附件一

行政院國家科學委員會補助專題研究計畫 □ 成 果 報 告 ■期中進度報告

低維奈米系統電傳輸與自組成性質之數值研究

計畫類別：■ 個別型計畫　□ 整合型計畫
計畫編號：NSC 97－2112－M－029－002－MY3
執行期間：　97 年 8 月 1 日至 100 年 7 月 31 日

計畫主持人：施奇廷
共同主持人：許經麥
計畫參與人員： 楊嘉會、鄭詠尹、王筑宣

成果報告類型(依經費核定清單規定繳交)：■精簡報告　□完整報告

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
□出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢
□涉及專利或其他智慧財產權，□一年□二年後可公開查詢

執行單位：東海大學物理系

中　華　民　國　99 年　5　月　31　日

研究進度與成果說明

本計畫為三年期計畫之第二年計畫，茲就本年度之研究成果、執行進度說明如下。

研究成果：

審查中之期刊論文四篇（論文全文附於文末）：

1. C. T. Shih, Y. Y. Cheng, S. Wells, Rudolf A. Römer, and C. L. Hsu, 2010, *Charge Transport in Cancer-Related Genes and Early Carcinogenesis*, submitted to *Computer Physics Communication*.
2. C. T. Shih, Hsuan-Wen Lin, Ann-Shyn Chiang, 2010, *Statistical Analysis and Modeling of the Temperature-Dependent Sleep Behavior of Drosophila*, submitted to Computer Physics Communication.
3. Chia-Hei Yang, Ching-Ling Hsu, Nan-Yow Chen, and Chi-Tin Shih, 2010, *Temporal Dynamics of Site Percolation in Nanoparticle Assemblies*, submitted to *Computer Physics Communication.*
4. Ann-Shyn Chiang, Chih-Yung Lin, Chao-Chun Chuang, Hsiu-Ming Chang, Chang-Huain Hsieh, Chang-Wei Yeh, Chi-Tin Shih, Jian-Jheng Wu, Guo-Tzau Wang, Yung-Chang Chen, Cheng-Chi Wu, Guan-Yu Chen, Yu-Tai Ching, Ping-Chang Lee, Chih-Yang Lin, Hui-Hao Lin, Chia-Chou Wu, Hao-Wei Hsu, Yun-Ann Huang, Jing-Yi Chen, Hsin-Jung Chiang, Chun-Fang Lu, Ru-Fen Ni, Chao-Yuan Yeh, & Jenn-Kang Hwang, 2010, *A mesoscopic atlas of brainwide wiring networks in Drosophila at single cell resolution*, submitted to *Science*.

研究進度：

1. 去年度在計算 DNA 傳輸性質的緊束縛模型中加入了代表 DNA 雙螺旋特性的電子躍遷項 $t_{\perp}$後，我們將這個新模型、以及原來的一維模型應用在 OMIM@NCBI（Online Mendelian Inheritance in Man, National Center for Biotechnology Information）與 HGMD

（Human Gene Mutation Database）兩個人類基因突變與疾病資料庫中的 35 個癌症相關的基因，發現仍能符合我們在 2008 年發表的論文（C. T. Shih *et al*., 2008, Phys. Rev. Lett. **100**, 018105）所提出的基因早期突變／修復的機制，此部分結果已投稿至 Computer Physics Communication。

2. 我們繼續收集了 OMIM 與 HGMD 內與所有人類疾病相關的基因（包含十個以上的致病點突變），發現符合我們的理論的基因超過 85%，顯示這個理論有可能是解釋早期致病機制的普遍理論，這部份結果已在撰寫中。

3. 我們以 Monte Carlo 模擬探討膠體溶液內的奈米金粒子在溶劑乾燥過程中，形成自我組成的網路結構的展透行為，決定了不同的溫度以及化學勢對展透臨界密度的影響，此部分結果已投稿至 Computer Physics Communication。

4. 在與清華大學腦科學中心的合作研究，我們分析了果蠅的睡眠行為實驗數據，發現溫度對睡眠行為有顯著的影響，並且提出一個理論模型來模擬其調控睡眠之神經元的活動，此部分結果已投稿至 Computer Physics Communication。

5. 此外，腦科學中心利用高解析度之果蠅腦神經元影像建立了「FlyCircuit」資料庫並據此建構了果蠅腦部神經元連結網路的草圖，我們利用統計物理方法分析了此網路的特性。此部分結果已投稿至 Science。

下年度的計畫

1. 由於模型為不精確之近似模型，故必需計算許多模型及隨機序列，並對所得結果進行更嚴密的統計分析。

2. 對人類的單核苷酸變異多性狀（Single Nucleotide Polymorphism）資料庫進行電性分析，比較其與致病突變之間的關係。

3. 藉由引入各種特徵參數（周長面積比、金粒子團簇分佈、歐拉特徵值等）探討膠體奈米金粒子的網路形成過程中，網路形貌隨時間的變化。

4. 隨著果蠅腦神經網路圖譜逐漸完成，對其進行標準之複雜網路分析，並且嘗試開發針對腦神經網路專屬的分析工具，加強與實驗生物學家之討論交流，讓此一跨領域研究平台更加完善。

Charge Transport in Cancer-Related Genes and Early Carcinogenesis

Chi-Tin Shih[a], Yun-Yin Cheng[a], Stephen A. Wells[b], Ching-Ling Hsu[c], Rudolf A. Römer[b]

[a] *Department of Physics, Tunghai University, 40704 Taichung, Taiwan and The National Center for Theoretical Sciences, 30013 Hsinchu, Taiwan*

[b] *Department of Physics and Centre for Scientific Computing, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, UK*

[c] *Department of Physics, Chung-Yuan Christian University, Chung-Li, Taiwan*

Abstract

The electronic transmission properties of DNA molecules are believed to play a significant role in many physical phenomena taking place in living organisms [1]. Here we study the charge transport (CT) properties of cancer-related genes, including some of the most important tumor suppressors. We find that the changes in averaged CT around the sites of pathogenic and cancerous mutations are statistically smaller than those on sites where pathogenic mutations have not been observed. The results suggested that CT might be an indicator to discriminate between pathogenic and non-pathogenic mutations at an early stage. Mutations which cause little change in CT may be more likely to occur, or more likely to be missed by damage-repair enzymes which probe CT, and are therefore more likely to persist and cause disease.

Keywords: deoxyribonucleicacid, charge transport, cancer, mutation *Revision* : 1.27

We study the relation between the point mutations and CT properties of some of the most important tumor suppressors, together with other cancer-associated genes. These genes regulate cell proliferation by monitoring various molecular signaling pathways. Disfunction of the tumor suppressors will cause abnormal cell proliferation and the development of cancers [2]. CT through DNA is inhibited at the damaged sites of the sequence, owing to misalignments of base pair π-stacking [1]. The base excision repair (BER) enzymes such as Endonuclease III and MutY are believed to efficiently locate the DNA base lesions or mismatches by probing the inhibition of the DNA-mediated CT due to the damages [3].

We have previously studied [4] the CT properties of the genomic sequence of the $p53$ tumor suppressor gene, which is known as the "guardian of the genome". The results show that on average the cancerous mutations of the gene yield smaller changes of the CT in contrast with non-cancerous mutations. Based on such a behavior of CT, we proposed a possible scenario of how cancerous mutations might circumvent the DNA damage-repair mechanism and survive to yield carcinogenesis [4]. Here we will add to this analysis by showing that the $p53$ gene with a randomly rearranged sequence does not show this effect at all. This test has also been applied to many other genetic sequences.

The native genetic sequences and mutations of cancer-related genes are retrieved from four different databases [5, 6]. The genomic sequence of a gene with length $\mathcal{N}$ base pairs (bps) is defined as $(s_1, s_2, \cdots, s_{\mathcal{N}})$. Each point mutation of a given gene is characterized by the set (k, s), where k and s are the position of the point mutation in the genomic sequence and the mutant nucleotide which replaces the nucleotide s_k of normal DNA, respectively.

The simplest model of coherent hole transport in DNA is given by an effective one-dimensional

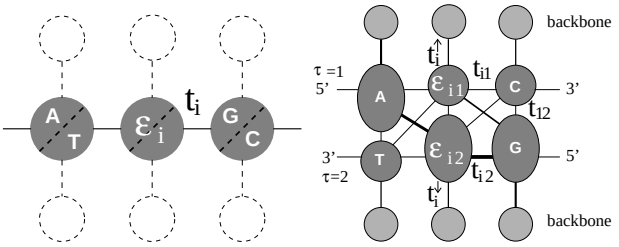


Figure 1: Schematic models for charge transport in DNA. The nucleobases are given as (dark grey) circles and ellipses. Electronic pathways are shown as solid lines of varying thickness to indicate variation in strength. The left model indicates the one-channel model where the sugar-phosphate backbone is ignored. In the right model, circles denote the smaller pyrimidines, ellipses are the large purines and (light grey) circles denote the sugar-phosphate backbone sites. Note that diagonal hopping between purines is favored, and between pyrimidines is disfavored, by the larger size of the purines.

Hückel-Hamiltonian for CT through nucleotide HOMO states [1], where each lattice point represents a nucleotide base (A,T,C,G) of the chain for $n = 1, \ldots, N$. In this tight-binding formalism, the on-site potentials ϵ_n are given by the ionization potentials $\epsilon_{\rm G} = 7.75e$V, $\epsilon_{\rm C} = 8.87e$V, $\epsilon_{\rm A} = 8.24e$V and $\epsilon_{\rm T} = 9.14e$V, at the nth site, cp. Fig. 1; $t_{n,n+1}$ is assumed to be nucleotide-independent with $t_{n,n+1} = 0.4e$V [1]

A model which is less coarse-grained is provided by the diagonal ladder model shown in Fig. 1. Here, both strands of DNA and the backbone are modelled explicitly and the different diagonal overlaps of the larger purines (A,G) and the smaller pyrimidines (C,T) are taken into account by suitable inter-strand couplings [7]. The intra-strand couplings are $0.35e$V between identical bases and $0.17e$V between different bases; the diagonal inter-strand couplings are $0.1e$V for purine-purine, $0.01e$V for purine-pyrimidine and $0.001e$V for pyrimidine-pyrimidine. Perpendicular couplings to the backbone sites are $0.7e$V, and perpendicular hopping across the hydrogen bond in a base pair is reduced to $0.005e$V.

The transmission coefficient $T(E)$ for a DNA sequence with length Nbps for different injection energy can be calculated for both models by using the transfer matrix method (TMM) [8, 9]. The position-dependent averaged transmission coefficient at the $j-$th base pair for transmission length L bps is defined as

$$\bar{T}_{j,L} = \frac{1}{L} \sum_{n=j-L+1}^{j} \int_{E_0}^{E_1} \frac{T_{n,L}(E)}{E_1 - E_0} dE \qquad (1)$$

where $T_{n,L}(E)$ is the transmission coefficient of a subsequence of the gene with length L and starting from the position n. n ranges from $j - L + 1$ to j such that the subsequence contains the $j-$th base pair. E_0 and E_1 are the lower and upper bounds of the incident energy of the carriers, e.g. for the $1D$ model used here, the values are 5.75 and $9.75e$V, respectively; for the diagonal model the bounds are 7 and $11e$V. Then we examine the difference between $T(E)$ of the normal and mutated genomic sequence of a point mutation [4]

$$\bar{\Delta}_{j,L}^{k,s} = \int_{E_0}^{E_1} \frac{|T_{j,L}(E) - T_{j,L}^{k,s}(E)|^2}{E_1 - E_0} dE \qquad (2)$$

where $T_{j,L}^{k,s}(E)$ is the transmission coefficient of the same segment of DNA but with the point mutation (k, s). $\Gamma(k, s; L)$ is the averaged effect of the point mutation (k, s) on CT properties for all subsequences containing the mutation and with length L. We make use of values of L between 10 and 60 bps.

There are a total of $3\mathcal{N}$ possible point mutations of a gene with $\mathcal{N}$bps, a set which we denote as M_a. The subset of pathogenic mutations is M_p. We calculate the averaged CT effect, $\Gamma(k, s; L)$ for the full set M_a and compare the distributions for M_a and M_p. Fig. 2 shows the distributions of Γ for the TP53 sequence of the $p53$ gene. For the 1-D model the full distribution is close to log-normal. For both the 1-D and diagonal models the distribution for pathogenic mutations is visibly shifted towards lower values compared to the full distribution. This difference disappears when the sequence is randomised.

We obtain a *local ranking* (LR) for each pathogenic mutation (k, s) by comparing its CT change $\Gamma(k, s; L)$ to thsoe for the other two potential mutations at the same position, obtaining a ranking $\gamma_{\rm LR}(k, s; L)$ for the mutation (k, s) of 1,2 or 3 for lowest, middle and highest CT change.

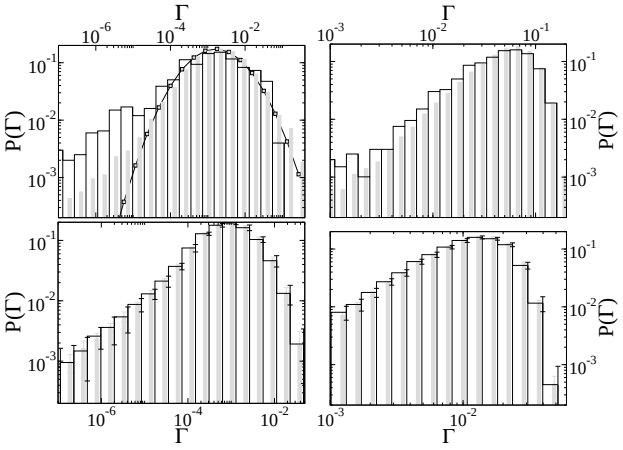


Figure 2: Distribution of the change in charge transport Γ for pathogenic (white bars) and all possible (grey bars) mutations for the $p53$ DNA strand with 20304 base pairs. The left/right columns show results for the 1D/diagonal models while the top/bottom rows indicate original/shuffled DNA sequences. For the randomly shuffled cases, error bars are calculated using 20 different sequence realizations. All results are for $L = 40$. The line for the 1-D case is a fit to a log-normal distribution.

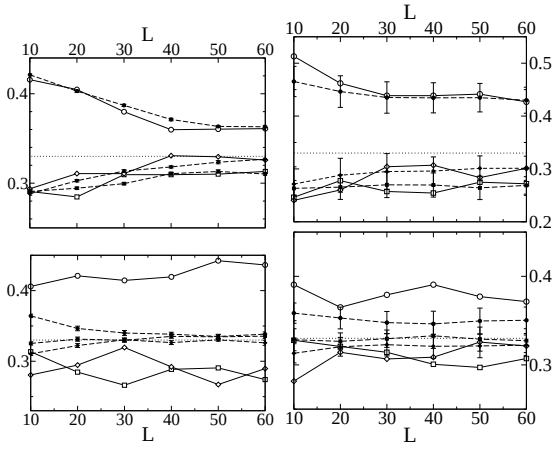


Figure 3: Distribution of the *local*, LR (top row) and *global*, GR (bottom row) ranking results of pathogenic mutations of $p53$ (open symbols) as a function of window lengths L. Circles indicate the proportion of pathogenic mutations with $\gamma = 1$; squares, $\gamma = 2$; diamonds, $\gamma = 3$. The closed symbols indicate averaged results for 20 randomly shuffled sequences. The left/right columns distinguish results for the 1D/diagonal models. The dotted horizontal lines show the 33% mark expected for a uniform random distribution.

Those k with more than one pathogenic mutations are excluded in the LR analysis. Similarly, we define a *global ranking* (GR) by ranking the normalized CT changes for *all possible* $3\mathcal{N}$ mutations. We assign global ranks $\gamma_{\rm GR}(k, s; L)$ of 1 for the lowest third of the global distribution of CT changes, 2 for the middle third and 3 for the top third. In Fig. 3 we show the incidences of the three possible local and global ranks for the cancerous mutations of $p53$. In local rankings, the incidence of pathogenic mutations with lowest CT change, $\gamma_{\rm LR} = 1$, is significantly higher than for $\gamma_{\rm LR} = 2, 3$. However, this result is barely distinguishable from the distributions of local rankings for shuffled sequences. In the global ranking, however, the incidence of pathogenic mutations with $\gamma_{\rm GR} = 1$ is higher than for $\gamma_{\rm GR} = 2, 3$ and the distribution for the native sequence *is* distinguishable from the distributions for shuffled sequences.

In total, we have performed these statistical tests for 35 cancer DNA sequences [5, 6]. The results [10] show that in about 95% of the DNA strands the pathogenic mutations are biased towards smaller changes in CT properties. Hence it seems very likely the scenario proposed here for early pathogenesis can be applied for an analysis of a wide range of genetic diseases.

This work was supported by the National Science Council in Taiwan (97-2112-M-029-002-MY3), the UK Leverhulme Trust (F/00215/AH) and the National Center for High-Performance Computing in Taiwan.

References

[1] Chaborty, T., editor, *Charge Migration in DNA: Perspectives from Physics, Chemistry and Biology*, Springer Verlag, Berlin, 2007.
[2] Sherr, C. J., Cell **116** (2004) 235.
[3] Rajski, S. R., Jackson, B. A., and Barton, J. K., Mutat. Res. **447** (2000) 49.
[4] Shih, C.-T., Roche, S., and Römer, R. A., Phys. Rev. Lett. **100** (2008) 018105.
[5] Steson, P. D. et al., Hum. Mutat. **21** (2003) 577, http://www.hgmd.cf.ac.uk/ac/index.php.
[6] Petitjean, A. et al., Hum. Mutat. **28** (2007) 622, http://www-p53.iarc.fr/index.html, R11.
[7] Wells, S. A., Shih, C.-T., and Römer, R. A., Int. J. Mod. Phys. B (2009) 4138.
[8] Roche, S., Phys. Rev. Lett. **91** (2003) 108101.
[9] Ndawana, M. L., Römer, R. A., and Schreiber, M., Europhys. Lett. **68** (2004) 678.
[10] Shih, C.-T., Cheng, Y.-Y., Wells, S. A., Hsu, C.-L., and Römer, R. A., (2010), unpublished.

Temporal Dynamics of Site Percolation in Nanoparticle Assemblies

Chia-Hei Yang[a], Ching-Ling Hsu[b], Nan-Yow Chen[c], Chi-Tin Shih[a,*]

[a] *Department of Physics, Tunghai University, Taichung 40704, Taiwan*
[b] *Department of Physics and Center for Nanotechnology, Chung-Yuan Christian University, Chung-li 32023, Taiwan*
[c] *Institute of Physics, Academia Sinica, Taipei 11529, Taiwan*

Abstract

This study examined a sub-monolayer disordered nanoparticle network film constructed from colloidal self-assembly using a two-dimensional site percolation model in a dynamical fashion. Applying free boundary conditions upon a lattice gas model, numerical results from Monte Carlo simulations showed that the 2D site-percolation threshold would reduce to 0.1, during time-evolving processes. The results are qualitatively consistent with the experiments. We also introduce computational homology of the morphological self-assembled patterns to inspect the temporal dynamics in quantification.

Keywords: Self-Assembly, Site Percolation, Monte Carlo Simulation

The morphology of deposited nanoparticles gives evidence of the solvent flow history prior to complete evaporation. Experimentally, colloidal nanoparticles assembly formed via deposition from solution frequently comprised patterns on multiple length scales resulting from various solvent de-wetting mechanisms. The morphological patterns are observed in the numerical study of two-dimensional lattice gas model that were first introduced by Rabani *et al* [1] via Monte Carlo simulations. Recently, a non-equilibrated system has recently been demonstrated for thin-film solutions of passivated nano-crystals during the irreversible evaporation of the solvent and capable of exhibiting complex transitory structures, even when equilibrium fluctuations are mundane.

The numerical simulation of percolation [2] has been well studied for decades since its first proposition in 1941. Conventionally, such investigations are mainly focused on simple-sampling mechanism only without any consideration of dynamical process. However, a dramatic example of a non-equilibrated system has recently been demonstrated for thin-film solutions of passivated nano-crystals during the irreversible evaporation of the solvent. It is to emphasis that such systems far from equilibrium can exhibit complex transitory structures, even when equilibrium fluctuations are mundane. Following this study, it is reported by Hsu *et al* [3] that the site percolation threshold in a two-dimensional disordered lattice gas model constructed from

*Corresponding author.
Email address: ctshih@thu.edu.tw (Chi-Tin Shih)

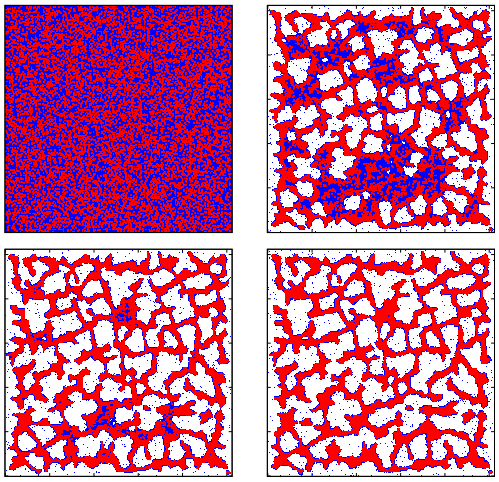


Figure 1: Four frames of nano patterns at the 0th, 200th, 300th and 500th of 500 Monte Carlo sweeps. One Monte Carlo sweep (MCS) is defined as one solvent cycle followed by one nanoparticle cycle. Lattice Size is 512×512 and nanoparticle size is 3×3. $T = 0.4$ and Coverage of nano cells is 0.30. Types of cells are labeled by Blue(liquid solvent cell area), Red (nano cell area) and White(dry cell area).

colloidal self-assembly is 40%, lower than the conventional value. Therefore, an intensified study with tactics from statistical physics on the subject self-assembly system is motivated.

The Hamiltonian of a lattice gas model in simulation[1] is:

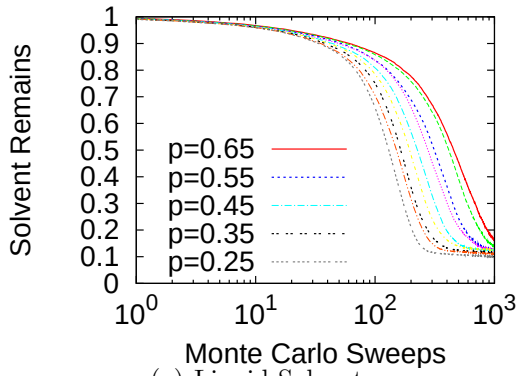
$$E = -\varepsilon_{ll} \sum_{i,j} l_i l_j - \varepsilon_{nl} \sum_{i,j} n_i l_j - \varepsilon_{nn} \sum_{i,j} n_i n_j + \mu \sum_i l_i. \quad (1)$$

Here, n, l are cell numbers of nano and solvent liquid. ε_{ll}, ε_{nl}, and ε_{nn} are, respectively, the values of interaction energy between two cells of liquid, liquid-nano, and nano-nano. $\mu = \mu_0$ is the value of chemical potential. To ensure that nanoparticles are well solvated by the liquid prior to solvent de-wetting, the values of attractive interaction energy are chosen as: $\varepsilon_{ll} \equiv 1$, $\varepsilon_{nl} = 1.5$, and $\varepsilon_{nn} = 2.0$.[4] By expressing thermal energy $k_B T$ and chemical potential μ in units of ε_{ll}, we consider only two independent energy scales. Fluctuations in solvent density are controlled by the Metropolis algorithm.[5] Nearest and next-nearest neighbors are all taken into account when calculating the energy change associated with a particular solvent transition.[6] For the convenience to compare with other studies which use the global energy in the nearest-neighbor model, here the values of the interaction energy summing from nearest and next-nearest neighbors are therefore multiplied by a re-normalization factor 58.6%.

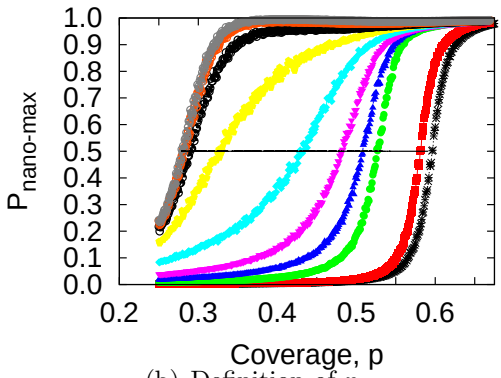
The evolved 2D site-percolation threshold reduces to 0.4 in Ref. [3], during time-evolving processes. However, after a long time of realization, the nanoparticle would aggregate toward the central part of the board and prevent percolation events. The late-time percolation threshold identified by crossing probability would be questionable. An alternative from the concept of percolation probability [2] is improvised. At every Monte Carlo sweep, we calculate the value of $P_{nano-max}$: the probability of an arbitrary nanoparticle belonging to the maximum nano cluster, and plot it as a sigmoid function through out all possible occupied probability of nanoparticles, in Fig. 2(b).

To represent the sigmoid curve position during the dynamic process in Fig. 2(b), we define the value of p_c at which value the probability of one arbitrary nano cell within the largest nano cluster is 0.5. It describes the overall behaviour of a nano self-assembly system in temporal dynamics, as in Fig. 3.

In this work we are concerned only with simulation morphologies before significant coarsening has occurred and after the initial



(a) Liquid Solvent



(b) Definition of p_c

Figure 2: Tracking the Sigmoid curve position of the nano clustering probability in time. $T = 0.4$ and $\mu = 2.25$. Lattice size is 512×512. We use open boundary conditions. Fig. 2(a) The fraction of liquid solvent for different values of nano coverage during the time evolution processes. Fig. 2(b) $P_{nano-max}$, the average probability that one arbitrary nano cell within the maximum nano cluster of 200 samples. The probability curve shifts from right at MCS=0 to left at MCS=400. The quantity of $P_{nano-max}$ is deviated from the definition of percolation probability in Ref. [2].

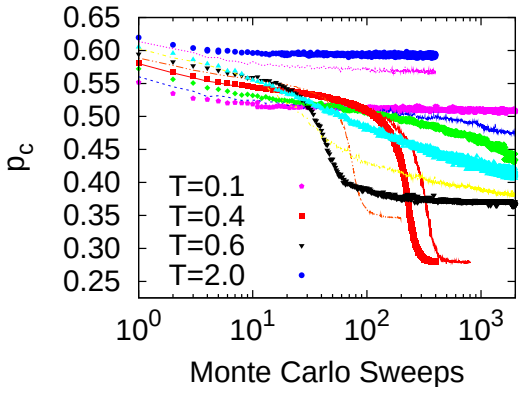


Figure 3: A set of temporal dynamics of p_c values in time evolution processes with different constant parameter of $T = 0.1$, 0.2, 0.3, 0.38, 0.4, 0.5, 0.6, 0.8, 1.0, 1.5 and 2.0. For instance, above the red-square curve of $T = 0.4$ is the greater-than-half probability of an arbitrary nano cell within the largest nano cluster obtained from Fig. 2(b). The numbers of realizations for each process are different from 20 to 200 due to limitation of computational time.

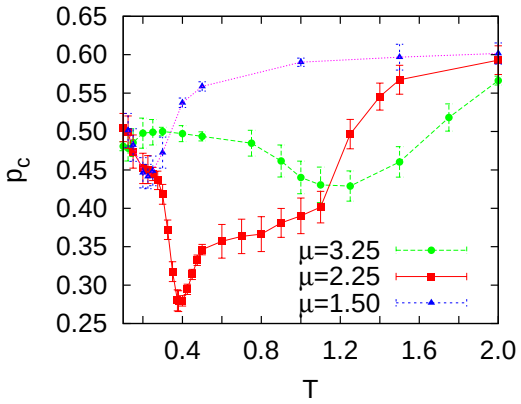
evaporation stage. This corresponds to less than 1000 Monte Carlo sweeps for most of the system, except some require up to 4000 Monte Carlo sweeps due to the dominations of liquid-evaporation mechanism of nucleation and growth. [4] We therefore terminate our simulation at four thousand Monte Carlo sweeps. With Metropolis algorithm, we explore the phase space in two different parameters: interaction energy ε_{ll} and chemical potential μ.

Having described the site graining threshold in temporal process and explored the phase space in two parameters: interaction ε_{ll} in Fig. 3 and chemical potential μ, the following observation are so concluded: The criteria to terminate the simulation is at the steady duration of the position of series of sigmoid curves of the percolation strength because the fractions of solvent remains in the evaporation process is highly dependent of the nano coverage in Fig. 2(a). In Fig. 4,

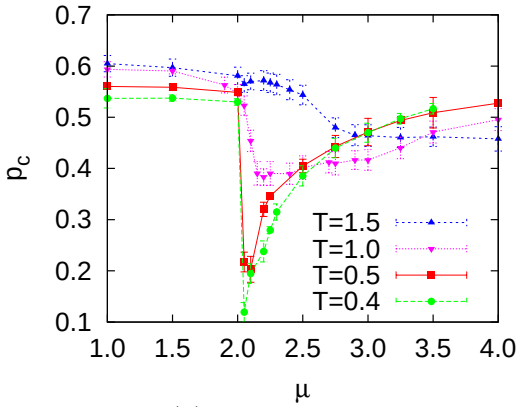
we use the value of p_c to trace the phase change. Above the curve, the self-assembly system is with higher probability to aggregate to a greater one of clusters to occupy a large portion of the board it is easier to percolate. Interestingly, our result in Fig. 4(a) indicates higher site percolation threshold are also shown at higher values of temperature. When the value of T is close to 2.0, the nano-particles are strongly perturbed by liquid solvent due to thermal fluctuation. Hence, the p_c values are close to 0.6, the percolation threshold of the typical site percolation. On the other hand, when the value of T decreases, the liquid solvent remains on the lattice board. Nanoparticles still aggregate but move slower, and change into a network-like cluster. As a result, the final p_c would gradually get closer to the case of the typical 2D site percolation model as the value of T is lower.

In Fig. 4(b), when the chemical potential μ is high, the liquid solvent gets harder to evaporate. The nanoparticle patterns are not able to sustain in evolution and the system starts a new simple sampling of nanoparticles when it reaches a steady state. That makes the system much more similar to a typical site percolation. So the value of p_c gets closer to 0.6 if the chemical potential μ increases. At the binodal line, $\mu = 2$, it corresponds to phase coexistence of the liquid and vapor. When $\mu > 2$ and is less the spinodal limit, liquid solvent remains locally metastable on the substrate surface and it leads to network-like nano morphologies. Naturally, in much lower nanoparticle density, the system would form giant clusters and easily percolate as μ approaching 2 from above.

A finite-size comparison of two p_c curves from 512×512 and 1024×1024 are studied at $T = 0.5$ and $\mu = 2$ with both realizations



(a) p_c versus T



(b) p_c versus μ

Figure 4: Behaviors of p_c in two different parameter spaces. Lattice size is 512×512. Fig. 4(a) The relation between p_c and T under constant chemical potential $\mu = 1.5$, 2.25 and 3.25. At $T = 0.4$ and $\mu = 2.25$, the minimum value of p_c is 0.279. At $\mu = 1.5$ the minimum of $p_c = 0.446$ is at $T = 0.2$. At $\mu = 3.25$ the minimum of $p_c = 0.429$ is at $T = 1.25$. Fig. 4(b) The relation between p_c and μ. Here the temperature value is in constant; $T = 0.4$, 0.5, 1.0, and 1.5. The minimum of p_c would be induced right above the binodal line, $\mu = 2$.

of 200 samples and their final steady p_c values are 0.3462 ± 0.0136 and 0.3529 ± 0.0128, respectively. Their dynamic behaviors in time are qualitatively in good agreement.

Generally, the idea of p_c value to describe a dynamic percolation system is feasible. It suggests that systems of drying-mediated nanoparticle assembly would aggregate into huge clusters in low nano coverage such as 0.119 at $T = 0.4$ when the chemical potential μ is right above its binodal line. Nevertheless, it must be considered with the physical limits since it neglects the condition of percolation.

[1] E. Rabani, D. Reichman, P. Geissler, and L. Brus, Nature, **426**, 271 (2003)

[2] D. Stauffer and A. Aharony, *Introduction to Percolation Theory*, revised second edition, Taylor and Francis (London, 1994).

[3] Ching-Ling Hsu, Szu-Ming Chu, Kiwi Wood, and Yi-Rong Yang, Phys. Stat. Sol. (a) **204**, 1856 (2007)

[4] A. Stannard, C. P. Martin, E. Pauliac-Vaujour, P. Moriarty, and U. Thiele, J. Phys. Chem. C, **112**, 15195 (2008)

[5] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, and A. H. Teller, and E. Teller, J. Chem. Phys., **21**, 1087 (1953)

[6] C. P. Martin, M. O. Blunt, and P. Moriarty, Nano Lett., **4**, 2389 (2004)

Statistical Analysis and Modeling of the Temperature-Dependent Sleep Behavior of Drosophila

Chi-Tin Shih[a,b,*], Hsuan-Wen Lin[c], Ann-Shyn Chiang[c,d]

[a] *Department of Physics, Tunghai University, Taichung 40704, Taiwan*
[b] *National Center for Theoretical Sciences, Hsinchu 30013, Taiwan*
[c] *Institution of Biotechnology, National Tsing-Hua University, Hsinchu 30013, Taiwan*
[d] *Brain Research Center, National Tsing-Hua University, Hsinchu 30013, Taiwan*

Abstract

The sleep behavior of drosophila is analyzed under different temperatures. The activity per minute of the flies is recorded automatically. Sleep for a fruit fly is defined as the periods without any activity and longer than 5 minutes. Several parameters such as total sleep time, circadian sleep profile, quality of sleep are analyzed. The sleep behaviors are significantly different for flies at different temperature. Interestingly, the durations of daytime sleep periods show a common scale-free power law distribution. We propose a stochastic model to simulate the dynamics of sleep-wake process to explain the distribution of daytime sleep.

Keywords: drosophila; sleep behavior; computational neuroscience

Sleep is an important part of animal behavior which is controlled by complex interactions in the neural circuits in the brains. The mechanism which regulates sleep behavior is still unclear. *Drosophila melanogaster* is an excellent model animal to study the complex "genetics $\leftrightarrow$ neuron $\leftrightarrow$ brain $\leftrightarrow$ behavior" network due to its short generation time, mass reproduction, ease of screening, and comprehensive genetic tools[1]. Sleep behavior has also been studied for a decade[2, 3]. Recent studies show that sleep is linked to many essential functions of brain like memory and plasticity[4, 5, 6].

In contrast to the mammals whose sleep can be defined by monitoring the brain activity by electroencephalography or other techniques, fruit fly's sleep can only be defined from their activity and response to external stimuli[7]. The activity of the fruit flies was measured by the drosophila activity monitor system (DAMS) developed by TriKinetics (Waltham MA). Each fly was loaded into a $65mm \times 5mm$ tube with fresh food in one end and then placed in the DAMS. An infrared beam was made to cross the tube perpendicularly. Each DAMS holds 32 tubes and records the infrared beam interruptions caused when a fly walked across the beam. The flies were in the tubes for 1 day adaptation and recorded for the following 3 days under 12-hour light (L) followed by 12-hour dark (D) period during the whole process to mimic the day-night cycles. The raw data recorded by DAMS were time series of the number of infrared beam interruptions for each minute. The sleep episodes were defined as the periods longer than 5 minutes without any activity[7, 8]. We developed a program to automatically analyze the raw data and calculate the characteristic parameters of fly's sleep.

Fig. 1 shows the sleep profile for three days for the female flies at different temperatures $T = 18$, 25, and 30°C. The favorable temperature for flies is roughly $22 \sim 27$°C. All three groups show circadian rhythm which is synchronized by the 24-hour

*Corresponding author.
Email address: ctshih@thu.edu.tw (Chi-Tin Shih)

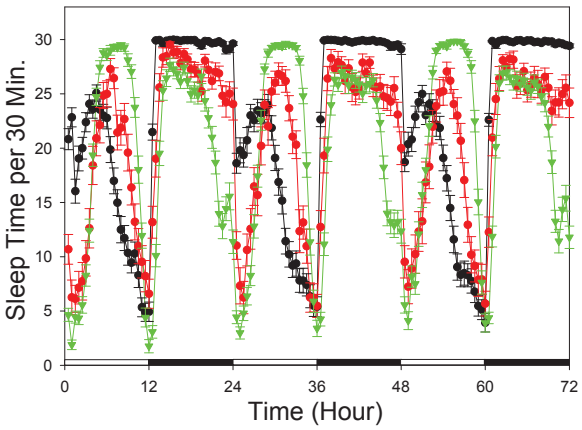


Figure 1: The averaged minutes of sleep for each 30-minute period for the fruit flies at temperature $T = 18°C$ (black, 83 flies), $T = 25°C$ (red, 34 flies), and $T = 30°C$ (green, 84 flies). The white and black bars on the horizontal axis represent the LD cycles of the environment.

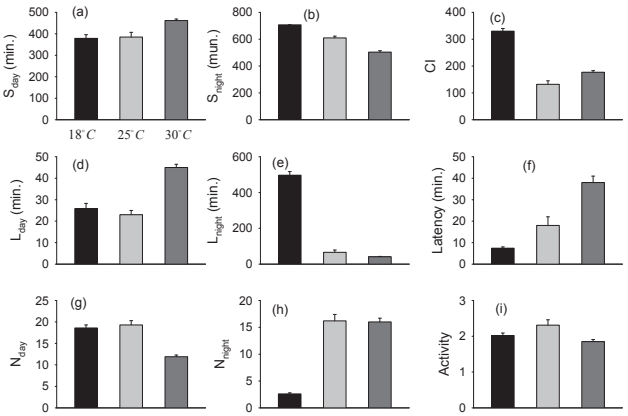


Figure 2: Characteristic parameters (see text) for flies at temperature $T = 18°C$ (black), $T = 25°C$ (light grey), and $T = 30°C$ (dark grey).

LD cycles. The error bars reflect the high degree of diversity of individual flies. However, the effect of temperature on sleep is still statistically significant.

To compare the sleep behavior for flies at different temperature in more detail, several characteristic parameters are calculated and shown in Fig. 2: total sleep in daytime (S_{day}) and nighttime (S_{night}) (Fig. 2(a) and (b), respectively); consolidation index CI (Fig. 2(c)) which is defined as $(\sum_i S_i^2)/(\sum_i S_i)$ where S_i is the length of $i-$th sleep episode in minutes and the summation runs over all sleep episodes; averaged sleep bout length in daytime (L_{day}) and nighttime (L_{night}) (Fig. 2(d) and (e), respectively); latency of nighttime sleep defined as the starting time of the first sleep episode in nighttime (Fig. 2(f)); sleep bout number in daytime (N_{day}) and nighttime (N_{night}) (Fig. 2(g) and (h), respectively); and the activity per waking minute (Fig. 2(i)). From Fig. 1 and Fig. 2, the temperature-dependent features of fly sleep are summarized as following:

1. Flies at 18°C slept almost all the night time.
2. Flies at 18°C and 25°C slept much more in night than in daytime. Their sleep quality is also better (less bout number and longer bout length) in night.
3. Total daytime sleep and nighttime sleep are about the same for flies at 30°C. However, their sleep quality is better in daytime.
4. The nighttime sleep for groups at 25 and 30°C is more fragile than that at 18°C.
5. The flies fall asleep faster in night for lower temperature.
6. Almost all the flies at 25°C and 30°C aroused when the light was turn on. On the other hand, many of the flies at 18°C kept on sleeping in the meantime.

Note that Fig. 2(i) shows that the activities (number of times which the fly passed the central infrared beam per waking minute) for the three groups are not significantly different. The result excludes the possibility that the effect of temperature can be just related to fly's activity and have nothing to do with sleep.

Although the sleep behavior showed diverse features for different temperature, we found they shared a common statistical property. We calculated the cumulative distribution function (CDF) of the daytime bout length L_{day} as

$$C(L_{day}) = \sum_{l=L_{day}}^{L_0} p(l) \qquad (1)$$

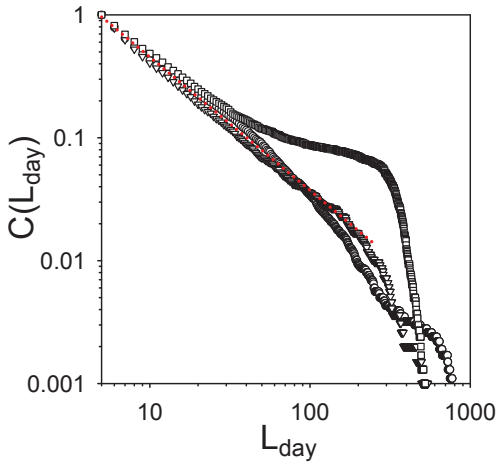


Figure 3: Cumulative distribution function $C(L_{day})$ of the daytime sleep length L_{day} for $T = 18°$C (circles), $T = 25°$C (triangles), and $T = 30°$C (squres). The data can be fitted by a power-law function (red dotted line).

where $p(l)$ is the probability of the length of a sleep episode in daytime is l minute and L_0 is the longest sleep duration. Fig. 3 shows $C(L_{day})$ for the three temperatures. The straight lines for $T = 18°$C and $T = 25°$C in this log-log plot can be fitted by a power-law decay function $C(L_{day}) = 5.52 \times L_{day}^{-1.08}$ up to $L_{day} \sim 200$. The flies in high temperature $T = 30°$C spent more daytime for sleep therefore its $C(L_{day})$ had fatter tail for large L_{day}. However, $C(L_{day})$ for the high temperature group can also be fitted by the same function for $L_{day} \leq 40$. Interestingly, the exponent is very close to the -1 of the Zipf's law in linguistics[9]. On the other hand, the CDFs for the nighttime bout length L_{night} at different temperature do not have such common property. To understand the common property at daytime, we use a simple model to simulate fly's sleep. Assume that $x(t)$ is the mean-field state to represent the overall activity of the population of neurons which regulate fly sleep. The fly is active as $x(t) \geq 0$ and quiescent for $x(t) < 0$ at time t. The dynamics of $x(t)$ is defined as $\Delta x(t) = x(t+1) - x(t) = \epsilon$ for $\Delta \geq x(t) \geq 0$ (active) and $\frac{b}{x} + \epsilon$ for $x(t) < 0$ (quiescent), where ϵ is an uncorrelated Gaussian random variable with zero mean and unity width. Δ and b are the tuning parameters in the simulation. b is the bias toward the active state when the fly is quiescent and Δ is a cutoff of $x(t)$. The autocorrelation of $x(t)$ comes from the bias b. The model has similar form as the one used in the study of human sleep in nighttime[10].

The time series of the activity is generated from the model to simulated fly's activity in the DAMS. We found when $b = 0.5$ and $\Delta = 10$ the CDF of sleep bout length is $C(L) = 5.45 \times L^{-1.09}$, which agrees very well with the $C(L_{day})$ observed in experiments. The key point of the model is that b, the bias toward the active state, is necessary for the flies should keep on the alert in daytime to react when life-threatening events occur.

In summary, the effect of temperature on fly's sleep behavior is studied by statistical methods and numerical simulation of a phenomenological mean-field model. The sleep phenotypes for different temperature are significantly different in circadian sleep profile, total sleep time, bout length, bout number in day and night, latency, and consolidation index. However, the different phenotypes have a common power-law CDF with exponent ~ -1 for daytime sleep durations. This can be explain that the flies have to keep on the alert in daytime to protect themselves.

This work was supported by the National Science Council in Taiwan (97-2112-M-029-002-MY3) and the National Center for High-Performance Computing in Taiwan.

[1] R. J. Greenspan, *Fly Pushing, The Theory and Practice of Drosophila Genetics*, 2nd Ed., Cold Spring Harbor Laboratory Press, New York (2004).
[2] J. C. Hendricks *et al.*, Neuron **25**, 129 (2003).
[3] P. J. Shaw, C Cirelli, R. J. Greenspan, and G. Tononi, Science **287**, 1834 (2000).
[4] G. Tononi and C. Cirelli, Sleep Med. Rev. **10**, 49 (2006).
[5] G. F. Gilestro, G. Tononi and C. Cirelli, Science **324**, 109 (2009).
[6] I. Ganguly-Fitzgerald, J. Donlea, P. J. Shaw, Science **313**, 1775 (2009)
[7] R. Andretic and P. J. Shaw, Methods in Enzymology **393**, 759 (2005).
[8] W. J. Joiner, A. Crocker, B. M. White, and A. Sehgal, Nature **441**, 757 (2006).
[9] G. K. Zipf, *Human Behavior and the Principle of Least-Effort*, Addison-Wesley (1949).
[10] C. C. Lo, L. A. N. Amaral, S. Havlin, P. Ch. Ivanov,

T. Penzel, J. H. Peter, and H. E. Stanley, Europhys. Lett. **57**, 625 (2002).
[11] C. C. Lo, T. Chou, T. Penzel, T. E. Scammell, R. E. Strecker, and H. E. Stanley, Proc. Nat. Acad. Sci. USA **101**, 17545 (2004).