_{i=1,\ldots,n}$, where $|\cdot|$ denotes the absolute value. Afterwards k, say 999, bootstrap samples $S_i = (S_i^{(1)}, S_i^{(2)})$ are generated by resampling with replacement from the estimated L-functions, where the subsample $S_i^{(j)}$ has the same sample size as the jth group ($j \in \{1, 2\}$). For all the 999 bootstrap samples a test statistic D_i is computed analogously to D^* by summing up the differences between the two function values. Finally, the D_i values of the bootstrap samples and the D^* value from the real sample are ordered by size. Given a test level of $\alpha = 0.05$, the difference between the two functions is considered to be significant if the rank of D^* in this series is larger than 950.

Table 1: Mean number and empirical standard deviation of capsids per image for wild type and deletion mutant

	wild		mutant	
	mean	std.dev.	mean	std.dev.
# B-capsid	27.00	18.32	119.67	65.22
# C-capsid	106.30	61.15	47.44	21.71

4 Results

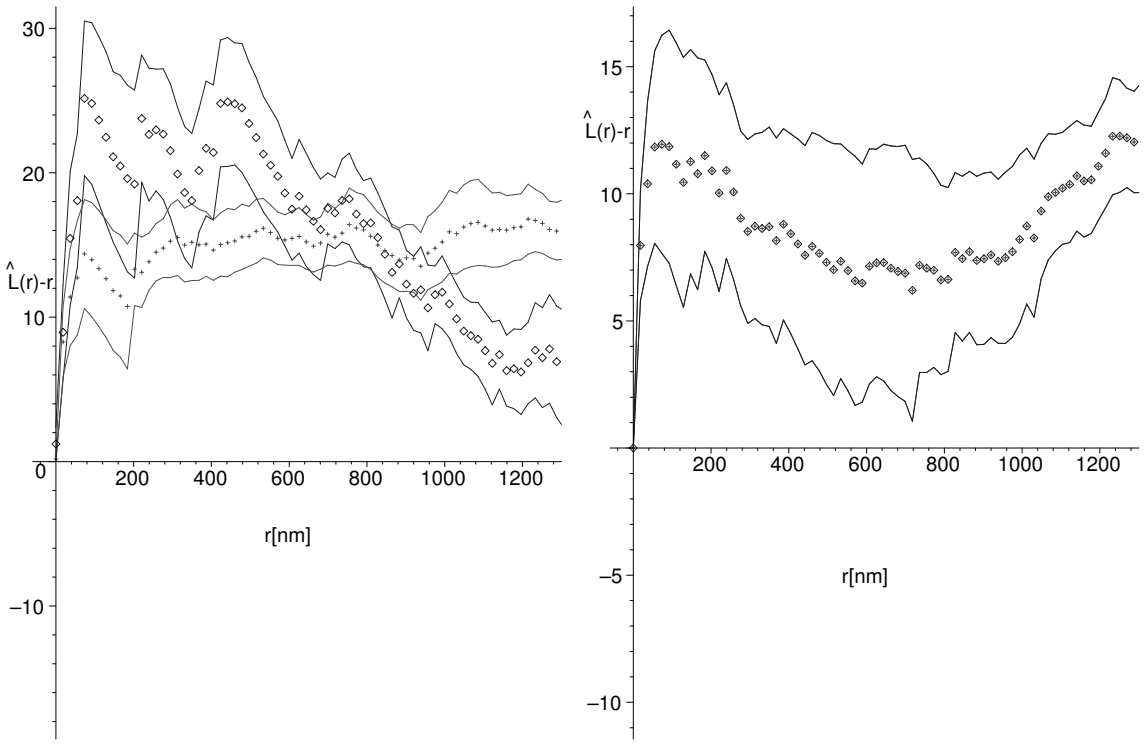
4.1 Different capsid ratios for wild type and deletion mutant

In Table 1 mean numbers per image as well as related empirical standard deviations of observed capsids are displayed. Altogether there were 10 sample images of the wild type and 9 images of the deletion mutant type regarded. In the case of the wild type there are about 4.5 times as many C-capsids than B-capsids leading to a significant rejection of the hypothesis of equal mean numbers ($p < 0.05$). With respect to the deletion mutant type an opposite relation between the numbers of capsids can be observed. In this case there are about 2.2 times as many B-capsids observable than C-capsids leading to a highly significant rejection of the hypothesis of equal mean numbers ($p < 0.01$). Note that for the comparison of the number of B-capsids and C-capsids within the same type of virus (wild type or deletion mutant) it does not matter if absolute numbers of points or intensities are regarded since the sampling window is identical in this case. With respect to mean total numbers of capsids as well as mean intensities no significant differences between the wild type and the mutant could be detected ($p > 0.05$).

4.2 Estimates of L-functions

In Figures 2 and 3 averaged estimates for the functions $L(r) - r$ are displayed together with corresponding (pointwise) standard errors. Note that we opted for displaying

$L(r) - r$ instead of the L-function itself since it might be easier interpretable with respect to deviations from the Poisson case, where it holds that $L(r) - r = 0$. By looking at the univariate $L(r) - r$-functions for the wild type shown in Figure 2a one can see that for small point-pair distances an attraction effect is visible. The region of attraction for a range of less than approx. 100 nm can be deduced from the fact that for such a range of r the slope of $\widehat{L}(r) - r$ is positive. In the wild type case this attraction effect is similar with respect to its strength and range for points generated by B-capsids compared to points generated by C-capsid. For the B-capsids it even seems to be a bit stronger than for the C-capsids, since here the function $\widehat{L}(r) - r$ for B-capsids runs above the same function for C-capsids for small r. With regard to the multivariate L-functions displayed in Figure 2b almost identical results are obtained leading to same conclusions as for the univariate case. The two underlying point processes for the B-capsids and the C-capsid seem to be far from being independent since in this case the averaged estimated multivariate functions $\widehat{L}(r) - r$ should be near 0 for all r. Furthermore, a range of attraction for an interpoint distance of less than 100 nm can be observed. Note that the theoretical equality of the two multivariate L-functions that is based on (3.6) is very well reflected by the estimates, therefore in particular only one interval of standard errors is given in this case. A different picture can be drawn in the case of the deletion mutant type. For the case of C-capsids a similar attraction effect for small point-pair distances of less than 100 nm is obtained that can be deduced from the fact that in this region the function $\widehat{L}(r) - r$ has a positive slope. On the contrary, in the case of B-capsids such an effect is much weaker with respect to the strength of the attraction (Figure 3a). This observation is enhanced by looking at the estimates of $L(r) - r$ in the multivariate case (Figure 3b). Here again we have a similar range of attraction (less than 100 nm) but the strength of the attraction has become very small compared to the wild type. Once more only one interval of standard errors is given for the multivariate case due to the equality of the L-functions based on (3.6).



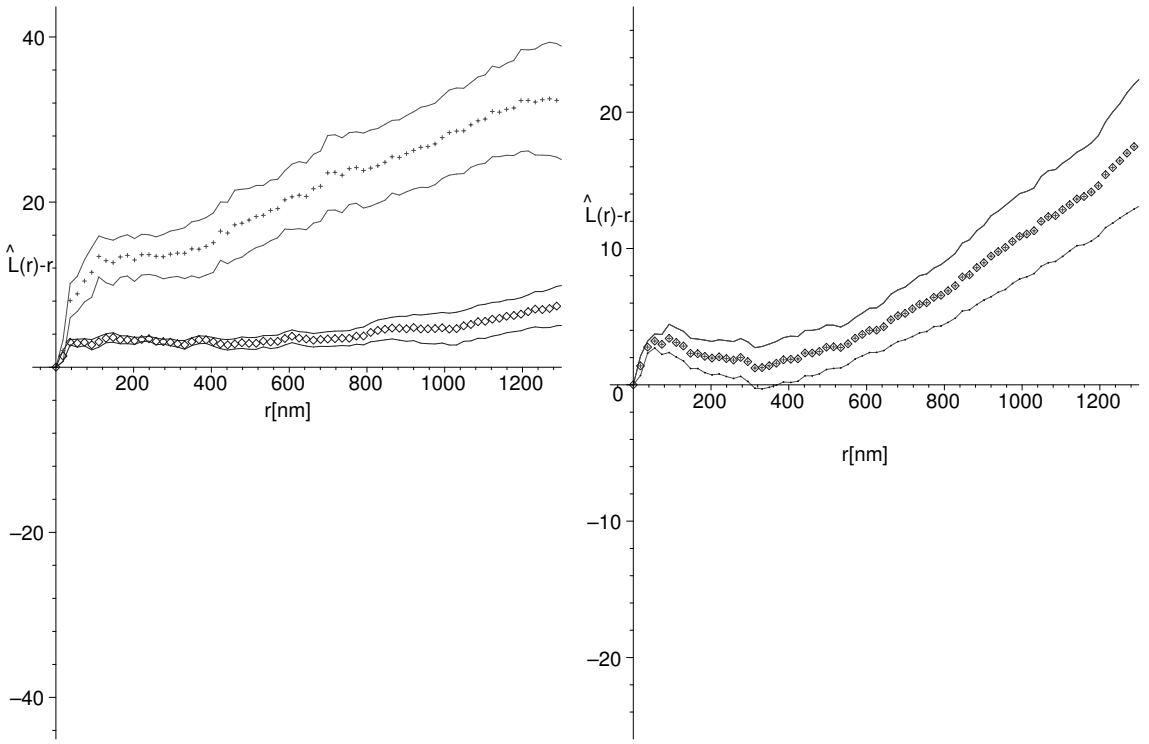
(a) *Univariate $\widehat{L}(r) - r$ functions for B-capsid ($\diamond$) and C-capsid (+)*

(b) *Multivariate $\widehat{L}(r) - r$ functions BC ($\diamond$) and CB (+)*

Figure 2: Averaged estimated functions $\widehat{L}(r) - r$ and standard errors for wild type

4.3 Estimates of nearest neighbor distance distribution functions

In Figures 4 and 5 the averaged estimated functions $\widehat{D}(r)$ are displayed which correspond to the L-functions analyzed in Section 4.2. Additionally the nearest neighbor distribution functions for the case of complete spatial randomness are displayed in order to facilitate an interpretation. Again a clustering behavior for small point-pair distances is clearly visible due to the fact that the slope of $\widehat{D}(r)$ is bigger than the slope of the corresponding theoretical function for the case of complete spatial randomness. For the wild type (Figure 4) this attraction seems to be slightly stronger for C-capsids than for B-capsids, while the estimated multivariate nearest neighbor distance distribution functions $\widehat{D}_{ij}(r)$ also indicate a clustering behavior, meaning that around a B(C)-capsid there is a tendency for clustering of C(B)-capsids. With regard



(a) *Univariate $\widehat{L}(r) - r$ functions for B-capsid ($\diamond$) and C-capsid (+)*

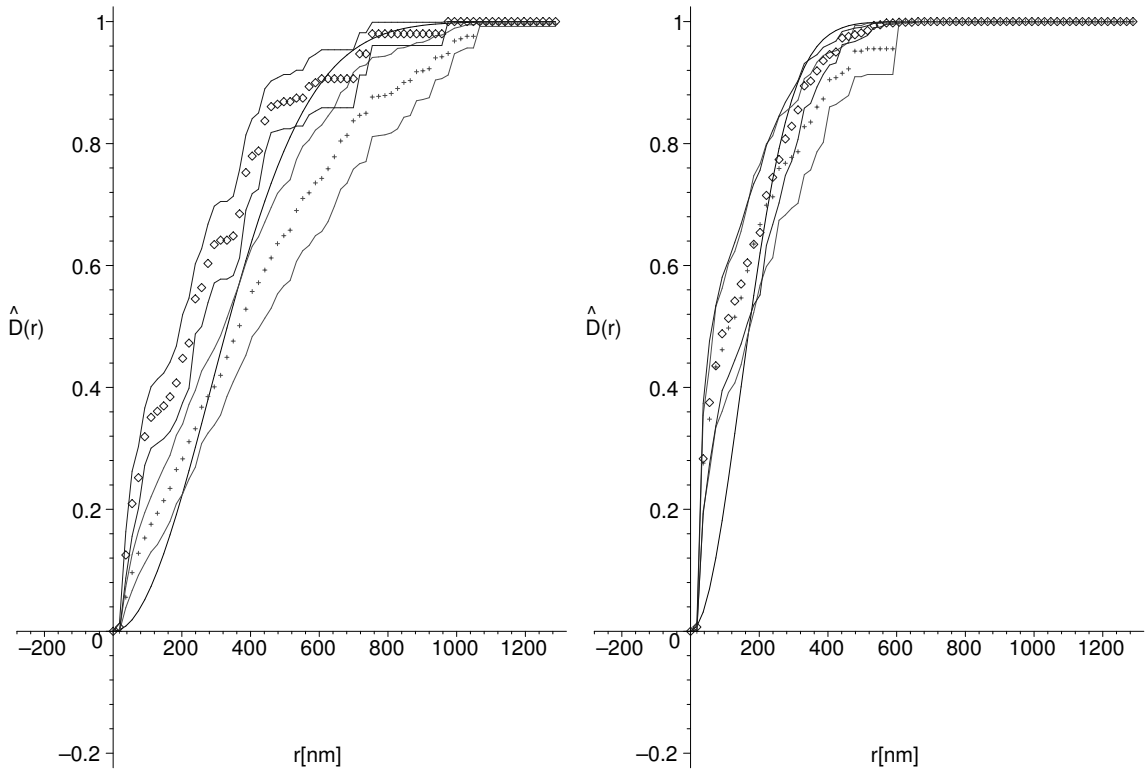
(b) *Multivariate $\widehat{L}(r) - r$ functions BC ($\diamond$) and CB (+)*

Figure 3: Averaged estimated functions $\widehat{L}(r) - r$ and standard errors for DM type

to the averaged estimated nearest neighbor distributions of the deletion mutant type (Figure 5) an observation similar to the observations made for the averaged estimates of the functions $L(r) - r$ can be stated. The attraction in the case of the C-capsids seems to be much stronger for small point-pair distances than for the B-capsids, where this effect has almost completely disappeared. By looking at the multivariate functions $\widehat{D}(r)$ the impression that the two underlying point processes show a dependence behavior is enhanced, since the averaged estimates are far away from the corresponding functions for the case of spatial independence of the two processes.

4.4 Investigations of other point process characteristics

In order to verify the results obtained in Sections 4.2 and 4.3 investigations have been performed for estimators of further point process characteristics like the pair



(a) $\widehat{D}(r)$ *functions for B-capsid* ($\diamond$) *and for C to B* (+)

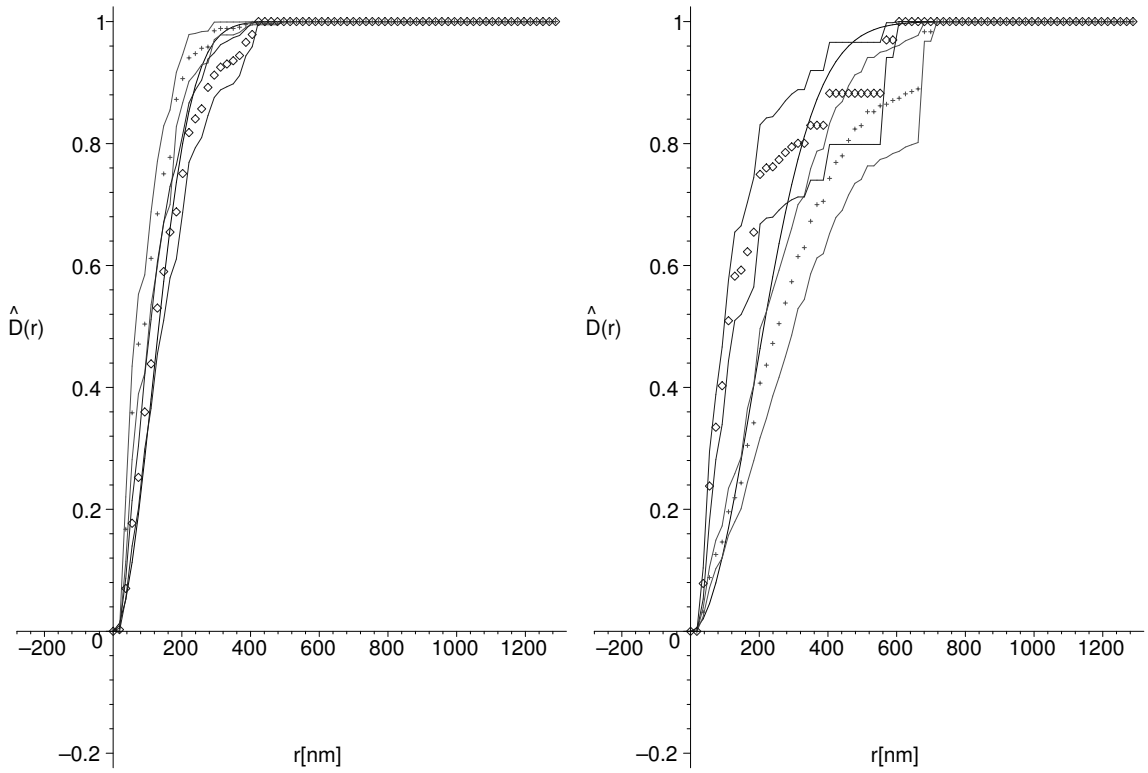
(b) $\widehat{D}(r)$ *functions for C-capsid* ($\diamond$) *and for B to C* (+)

Figure 4: Averaged estimated functions $\widehat{D}(r)$, standard errors, and corresponding theoretical distribution functions in the case of complete spatial randomness for wild type

correlation function $g(r)$, the spherical contact distribution function $H_s(r)$, the J-function $J(r)$ and the I-function $I(r)$ (van Lieshout and Baddeley, 1999). Here, results turned out to be quite similar compared to the results obtained for estimations of the L-function and the nearest neighbor distance distribution $D(r)$, both in the univariate as well as in the multivariate case. Therefore these results are omitted here.

4.5 Bootstrap tests on equality of functions

The L–functions and corresponding pointwise standard errors displayed in Figures 2 and 3 provide the impression that the averaged univariate functions $L(r) - r$ of B-capsids and C-capsids are significantly different for the mutant type, whereas there are no significant differences for the wild type. In order to formally test such hypotheses



(a) $\widehat{D}(r)$ *functions for B-capsid (⋄) and for C to B (+)*

(b) $\widehat{D}(r)$ *functions for C-capsid (⋄) and for B to C (+)*

Figure 5: Averaged estimated functions $\widehat{D}(r)$, standard errors, and corresponding theoretical distribution functions in the case of complete spatial randomness for DM type

bootstrap tests were performed on the equality of the averaged univariate functions $L(r) - r$ of B-capsids and C-capsids for the wild type and for the mutant type (cmp. Section 3.5). We used $k = 999$ simulated bootstrap samples and values of r_i between 20 nm and 1000 nm with a stepwidth of 20 nm. The outcome of the simulation tests shows that in the case of the mutant type there is a highly significant rejection ($p < 0.01$) of the hypothesis that the functions $L(r) - r$ are identical for B-capsids and C-capsids, whereas for the wild type the same hypothesis can not be rejected ($p > 0.2$).

5 Discussion

The virological interpretation of these results is difficult, since the molecular mechanism underlying nucleocapsid clustering is unknown as well as the exact function of pM97 in the viral morphogenesis. The steps of nucleocapsid morphogenesis are vDNA replication, procapsid formation, cleavage of the concatemeric vDNA to unit genomes, and packaging of the genomes into the capsid (Sheaffer et al., 2001). Thus, the formation of the mature C-capsid is completed by cleavage and packaging of the viral DNA into the B-capsid and sealing of the capsid (Gao et al., 1994; Church and Wilson, 1997; Chee et al., 1990; Rigoutsos et al., 2003).

The influence of the putative kinase encoded by the MCMV M97 on the processes of nucleocapsid maturation is unclear. For HCMV it has been shown that the homologous pUL97 phosphorylates the viral processivity factor UL44 (Krosky et al., 2003, Marschall et al., 2003) and also targets the nuclear lamina (Marschall et al., 2005). In MCMV deletion of M97 apparently leads to a strong reduction of nuclear C-capsids in favor of B-capsids and to an observable decrease of the nucleocapsid clustering. Since nuclei infected with the M97-deleted MCMV apparently accumulate packaging-competent B-capsids three hypotheses to explain this phenotype are possible: i) pM97 also has to phosphorylate one or more proteins involved in cleavage or packaging of the vDNA into the B-capsid. ii) The missing phosphorylation of M44 (positional homolog of UL44) may lead to the generation of shorter concatemers or even fragmented viral genomes leading to a reduced amount of packaging competent vDNA. iii) A different reason for clustering could be that the attraction of nucleocapsids is not dependent on DNA but on a putative nucleoskeleton that might be relevant for herpesvirus nuclear processes (Simpson-Holley et al. 2005, Forest et al, 2005). Accordingly, M97 could interact with a component involved in this process. Referring to a DNA-mediated clustering, a defect in capsid formation or sealing of the loaded capsid seems not very probable, since this would lead to an accumulation of unstable procapsids or empty

A-capsids, respectively. According to the first hypothesis, a failure of DNA-processing would also lead to an accumulation of the replicated vDNA, while the second hypothesis would call for shorter or fragmented concatemers and genomes. In favor of the second hypothesis it could be speculated that the long concatemeric vDNA produced by the wild type MCMV may form binding sites for several B-capsids, thus giving rise to the observed clusters. Accordingly, the generation of short concatemers or genome fragments that contain less binding sites and small diffusible fragments competing for binding sites would reduce capsid clustering. This hypothesis could be confirmed by pulse-field gel electrophoresis and Southern blotting of the replicated genomes. In case that clustering is mediated by the putative nucleoskeleton, the skeletal structure would have to be shown. Unfortunately, the direct visualisation of this putative nucleoskeleton was not yet possible.

Note that the total amount of ignored capsids was rather small (approx. 5%). It consisted of A–capsids (approx. 1%) and of capsids that could not be clearly assigned to one class (approx. 4%). Since the A–capsids may be regarded as by–products unable to undergo futher maturation we think that it is justified to ignore them for the analysis. Finally, the amount of capsids that could not be clearly assigned to a class has been so small (approx. 2–4 per image) such that we decided to ignore them in our analysis.

Throughout this paper we assumed stationarity and isotropy for the regarded point patterns. Some formal statisticial tests have been performed in order to check whether there is stationarity and isotropy e.g. with respect to the distance to the center of the cell nucleus. Although the results of such tests are providing hints in favor of these assumptions, we would like to regard these two properties only as working hypotheses. This is due to the fact that the numbers of points per sampling region might not be sufficiently large enough for definitive statements with respect to stationarity and isotropy.

A similar argument as for the cases of stationarity and isotropy might be provided with

regard to possible tests for complete spatial randomness. Apart from this fact, it was not the main focus of this paper to check whether the point patterns are completely spatially random but to quantify differences between different cell types and different capsids within the same cell type. Therefore we decided to omit tests for complete spatial randomness and instead applied tests that are based on bootstrapping and Monte-Carlo simulation in order to check whether two functionals are equal or not.

Although the point patterns are finite and bounded it should be noticed that we assume that they are realizations of stationary point processes restricted to a bounded sampling region. It is quite common to assume that observed realizations are generated by unbounded stationary point processes restricted to a bounded sampling region, because very often data behaves in rather different and non-stationary way outside of a finite sampling region (Beil et al., 2005; Diggle et al., 2000; Schladitz et al., 2003).

Other types of edge-correction techniques might be applicable to this data as well as other types of estimators apart from the Horvitz-Thompson type estimators used in this paper. For example estimators of Kaplan-Meier type might be used (Baddeley and Gill, 1997). Especially in the case of B-capsids for the wild type the amount of points per sample region does not seem to be sufficiently large in order to obtain reliable results for individual samples. This relatively small amount of points together with a natural variability of the biological data is the main reason that instead of regarding estimates for individual regions averages for the two different groups are regarded. To justify the consideration of averaged estimates of the point process characteristics for the two different groups, it is assumed that point patterns for the same group are realizations of independent and identically distributed point processes. Such a technique of regarding averages instead of individual estimates usually improves the accuracy of the estimators (Stoyan et al., 1995). The averaging of individual estimates for point process characteristics in this paper was performed arithmetically. Other methods of averaging are of course thinkable, e.g., by using weights that are proportional to the numbers of points per sample or to the estimated intensities (Diggle

et al., 1991; Diggle et al., 2000). Suffice it here to say that we also considered other methods, but obtained similar results and therefore decided to stick to arithmetic means for simplicity reasons.

Other ways of analyzing the occurring point patterns are of course thinkable. For example a fitting of parameters of an underlying Gibbs point process model by usage of the pseudolikelihood might be useful (Baddeley et al., 2006). Such a method could provide additional knowledge about the behavior of the observed point patterns as well as validate the findings of this study.

In summary, we show that by combining electron microscopic imaging methods with statistical analysis we are able to quantitatively approach structural phenotypes based on unknown molecular interactions and thus generate new hypotheses on nucleocapsid maturation processes, but which have to be proven by further molecular and electron microscopic analyses.

Acknowledgements

The authors are grateful to Viktoria Idt for her assistance in the computations that led to the statistical analysis part. Furthermore, we thank M. Beil, S. Eckel, T. Mattfeldt and the two referees for helpful comments and suggestions.

References

Azzeh, M., Honigman, A., Taraboulos, A., Wolf, D. G. (2005) *The role of human cytomegalovirus UL97 kinase in cellular localization of viral tegument proteins.* In 30th International Herpesvirus Workshop. Turku, Finnland.

Baddeley, A. J. (1998) *Spatial sampling and sensoring.* Chapter 2 of *Current Trends in Stochastic Geometry and its Applications* edited by Kendall, W. S., Lieshout, M. N. M. van, Barndorff-Nielsen, O. E., Chapman and Hall, London, New York.

Baddeley, A. J., Gill, R. D. (1997) *Kaplan-Meier estimators of distance distribution for spatial point processes.* The Annals of Statistics **25**, No. 1, 263-292.

Baddeley, A. J., Gregori, P., Mateu, J., Stoica, R., Stoyan, D. (2006) *Case Studies in Spatial Point Process Modelling.* Springer Lecture Notes in Statistics **185**. Springer, Berlin.

Baddeley, A. J., Moyeed, R. A., Howard, C. V., Boyde, A. (1993) *Analysis of a three-dimensional point pattern with replication.* Applied Statistics **42**(4), 641-668.

Beil, M., Fleischer, F., Paschke, P., Schmidt, V. (2005) *Statistical analysis of 3D centromeric heterochromatin structure in interphase nuclei.* Journal of Microscopy **217**, 60-68.

Buser, C., Walther, P., Mertens, T., Michel, D. (2007) *Cytomegalovirus primary envelopment occurs at large infoldings of the inner nuclear membrane.* Journal of Virology (in press).

Chee, M. S., Lawrence, G. L. Barrell, B. G. (1989) *Alpha-, beta- and gammaherpesviruses encode a putative phosphotransferase.* Journal of General Virology **70**(5), 1151-1160.

Church, G. A., Wilson, D. W. (1997) *Study of herpes simplex virus maturation during a synchronous wave of assembly.* Journal of Virology **71**(5), 3603-3612.

Diggle, P., J. (2003) *Statistical Analysis of Spatial Point Patterns.* Arnold, London.

Diggle, P., J., Lange N. , Benes F. B. (1991) *Analysis of variance for replicated spatial point patterns in clinical neuroanatomy.* Journal American Statistical Association **86**, 618-625.

Diggle, P. J., Mateu. J., Clough, H. E. (2000) *A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns.* Advances in Applied Probability **32**, 331-343.

Forest, T., Barnard, S., Baines, J. D. (2005) *Active intranuclear movement of herpesvirus capsids.* Nature Cell Biology **7**(4), 429-431.

Gao, M., Matusick-Kumar, L., Hurlburt, W., DiTusa, S. F., Newcomb, W. W., Brown, J. C., McCann, P. J., Deckman, I., Colonno, R. J. (1994) *The protease of herpes simplex virus type 1 is essential for functional capsid formation and viral growth.* Journal of Virology **68**(6), 3702-3712.

Hanisch, K.-H. (1984) *Some remarks on estimators of the distribution function of nearest-neighbor distance in stationary spatial point patterns.* Statistics **15**, 409-412.

Howard, C. V., Reid, S., Baddeley, A. J., Boyde, A. (1985) *Unbiased estimation of particle density in the tandem-scanning reflected light microscope.* Journal of Microscopy **138**, 203-212.

Krosky, P. M., Baek, M. C., Coen, D. M. (2003) *The human cytomegalovirus UL97 protein kinase, an antiviral drug target, is required at the stage of nuclear egress.* Journal of Virology **77**(2), 905-914.

Krosky, P. M., Baek, M. C., Jahng, W. J., Barrera, I., Harvey, R. J., Biron, K. K., Coen, D. M., Sethna, P. B. (2003) *The human cytomegalovirus UL44 protein is a substrate for the UL97 protein kinase.* Journal of Virology **77**(14), 7720-7727.

Lieshout, M. N. M. van, Baddeley, A. J. (1999) *Indices of dependence between types in multivariate point patterns.* Scandinavian Journal of Statistics **26**, 511-532.

Marschall, M., Freitag, M., Suchy, P., Romaker, D., Kupfer, R., Hanke, M., Stamminger, T. (2003) *The protein kinase pUL97 of human cytomegalovirus interacts with and phosphorylates the DNA polymerase processivity factor pUL44.* Virology **311**(1), 60-71.

Marschall, M., Marzi, A., Aus dem Siepen, P., Jochmann, R., Kalmer, M., Auerochs, S., Lischka, P., Leis, M., Stamminger, T. (2005) *Cellular p32 recruits cytomegalovirus kinase pUL97 to redistribute the nuclear lamina.* Journal of Biology Chemistry **280**, 33357-33367.

Mattfeldt, T., Eckel, S., Fleischer, F., Schmidt, V. (2006) *Statistical analysis of reduced pair correlation functions of capillaries in the prostate gland.* Journal of Microscopy **223**, 107-119.

Mayer, J., Schmidt, V., Schweiggert, F. (2004) *A unified simulation framework for spatial stochastic models.* Simulation Modelling Practice and Theory **12**, 307-326.

Mettenleiter, T.C. (2002) *Herpesvirus assembly and egress.* Journal of Virology **76**(4), 1537-1547.

Murphy, E., Yu, D., Grimwood, J., Schmutz, J., Dickson, M., Jarvis, M. A. Hahn, G., Nelson, J. A., Myers, R. M., Shenk, T. E. (2003) *Coding potential of laboratory and clinical strains of human cytomegalovirus.* Proceedings of the National Academy of Sciences USA, **100**(25), 14976-14981.

Rigoutsos, I., Novotny, J., Huynh, T., Chin-Bow, S. T., Parida, L., Platt, D. Coleman, D., Shenk, T. (2003) *In silico pattern-based analysis of the human cytomegalovirus genome.* Journal of Virology **77**(7), 4326-4344.

Schladitz, K., Särkkä, A., Pavenstädt, I., Haferkamp, O., Mattfeldt, T. (2003) *Statistical analysis of intramembranous paricles using fracture specimens.* Journal of Microscopy **211**, 137-153.

Sheaffer, A. K., Newcomb, W. W., Gao, M., Yu, D., Weller, S. K., Brown, J. C. Tenney, D. J. (2001) *Herpes simplex virus DNA cleavage and packaging proteins associate with the procapsid prior to its maturation.* Journal of Virology **75**(2), 687-698.

Simpson-Holley, M., Colgrove, R. C., Nalepa, G., Harper, J. W., Knipe, D. M. (2005) *Identification and functional evaluation of cellular and viral factors involved in the alteration of nuclear architecture during herpes simplex virus 1 Infection.* Journal of Virology **79**(20), 12840-12851.

Stoyan, D., Kendall, W., S., Mecke, J. (1995) Stochastic Geometry and its Applica-

tions. J. Wiley & Sons, Chichester.

Stoyan, D., Stoyan, H. (2000) *Improving ratio estimators of second order point process characteristics.* Scandinavian Journal of Statistics **27**, 641-656.

Stoyan, D., Stoyan, H., Tscheschel, A., Mattfeldt, T. (2001) *On the estimation of distance distribution functions for point processes and random sets.* Image Analysis and Stereology **20**, 65-69.

Wagner, M., Michel, D., Schaarschmidt, P., Vaida, B., Jonjic, S., Messerle, M., Mertens, T., Koszinowski, U. (2000) *Comparison between human cytomegalovirus pUL97 and murine cytomegalovirus (MCMV) pM97 expressed by MCMV and vaccinia virus: pM97 does not confer ganciclovir sensitivity.* Journal of Virology **74**(22), 10729-10736.

Wolf, D. G., Courcelle, C. T., Prichard, M. N., Mocarski, E. S. (2001) *Distinct and separate roles for herpesvirus-conserved UL97 kinase in cytomegalovirus DNA synthesis and encapsidation.* Proceedings of the National Academy of Sciences USA **98**(4), 1895-1900.