ELSEVIER

Mathematical Biosciences 162 (1999) 85–101

Mathematical Biosciences
an international journal

www.elsevier.com/locate/mathbio

# Modeling the low-LET dose-response of BCR–ABL formation: predicting stem cell numbers from A-bomb data

Tomas Radivoyevitch *, David G. Hoel

*Department of Biometry and Epidemiology, Medical University of South Carolina, Charleston, SC 29425, USA*

Received 23 February 1999; received in revised form 29 June 1999; accepted 21 July 1999

**Abstract**

Formation of the BCR–ABL chromosomal translocation $t(9;22)(q34;q11)$ is essential to the genesis of chronic myeloid leukemia (CML). An interest in the dose-response of radiation induced CML therefore leads naturally to an interest in the dose-response of BCR–ABL formation. To predict the BCR–ABL dose-response to low-linear energy transfer (LET) ionizing radiation, three models valid over three different dose ranges are examined: the first for doses greater than 80 Gy, the second for doses less than 5 Gy and the third for doses greater than 2 Gy. The first of the models, due to Holley and Chatterjee, ignores the accidental binary eurejoining of DNA double-strand break (DSB) free ends ('eurejoining' refers to the accidental restitution of DSB free ends with their own proper mates). As a result, the model is valid only in the limit of high doses. The second model is derived directly from cytogenetic data. This model has the attractive feature that it implicitly accounts for single-track effects at low doses. The third model, based on the Sax–Markov binary eurejoining/misrejoining (SMBE) algorithm, does not account for single-track effects and is therefore limited to moderate doses greater than approximately 2 Gy. Comparing the second model to lifetime excess CML risks expected after 1 Gy, estimates of the number of hematopoietic stem cells capable of causing CML were obtained for male and female atomic bomb survivors in Hiroshima and Nagasaki. The stem cell number estimates lie in the range of $5 \times 10^7$–$3 \times 10^8$ cells. 

* Corresponding author. Tel.: +1-803 766 7064; fax: +1-803 876 1126.
*E-mail address:* radivot@musc.edu (T. Radivoyevitch)

0025-5564/99/$ - see front matter 
PII: S0025-5564(99)00039-5

*Keywords:* BCR–ABL; Dose-response; CML; Ionizing radiation; Stem cells

## 1. Introduction

Among the various kinds of chromosomal aberrations induced by ionizing radiation [1,2], translocations are extremely relevant to the etiology of cancer [3]. Two translocation-mediated cancers, chronic myeloid leukemia (CML) [1] and acute promyelocytic leukemia (APL), are of particular interest because their associations with the specific translocations BCR–ABL and PML–RARA are not only exceptionally strong [4,5], but causal too – mice transgenic for the chimeric protein products of these translocations have an increased incidence of the corresponding leukemias [6–10]. Between these two cancers, we focus here on CML rather than APL because there appears to be more data available for CML. For example, whereas CML risks in A-bomb survivors were modeled by Preston et al. [11], APL risks were not, presumably because not enough APLs were induced among the survivors. The induction of CMLs among A-bomb survivors can be viewed as proof that ionizing radiation creates BCR–ABL, i.e., that it creates the Philadelphia translocation t(9;22) between the BCR gene on chromosome 9 and the ABL gene on chromosome 22; this conclusion has been corroborated through reverse-transcriptase polymerase chain reaction (RT-PCR) experiments on heavily irradiated human tissue cultures [12,13].

The expected number of BCR–ABL translocations in a human $G_0/G_1$ cell irradiated with a low-linear energy transfer (LET) dose $D$, denoted by $E(ba|D)$, will be derived for three different dose ranges: in the limit of high doses ($D \geqslant 80$ Gy; Section 4), we rederive Holley and Chatterjee's result that $E(ba|D)$ is linear in dose [14]; for doses less than 5 Gy (Section 4), we show that under our assumptions $E(ba|D)$ is proportional to the expected number of translocations $E(t|D)$ determined cytogenetically; and in the case of low-LET doses large enough that one-track action can be ignored ($D \geqslant 2$ Gy; Section 6), $E(ba|D)$ is derived from the quadratic-linear total misrejoining dose-response of the Sax–Markov binary eurejoining/misrejoining (SMBE) model [15]. We then estimate $N$, the number of stem cells capable of causing CML, by equating $N \cdot E(ba|D = 1)$ to estimates of the excess lifetime risks of CML after 1 Gy (Section 7). Our estimates of $N$ range from $5 \times 10^7$ to $3 \times 10^8$ stem cells. Stem

[1] Abbreviations: ABL – Abelson; APL – acute promyelocytic leukemia; BCR – breakpoint cluster region; CML – chronic myeloid leukemia; DSB – double-strand break; kbp – kilo-basepairs; LET – linear energy transfer; LQL – linear-quadratic-linear; Mb – Mega-basepairs; RT-PCR – reverse transcriptase polymerase chain reaction; SMBE – Sax–Markov binary eurejoining/misrejoining.

cell number estimates such as these are pertinent to biologically-based epidemiological models of cancer risk [16,17].

## 2. Background

### 2.1. Linear energy transfer

A photon is a wave packet of orthogonal electric and magnetic fields where the frequency of the wave determines the photon's energy and an envelope multiplies the wave amplitude thus localizing it spatially. Charged particles on the other hand, produce constant electric fields in the moving frame and spiked electric fields in the stationary frame. In almost all cases, the interaction between radiation and matter is through the coupling of an electric field with the dipole moment of a target molecule (the exception is the neutron, but even here, the neutron first undergoes the rare event of actually hitting a nucleus and the secondaries, charged protons and electrons, are responsible for the biological effects). For a specific particle and energy, the LET is defined as the *average* linear density of energy depositions over the remainder of the particle track [18]. (A particle track is the set of ionizing event positions produced as a consequence of one incoming high energy particle.) It is known that charged particles have a higher LET than photons and that high energy photons have a lower LET than low energy photons. One explanation for both of these facts is that slower moving electric fields couple more strongly to molecular dipole moments than do highly oscillatory fields. This would also explain why protons (being slow) have higher LET than electrons, and why slow electrons have higher LET than fast electrons. We are concerned here only with low LET radiation, e.g., X-rays and $\gamma$-rays.

### 2.2. DNA double-strand breaks

Ionizing radiation creates a wide variety of reactive chemical species, the most important of which are hydroxyl radicals in the neighborhood of DNA [19]. Upon striking the DNA, hydroxyls can cause base-adducts which lead to mutations, or they can react with the ribose moiety to create a single strand break (SSB) in the DNA backbone. If two SSBs occur within about 10 base-pairs (bp) of each other, but on opposite strands of the DNA, a double strand break (DSB) is likely to result. Unlike chemicals that cause single strand breaks uniformly throughout the genome, ionizing radiation delivers energy along particle tracks so that SSBs created by the same track become correlated in space. This spatial correlation of SSBs makes ionizing radiation particularly effective at producing DSBs. It is now well accepted that DSBs are the main intermediate responsible for the biological effects of ionizing radiation [20].

By definition, one Gray of absorbed dose equals one joule of energy absorbed into one kilogram of mass. Because DSB formation is a single-track phenomenon (i.e., it is highly unlikely that a DSB would be created by two SSBs from two separate particle tracks), the number of DSBs per base pair is directly proportional to dose and independent of dose-rate. The proportionality constant, known as the DSB yield $Y$, has been estimated to be approximately 0.5 DSBs per 100 Mb per Gray [21,22].

### 2.3. Misrejoining DSBs

We consider DSBs to be either active or inactive, see Ref. [23]. Active DSBs have two separated DSB free ends moving within a common subspace of the cell's nucleus (matrix attachment regions upstream and downstream of a particular DSB are assumed to be anchored near one another in the nuclear matrix). Inactive DSBs, on the other hand, never separate into free ends (they can be thought of as protein 'splinted') and always repair correctly; inactive DSBs are not relevant to this paper. An active DSB free end can *misrejoin* with the free end of a different active DSB, provided that both active DSBs lie within the same region of the nucleus, or it can *accidentally eurejoin* with its own proper mate (the word 'eurejoin' is used to distinguish this process from the 'repair' of inactive DSBs; 'accidental binary eurejoining', 'binary eurejoining' and 'eurejoining' all refer to the same process), or it can remain *unrejoined*, perhaps being stabilized by the growth of a telomere. Though we believe there is only one type of active DSB, its fate being determined by chance alone, it will be useful to identify as 'reactive' DSBs those active DSBs which do eventually misrejoin.

Intuitively, the probability that two active DSB free ends misrejoin should increase as the initial separation distance between the active DSBs decreases. This notion can been represented by either a distance model in which some function $g(r)$ equals the probability that two active DSBs an initial distance $r$ apart misrejoin, or by a site model which approximates the distance model by treating the nucleus as $\eta$ isolated nuclear subvolumes, or sites. In the site model (our focus here), DSBs within the same site misrejoin with a probability that is independent of their initial separation, and DSBs within different sites never misrejoin.

### 2.4. The SMBE model

The SMBE model [15] is a stochastic site model representation of active DSBs rejoining according to Sax's mechanism [24]. Without any adjustable parameters (i.e., based on combinatorial arguments alone), the SMBE model predicts the steady state statistical distribution of DSB misrejoinings per site given the initial statistical distribution of active DSBs per site. The SMBE

model is connected to experimental measurements (i.e., misrejoinings per cell and the dose delivered in Gray) through the number of sites per cell nucleus, $\eta$, and the expected number of active DSBs per Gray per cell, $\gamma$. With these definitions, $\eta$ times the average number of misrejoinings per site equals the average number of misrejoinings per cell, and the average number of active DSBs per site is $\gamma D/\eta$ where $D$ is the dose in Gray. For low-LET radiation, $\gamma D/\eta$ can be assumed to be the Poisson mean of the initial distribution of active DSBs per site (for a particular LET the dose must be high enough that the Poisson approximation holds, see Section 8). Estimates of $\gamma$ and $\eta$ are derived in Section 6.

In the limit of high doses, the expected final number of SMBE misrejoinings per cell approaches the initial number of active DSBs per cell minus half the number of sites, i.e., $M(\infty) = \gamma D - \eta/2$, (see Ref. [15]). The first of the BCR–ABL dose-response models described below corresponds to this asymptotic form of the SMBE model. The second model, though derived directly from low-to-moderate dose cytogenetic data, could also be guised in the SMBE framework. To be specific, the low-dose linearity of the cytogenetic data can be emulated by the SMBE model using over-dispersed Poisson–Poisson initial active DSB distributions (see Section 8). The third BCR–ABL model corresponds to the SMBE model with Poisson initial conditions.

### 2.5. BCR–ABL and CML

The Philadelphia (Ph) chromosome, known to be associated with CML for nearly 40 years [25], is produced by a translocation between the ABL gene on chromosome 9 and the BCR gene on chromosome 22. The result of this translocation is a BCR–ABL chimeric protein with an elevated tyrosine kinase activity that appears to be central to CML carcinogenesis [8,9]. Further supporting a causal relation between BCR–ABL and CML, different forms of this tyrosine kinase, arising from translocations in different introns, have been found to be associated with distinct types of $Ph^+$ leukemias [4,10]. It is interesting to note that, while the position of a translocation within a particular intron should be relatively insignificant since introns are not expressed, the sizes of intron sequences should be very significant since they will influence the translocation probability and thus, presumably, the incidence of a corresponding translocation-mediated cancer; this notion is consistent with CML being more common than APL.

## 3. Assumptions

A basic assumption used in each of the three models below is that genetic loci incur misrejoining events with a probability proportional to the their target

size. The third model (Section 6) also requires that nuclei can be approximated as many equal-sized isolated nuclear subvolumes called sites; the introduction of sites provides consistency between moderate dose cytogenetic data [1,2] and high dose pulsed-field gel data [26], (see also Ref. [15]). We show that introducing sites has no consequences for the first two models (Sections 4 and 5) if we assume, and we shall throughout, that each site is equally likely to contain any of the genetic loci. To estimate the number of stem cells capable of causing CML, we assume in Section 7 that the dose-response of BCR–ABL formation is related to the CML dose-response by an unknown scale factor – the number of target stem cells at risk of transformation. The following parameter values are also assumed: a low-LET DSB yield $Y$ of 0.0054 DSB/(Gy Mb) [21,22], a BCR target size $N_{\mathrm{BCR}}$ of 5.8 kbp [4], an ABL target size $N_{\mathrm{ABL}}$ of 300 kbp [4] and a total human genome size $\Gamma$ of 3200 Mb [27].

## 4. Doses $\geqslant$ 80 Gy

When a reactive DSB strikes the BCR locus only one of the two DSB free ends is capable of participating in the formation of BCR–ABL. We denote this end by BCR*, or simply $b$. Similarly, the free end *ABL, or $a$, is defined as a reactive DSB free end which can join with BCR* to form BCR–ABL, or $ba$. Using this notation, we now follow the strategy of Holley and Chatterjee [14] to arrive at the BCR–ABL dose-response $E(ba|D)$ for doses high enough ($D \geqslant 80$ Gy) that a constant fraction $f = 0.25$ of the DSBs misrejoin [26]. Namely, we shall derive $E(ba|D)$ as the expected number of BCR* free ends multiplied by the probability that one of them rejoins with *ABL.

The expected number of BCR* free ends resulting from a dose $D$ is

$$E(b|D) = 2N_{\mathrm{BCR}}fYD = 2(0.0058)(0.25)(0.0054)D = 0.0000156D, \qquad (1)$$

where the 2 comes from two available alleles and $N_{\mathrm{BCR}}fYD$ is the number of reactive DSBs expected to arise from one BCR locus. Given that BCR* exists, the probability it misrejoins with *ABL is equal to the expected number of *ABL free ends $E(a|b,D)$ divided by the expected number of reactive DSB free ends $2E(m|b,D)$,

$$P(ba|b) = \frac{E(a|b,D)}{2E(m|b,D)} = \frac{(2N_{\mathrm{ABL}})fYD}{2(2\Gamma)fYD} = \frac{N_{\mathrm{ABL}}}{2\Gamma} = \frac{0.3}{6400}, \qquad (2)$$

where high doses allowed us to replace $E(a|b,D)$ and $E(m|b,D)$ by $E(a|D)$ and $E(m|D)$, i.e., the consumption of one reactive DSB to form BCR* can be disregarded since at 80 Gy we expect on the order of 1000 reactive DSBs [26]. The BCR–ABL dose-response at high doses is thus

$$E(ba|D) = P(ba|b)E(b|D) = \frac{N_{\mathrm{ABL}}}{2\Gamma}(2N_{\mathrm{BCR}}fYD) \tag{3}$$

$$= \left(\frac{0.3}{6400}\right)(0.0000156D) = 7.3 \times 10^{-10}D. \tag{4}$$

This equation predicts 7.3 BCR–ABL misrejoinings in $10^8$ cells irradiated with 100 Gy, in very close agreement with the value of 6.6 found by Holley and Chatterjee [14] using different parameter values, though the same technique, and the value of 5 found experimentally [12].

## 5. Doses $\leqslant$ 5 Gy

In this section we restrict our analysis to doses less than 5 Gy. Since BCR–ABL misrejoinings can be assumed to arise as translocations in this dose-range, the dose-response $E(ba|D)$ equals the expected number of translocations $E(t|D)$ multiplied by the probability that a particular translocation is a BCR–ABL translocation, $P(ba|t)$. An expression for $P(ba|t)$ is found as

$$P(ba|t) = \left(\frac{2N_{\mathrm{BCR}}}{2\Gamma}\right)\left(\frac{2N_{\mathrm{ABL}}}{2\Gamma - N_{\mathrm{ch22}}}\right) + \left(\frac{2N_{\mathrm{ABL}}}{2\Gamma}\right)\left(\frac{2N_{\mathrm{BCR}}}{2\Gamma - N_{\mathrm{ch9}}}\right) \tag{5}$$

$$= \left(\frac{N_{\mathrm{BCR}}}{\Gamma}\right)\left(\frac{N_{\mathrm{ABL}}}{\Gamma - N_{\mathrm{ch22}}/2}\right) + \left(\frac{N_{\mathrm{ABL}}}{\Gamma}\right)\left(\frac{N_{\mathrm{BCR}}}{\Gamma - N_{\mathrm{ch9}}/2}\right) \tag{6}$$

$$= \left(\frac{5.8}{3.2 \times 10^6}\right)\left(\frac{300}{3.172 \times 10^6}\right) + \left(\frac{300}{3.2 \times 10^6}\right)\left(\frac{5.8}{3.128 \times 10^6}\right)$$

$$= 3.5 \times 10^{-10} \tag{7}$$

$$P(ba|t) \approx \frac{2N_{\mathrm{BCR}}N_{\mathrm{ABL}}}{\Gamma^2} = 3.4 \times 10^{-10}, \tag{8}$$

where $N_{\mathrm{ch9}} = 145$ Mb and $N_{\mathrm{ch22}} = 56$ Mb are the sizes of chromosomes 9 and 22, respectively [27]. To understand Eq. (5), note that the two summands correspond to the two ways in which the two misrejoinings of the translocation can fall on both BCR and ABL. The first summand is the probability that the first of the misrejoinings arises in BCR, multiplied by the probability that the second misrejoining arises in ABL; since the two misrejoinings form a translocation we know that the second misrejoining must not lie on the same chromosome as the first. The second term corresponds to the event that the first misrejoining arises in ABL and the second in BCR. [2]

[2] As a double check on Eq. (5), note that if the loci had the same size $N_x$ and if the chromosome 9 and 22 sizes were neglected relative to $2\Gamma$, Eq. (5) becomes $2N_x^2/\Gamma^2 = (4N_x/2\Gamma)(2N_x/2\Gamma)$, in agreement with the first reactive DSB having twice the available target as the second.

Assuming that translocations arise as frequently as dicentrics [28], a topic of debate [29], the expected number of translocations in a $\gamma$-irradiated cell is

$$E(t|D) = \alpha D + \beta D^2 = 0.02D + 0.06D^2, \tag{9}$$

where $D$ is the dose in Gy and the parameter estimates $\alpha = 0.02$ and $\beta = 0.06$ are for dicentrics [1,2]. The predicted BCR–ABL dose-response for doses less than 5 Gy is therefore

$$E(ba|D) = P(ba|t)E(t|D) = 3.5 \times 10^{-10}\left(0.02D + 0.06D^2\right) \tag{10}$$

or

$$E(ba|D) = 7.0 \times 10^{-12}D + 2.1 \times 10^{-11}D^2. \tag{11}$$

A 1 Gy $\gamma$-ray exposure should therefore cause a cell to expect $2.8 \times 10^{-11}$ BCR–ABL translocations, i.e., after 1 Gy we would need to search through $3.7 \times 10^{10}$ cells to find, on average, one BCR–ABL translocation.

## 6. Doses $\geqslant$ 2 Gy

In this section the SMBE model [15] is used to predict the BCR–ABL dose-response curve for low-LET doses greater than about 2 Gy. The SMBE model assumes that the number of active DSBs is a constant fraction, $f$, of the total DSBs. Since eurejoinings are relatively rare at high doses, $f$ at all doses can be estimated as the fraction of DSBs that misrejoin at high doses, which we take as $f = 0.25$ from Löbrich et al. [26]. (In our previous work [15] unrejoinable DSBs were considered as misrejoinings so $f = 0.33$ was used. Here, since unrejoinable active DSBs cannot create BCR–ABL, we ignore them, i.e., the term 'active DSBs' will hereafter refer to those active DSBs which eventually either eurejoin or misrejoin.) The number of active DSBs per human $G_0/G_1$ cell per Gy, $\gamma$, is then

$$\gamma = fY(2\Gamma) = (0.25)(0.0054)(6400) = 8.75. \tag{12}$$

In order to make the SMBE model consistent with both the moderate dose parameter $\beta = 0.06$ in Eq. (9) and the high dose parameter $\gamma = 8.75$ in Eq. (12), the nucleus is treated as a collection of many equal-sized isolated nuclear subvolumes, called sites. To determine the number of sites per cell, $\eta$, we match the low dose limiting form of the SMBE misrejoining dose-response [15], $2\gamma^2D^2/3\eta$, with four times the quadratic component of measured dicentrics (two misrejoinings per dicentric and two misrejoinings per translocation)

$$\frac{2\gamma^2}{3\eta} = 4\beta \Rightarrow \eta = \frac{2\gamma^2}{12\beta} = \frac{2(8.75)^2}{12(0.06)} = 207. \tag{13}$$

We assume that the $\eta$ nuclear sites are each equally likely to contain any particular genetic locus, and that each locus lies completely within one site. The probability that a cell contains a BCR–ABL misrejoining is then $\eta$ times the probability that a particular site contains a BCR–ABL misrejoining. We can therefore restrict our focus to the activities of a single site.

Before embarking on SMBE calculations, let us first determine if introducing sites has any impact on our previous sections. Converting Eqs. (1) and (2) to single site expressions we obtain

$$E_1(b|D) = 2N_{\mathrm{BCR}} fYD/\eta = E(b|D)/\eta \tag{14}$$

and

$$P_1(ba|b) = \frac{E(a|b)/\eta}{2E(m|b)/\eta} = \frac{(2N_{\mathrm{ABL}})fYD}{2(2\Gamma)fYD} = \frac{N_{\mathrm{ABL}}}{2\Gamma} = P(ba|b), \tag{15}$$

so

$$E(ba|D) = \eta E_1(b|D)P_1(ba|b) = E(b|D)P(ba|b). \tag{16}$$

Thus sites have not changed the high dose limiting form of $E(ba|D)$ given by Eq. (4). For low to moderate doses (Section 5), since $E_1(t|D) = E(t|D)/\eta$ implies $E(ba|D) = \eta P_1(ba|t)E_1(t|D) = P_1(ba|t)E(t|D)$, any changes that sites may introduce in $E(ba|D)$ must arise through differences between $P_1(ba|t)$ and $P(ba|t)$. In Section 5 the entire nucleus was one site, so we knew both alleles of BCR and ABL were present in the site. Now we need the probability that BCR and ABL are within the site, and, given that both loci as well as a translocation are in the site, we need the probability that the two translocation derived misrejoinings arise on both BCR and ABL. Of the four loci of interest (2 alleles of BCR and ABL), the probability that more than two loci are in one site is on the order of $1/\eta^3$ and can be ignored. There are four ways in which one of the two BCR alleles can be paired with one of the two ABL alleles, so the probability that a particular site contains BCR and ABL is $4/\eta^2$. Thus

$$\begin{aligned} P_1(ba|t) &= \left(\frac{4}{\eta^2}\right)\left[\left(\frac{N_{\mathrm{BCR}}}{2\Gamma/\eta}\right)\left(\frac{N_{\mathrm{ABL}}}{(2\Gamma - N_{\mathrm{ch22}})/\eta}\right)\right. \\ &\quad \left. + \left(\frac{N_{\mathrm{ABL}}}{2\Gamma/\eta}\right)\left(\frac{N_{\mathrm{BCR}}}{(2\Gamma - N_{\mathrm{ch9}})/\eta}\right)\right] && (17) \\ &= \left(\frac{2N_{\mathrm{BCR}}}{2\Gamma}\right)\left(\frac{2N_{\mathrm{ABL}}}{(2\Gamma - N_{\mathrm{ch22}})}\right) + \left(\frac{2N_{\mathrm{ABL}}}{2\Gamma}\right)\left(\frac{2N_{\mathrm{BCR}}}{(2\Gamma - N_{\mathrm{ch9}})}\right) && (18) \\ &= \left(\frac{N_{\mathrm{BCR}}}{\Gamma}\right)\left(\frac{N_{\mathrm{ABL}}}{(\Gamma - N_{\mathrm{ch22}}/2)}\right) + \left(\frac{N_{\mathrm{ABL}}}{\Gamma}\right)\left(\frac{N_{\mathrm{BCR}}}{(\Gamma - N_{\mathrm{ch9}}/2)}\right) && (19) \\ &= P(ba|t). && (20) \end{aligned}$$

Table 1
The SMBE misrejoining distribution per site resulting from Poisson active DSBs

| | Poisson active DSBs | SMBE misrejoinings | Poisson misrejoinings |
|---|---|---|---|
| Mean | 0.1000000000 | 0.0064074527 | 0.0064074527 |
| Dispersion[a] | 1.0000000000 | 2.0346164365 | 1.0000000000 |
| $p(0)$ | 0.9048374180 | 0.9968393119 | 0.9936130311 |
| $p(1)$ | 0.0904837418 | 0 | 0.0063665285 |
| $p(2)$ | 0.0045241870 | 0.0030768796 | 0.0000203966 |
| $p(3)$ | 0.0001508062 | 0.0000815853 | 0.0000000435 |
| $p(4)$ | 0.0000037701 | 0.0000021784 | 0.0000000000 |
| $p(5)$ | 0.0000000754 | 0.0000000438 | 0.0000000000 |
| $p(6)$ | 0.0000000012 | 0.0000000007 | 0 |
| $p(7)$ | 0.0000000000 | 0.0000000000 | 0 |

[a] The dispersion is defined as the variance divided by the mean.

The introduction of sites has therefore left the results of previous sections unchanged. Keep in mind, however, that proximity effects [30] are implicitly present in the translocation dose-response $E(t|D)$.

Returning now to the SMBE model, we shall assume that after an acute dose of low-LET irradiation, active DSBs per site are Poisson distributed with mean $\mu = (8.75/207)D = 0.0417D$. The SMBE misrejoining distribution (per site) resulting from a Poisson distribution of active DSBs with mean $\mu_1 = 0.1$ per site ($\approx$2 Gy) is compared to a Poisson misrejoining distribution in Table 1. The results indicate that the SMBE misrejoining distribution is broader than a Poisson distribution with the same mean. Furthermore, the SMBE misrejoining dispersion (variance divided by the mean) plotted as a function of dose (Fig. 1) differs considerably from unity. Thus, though others have assumed misrejoinings to be Poisson [14], we shall not.

The BCR–ABL dose-response can be formed as

$$E(ba|D) = \eta \sum_{m=0}^{\infty} E_1(ba|m)p_1(m|D) = \eta \sum_{m=2}^{\infty} E_1(ba|m)p_1(m|D), \tag{21}$$

where $E_1(ba|1) = E_1(ba|0) = 0$ and $p_1(m|D)$ is the distribution of misrejoinings per site at dose $D$. The expected number of BCR–ABL misrejoinings in a site with exactly $m$ misrejoinings, $E_1(ba|m)$, is equal to the probability that BCR and ABL are both in a particular site, multiplied by the probability that a reactive DSB landed on each of these loci, multiplied by the probability that BCR* rejoins with *ABL rather than some other free end. [3] Thus

[3] Although the three factors are probabilities, and thus so too their product, *ba* is an extremely rare event so $E_1(ba|m)$ is very well approximated by $P_1(ba|m)$.

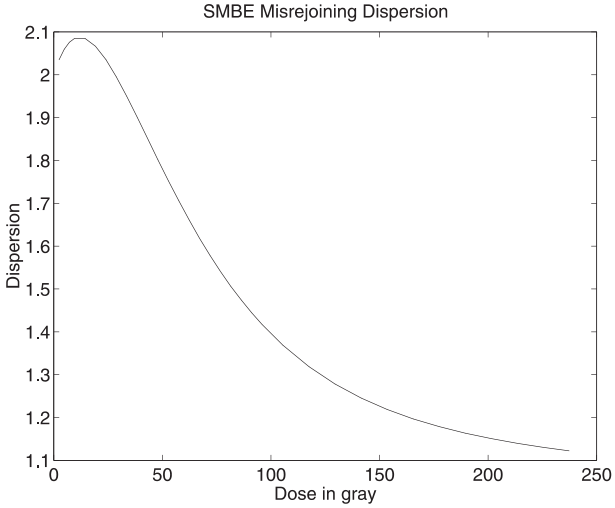


Fig. 1. The dispersion (variance divided by the mean) of the SMBE misrejoining distribution deviates from unity for a substantial range of doses. Here $\eta = 207$ sites, $\gamma = 8.75$ active DSBs per gray per cell, and the initial distribution of active DSBs per site is Poisson with mean $\gamma D/\eta$.

$$E_1(ba|m) = \left(\frac{4}{\eta^2}\right)\left(\frac{mN_{\mathrm{BCR}}}{2\Gamma/\eta}\,\frac{(m-1)N_{\mathrm{ABL}}}{2\Gamma/\eta}\right)\left(\frac{1}{2(m-1)}\right) \tag{22}$$

$$= \left(\frac{mN_{\mathrm{BCR}}}{\Gamma}\,\frac{(m-1)N_{\mathrm{ABL}}}{\Gamma}\right)\left(\frac{1}{2(m-1)}\right), \tag{23}$$

where Eq. (23) was previously derived by Holley and Chatterjee for cells without sites [14], i.e., $E_1(ba|m) = E(ba|m)$. The first factor in Eq. (22), the probability that both BCR and ABL are in a particular site, is derived as in Eq. (17). The second factor of Eq. (22) is understood as follows. Imagine a dart board with area $2\Gamma/\eta$ and two separate bull's eyes with areas $N_{\mathrm{BCR}}$ and $N_{\mathrm{ABL}}$. Throwing $m$ darts (reactive DSBs) on the board while focusing on BCR, the probability that one of these $m$ reactive DSBs hits BCR is $(mN_{\mathrm{BCR}})/(2\Gamma/\eta)$. Given that one of the reactive DSBs does hit BCR, there are $m-1$ reactive DSBs remaining which may have hit ABL. Viewing these as being rethrown at the ABL target, the probability of hitting ABL is thus $((m-1)N_{\mathrm{ABL}})/(2\Gamma/\eta)$. The product of these terms is the probability that both BCR* and *ABL are created in a site with exactly $m$ misrejoinings and exactly one (BCR, ABL) pair. The third factor of Eq. (22) is the probability that BCR* finds *ABL out of the $2(m-1)$ reactive DSB free ends capable of recombining with BCR* (by definition of $m$, BCR* cannot recombine with *BCR). The BCR–ABL dose-response is therefore

$$E(ba|D) = \eta \sum_{m=2}^{\infty} \left( \frac{mN_{\mathrm{BCR}}}{\Gamma} \frac{(m-1)N_{\mathrm{ABL}}}{\Gamma} \right) \left( \frac{1}{2(m-1)} \right) p_1(m|D) \tag{24}$$

$$= \eta \sum_{m=2}^{\infty} \left( \frac{mN_{\mathrm{BCR}}N_{\mathrm{ABL}}}{2\Gamma^2} \right) p_1(m|D) = \left( \frac{\eta N_{\mathrm{BCR}}N_{\mathrm{ABL}}}{2\Gamma^2} \right) \overline{m}_1(D) \tag{25}$$

$$= \left( \frac{N_{\mathrm{BCR}}N_{\mathrm{ABL}}}{2\Gamma^2} \right) \overline{m}(D) \tag{26}$$

$$= \left( \frac{(0.0058)(0.3)}{2(3200)^2} \right) \overline{m}(D) = 8.5 \times 10^{-11} \overline{m}(D), \tag{27}$$

i.e., proportional to the expected total misrejoinings per cell. Note that $E(ba|D)$ depends on the site number only to the extent that $\overline{m}(D)$ depends on the site number.

Let us now examine Eq. (27) using analytic forms of the SMBE model [15] which arise in the limit of large $\mu(D)$,

$$\overline{m}(D) = \mu(D) - \frac{\eta}{2} = \gamma D - \frac{\eta}{2} \tag{28}$$

$$= fY(2\Gamma)D - \frac{\eta}{2}, \tag{29}$$

and in the limit of small $\mu(D)$,

$$\overline{m}(D) = \frac{2[\mu(D)]^2}{3\eta} = 4\beta D^2. \tag{30}$$

Applying Eq. (29) to Eq. (26) we have

$$E(ba|D) = \left( \frac{N_{\mathrm{BCR}}N_{\mathrm{ABL}}}{2\Gamma^2} \right) \overline{m}(D) = \left( \frac{N_{\mathrm{BCR}}N_{\mathrm{ABL}}}{2\Gamma^2} \right) 4\beta D^2$$

$$= \left( \frac{2N_{\mathrm{BCR}}N_{\mathrm{ABL}}}{\Gamma^2} \right) \beta D^2, \tag{31}$$

which also equals the product of Eq. (8) and the quadratic component of Eq. (9), i.e., the SMBE model is consistent with the cytogenetic model of Eq. (11). Applying Eq. (28) to Eq. (26),

$$E(ba|D) = \left( \frac{N_{\mathrm{BCR}}N_{\mathrm{ABL}}}{2\Gamma^2} \right) \overline{m}(D) = \left( \frac{N_{\mathrm{BCR}}N_{\mathrm{ABL}}}{2\Gamma^2} \right) (fY(2\Gamma)D - \eta/2). \tag{32}$$

The slope in this equation agrees with the high dose model given by Eq. (3). The result here, however, also provides the non-zero dose-intercept of the high dose linear asymptote; at 100 Gy Eq. (32) is in slightly better agreement with Ito et al. [12] than Eq. (4).

Applying the SMBE $\overline{m}(D)$ to Eq. (27) for $D \geqslant 2.4$ Gy ($\mu_1 \geqslant 0.1$) and using Eq. (11) for $D \leqslant 2.4$ Gy, we computed the BCR–ABL dose-response $E(ba|D)$ as shown in Fig. 2. The linear-quadratic-linear (LQL) nature of $E(ba|D)$ is evident from these plots. The slight mismatch between the models at 2.4 Gy (Fig. 2, Right Panel) arises because the cytogenetic model includes a linear term not present in the SMBE model, and because the SMBE model undershoots its low-dose quadratic asymptote at $\mu_1 = 0.1$ (i.e., in Table 1 $\overline{m}_1$ equals 0.0064 rather than 0.0067).

## 7. Stem cell estimates

To estimate $N$, the number of hematopoietic stem cells capable of causing CML, we assume that the lifetime excess risk of CML per person has a dose-response $E(cml|D) = N \cdot E(ba|D)$, or equivalently, that a BCR–ABL translocation in one of the $N$ cells is necessary and sufficient to eventually cause CML. Applying this assumption to Eq. (11) we have

$$E(cml|D) = 7.0 \times 10^{-12} N\left(D + 3D^2\right). \tag{33}$$

This equation will subsequently be equated to CML risks from A-bomb data and solved to yield estimates of $N$.

The A-bomb data we use consists of the following list of curve fits for excess CML risks [11]

$$E(cml|D,t) = \begin{cases} 0.17D\,\mathrm{e}^{-0.21(t-25)} & \text{Hiroshima males} \\ 0.05D\,\mathrm{e}^{-0.21(t-25)} & \text{Nagasaki males} \\ 0.70D\,\mathrm{e}^{-0.03(t-25)} & \text{Hiroshima females} \\ 0.20D\,\mathrm{e}^{-0.03(t-25)} & \text{Nagasaki females} \end{cases} \tag{34}$$

where $t$, the number of years since the exposure, is restricted to $t \geqslant 5$ since data collection did not begin until 1950. To approximate the risk of CML in the time interval $0 \leqslant t \leqslant 5$, we set $E(cml|D,t) = 0$ for times $0 \leqslant t < 2$ and $E(cml|D,t) = E(cml|D,5)$ for times $2 \leqslant t \leqslant 5$. Evaluating $E(cml|D) = \int_0^\infty E(cml|D,t)\,\mathrm{d}t$, we found that the lifetime excess risks for CML are $(88/10^4)D$ and $(26/10^4)D$ for males in Hiroshima and Nagasaki, and $(46/10^4)D$ and $(13/10^4)D$ for females in Hiroshima and Nagasaki, respectively. Equating the magnitudes of these linear functions at 1 Gy to the magnitude of the linear-quadratic CML dose-response (Eq. (33)) at 1 Gy, we estimate $N$ to be $3.1 \times 10^8$ and $9.2 \times 10^7$ for Hiroshima and Nagasaki males, and $1.6 \times 10^8$ and $4.7 \times 10^7$ for Hiroshima and Nagasaki females, respectively. Thus, assuming that the models are reasonable, the true number of stem cells should have an order of magnitude consistent with the range of $5 \times 10^7$–$3 \times 10^8$.

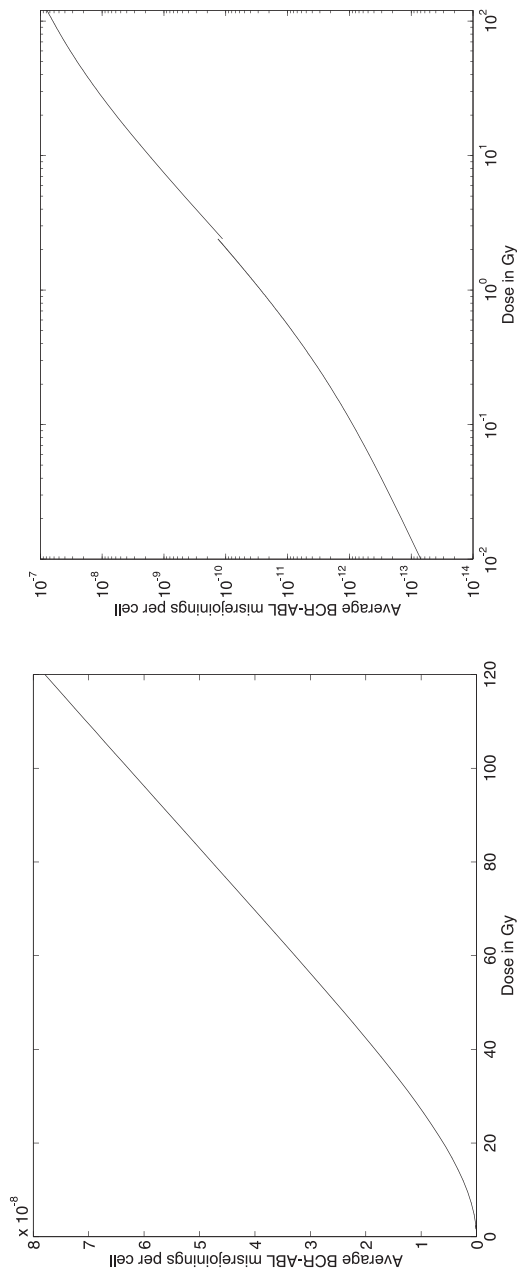


Fig. 2. The expected number of BCR-ABL misrejoinings per cell is shown in these plots as a function of dose. For doses greater than 2.4 Gy ($\mu_1 = 0.1$) the dose-response $E(ba|D)$ was evaluated using Eq. (27) with $\overline{m}(D)$ determined by the SMBE model. For doses less than 2.4 Gy $E(ba|D)$ was determined by the cytogenetic model of Eq. (11). The slight discontinuity at 2.4 Gy is described in the text. For the SMBE model we used $\eta = 207$ sites, $\gamma = 8.75$ active DSBs per gray per cell, and an initial distribution of active DSBs per site that is Poisson with mean $\gamma D/\eta$.

## 8. Discussion

The BCR–ABL dose-response (Fig. 2) was described in terms of three models. The first of these, valid in the limit of very high doses, approximates the third model in the same limit of high doses. The second model, valid for low to moderate doses, uses minimal assumptions to convert translocation (actually dicentric) dose-response data directly into a predicted CML dose-response. The second model is linked to the first model by a third model valid for moderate doses and higher. The three models were shown to be self-consistent. This was expected since the models are based on the same assumptions, and, at their respective dose boundaries, the same data.

In principle it is possible to extend the third model to cover the entire dose range. The extension involves assigning an LET 'equivalent' to the radiation, defining a Poisson active DSB distribution per site for a single particle track using the LET equivalent and an assumption of constant LET across the length of each track, and replacing the initial active DSB distribution (per site) with an over-dispersed compound Poisson–Poisson distribution (the Poisson number of tracks per site being compounded with the Poisson number of active DSBs per site per track). The result would be one model covering the entire dose range. However, the overall dose-response would be no different than for the two separate models since the LET equivalent in the third model would be identified by forcing its dose-response to agree with the measured translocation dose-response $\alpha D + \beta D^2$ (Eq. (9)) used in the second model, i.e., we would just be encoding the $\alpha$ parameter of the second model into a new LET-equivalent parameter of the third model.

By comparing excess lifetime CML risks among A-bomb survivors to the BCR–ABL dose-response of the second model, we concluded that the CML target stem cell number is on the order of $5 \times 10^7$–$3 \times 10^8$ cells. These estimates are low compared to Holmberg's [31] initial assumption that $N = 1 \times 10^9$ stem cells, which he then used with the same A-bomb data to conclude that BCR–ABL induction could indeed be the primary event in CML formation. Since Holmberg's conclusion corresponds to our assumption, it seems reasonable that our conclusion regarding $N$ should be consistent with Holmberg's assumption regarding $N$. Why the discrepancy? A factor of three can be accounted for by Holmberg's choice of $N_{\mathrm{BCR}} = 3$ kbp, $N_{\mathrm{ABL}} = 200$ kbp and $\Gamma = 2950$ Mb, compared to our choices of $N_{\mathrm{BCR}} = 5.8$ kbp, $N_{ABL} = 300$ kbp and $\Gamma = 3200$ Mb, and an additional factor of two arises from Holmberg's apparent mistake of using $P(ba|t) = N_{\mathrm{BCR}} N_{\mathrm{ABL}}/\Gamma^2$ instead of $P(ba|t) = 2N_{\mathrm{BCR}} N_{\mathrm{ABL}}/\Gamma^2$, as in Eq. (8). Taking these differences into account, the discrepancy vanishes. Note that while Holmberg assumed males and females have the same $N$, we predict that $N$ in females is about 60% the number expected in males.

## Acknowledgements

We thank W.R. Holley, R.J. Preston and T. Mizuno for discussions. This publication is supported in part by funds from the US Department of Energy cooperative agreement DE–FC02–98CH10902.

## References

[1] M. Bauchinger, Quantification of low-level radiation exposure by conventional chromosome aberration analysis, Mut. Res. 339 (1995) 177.
[2] D.C. Lloyd, A.A. Edwards, Chromosome aberrations in human lymphocytes: effect of radiation quality, dose and dose rate, in: T. Ishihara, M. Sasaki (Eds.), Radiation-Induced Chromosome Damage in Man, Liss, New York, 1983, p. 23.
[3] T.H. Rabbits, Chromosomal translocations in human cancer, Nature 372 (1994) 143.
[4] J.V. Melo, The diversity of BCR–ABL fusion proteins and their relationship to leukemia phenotype, Blood 88 (1996) 2375.
[5] A.T. Look, Oncongenic transcription factors in the acute human leukemias, Science 278 (1997) 1059.
[6] D. Brown, S. Kogan, E. Lagasse, I. Weissman, M. Alcalay, P.G. Pelicci, S. Atwater, J.M. Bishop, A PML/RAR$\alpha$ transgene initiates murine acute promyelocytic leukemia, Proc. Nat. Acad. Sci. USA 94 (1997) 2251.
[7] L.Z. He, C. Triboli, R. Rivi, D. Peruzzi, P.G. Pelicci, G.C.V. Soares, P.P. Pandolfi, Acute leukemia with promyelocytic features in PML/RAR$\alpha$ transgenic mice, Proc. Nat. Acad. Sci. USA 94 (1997) 5302.
[8] G.Q. Daley, R.A.V. Etten, D. Baltimore, Induction of chronic myelogenous leukemia in mice by the P210$^{bcr/abl}$ gene of the Philadelphia chromosome, Science 247 (1990) 824.
[9] N. Heisterkamp, G. Jenster, J. TenHoeve, D. Zovich, P.K. Pattengale, J. Groffen, Acute Leukemia in bcr/abl transgenic mice, Nature 340 (1990) 251.
[10] J.W. Voncken, V. Kaartinen, P.K. Pattengale, J.G.W.T.V. Germeraad, N. Heisterkamp, BCR/ABL P210 and P190 cause distinct leukemia in transgenic mice, Blood 86 (1995) 4603.
[11] D.L. Preston, S. Kusumi, M. Tomonaga, S. Izumi, E. Ron, N.K.A. Kuramoto, H. Dohy, T. Matsui, H. Nonaka, D.E. Thompson, M. Soda, K. Mabuchi, Cancer incidence in atomic bomb survivors. Part III: leukemia, lymphoma and multiple myeloma, 1950–1987, Radiat. Res. 137 (1994) S68.
[12] T. Ito, T. Seyama, T. Mizuno, T. Hayashi, K.S. Iwamoto, N.N.K. Dohi, M. Akiyama, Induction of BCR–ABL fusion genes by in vitro x-irradiation, Jpn. J. Cancer Res. 84 (1993) 105.
[13] T. Mizuno, S. Kyoizumi, T. Suzuki, K.S. Iwamoto, T. Seyama, Continued expression of a tissue specific activated oncogene in the early steps of radiation-induced human thyroid carcinogenesis, Oncogene 15 (1997) 1455.
[14] W.R. Holley, A. Chatterjee, A computational approach to the relationship between radiation induced double strand breaks and translocations, Basic Life Sci. 63 (1994) 251.
[15] T. Radivoyevitch, D.G. Hoel, A.M. Chen, R.K. Sachs, Misrejoining of double-strand breaks after X-irradiation: relating moderate to very high doses by a Markov model, Radiat. Res. 149 (1998) 59.
[16] M. Kai, E.G. Luebeck, S.H. Moolgavkar, Analysis of the incidence of solid cancer among atomic bomb survivors using a two-stage model of carcinogenesis, Radiat. Res. 148 (1997) 348.

[17] S.H. Moolgavkar, A.G. Knudson, Mutation and cancer: A model of human carcinogenesis, J. Nat. Cancer Inst. 66 (1981) 1037.
[18] A.M. Kellerer, Fundamentals of microdosimetry, in: K. Kase, B. Bjarngard, F. Attix (Eds.), The Dosimetry of Ionizing Radiation, Academic Press, Orlando, FL, 1985, pp. 77–162.
[19] C. von Sonntag, The Chemical Basis of Radiation Biology, Taylor and Francis, London, 1987.
[20] J.F. Ward, The yield of DNA double-strand breaks produced intracellularly by ionizing radiation: a review, Int. J. Radiat. Biol. 57 (1990) 1141.
[21] M. Lobrich, S. Ikpeme, J. Kiefer, Measurement of DNA double-strand breaks in mammalian cells by pulsed-field gel electrophoresis: a new approach using rarely cutting restriction enzymes, Radiat. Res. 138 (1994) 186.
[22] B. Cedervall, R. Wong, N. Albright, J. Dynlacht, W.C. Dewey, Methods for the quantification of DNA double-strand breaks determined from the distribution of DNA fragment sizes measured by pulsed-field gel electrophoresis, Radiat. Res. 143 (1995) 8.
[23] T. Radivoyevitch, D.G. Hoel, P.J. Hahnfeldt, B. Rydberg, R.K. Sachs, Recent data obtained by pulsed-field gel electrophoresis suggest two types of double-strand breaks, Radiat. Res. 149 (1998) 52.
[24] K. Sax, An analysis of X-ray induced chromosomal aberrations in tradescantia, Genetics 25 (1940) 41.
[25] P.C. Nowell, D.A. Hungerford, A minute chromosome in human chronic granulocytic leukemia, Science 132 (1960) 1497.
[26] M. Lobrich, B. Rydberg, P.K. Cooper, Repair of X-ray induced DNA double-strand breaks in specific NotI restriction fragments in human fibroblasts: joining of correct and incorrect ends, Proc. Nat. Acad. Sci. 92 (1995) 12050.
[27] N.E. Morton, Parameters of the human genome, Proc. Nat. Acad. Sci. USA 88 (1991) 7474.
[28] J.N. Lucas, A.M. Chen, R.K. Sachs, Theoretical predictions on the equality of radiation-produced dicentrics and translocations detected by chromosome painting, Int. J. Radiat. Biol. 69 (1996) 145.
[29] S. Knehr, H. Zitzelberger, H. Braselmann, U. Nahrstedt, M. Bauchinger, Chromosome analysis by fluorescence in situ hybridization: Further indications for a non-DNA-proportional involvement of single chromosomes in radiation-induced structural aberrations, Int. J. Radiat. Biol. 70 (1996) 385.
[30] R.K. Sachs, A.M. Chen, D.J. Brenner, Review: Proximity effects in the production of chromosome aberrations by ionizing radiation, Int. J. Radiat. Biol. 71 (1997) 1.
[31] M. Holmberg, Is the primary event in radiation-induced chronic myelogenous leukemia the induction of the $t(9;22)$ translocation? Leukemia Res. 16 (1992) 333.