Analyzing Meiotic DSB Interference by Combining Southern Blotting and Microarray Analysis

Hiroshi Toyoizumi
Graduate School of Accouting, Waseda University,
Department of Applied Mathematics, Waseda University,
Tokyo, Japan
e-mail: toyoizumi@waseda.jp

Hideo Tsubouchi
MRC Genome Damage and Stability Centre,
University of Sussex,
Brighton, United Kingdom
e-mail: H.Tsubouchi@sussex.ac.uk

***Abstract*—It is well-known that crossover formation events on a chromosome interfere with each other during meiosis, and this interference affects the distribution of genetic exchanges on a chromosome in sexual reproduction. However, due to the technical difficulties, it is unknown if meiotic double strand break (DSB) formation, the initiating event of meiotic recombination, shows interference. We discuss a method that employs probability theory of survival analysis in conjunction with: chromosome fragment distribution, detected by Southern blotting; and genome-wide DSB intensity maps, obtained by microarray analysis. We show that this method is a promising tool to analyse DSB interference.**

***Keywords*-meiosis; DSB; Southern blotting; crossover interference; non-homogenous Poisson process; survival analysis.**

I. Introduction

Meiosis is a specialized cell cycle essential for producing gametes in sexual reproduction [1]. During meiosis, DNA double-strand breaks (DSBs) are programmed to be formed and induce homologous recombination. The homologous recombination mechanism facilitates recognition of homologous chromosomes and establishes physical connections between them via crossovers. The study of meiotic DSB formation and homologous recombination is important because these events are not only central to the life cycle of sexually reproducing organisms, but they are also a driving force for the production of genetic diversity.

Crossover interference is a phenomenon that is known to influence the distribution of crossovers such that the presence of a crossover reduces the likelihood of another crossover forming nearby [2]–[4]. However, it is not known if there is a similar interference mechanism operating at the level of DSB formation, since unlike the case of crossovers, DSB interference can not be observed directly [2].

The formation of meiotic DSBs and crossovers is controlled by many complex biological processes and the mechanism has been intensively studied using various methods [5]–[7]. There are two popular ways to analyse DSB formation: Southern blotting and whole genome mapping obtained by microarray analysis. Both methods fall short of analysing potential DSB interference, but by combining both methods we can analyze the strength of DSB interference. Recently, by assuming that there is no strong DSB interference, we derived an algorithm to estimate the number of DSBs from the experimental results of Southern blotting [8]. In this paper, we show that, by reversing the logic, we can check for the presence of DSB interference. In a large part of past linkage analysis, genetic recombination, which is the consequence of DSB formation, has been assumed to be a non-homogenous Poisson process [3], [9]. In the context of survival analysis with the partial observation, the non parametric Nelson-Aalen estimator has been intensively used to estimate the cumulative hazard rate from censored data [10]–[12]. By simply comparing the two DSB intensities, the one obtained from microarray analysis and the other based on the Nelson-Aalen estimator, which will coincide if DSB formation follows a non-homogenous Poisson process, we can analyze the DSB interference on a chromosome.

II. Southern Blotting and Whole Genome Mapping by Microarray Analysis

A. Southern Blotting and Distribution Function of First Break

A common molecular biology technique called Southern blotting (see Fig. 1) enables detection of one unique location in the genome, making it possible to examine DSB formation per given chromosome. DSB formation can then be studied in mutants of interest, such as those that form but do not repair DSBs [13], [14]. In Southern blotting, the total genomic DNA prepared from cells introduced into meiosis is separated according to size by gel electrophoresis. The separated DNA molecules are then transferred to a nylon membrane on which broken chromosome fragments carrying one end of a chromosome to be examined are probed with a radioactive oligonucleotide. Thus, when a chromosome is intact (i.e., no DSBs are formed), only a single band appears at the location corresponding to the size of the whole chromosome. Once DSBs are formed, chromosomal DNA is broken and smaller molecular pieces appear accordingly, producing numerous bands below the position of the intact chromosome. Although this method is suitable for determining the location of DSBs along chromosomes, the strength of the signal at a given location

 ISBN: 978-1-61208-260-8

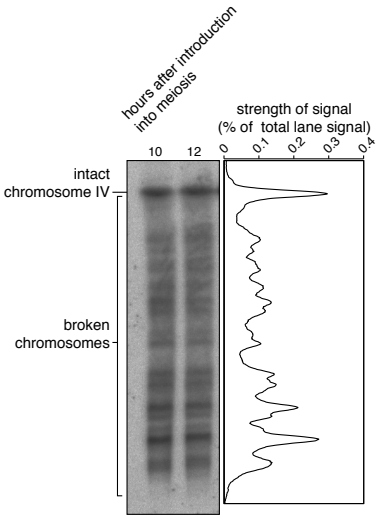

Figure 1. An example of DSB detection by Southern blotting. Mutant budding yeast diploid cells (the *sae2* mutant) were introduced into meiosis and cells were harvested at 10 and 12 hours after introduction into meiosis. The Southern blotting was used to detect chromosome IV. Lane profiles of 10 and 12 hours in each mutant background were normalized and averaged to obtain the profiles shown on the right [13], [14]. This figure is adopted from [8].

does not necessarily correlate with the actual number of DSBs formed there. This is because, when two or more DSBs are formed on a chromosome, only the DNA fragment carrying the chromosome end hybridised to the probe is detected, while others are invisible. In other words, we can only detect the first DSB from one end of the chromosome using this approach.

Let T_1 be the size of the first DSB from the left end of a chromosome. Define the probability distribution of T_1 as $F(x) = P\{T_1 \leq x\}$, which is assumed to be continuous and differentiable for simplicity. Here, the probability measure P can be regarded as the sample ratio of Southern blotting. Thus, by using these sample ratio values obtained from Southern blot analysis, we can estimate the distribution $F(x)$.

B. Microarray Analysis and DSB intensity

DSBs formed during meiosis are exonucleolytically resected from their ends, producing 3'-ended single-stranded DNA (ssDNA). These ssDNA molecules can be selectively recovered by using Benzoyl naphthoyl DEAE (BND) cellulose. Based on microarray analysis of break-associated ssDNA isolated by BND cellulose enrichment, the intensity and distribution of meiotic DSBs were measured [15].

Let t be the position from left end point on a chromosome and $N(t)$ be the number of breaks in this region $(0, t)$. The intensity of DSBs is defined by

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{E[N(t + \Delta t) - N(t)]}{\Delta t}, \quad (1)$$

where the expectation can be regarded as the sample averages of the signal obtained from the microarray analysis. This method is very effective and precise about the location of DSBs, and is used to locate the hot spots and cold spots on chromosomes. However, this analysis cannot reveal which DSBs occurred on a particular chromosome, thus we cannot test for DSB interference.

III. Interference Function and Hazard Rate Function

Interference is often measured by the coincidence function [3], [9]. Let $E(t_1, t_2)$ be the event that a break is in the interval $[t_1, t_2)$ on a chromosome. The coincidence function is then defined by

$$c(s, t) = \frac{P\{E[s, s + \Delta t], E[t, t + \Delta t]\}}{P\{E[s, s + \Delta t]\} P\{E[t, t + \Delta t]\}}, \quad (2)$$

for some small positive Δt. The formation of DSBs is said to have positive interference when a break constrains the formation of other breaks. Thus, $P\{E[s, s + \Delta t], E[t, t + \Delta t]\}$ is smaller than the product probability $P\{E[s, s + \Delta t]\} P\{E[t, t + \Delta t]\}$ and $c(s, t) < 1$. If there is no interference and any two breaks occur independently, the numerator and denominator of (2) coincide, and $c(s, t) = 1$. It is well-known that, in the case of crossovers, $c(s, t) < 1$ and there is positive interference among crossovers.

However, the coincidence function is not useful for analysing DSB interference, because neither Southern blotting nor whole genome mapping can identify two breaks occurring on a single chromosome. Hence, $P\{E[s, s + \Delta t], E[t, t + \Delta t]\}$ cannot be estimated.

Here, we propose a similar but different indicator for DSB interference based on the intensity of DSBs and the hazard rate function of the first break. Let $h(x)$ be the hazard rate function of the first break T_1, which is defined as

$$h(x)dx = P\{x \leq T_1 < x + dx | T_1 \geq x\} = \frac{f(x)dx}{1 - F(x)}, \quad (3)$$

where $f(x) = dF(x)/dx$ is the density of the random variable T_1. We define a new function $\tilde{c}(t)$ by

$$\tilde{c}(t) = \frac{h(t)}{\lambda(t)} = \frac{f(t)}{\lambda(t)(1 - F(t))}. \quad (4)$$

Roughly speaking, the coincidence function $c(s, t)$ measures the tendency of double breaks around t and s, while $\tilde{c}(t)$ measures the tendency of breaks at the position t given no breaks in $[0, t)$. When DSB formation shows positive interference, $E[N(t + dt) - N(t) | N(t) > 0] < E[N(t +$

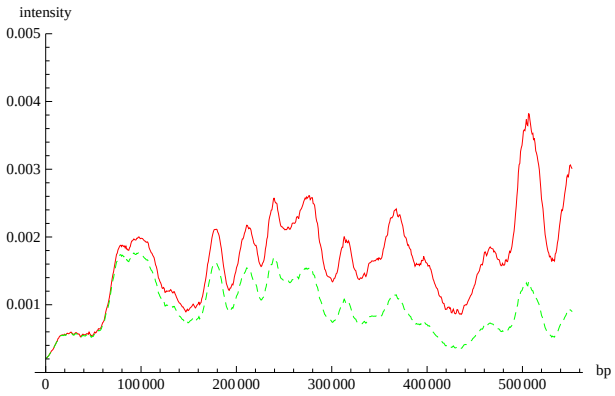

Figure 2. The hazard rate function $h(t)$ (red line) and the size density $f(t)$ (green dashed line) deduced from Southern blotting on chromosome no. 11. The data is adopted from [13].

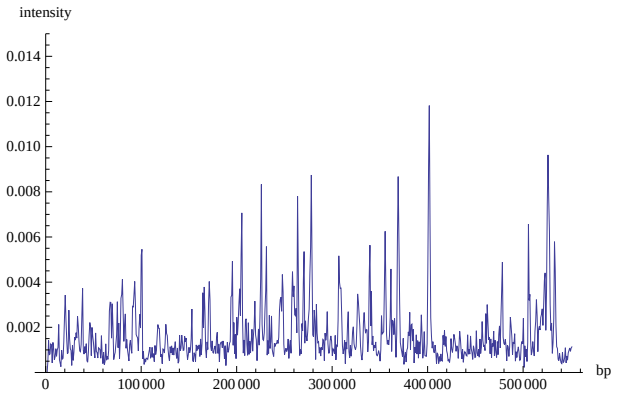

Figure 3. The intensity function $\lambda(t)$ on chromosome no. 11 deduced from whole genome analysis. The data is adopted from [15].

$dt) - N(t)]$. Thus, we have

$$\begin{aligned}\lambda(t)dt &= E[N(t+dt) - N(t)|N(t) = 0]P\{N(t) = 0\} \\ &\quad + E[N(t+dt) - N(t)|N(t) > 0]P\{N(t) > 0\} \\ &< h(t)dt(1 - F(t)) + \lambda(t)dtF(t),\end{aligned}$$

and $\lambda(t) < h(t)$, and thus $\tilde{c}(t) > 1$. The hazard rate function $h(t)$ coincides with the intensity $\lambda(t)$ and $\tilde{c}(t) = 1$ when there is no DSB interference and $N(t)$ is a non-homogeneous Poisson process (see [8] for example).

IV. Preliminary Results

Using existing data of Southern blotting and genome wide DSB intensity maps, we have obtained some preliminary results about DSB interference. Fig. 2 depicts the hazard rate function $h(t)$ deduced from the DSB size distribution of Southern blotting [13] (see [12] for the estimation method based on Nelson-Aalen estimator). Also, Fig. 3 is the intensity $\lambda(t)$ from the microarray analysis of [15]. Since the two datasets come from completely different experiments, the signal levels are not adjusted. Here, we use a simple normalization to rescale the intensity $\lambda(t)$ to satisfy

$$\int_0^c h(t)dt = \int_0^c \lambda(t)dt, \tag{5}$$

which should be valid when there is no DSB interference, since in that case, $h(t) = \lambda(t)$. The resulted rescaled intensity $\lambda(t)$ is compared with the hazard rate function $h(t)$ in Fig. 4. We use the moving average of 20 signals of the original data of $\lambda(t)$ for noise filtration.

The two curves $h(t)$ and $\lambda(t)$ are agreeable with the locations of hot spots and cold spots. However, $h(t)$ is smaller on the left hand side of the chromosome, and especially at around 400 Kbp, where we see a significant difference between $\lambda(t)$ and $h(t)$. Note that $f(t)$ corresponds to the intensity for the strongest positive DSB interference and only one break exists on the chromosome. The interference

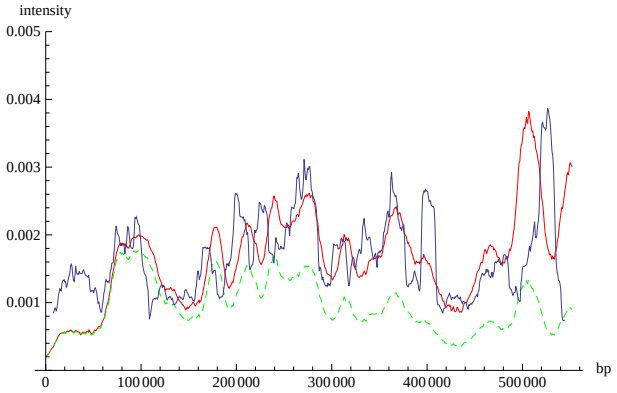

Figure 4. The hazard rate function $h(t)$ (red), the smoothed and scaled intensity function $\lambda(t)$ (blue) and the size density function $f(t)$ (green dashed).

function $\tilde{c}(t)$ in Fig. 5 also shows no or very small positive interference (less than $\tilde{c}(t) < 1.5$). Most of the variance of $\tilde{c}(t)$ can be explained by the poor calibration of two datasets obtained from different experiments. Thus, we can conclude that there is no or very weak positive DSB interference on chromosome no. 11 of budding yeast in the sae2 mutant background.

V. Conclusion

In this paper, we discuss the possibility of analyzing DSB interference by combining the results from Southern blotting and microarray analysis. We also included some preliminary results about DSB interference, but we should point out that the results obtained by applying our method are not final and we need further analysis to discuss the existence of DSB interference.

 ISBN: 978-1-61208-260-8

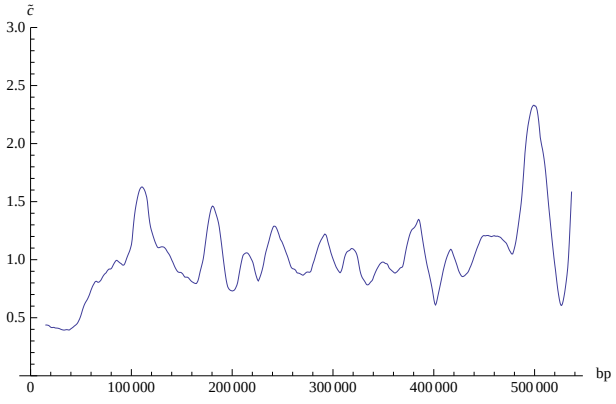

Figure 5. The interference function $\tilde{c}(t)$. $\tilde{c}(t)$ is supposed to be 1 when there is no interference.

REFERENCES

[1] M. Petronczki, M. F. Siomos, and K. Nasmyth, "Un ménage àquatre: The molecular biology of chromosome segregation in meiosis," *Cell*, vol. 112, no. 4, pp. 423–440, 2 2003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0092867403000837 Accessed March 2013

[2] L. Berchowitz and G. Copenhaver, "Genetic interference: don't stand so close to me," *Current Genomics*, vol. 11, no. 2, p. 91, 2010.

[3] M. S. McPeek and T. P. Speed, "Modeling interference in genetic recombination," *Genetics*, vol. 139, no. 2, pp. 1031–1044, 1995. [Online]. Available: http://www.genetics.org/content/139/2/1031.abstract Accessed March 2013

[4] S. Karlin and U. Liberman, "A natural class of multilocus recombination processes and related measures of crossover interference," *Advances in Applied Probability*, vol. 11, no. 3, pp. 479–501, 1979.

[5] S. Keeney, *Mechanism and control of meiotic recombination initiation*. Academic Press, 2001, vol. Volume 52, pp. 1–53. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0070215301520086 Accessed March 2013

[6] W. Crismani, C. Girard, N. Froger, M. Pradillo, J. L. Santos, L. Chelysheva, G. P. Copenhaver, C. Horlow, and R. Mercier, "Fancm limits meiotic crossovers," *Science*, vol. 336, no. 6088, pp. 1588–1590, 06 2012. [Online]. Available: http://www.sciencemag.org/content/336/6088/1588.abstract Accessed March 2013

[7] A. Lorenz, F. Osman, W. Sun, S. Nandi, R. Steinacher, and M. C. Whitby, "The fission yeast fancm ortholog directs non-crossover recombination during meiosis," *Science*, vol. 336, no. 6088, pp. 1585–1588, 06 2012. [Online]. Available: http://www.sciencemag.org/content/336/6088/1585.abstract Accessed March 2013

[8] H. Toyoizumi and H. Tsubouchi, "Estimating the number of double-strand breaks formed during meiosis from partial observation." *J Comput Biol*, vol. 19, no. 12, pp. 1277–1283, Dec 2012. [Online]. Available: http://dx.doi.org/10.1089/cmb.2012.0209 Accessed March 2013

[9] J. Haldane, "The combination of linkage values and the calculation of distances between the loci of linked factors," *J. Genet*, vol. 8, no. 29, p. 309, 1919.

[10] P. K. Andersen, Ø. Borgan, N. L. Hjort, E. Arjas, J. Stene, and O. Aalen, "Counting process models for life history data: A review [with discussion and reply]," *Scandinavian Journal of Statistics*, vol. 12, no. 2, pp. 97–158, 01 1985. [Online]. Available: http://www.jstor.org/stable/4615980 Accessed March 2013

[11] P. Andersen, *Statistical models based on counting processes*. Springer Verlag, 1993.

[12] T. Fleming and D. Harrington, *Counting processes and survival analysis*. Wiley Online Library, 1991, vol. 8.

[13] S. Farmer, E.-J. E. Hong, W.-K. Leung, B. Argunhan, Y. Terentyev, N. Humphryes, H. Toyoizumi, and H. Tsubouchi, "Budding yeast pch2, a widely conserved meiotic protein, is involved in the initiation of meiotic recombination," *PLoS ONE*, vol. 7, no. 6, p. e39724, 06 2012. [Online]. Available: http://dx.doi.org/10.1371%2Fjournal.pone.0039724 Accessed March 2013

[14] S. Farmer, W.-K. Leung, and H. Tsubouchi, "Characterization of meiotic recombination initiation sites using pulsed-field gel electrophoresis," *Methods in Molecular Biology*, vol. 745, pp. 33–45, June 2011. [Online]. Available: http://www.springerprotocols.com/Abstract/doi/10.1007/978-1-61779-129-1_3 Accessed March 2013

[15] C. Buhler, V. Borde, and M. Lichten, "Mapping meiotic single-strand dna reveals a new landscape of dna double-strand breaks in saccharomyces cerevisiae," *PLoS Biol*, vol. 5, no. 12, p. e324, 12 2007. [Online]. Available: http://dx.doi.org/10.1371%2Fjournal.pbio.0050324 Accessed March 2013

 ISBN: 978-1-61208-260-8