



SCIENTIFIC REPORTS



OPEN

In Silico Non-Homologous End Joining Following Ion Induced DNA Double Strand Breaks Predicts That Repair Fidelity Depends on Break Density

N. T. Henthorn¹ , J. W. Warmenhoven¹, M. Sotiropoulos¹ , R. I. Mackay², N. F. Kirkby^{1,3}, K. J. Kirkby^{1,3} & M. J. Merchant^{1,3}

This work uses Monte Carlo simulations to investigate the dependence of residual and misrepaired double strand breaks (DSBs) at 24 hours on the initial damage pattern created during ion therapy. We present results from a nanometric DNA damage simulation coupled to a mechanistic model of Non-Homologous End Joining, capable of predicting the position, complexity, and repair of DSBs. The initial damage pattern is scored by calculating the average number of DSBs within 70 nm from every DSB. We show that this local DSB density, referred to as the cluster density, can linearly predict misrepair regardless of ion species. The models predict that the fraction of residual DSBs is constant, with 7.3% of DSBs left unrepaired following 24 hours of repair. Through simulation over a range of doses and linear energy transfer (LET) we derive simple correlations capable of predicting residual and misrepaired DSBs. These equations are applicable to ion therapy treatment planning where both dose and LET are scored. This is demonstrated by applying the correlations to an example of a clinical proton spread out Bragg peak. Here we see a considerable biological effect past the distal edge, dominated by residual DSBs.

The induction of DNA double strand breaks (DSBs) is the mechanism by which radiation kills cells; with DSBs either repaired, misrepaired, or left unrepaired (residual). Misrepaired DSBs, considered potentially lethal lesions¹, can lead to chromosome-type or chromatid-type aberrations, which are experimentally measurable². These aberrations cause genomic instability, contributing to the chance of cell death³ and development of secondary cancers. Residual DSBs are likely to cause cell cycle arrest if detected by the various cell cycle checkpoints. The unresolved residual DSBs will cause cell death, and residuals that are missed by the G2/M checkpoint will cause the cell to undergo mitotic catastrophe. For this reason residual DSBs are conventionally considered to contribute substantially to the likelihood of cell death³.

A difference in the cell killing effectiveness between radiation qualities has previously been noted in the literature^{4,5}, where a greater effect is seen for higher linear energy transfer (LET) radiations. Clinically it is important to understand this difference as it impacts the required dose for tumour control between radiation qualities, such as photons and protons. In practice, this is accounted for by use of the proton relative biological effectiveness (RBE), which is the ratio of proton dose relative to photon dose to achieve iso-effectiveness. An RBE value of 1.1 is in clinical use to give the physical proton dose that is biologically equivalent to the required photon dose^{6,7}. However, it has been shown that the proton RBE is variable depending on a number of factors including dose, fraction, α/β ratio, and LET⁸. Marshall *et al.*⁹ have recently shown that consideration of RBE variation along a proton Spread Out Bragg Peak (SOBP) results in increases in both the delivered dose and range when compared to a

¹Division of Molecular and Clinical Cancer Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK. ²Christie Medical Physics and Engineering, The Christie NHS Foundation Trust, Manchester, UK. ³The Christie NHS Foundation Trust, Manchester, UK. N. T. Henthorn and J. W. Warmenhoven contributed equally to this work. Correspondence and requests for materials should be addressed to J.W.W. (email: john.warmenhoven@postgrad.manchester.ac.uk)

constant RBE. To fully exploit the beneficial dose deposition characteristics of protons over photons, the variable RBE must be understood and incorporated into Treatment Planning Systems (TPS).

RBE has been linked to LET, where the LET of a particle is defined as the energy deposited per unit length. An increase in LET is equivalent to an increase in the number of ionisations in the same distance. In the context of particle therapy, a higher LET causes a greater number of DSBs¹⁰ as well as a higher complexity of DSB¹¹. The proximity between DSBs is also dependent on the LET, with high LET leading to more proximal breaks due to the closer spacing between energy depositions¹². Assuming irradiation in the same cell line and cell cycle phase, i.e. same α/β or simplistically “the same biology”, then the differences in the success or failure of the cell’s DNA damage response (DDR) depends only on these three DSB factors; the number of, complexity of, and proximity of DSBs, regardless of the radiation quality used. Similar statements have been expressed in other work, for example the Local Effects Model (LEM)^{13,14}.

Break complexity and proximity have both been suggested to influence misrepair¹⁵. It has been proposed that either a class of DSB exists that has a propensity for causing aberrations^{16,17}, or that misrepair occurs through pairwise interaction of DSBs dependent on their proximity¹⁸. The influence of DSB proximity on chromosome aberrations has been reviewed previously¹⁹ and has been investigated experimentally²⁰. Determining the distance over which these DSB interactions occur is of importance for mechanistic models of DNA repair^{21,22}. However, predictions for the value of this DSB separation varies within the literature; with values ranging from 0.25 μm to 1.30 μm ^{12,23–27}. The motion of DSB ends has a large impact on this value and is a debated topic in the literature. Some experimental work has shown that DSB ends in metazoan have limited motion in the nucleus and are unable to explore the entire volume^{28,29}, travelling less than 0.5 μm in a 10 μm diameter nucleus³⁰. On the other hand, evidence has also been published to suggest greater mobility of DSB ends³¹, particularly along the tracks of damage caused by heavy ions^{31,32}. More directed motion has also been suggested by Neumaier *et al.*, reporting the observation of repair centres in human cells³³. To fully understand the consequence of how DSB end mobility leads to chromosome aberrations a more detailed repair model must be used.

Averbeck *et al.*³⁴ have shown that DSBs induced across the carbon Spread Out Bragg Peak (SOBP) are efficiently repaired, although there is an increase in cell death. The authors attribute this increase to a greater number of misrepaired DSBs, highlighting the role of misrepair in RBE. It is known that, for the same cell line, the RBE with different ion species at the same LET (iso-LET) is substantially different³⁵. This implies that the change in RBE of ions arises from differences in track structure³⁶, the energy deposition pattern at the nanoscale. This can explain both DSB complexity and proximity, information that is not fully encompassed by the LET alone. The concept of track structure leading to biological outcomes is a core principle of the emerging field of nanodosimetry³⁷. To our knowledge, no study has assessed the difference in misrepair induction between ion species and across a range of LET relevant to the clinical setting.

In this work, Monte Carlo simulations are used to predict the damage and repair of cells irradiated by protons, alpha particles, and carbon ions at a range of iso-LET. By scoring the average number of DSBs within close proximity of each other we are able to show a linear relation to misrepair across all the radiation qualities. We believe that this local DSB density, which we refer to as the cluster density, can be used as a predictor of chromosome aberrations. The residual DSBs at 24 hours of repair are also investigated. Our model predicts a constant fraction of DSBs left unrepaired, regardless of ion type, dose, or LET. From analysis of the results, we have developed correlations that are capable of predicting the yield of residual and misrepaired DSBs for our model. These correlations express the impact that the physics of the beam, dose and LET, has on the early biological effect, misrepaired and residual DSBs at 24 hours. The trends displayed by these correlations for misrepair and residual DSBs have a relevance to proton therapy planning.

It is recognised that the work contained here covers a wide range of fields which not all readers may be familiar with. We include a table summarising the field specific terms used (Supplement 1).

Methods

Simulation of DSBs. This work uses the Monte Carlo based toolkit Geant4 (10.02 patch 01)³⁸, with the default G4EmDNAPhysics list. Incerti *et al.* have presented and validated the parameters of this physics list³⁹, which has been used in a number of published DNA damage simulations^{40–42}. The Geant4-DNA extension accurately models event-by-event particle tracking down to low energies, simulating each track as a series of steps determined by physical interactions.

We use this toolkit to simulate the transport of different ion species across a water medium representing a simplified cell model. Using water as a surrogate for biological material has become a standard assumption in Monte Carlo studies of DNA damage⁴³. The cell geometry consists of a 5 μm diameter spherical nucleus, as might be typical of a lymphocyte⁴⁴, in the centre of a 10 μm box, used as a surrogate for the cellular cytoplasm. A uniform average dose is delivered to the cell through methods described in Supplement 2. DSBs are calculated according to the principles of nanodosimetry through assessing the clustering of energy depositions within the nucleus. This work follows the methodology of Francis *et al.*⁴⁵, where assumptions are made in order to convert energy depositions into strand breaks. The methodology assumes that a fraction of the nucleus is occupied by the sensitive material (DNA backbones). We determined the sensitive fraction of the nucleus as 15% by fitting the predicted initial yield of DSBs to literature data^{40,42,46–48}, Supplement 3. The cell model implicitly considers indirect damage by fitting the sensitive fraction to literature data that includes indirect effects. In our model energy depositions in the sensitive volume above 37.5 eV are guaranteed to cause strand damage, whilst energy depositions below 5 eV cannot cause strand damage, with the probability increasing linearly across this range⁴⁶. An accepted damage site is randomly assigned to strand 1 or 2 of the double helix and the position is recorded. Following the simulation of all primary particles, to achieve the required dose, the list of damage sites is analysed by a clustering algorithm based on a modified implementation of the Density-Based Spatial Clustering of Applications with Noise (DBSCAN)⁴⁹

algorithm. The algorithm determines DSBs by clustering damage sites that are on opposite strands and separated by a maximum distance of 3.32 nm, equivalent to the separation of 1 helical turn (10 bp)⁵⁰.

A separate simulation of the chromatin fibre is used to determine the structure of individual DSBs, most of the details of which were reported in our previous work⁵¹. We build a model of the double helix, with backbones and bases modelled as quarter cylinders and half cylinders⁵² with volumes of 0.28 nm³ and 0.13 nm³ respectively. The double helix is wound around cylindrical histones in 1.65 left-handed turns to form the nucleosome, which are then arranged in the solenoid chromatin conformation⁵³. The chromatin is irradiated with primary ions matching the energy of those used in the cell model. For each incident ion, energy depositions are cumulatively scored in the DNA volumes and the same energy dependent probability of damage induction is applied (5–37.5 eV). Indirect damage is included with use of the Geant4-DNA chemistry modules⁵⁴. Hydroxyl radicals crossing a DNA backbone or base are assigned a probability of inducing damage, with the probability fitted to produce 65% strand damage due to indirect effects when the fibre is irradiated by a 60-Co source⁴⁶. DNA volumes damaged by direct and/or indirect effects are recorded per primary particle. The damaged volumes are analysed by an improved clustering algorithm, since both the genomic separation between damage sites and their strand is known. This removes any discrepancy that may arise due to damaged backbones within 3.32 nm but separated by more than 10 bp. A damaged base is included in the cluster if it is between 3 bp of the extreme ends of the backbones involved in the break¹¹. Damaged bases directly attached to a damaged backbone are not included in the cluster since it is assumed that these will be removed along with the backbone during repair. The process is repeated for at least 10⁶ primary ions recording details of the induced DSBs per primary to develop a break complexity library.

The results of the cell and fibre model are combined to give the positions of DSBs, calculated as the central coordinate of the damage sites in the cluster, and the break complexity, defined as the number of extra SSBs and base lesions included in the break. This damage data is then passed to our repair model.

Non-Homologous End Joining Repair Model. The geometric distribution and complexity of DSBs are used to convert each DSB into two pseudo-molecules representing exposed DNA strand ends. Ends are allowed to move within the cell nucleus by sub-diffusive motion, where mean squared displacement scales with time as t^α , with $\alpha < 1$. This has been shown experimentally to better model DSB end mobility compared to Brownian motion²⁹. We have assumed the continuous time random walk implementation of sub-diffusion⁵⁵. In this method DSB ends are assigned a random waiting time drawn from an exponential distribution, during which they do not move. After this waiting time, the end is displaced in a random direction with length drawn from a Gaussian distribution, and a new waiting time is assigned. Any proposed step leaving the nuclear envelope is rejected and resampled.

Simultaneous to their motion, DSB ends undergo a series of stochastic time constant based state changes. Each change of state represents the recruitment and action of a particular protein or complex of proteins, implicitly assumed to be randomly distributed throughout the nucleus. Time constants are fitted to reproduce the known recruitment kinetics of these proteins or complexes of proteins. We have chosen to model only the canonical non-homologous end joining (c-NHEJ) process as it is computationally less intensive; the impact of this is discussed later. The c-NHEJ process has an agreed core group of proteins which bind sequentially during repair⁵⁶. DSB ends are initially placed with no associated proteins. The ends then either change to an inhibited state or to a state with Ku70/80 attached. The inhibited state represents attachment of proteins from competing processes occupying the DSB end and prevent Ku70/80 from binding. Once a DSB end has recruited Ku70/80 it cannot freely dissociate back to a naked DSB end⁵⁷, but must progress to a state with DNA-PKcs attached. Once a DSB end has recruited Ku70/80 and DNA-PKcs (DNA-PK complex) it cannot dissociate to a previous state⁵⁸. Two ends in the DNA-PK complex state, that are within 25 nm, can react to form a long range synaptic complex⁵⁹. In this complex DNA-PKcs can cross/auto-phosphorylate, and in turn phosphorylate Ku70/80⁵⁷. The synaptic complex can either dissociate to two DSB ends with no attached proteins, or move to a short range synaptic complex state⁵⁹. Dissociation occurs with a short time constant resulting in only transient formation of the long-range complex⁵⁹. We have used this model to reproduce results of fluorescent recovery after photo-bleaching (FRAP) experiments reported in the literature⁵⁸. Once in a short range synaptic complex state the repair pathway cleans all extra lesions associated to the break before final ligation can occur, with time constants taken from literature²⁶. The final ligation step represents the actions of XLF, XRCC4, DNA-Ligase IV, and polymerases. The c-NHEJ repair model allows us to investigate correctly and incorrectly repaired DSBs, as well as the number of residual DSBs for a given repair time. The full details of this repair model will be published separately at a later date.

Data availability. The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Results

Misrepair and LET. The cell model was irradiated with 1, 2, or 5 Gy of protons, alphas, and carbon-12⁶⁺ across a range of track averaged linear energy transfer (LET). The NHEJ model was run for a repair time of 24 hours, tracking motion, protein recruitment, and the repair of DSB ends. The fraction of DSBs that misrepaired was recorded, calculated as a fraction of the DSBs that are joined at 24 hours. Figure 1a shows a difference in misrepair between the three ions investigated, with the difference becoming more pronounced at higher values of LET. Similar values for the fraction misrepaired are seen between different doses, though there is a difference in the absolute number misrepaired. Carbon-12⁶⁺ and alpha particles show relatively similar behaviour, however, protons have a higher fractional misrepair. We investigate this difference in terms of the repair pathway and the ion induced damage patterns.

Within the c-NHEJ model only two DNA-PK complexes can undergo synapsis, causing possible misrepair. Ku70/80 has a high affinity for naked DSB ends, and once loaded rapidly recruits DNA-PKcs^{60–62}. The recruitment

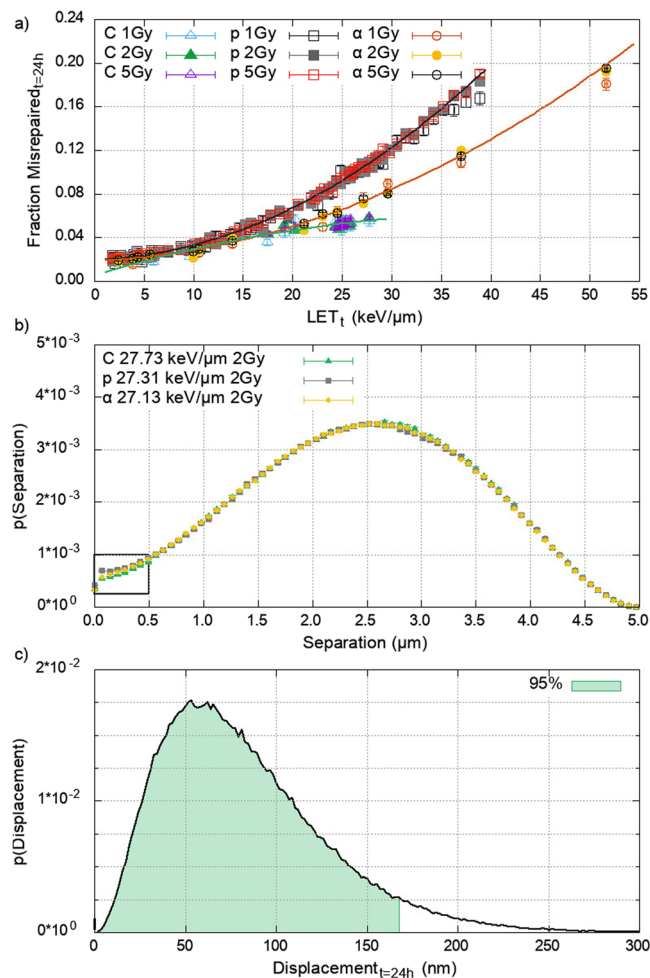


Figure 1. (a) The average misrepair of DSBs as a fraction of total DSBs joined at 24 hours at a range of LET_t for protons (p), alphas (α), and carbon-12⁶⁺ (C). A dose independent difference in misrepair between iso-LET_t ions is shown. Carbon-12⁶⁺ shows anomalous behaviour between LET_t of 20–30 keV/μm, discussed later. Error bars show the standard error in the mean between 200 repeat simulations. Lines show second order polynomial fits to guide the eye. (b) The DSB separation probability density function for iso-LET_t protons (p), alpha (α), and carbon-12⁶⁺ (C) in a 5 μm diameter spherical nucleus. The distributions show the probability that another DSB is located at a given separation from any DSB. Differences are only seen at small separations (inset), due to separations between intra-track DSBs. (c) The DSB end displacement at 24 hours. Showing that 95% of DSB ends move 168 nm or less in 24 hours. This limits DSB interactions to nearby neighbours where the DSB separation distribution is different between the ions.

of Ku70/80 is largely unaffected by break complexity⁶², and both Ku70/80 and DNA-PKcs are abundant in the cell making depletion of local concentrations near composite break sites unlikely⁶³. Therefore, in the model there is a similar and rapid formation of DNA-PK complexes at break sites, regardless of the radiation quality used. As such this cannot explain the difference in misrepair observed.

Different particles at iso-LET_t produce different track structures as they pass through the nucleus. The resultant difference in the energy deposition pattern leads to differences in DSB complexity. To determine the influence of break complexity on misrepair for a single ion species, the repair model was rerun for two cases; one populated with only simple DSBs and one populated with only complex DSBs. The predicted misrepair fraction for either case is not different from the data in Fig. 1a (Supplement 4). Furthermore, we investigate the effect that break complexity has on our ability to differentiate misrepair between ions species, using the same cases as above. For the data in Fig. 1a, the differences in misrepair are statistically indistinguishable below an LET_t of 5 keV/μm. Neither case alters this value, showing that within our model the simulated break complexity does not impact misrepair.

A further implication of the different energy deposition patterns is a difference in the spatial distribution of the DSBs themselves. We assess proximity between DSBs by determining the probability that a DSB is within a radial distance from another DSB. Figure 1b shows the probability density functions (PDF) of DSB separation for the three ion species at an iso-LET_t equivalent to the distal edge of the proton SOBP (≈27 keV/μm). A non-zero probability for separations of 0 μm is shown since it includes DSBs that are separated between 0 and 10 nm.

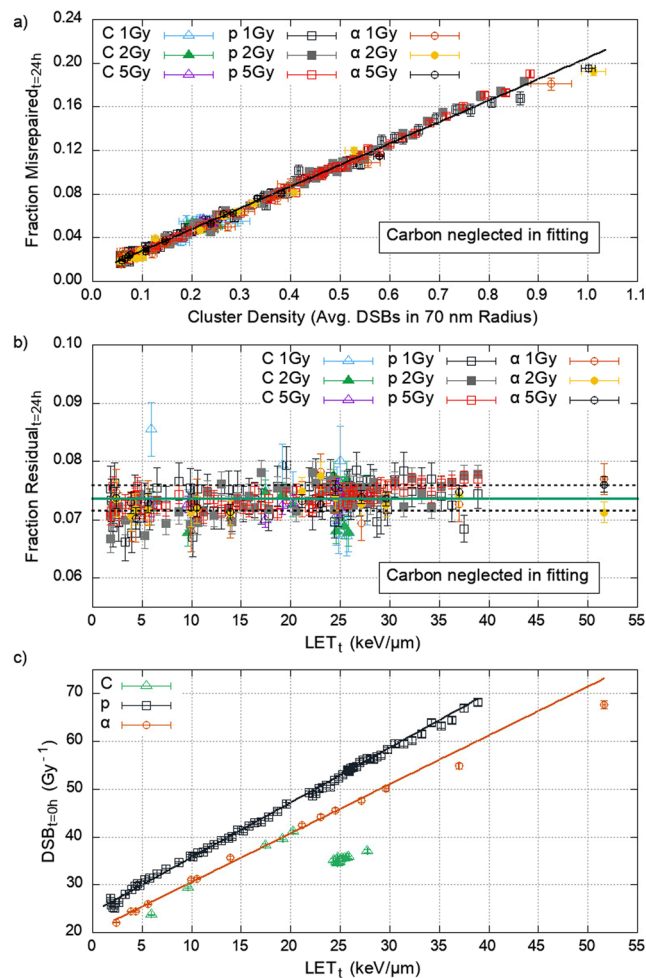


Figure 2. (a) The average number of DSBs separated by less than 70 nm from any given DSB (cluster density) and the corresponding misrepair at 24 hours for protons (p), alphas (α), and carbon-12⁶⁺ (C). Showing a linear relation ($R^2 > 0.9$) with fitting parameters presented in Table 1. Error bars represent the standard error in the mean between multiple repeat simulations (2500 for cluster density, 200 for misrepair). (b) Fraction of DSBs left unrepaired following 24 hours of repair across the LET_t range investigated. The solid line shows a constant residual fraction of 0.073. Dashed lines show the standard deviation in this value. Error bars are the standard error in the mean for 200 repeats. (c) The predicted number of DSBs per unit dose for protons, alphas and carbon-12⁶⁺ as a function of LET_t, with lines showing linear fits for protons and alphas. There is a clear discontinuity in carbon-12⁶⁺ data, with lines showing linear fits for protons and alphas. There is a clear discontinuity in carbon-12⁶⁺ data, its data is not used to determine any correlations but is included in the plots for interest only.

Similarities at larger DSB separations are seen, indicating that it is equally probable to have DSBs separated by a value greater than around 0.5 μm. Below 0.5 μm a difference can be seen between the ions, due to separations of intra-track DSBs. Here, we see a greater probability of proximal DSBs in the proton-irradiated cell compared to the alpha and carbon-12⁶⁺ case. The importance of this scale is further established in Fig. 1c. This shows that the end displacement in 24 hours, 95% moving less than 168 nm, limits interactions to nearby breaks only and coincides with the DSB separations for which the iso-LET_t ions are different. As such break proximity can explain differences in misrepair.

Cluster Density, Misrepair, and Residual DSBs. To quantify the initial damage patterns, the average neighbouring DSBs within a given radius for each DSB is determined, which we refer to as the “cluster density”. If motion of DSB ends is predominantly responsible for misrepair, then this value scales with the number of potential incorrect partners a DSB end has available for interaction. The cell model was irradiated with the same radiation qualities as shown in Fig. 1a, and the average cluster density was calculated for 2500 repeat simulations. We find that the cluster density has the strongest correlation to misrepair when calculated with a radial distance of 70 nm, giving a minimum in χ^2 . However, the strength of this correlation is not greatly decreased by deviations of ± 30 nm from this value (Supplement 5). Figure 2a shows a strong linear relation between cluster density and misrepair. This shows that the local DSB concentration (cluster density) in the initial damage pattern is predictive

Parameter	Proton \pm %	Alpha \pm %
a	0.1966 \pm 0.4	0.1966 \pm 0.6
b	0.008 \pm 3.4	0.008 \pm 3.4
c	0.0736 \pm 0.2	0.0736 \pm 0.2
d	1.149 \pm 1.0	1.02 \pm 2.5
e	24.10 \pm 0.6	20.2 \pm 1.3
f	4.879E-4 \pm 0.8	3.25E-3 \pm 5.0
g	2.84E-3 \pm 4.7	1.61E-3 \pm 33
h	5.13E-2 \pm 1.6	5.51E-2 \pm 5.7

Table 1. The correlation fitting parameters developed in this work for protons and alphas with their asymptotic standard error shown as a percentage error.

of the misrepair at 24 hours in our model. Based on our results, we suggest this physico-marker can be used to indicate the impact that beam parameters have on biological response. Cluster density can be estimated from the ion LET_t through a fitted second order polynomial (Supplement 6), with parameters given in Table 1. Generally, higher LET_t results in a higher cluster density, with protons producing a higher cluster density relative to iso-LET_t alphas (a similar trend as seen in Fig. 1a).

In the context of radiotherapy, residual DSBs are another biological endpoint of interest. Our model predicts that the fraction of residual DSBs following 24 hours of repair is constant across the radiation qualities investigated, with an average of 7.3% breaks left unrepaired. Figure 2b shows that this constancy holds across the ion species, dose, and LET_t range investigated.

To apply the misrepair and residual correlations to a clinically relevant case, and to make predictions of absolute yields, the number of DSBs is determined as a function of LET_t and Dose. For the clinically relevant LET_t range of 0–20 keV/μm we observe a linear correlation to initial DSB yield for protons and alphas (Fig. 2c). This linear trend has been reported in the literature for both ions^{46,48,64}. The DSB yield for alpha particles begins to deviate from linearity above ≈30 keV/μm. However, in the rest of this work we only use the correlation between the bounds of its linearity. The constant fraction of residual DSBs at 24 hours combined with the linear dependence of initial DSB yield on LET_t leads to a linear relationship between the yield of residuals and LET_t (Supplement 8); this has been experimentally observed for 53BP1 foci under similar conditions by Chaudhary *et al.*⁶⁵, although at a systematically lower yield. Potential causes for this are discussed later. The parameters for all correlations are given in Table 1.

The DSB yield for carbon ions is linear up to ≈20 keV/μm, after which there appears to be a discontinuity, also noticed by other users of Geant4-DNA (personal communication S. McMahon). The behaviour originates from a switching between relativistic and classical calculations in the Geant4-DNA models. This may have implications for studies simulating clinical carbon ions (Supplement 7). It is possible to use only relativistic calculations, though this has not been validated by the Geant4-DNA collaboration, and further implications are unknown. As such we have chosen to use the default Geant4-DNA models and not include carbon-12⁶⁺ data in any of the correlations.

Residual and Misrepaired DSB Yields. Here we present the correlations from Fig. 2a,b and c. We convert the correlations of fractional misrepair and residual DSBs into absolute yields, shown in equations (4) and (6). In order to apply the correlations to an example of a clinical proton SOBP, we determine the dependency on dose and LET_t, shown in equations (5) and (7).

$$F_{Mis.} = a \cdot Cluster\ Density + b \quad (1)$$

$$F_{Res.} = c \quad (2)$$

$$N_{DSB} = D \cdot (d \cdot L + e) \quad (3)$$

$$Misrepair\ Yield = N_{DSB} \times \{a \cdot Cluster\ Density + b\} \times \{1 - c\} \quad (4)$$

$$D \cdot (d \cdot L + e) \times \{a \cdot [f \cdot L^2 + g \cdot L + h] + b\} \times \{1 - c\} \quad (5)$$

$$Residual\ Yield = N_{DSB} \times c \quad (6)$$

$$D \cdot (d \cdot L + e) \times c \quad (7)$$

Equations 1–7: F_{res} is the fraction of DSBs that are unresolved at 24 hours, F_{mis} is the fraction of DSBs that are misrepaired at 24 hours, N_{DSB} is the initial number of DSBs induced by the exposure, Cluster Density is the average number of neighbouring DSBs within a 70 nm radius, L is the track averaged linear energy transfer in units of keV/μm, D is the dose in units of Gy, and a-h are parameters determined through best fitting, the values of which are reported in Table 1 for protons and alphas.

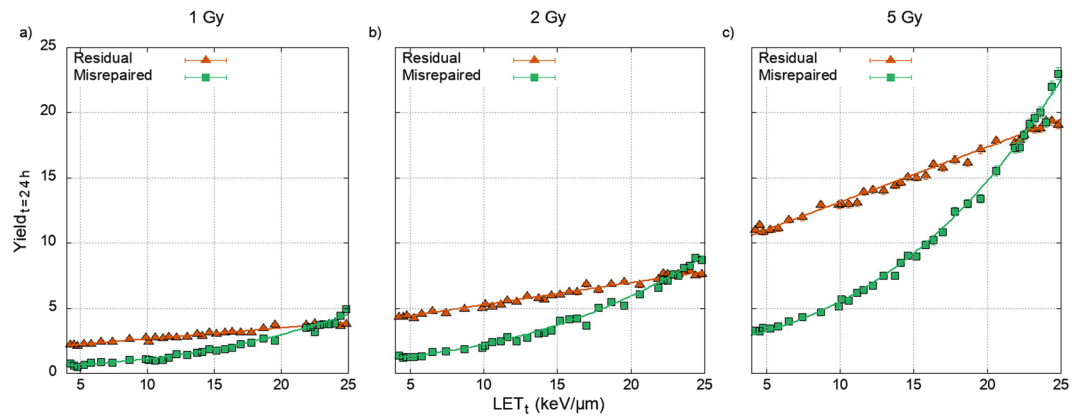


Figure 3. The correlation predicted yield of misrepaired, equation (5), and residual, equation (7), DSBs across the clinically relevant proton LET_t range, with a constant dose of (a) 1 Gy, (b) 2 Gy, and (c) 5 Gy. Points show values measured within the simulation. Error bars in the yield are the standard error in the mean of 200 repeats. Error bars in the LET_t are the standard error in the mean for 50,000 repeats. Both yield and LET_t error bars are too small to be seen on this scale.

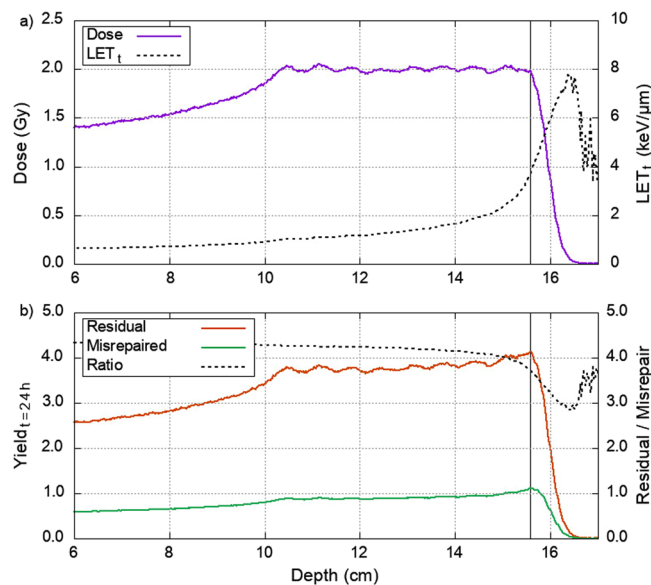


Figure 4. (a) The depth dose (solid) and LET_t (dashed) profile of an example proton SOBP with 2 Gy across the plateau comprised of nine pristine Bragg peaks. (b) The predictions of residual and misrepaired DSB yields for the same proton SOBP, calculated with equations (7) and (5) respectively. The ratio of residual DSBs to misrepaired DSBs is shown as a dashed line. The solid black vertical line on both graphs denotes the depth where the physical dose begins to fall off (15.58 cm).

Figure 3 shows how the yield of residual and misrepaired DSBs predicted by equations (5) and (7) increases across a clinically relevant LET_t range⁶⁵ whilst keeping dose constant. The plotted data points show the results of our model and show good agreement with the correlations derived earlier. There is a difference in behaviour between the predicted yield of residual and misrepaired DSBs, with misrepaired DSBs becoming more important at high LET_t .

The proton SOBP depth dose and LET_t profile is simulated using the Geant4 “QGSP_BIC” physics list (Fig. 4a). Here, the SOBP is simulated from the combination of nine pristine Bragg peaks, where the maximum proton energy is 150 MeV. Using equations (5) and (7), the expected yields of misrepaired and residual DSBs per nucleus is calculated with proton depth (Fig. 4b). Due to low proton fluence there is increasing noise in LET_t at the distal end of Fig. 4a. However, since the dose in this region is negligible the noise in LET_t doesn't impact our predictions. The undulation of the plateau is due to combination of the 9 pristine Bragg peaks. Figure 4b shows a gradual increase in residual and misrepaired DSBs across the SOBP, with a more pronounced peak at the distal end. After this, the yields of residual and misrepaired DSBs fall off at a slower rate than physical dose, resulting in biological effect at low dose regions past the SOBP. Interestingly, the peak of residual DSBs coincides with

the distal edge of the dose plateau whilst the peak of misrepaired DSBs is situated at a slightly increased depth. Combined with Fig. 3 this demonstrates the higher sensitivity that misrepair events have to LET_t . For this 2 Gy SOBP misrepaired DSBs are always predicted at a lower yield than residual DSBs; however, the plotted ratio shows that the relative importance of misrepair is highest at the distal edge.

Discussion

In this work, we demonstrate plausible mechanisms which, combined, lead to the conclusion that the early biological effect of radiation depends on the extent to which DSBs are clustered, shown with the cluster density.

The fraction of misrepaired DSBs can be predicted by cluster density; independent of ion type, dose, LET_t , and break complexity. The cluster density produced by different ion species is not the same for the same LET_t values (Fig. 1b), and therefore produce different early biological effect (Fig. 1a). We suggest that this can explain the experimentally observed differences in cell kill and mutation caused by different ion species at iso- LET^{35} . This cluster density dependent response is the result of three factors: the time required for an exposed DNA end to be prepared for re-joining, the motion of these broken ends during this time, and the fact that formation of synaptic complexes depends on co-localisation of two DSB ends. The diffusion of breaks will likely result in separation of initially co-localised correct partners and, in regions of high DSB density, increases the possibility of co-localising with an incorrect partner. We quantify this in terms of cluster density. The cluster density can be estimated from LET_t for a given ion type (Supplement 6) and is dose independent. Only intra-track DSBs are proximal enough to contribute to the cluster density. We do not see any considerable numbers of DSBs formed as a result of inter-track energy depositions, which we have investigated up to 5 Gy (data not shown).

Our model predicts that the fraction of DSBs left unrepaired is constant; independent of ion type, dose, and LET_t (Fig. 2b). This arises due to the tight confinement of DSB ends around their initial break site, Fig. 1c. Only a small fraction of DSB ends are able to escape this confinement, and are then unlikely to meet another end. The escaping fraction is solely a product of the sub-diffusive motion and therefore not dependent on the beam parameters. The number of initial breaks has a linear relationship with dose and with LET_t , we therefore predict that the absolute yield of residual DSBs scales linearly with these factors. This is in agreement with current clinical practice, where treatment plans use dose as a surrogate for cell kill. Furthermore, it lends weight to the concept of LET optimisation.

Individual NHEJ pathways are thought to respond to different break complexities⁶⁶. Our repair model only contains the DNA-PKcs dependent c-NHEJ pathway, capable of resolving simple and complex breaks. For this reason the kinetics of our model were fitted to experimental data from laser generated complex DSBs⁵⁸. The model omits DNA-PKcs independent pathways, which resolve simple breaks and have been suggested to complete faster⁶⁶. However, the above mechanism of misrepair is still applicable for these pathways, since they also require time dependent recruitment of proteins to mobile ends before synapsis between co-localised partners. Differences would arise in the time required to prepare ends for re-joining, leading to a difference in the radius at which the cluster density should be determined. This means that our cluster density radius is likely an misestimation for fully NHEJ capable cells. The effect of this is somewhat dampened by the fact that some proportion of these simple DSBs can be resolved by c-NHEJ. We therefore do not expect the addition of DNA-PKcs independent repair pathways to substantially impact the results presented.

The recruitment kinetics of Ku70/80 and DNA-PKcs are fitted to laser irradiated *in vitro* Xrs6 and V3 cells with c-NHEJ pathways⁵⁸. Repair kinetics are fitted to proton irradiated *in vitro* AG01522 cells⁶⁵. As such the response of our model is not representative of a specific cell type. However, it can be shown that the mechanism of cluster density dependent misrepair holds for any NHEJ capable cell type, following the same justification as above. The induction of DNA damage in this model is fitted to the DSB yields of other *in silico* models for a range of cell types^{40,42,46–48}, with a constant nucleus size of 5 μ m diameter. Therefore, any deviation from this size will change the DSB density, as the same number of DSBs will be distributed across a different volume. The initial damage is therefore also not representative of any specific cell type. A misestimation in the initial DSB yield would lead to misestimation in the yields of residual and misrepaired DSBs. A misestimation in the DSB density would lead to changes in the cluster density, and consequently in the fraction of misrepaired DSBs. However, potential misestimation only changes the initial yield, not the mechanism of cluster density dependent misrepair. Given all of the above, we would expect the correlations detailed in this work to hold for both end points in any NHEJ capable cell, but with cell specific changes in the predicted yields. This change in yield, but not in trends, could explain the systematically lower yield of residual DSBs compared to the experimental results of Chaudhary *et al.*⁶⁵.

The model does not include the Homologous Recombination (HR) pathway which is an “error free” pathway available to cells occupying the relevant cell cycle stages. It has been shown that NHEJ is a faster repair pathway than HR⁶⁷, and is dominant throughout the cell cycle⁶⁷. The core c-NHEJ proteins, DNA-PKcs and Ku70/80, have similar, rapid, recruitment kinetics throughout the cell cycle³⁸. Phosphorylation of these recruited proteins causes dissociation from DSB ends, in our model this has a time constant of 95 seconds, providing an opportunity for processing by HR⁵⁸. However, our model shows that the timescale of interest for predicting misrepair is much less than this (Supplement 9) and therefore, we do not expect addition of HR to cause a notable deviation in cellular behaviour from current predictions.

We have implemented sub-diffusion in our model by a continuous time random walk method with no form of tethering. This results in limited motion of DSB ends on the order of 100 nm at 24 hours, similar to that which has been reported for live cell experiments²⁸. This limited motion means that the size of the nucleus, on the order of microns, does not affect the availability of potential partners. In V(D)J recombination the sub-diffusive motion is better described by fractional Langevin motion (fLm)⁶⁸. If fLm were to apply to DSB repair in general, broken ends would be more strictly confined around their formation site. However, the confinement scale proposed would allow for an exploration volume that is larger than the volume of interest for misrepair, which we show to have a radius of around 70 nm (Fig. 2a). Therefore, we do not expect it to influence misrepair substantially.

The scale of motion associated with DSB ends *in vivo* is a debated topic within the literature. The only assumption made regarding this in our model is that the motion is sub-diffusive. The magnitude of the mobility of these ends was then scaled in order to reproduce the endpoint of residual breaks reported in literature⁶⁵. The final scale of motion, and the fact that it agrees with the more conservative experimental results, is therefore an emergent property. To achieve greater mobility whilst retaining agreement to the literature reported residual yields would require a different mechanism. One possible mechanism that has previously been proposed, and observed experimentally by some³³, is the formation of repair centres. This directed motion could result in large displacements of DSB ends whilst still maintaining their proximity, a necessary component for resolution of breaks.

This work uses a detailed short segment of the chromatin fibre to determine break complexity. The model includes damage from both direct and indirect effects in order to predict the DSB complexity. However, we do not consider the impact of some chemical effects such as the oxygen fixation hypothesis^{69,70}. Inclusion of these processes would likely yield higher residuals at 24 hours due to a decrease in repair efficacy. Furthermore, our process of converting damage sites into DSBs is a simplification that does not account for deletions of very short DNA segments from complex breaks. If these segments are of insufficient length to recruit repair proteins⁷¹ they are both irreparable and experimentally unmeasurable in fluorescent repair kinetics experiments⁷². We therefore would expect to underestimate the true yield of residuals.

We have discussed reasons that lead us to believe the mechanisms behind residual and misrepaired DSBs are applicable to all NHEJ capable cells. This would result in conserved behaviour between cell types, but with cell specific yields of residuals and misrepaired DSBs. We therefore propose that the general behaviour predicted by our model from the interaction of these mechanisms is representative of reality.

The components of the biological response, residual and misrepaired DSBs, are correlated to our results and combined to give equations (4) and (6). These correlations show how the physical descriptors of the incident beam that are relevant to biological response can be summed up using cluster density. The mechanisms we have proposed for misrepair and residuals respond differently to LET_T, Fig. 3, however both scale equally with dose. This means that our model predicts no change in the ratio of cells undergoing misrepair for either hypo or hyper fractionation. However, it also means that regardless of dose, misrepair will be dominant at LET_T corresponding to the dose fall-off region of the proton SOBP, >22 keV/μm, Fig. 3b. Assuming that misrepair is a marker for chromosome aberration this potentially means areas of higher genomic instability are created at the distal edge of the treatment volume. A second clinically relevant trend is predicted; using the 2 Gy case as an example, Fig. 3b, a change in LET_T from 20 keV/μm to 10 keV/μm reduces the average yield of misrepaired DSBs from 6.1 to 2.1, whilst only decreasing the average yield of residual DSBs from 6.8 to 5.1. Treatment plans extend the prescribed dose region beyond the gross tumour volume (GTV) partly in order to treat microscopic spread. In these regions, and bordering areas, it would seem that induction of genomic instability could be limited by moving high LET components into the tumour, whilst a similar cell kill potential could be achieved by maintaining dose.

Predictions of our model can be derived from more conventionally scored parameters in proton therapy, equations (5) and (7). The predicted biological response, shown in Fig. 4, follows similar trends to the phenomenological predictions of RBE across a proton SOBP^{9,73}. In both cases there is a gradual increase in response across the dose plateau with a pronounced peak at the distal edge. As such it lends weight to the concept that RBE is driven by a combination of residual and misrepaired DSBs. Due to both residual and misrepaired DSB yields falling off at a slower rate than physical dose, the model predicts considerable biological effect past the distal edge of the SOBP. Misrepair falls off at a slower rate than residual DSBs, however the latter still dominates the yield. Therefore, we suggest that the biologically extended range reported by Marshall *et al.*⁹, and others, is most likely due to the presence of residual DSBs. For the proton SOBP presented here, misrepaired DSBs are predicted to peak at a slightly increased depth compared to residual DSBs, although the yields always remain lower. During dose escalation, this peak will be the first region where misrepair starts to dominate the biological response. Since the misrepair peak is deeper than the physical dose peak, this could have implications clinically.

Methods of LET optimisation have been suggested for clinical use, with the justification that RBE is dependent on LET^{73–75}. The difference in response of residual and misrepaired DSBs to LET_T further emphasises the benefits of this type of planning, but suggests that prediction of biological outcome is not a straightforward combination of dose and LET. Therefore, if the benefits of proton therapy are to be fully exploited, the mechanisms we have proposed should be considered at the treatment planning stage. This could allow the manipulation of not only RBE in general but also the dominant pathway taken to cell death; a difference which could possibly be exploited clinically.

References

1. Curtis, S. B. Lethal and Potentially Lethal Lesions Induced by Radiation—A Unified Repair Model. *Radiat. Res.* **106**, 252–270 (1986).
2. Cornforth, M. N. & Bedford, J. S. In *Advances in radiation biology* (eds Lett, J. T. & Sinclair, W. K.) **17**, 423–496 (San Diego: Academic, 1993).
3. Kunderát, P. & Stewart, R. D. On the biophysical interpretation of lethal DNA lesions induced by ionising radiation. *Radiat. Prot. Dosimetry* **122**, 169–172 (2006).
4. Joiner, M. *Basic Clinical Radiobiology* Edited by. 375, <https://doi.org/10.1201/b13224> (2009).
5. Hall, E. J. & Giaccia, A. J. *Radiobiology for the Radiologist*. (Lippincott Williams & Wilkins, 2006).
6. Paganetti, H. *Proton therapy physics*. (CRC Press, 2011).
7. ICRU. ICRU report 78 - prescribing, recording, and reporting proton-beam therapy. *J. ICRU* **7** (2007).
8. Paganetti, H. Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer. *Phys. Med. Biol.* **59**, R419–72 (2014).
9. Marshall, T. I. *et al.* Investigating the implications of a variable RBE on proton dose fractionation across a clinical pencil beam scanned spread-out Bragg peak. *Int. J. Radiat. Oncol.* **95**, 70–77 (2016).
10. Baiocco, G. *et al.* The origin of neutron biological effectiveness as a function of energy. *Sci. Rep.* **6**, 34033 (2016).
11. Watanabe, R., Rahmanian, S. & Nikjoo, H. Spectrum of Radiation-Induced Clustered Non-DSB Damage - A Monte Carlo Track Structure Modeling and Calculations. *Radiat. Res.* **183**, 525–40 (2015).

12. Rydberg, B., Cooper, B., Cooper, P. K., Holley, W. R. & Chatterjee, A. Dose-dependent misrejoining of radiation-induced DNA double-strand breaks in human fibroblasts: experimental and theoretical study for high- and low-LET radiation. *Radiat. Res.* **163**, 526–534 (2005).
13. Elsässer, T. *et al.* Quantification of the relative biological effectiveness for ion beam radiotherapy: Direct experimental comparison of proton and carbon ion beams and a novel approach for treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.* **78**, 1177–1183 (2010).
14. Friedrich, T., Scholz, U., Elsässer, T., Durante, M. & Scholz, M. Calculation of the biological effects of ion beams based on the microscopic spatial damage distribution pattern. *Int. J. Radiat. Biol.* **88**, 103–107 (2012).
15. Ballarini, F. & Carante, M. P. Chromosome aberrations and cell death by ionizing radiation: Evolution of a biophysical model. *Radiat. Phys. Chem.* **128**, 18–25 (2016).
16. Sachs, R. K. & Brenner, D. J. Effect of LET on Chromosomal Aberration Yields. I. Do Long-lived, Exchange-prone Double Strand Breaks Play a Role? *Int. J. Radiat. Biol.* **64**, 677–688 (1993).
17. Schipler, A. & Iliakis, G. DNA double-strand-break complexity levels and their possible contributions to the probability for error-prone processing and repair pathway choice. *Nucleic Acids Res.* **41**, 7589–7605 (2013).
18. Cornforth, M. N. Testing the Notion of the One-Hit Exchange. *Radiat. Res.* **121**, 21–27 (1990).
19. Sachs, R. K., Chen, a. M. & Brenner, D. J. Review: proximity effects in the production of chromosome aberrations by ionizing radiation. *Int. J. Radiat. Biol.* **71**, 1–19 (1997).
20. Roukos, V. *et al.* Spatial dynamics of chromosome translocations in living cells. *Science (80-.)*. **341**, 660–664 (2013).
21. McMahon, S. J., Schuemann, J., Paganetti, H. & Prise, K. M. Mechanistic Modelling of DNA Repair and Cellular Survival Following Radiation-Induced DNA Damage. *Sci. Rep.* **6**, 33290 (2016).
22. McMahon, S. J., McNamara, A. L., Schuemann, J., Paganetti, H. & Prise, K. M. A general mechanistic model enables predictions of the biological effectiveness of different qualities of radiation. *Sci. Rep.* **7**, 10790 (2017).
23. Holley, W. R., Mian, I. S., Park, S. J., Rydberg, B. & Chatterjee, A. A Model for Interphase Chromosomes and Evaluation of Radiation-Induced Aberrations. *Radiat. Res.* **158**, 568–580 (2002).
24. Edwards, A. A., Moiseenko, V. V. & Nikjoo, H. On the mechanism of the formation of chromosomal aberrations by ionising radiation. *Radiat. Environ. Biophys.* **35**, 25–30 (1996).
25. Chen, A. M., Lucas, J. N., Hill, F. S., Brenner, D. J. & Sachs, R. K. Proximity effects for chromosome aberrations measured by FISH. *Int. J. Radiat. Biol.* **69**, 411–420 (1996).
26. Friedland, W., Jacob, P. & Kundrát, P. Stochastic simulation of DNA double-strand break repair by non-homologous end joining based on track structure calculations. *Radiat. Res.* **173**, 677–688 (2010).
27. Friedland, W. & Kundrát, P. Track structure based modelling of chromosome aberrations after photon and alpha-particle irradiation. *Mutat. Res. - Genet. Toxicol. Environ. Mutagen.* **756**, 213–223 (2013).
28. Soutoglou, E. *et al.* Positional stability of single double-strand breaks in mammalian cells. *Nat. Cell Biol.* **9**, 675–82 (2007).
29. Girst, S. *et al.* Subdiffusion Supports Joining Of Correct Ends During Repair Of DNA Double-Strand Breaks. *Sci. Rep.* **3**, 1–6 (2013).
30. Iarovaia, O. V. *et al.* Dynamics of double strand breaks and chromosomal translocations. *Mol. Cancer* **13**, 249 (2014).
31. Aten, J. A. *et al.* Dynamics of DNA Double-Strand Breaks Revealed by Clustering of Damaged Chromosome Domains. *Science (80-.)*. **303**, 92 LP–95 (2004).
32. Jakob, B., Splinter, J., Durante, M. & Taucher-Scholz, G. Live cell microscopy analysis of radiation-induced DNA double-strand break motion. *Proc. Natl. Acad. Sci.* **106**, 3172–3177 (2009).
33. Neumaier, T. *et al.* Evidence for formation of DNA repair centers and dose-response nonlinearity in human cells. *Proc. Natl. Acad. Sci.* **109**, 443–448 (2012).
34. Averbek, N. B. *et al.* Efficient Rejoining of DNA Double-Strand Breaks despite Increased Cell-Killing Effectiveness following Spread-Out Bragg Peak Carbon-Ion Irradiation. *Front. Oncol.* **6**, 1–8 (2016).
35. Belli, M. *et al.* Inactivation and mutation induction in V79 cells by low energy protons: re-evaluation of the results at the LNL facility. *Int. J. Radiat. Biol.* **63**, 331–337 (1993).
36. Hill, M. A. Fishing for radiation quality: Chromosome aberrations and the role of radiation track structure. *Radiat. Prot. Dosimetry* **166**, 295–301 (2015).
37. Rabus, H. & Nettelbeck, H. Nanodosimetry: Bridging the gap to radiation biophysics. *Radiat. Meas.* **46**, 1522–1528 (2011).
38. Agostinelli, S. *et al.* GEANT4 - A simulation toolkit. *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* **506**, 250–303 (2003).
39. Incerti, S. *et al.* Comparison of GEANT4 very low energy cross section models with experimental data in water. *Med. Phys.* **37**, 4692–4708 (2010).
40. Meylan, S. *et al.* Simulation of early DNA damage after the irradiation of a fibroblast cell nucleus using Geant4-DNA. *Sci. Rep.* **7**, 11923 (2017).
41. Bueno, M., Schulte, R., Meylan, S. & Villagrasa, C. Influence of the geometrical detail in the description of DNA and the scoring method of ionization clustering on nanodosimetric parameters of track structure: a Monte Carlo study using Geant4-DNA. *Phys. Med. Biol.* **60**, 8583–8599 (2015).
42. Pater, P. *et al.* Proton and light ion RBE for the induction of direct DNA double strand breaks. *Med. Phys.* **43**, 2131–2140 (2016).
43. Friedland, W., Dingfelder, M., Kundrát, P. & Jacob, P. Track structures, DNA targets and radiation effects in the biophysical Monte Carlo simulation code PARTRAC. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* **711**, 28–40 (2011).
44. Rosenbluth, M. J., Lam, W. A. & Fletcher, D. A. Force microscopy of nonadherent cells: a comparison of leukemia cell deformability. *Biophys. J.* **90**, 2994–3003 (2006).
45. Francis, Z., Villagrasa, C. & Clairand, I. Simulation of DNA damage clustering after proton irradiation using an adapted DBSCAN algorithm. *Comput. Methods Programs Biomed.* **101**, 265–270 (2011).
46. Friedland, W., Jacob, P., Bernhardt, P., Paretzke, H. G. & Dingfelder, M. Simulation of DNA Damage after Proton Irradiation. *Radiat. Res.* **159**, 401–410 (2003).
47. Nikjoo, H., O'Neill, P., Terrissol, M. & Goodhead, D. T. Quantitative modelling of DNA damage using Monte Carlo track structure method. *Radiat. Environ. Biophys.* **38**, 31–38 (1999).
48. Nikjoo, H., Neill, P. O., Wilson, W. E. & Goodhead, D. T. *Computational Approach for Determining the Spectrum of DNA Damage Induced by Ionizing Radiation.* **583**, 577–583 (2001).
49. Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. A Density-based Algorithm for Discovering Clusters a Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining* 226–231 (AAAI Press, 1996).
50. Drew, H. R. *et al.* Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. USA* **78**, 2179–83 (1981).
51. Henthorn, N. T. *et al.* Nanodosimetric Simulation of Direct Ion-Induced DNA Damage Using Different Chromatin Geometry Models. *Radiat. Res.* **0**, RR14755.1 (2017).
52. Bernal, M. a. & Liendo, J. a. An investigation on the capabilities of the PENELOPE MC code in nanodosimetry. *Med. Phys.* **36**, 620–625 (2009).
53. Finch, J. T. & Klug, a. Solenoidal model for superstructure in chromatin. *Proc. Natl. Acad. Sci. USA* **73**, 1897–1901 (1976).
54. Karamitros, M. *et al.* Diffusion-controlled reactions modeling in Geant4-DNA. *J. Comput. Phys.* **274**, 841–882 (2014).
55. Metzler, R. & Klafter, J. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.* **339**, 1–77 (2000).

56. Yang, K., Guo, R. & Xu, D. Non-homologous end joining: advances and frontiers. *Acta Biochim. Biophys. Sin. (Shanghai)*. **48**, 632–640 (2016).
57. Lee, K.-J. *et al.* Phosphorylation of Ku dictates DNA double-strand break (DSB) repair pathway choice in S phase. *Nucleic Acids Res.* **44**, 1732–1745 (2016).
58. Uematsu, N. *et al.* Autophosphorylation of DNA-PKCS regulates its dynamics at DNA double-strand breaks. *J. Cell Biol.* **177**, 219–229 (2007).
59. Graham, T. G. W., Walter, J. C. & Loparo, J. J. Two-Stage Synapsis of DNA Ends during Non-homologous End Joining. *Mol. Cell Biol.* **61**, 850–858 (2017).
60. Sun, J., Lee, K. J., Davis, A. J. & Chen, D. J. Human Ku70/80 protein blocks exonuclease 1-mediated DNA resection in the presence of human Mre11 or Mre11/Rad50 protein complex. *J. Biol. Chem.* **287**, 4936–4945 (2012).
61. Walker, J. R., Corpina, R. A. & Goldberg, J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* **412**, 607–14 (2001).
62. Dynan, W. S. & Yoo, S. Interaction of Ku protein and DNA-dependent protein kinase catalytic subunit with nucleic acids. *Nucleic Acids Res.* **26**, 1551–1559 (1998).
63. Mimori, T., Hardin, J. A. & Steitz, J. A. Characterization of the DNA-binding protein antigen Ku recognized by autoantibodies from patients with rheumatic disorders. *J. Biol. Chem.* **261**, 2274–2278 (1986).
64. Bernal, M. a. *et al.* The invariance of the total direct DNA strand break yield. *Med. Phys.* **38**, 4147–4153 (2011).
65. Chaudhary, P. *et al.* Variations in the Processing of DNA Double-Strand Breaks Along 60-MeV Therapeutic Proton Beams. *Int. J. Radiat. Oncol. Biol. Phys.* **95**, 86–94 (2016).
66. Li, Y., Reynolds, P., O'Neill, P. & Cucinotta, F. A. Modeling damage complexity-dependent non-homologous end-joining repair pathway. *PLoS One* **9** (2014).
67. Mao, Z., Bozzella, M., Seluanov, A. & Gorbunova, V. DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle* **7**, 2902–2906 (2008).
68. Lucas, J. S., Zhang, Y., Dudko, O. K. & Murre, C. 3D trajectories adopted by coding and regulatory DNA elements: First-passage times for genomic interactions. *Cell* **158**, 339–352 (2014).
69. Grimes, D. R. & Partridge, M. A mechanistic investigation of the oxygen fixation hypothesis and oxygen enhancement ratio. *Biomed. Phys. Eng. Express* **1**, 1–22 (2016).
70. Ewing, D. The Oxygen Fixation Hypothesis: A Reevaluation. *Am. J. Clin. Oncol.* **21** (1998).
71. Lieber, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79**, 181–211 (2010).
72. Friedland, W., Jacob, P. & Kundrat, P. Mechanistic Simulation of Radiation Damage To Dna and Its Repair: on the Track Tow Ards Systems. *Radiat. Prot. Dosimetry* **143**, 542–548 (2011).
73. McNamara, A. L., Schuermann, J. & Paganetti, H. A phenomenological relative biological effectiveness (RBE) model for proton therapy based on all published *in vitro* cell survival data. *Phys. Med. Biol.* **60**, 8399–8416 (2015).
74. Underwood, T. & Paganetti, H. Variable Proton Relative Biological Effectiveness: How Do We Move Forward? *Int. J. Radiat. Oncol. Biol. Phys.* **95**, 56–58 (2016).
75. Grassberger, C., Trofimov, A., Lomax, A. & Paganetti, H. Variations in linear energy transfer within clinical proton therapy fields and the potential for biological treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.* **80**, 1559–1566 (2011).

Acknowledgements

N.T.H. and J.W.W. would like to acknowledge financial support from EPSRC (grant No.: EP/J500094/1). The computational element of this research was achieved using the Condor High Throughput Computing facility at the University of Manchester.

Author Contributions

N.T.H. developed the DNA damage model. J.W.W. developed the DNA repair model. N.T.H. and J.W.W. gathered and analysed the data. M.S. developed the dose and LET depth code used in Fig. 4. N.T.H. and J.W.W. drafted the manuscript based on discussions with M.S., R.I.M., K.J.K., N.F.K., and M.J.M. All authors reviewed and approved the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-21111-8>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018