

平成31年度放射線対策委託費
(放射線安全規制研究戦略的推進事業費)
放射線安全規制研究推進事業

染色体線量評価手法の標準化に向けた 画像解析技術に関する調査研究

国立研究開発法人量子科学技術研究開発機構
数藤由美子 (主任研究者)

研究概要(1)

課題名 「染色体線量評価手法の標準化に向けた画像解析技術に関する調査研究」

研究期間:平成30年度～平成31年度(2年間)

背景・目的

- 染色体解析、特に末梢血リンパ球培養による**二動原体分析法**は、鋭敏で最も信頼性の高い生物学的線量評価法として国際的に標準化されている。細胞培養から染色体画像撮影までは自動化が成し遂げられた。しかしながら、画像判定は依然として人間の目視によるため、**染色体判定基準の機関間および判定者間の変動が顕著**であること、また、**結果を得るまでの時間と労力が大きい**ことなどから、大規模放射線事故時に際し複数の機関が分析を行う場合に問題を生じることが予測される。
- 本研究(フィージビリティ・スタディ)では、染色体判定者の主観や習熟度の差に影響されない効率の良い自動画像分析技術として、**人工知能(AI)技術導入**の有用性や実現可能性を試験・評価し、**大規模放射線事故(特に初期のトリアージ段階)の被ばく医療診断支援**への適用を検討する。

期待される成果

- 染色体のAI自動解析技術の導入により、**生物学的線量評価の標準化・迅速化**(1検体につき実質30時間以上要していた目視検査が僅か10分程度で可能)が達成される。これにより**大規模放射線事故時における被ばく医療診断支援に大きく貢献**できる。

研究概要(2)

実施状況(当初の計画以上に進捗)

今年度(最終年度)は以下を実施

① 関連情報の収集・更新(随時)

- ✓ 国際会議3件、国内学会3件、国際学会1件、英国PHE研究会合(統計解析手法の情報収集を含む)
- ✓ 企業展示会(AI・人工知能EXPO)、AI関連企業ヒアリング7社、染色体検査会社ヒアリング1社→ AI技術導入試験の外部委託先候補を選定(入札により5月に1社決定)

② テスト用画像データ(0~5.0 Gy照射血由来)準備(前年度分を含め計14,688枚)

③ AI技術導入試験Phase 1の実施 (プレ試験 5~6月、本試験 9~12月)


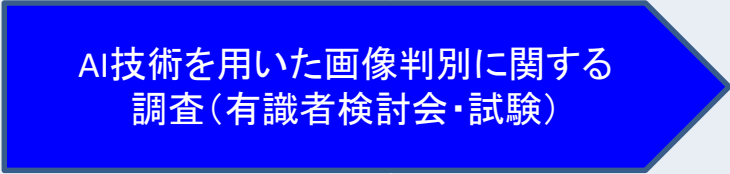




④ 検討委員会の開催(画像検討会 8回、各分野専門の検討委員による研究会合 3回)

⑤ 今後の研究課題の明確化(12月)

- ✓ **基盤技術の構築**: 深層学習手法を導入した**学習モデルの試作**、**AIに適した標本・画像作成法の確立**[実際の運用で必要となる多検体対応の細胞培養法の確立(論文受理)を含む]、線量推定法の検討(最尤法・ベイズ法)
- ✓ **AI導入の評価**: **試作モデル**で約50細胞の画像データ(**トリアージ・レベル**の線量評価対応の標準観察数)から、2 Gy未満と2 Gy以上の被ばくを判別できた(95%信頼区間の上限値をもとに線量推定)。
- ✓ **線量推定作業の高度の効率化**: 1検体当たりの画像判定(1000細胞として)が、**実質30時間以上(目視)から10分(AI)へ**。

⑥ 本格開発研究準備開始、線量推定試験追加試験実施(1月~3月)

【ロードマップ】

実施項目	平成30年度	平成31年度
調査研究1: 染色体線量評価における画像解析技術開発の情報収集		
調査研究2: 人工知能学習技術をはじめとする画像判別手法に関する調査 (サブテーマ: 数理統計処理に関する調査)	  情報収集開始  試験の準備・実施	 (有識者検討会)
総括		 (課題抽出・提言)

【研究体制】

全体総括
数藤由美子
(量研)

調査研究1 画像解析技術開発の情報収集

全般/数藤由美子・高島良生・
穂山美穂・栗原治(量研)
AI・IT関連/長田直樹・弓至(北海道大)

【補足】

主任研究者	数藤
研究分担者	無し
研究協力者	検討委員
研究参加者	量研スタッフ
外部委託	(株)エルピクセル

調査研究2 人工知能(深層学習)導入試験

標本開発、画像・教師データの検討・作成:
数藤由美子(量研)、平井百樹(東京大)
AI導入試験:
長田直樹・弓至(北海道大)
外部委託(エルピクセル株式会社)
線量推定試験:
数藤由美子(量研)、長田直樹・弓至(北海道大)
検討委員:
AI学術専門家/長田直樹・弓至(北海道大)、
瀧川一学(理研/北海道大)、野里博和(産総研)
統計学/江藤亜紀子(保健医療科学院)
災害シミュレーション専門家/市川学(芝浦工大)
染色体技術専門家/平井百樹(東京大)

【今年度実施概要1 背景と目的】

■ 染色体解析、特に末梢血リンパ球培養による**二動原体分析法**は、鋭敏で最も信頼性の高い生物線量評価法として国際的に標準化されている。細胞培養から染色体画像撮影までは自動化が成し遂げられた。ポアソン分布する染色体異常の画像判定は観察細胞数が大きい、分析の自動化は未だ完成されず、人力を頼っている。**染色体判定基準のラボ間・判定者間の変動は大きく**(図1)、結果を得るまでの**時間と労力が大きい**(1検体当たり1000細胞観察で、高度熟練検査者でも総目視時間30時間以上、労働期間にして3日以上かかる)ことから、**実際の事故対応に際しては大きな問題をはらむことが予測される**。これは実際に我々が2011年3~6月、福島原発事故緊急作業員12名の検査を実施した経験から得た教訓でもある[Y. Suto et al., Health Physics, 2013]。**画像判定の標準化・迅速化は長年の大きな課題**である。定性的な一般染色体検査(核型分析)と異なり膨大な観察細胞数が必要な二動原体分析でのAI技術導入の効果は絶大である。

■ 本フィージビリティ・スタディでは、染色体検査者の主観や習熟度の差に影響されない効率の良い自動画像分析技術として、**AI技術の導入の有用性・実現可能性を試験・評価する**。大規模放射線事故、特に初期の**トリアージ段階での、被ばく医療診断支援への適用が期待される**。

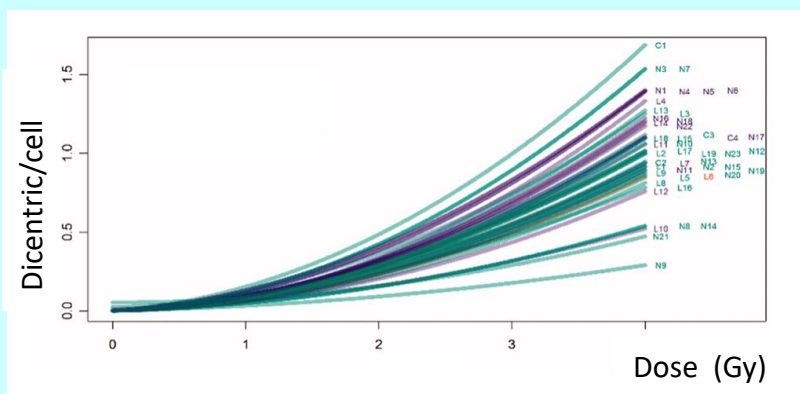
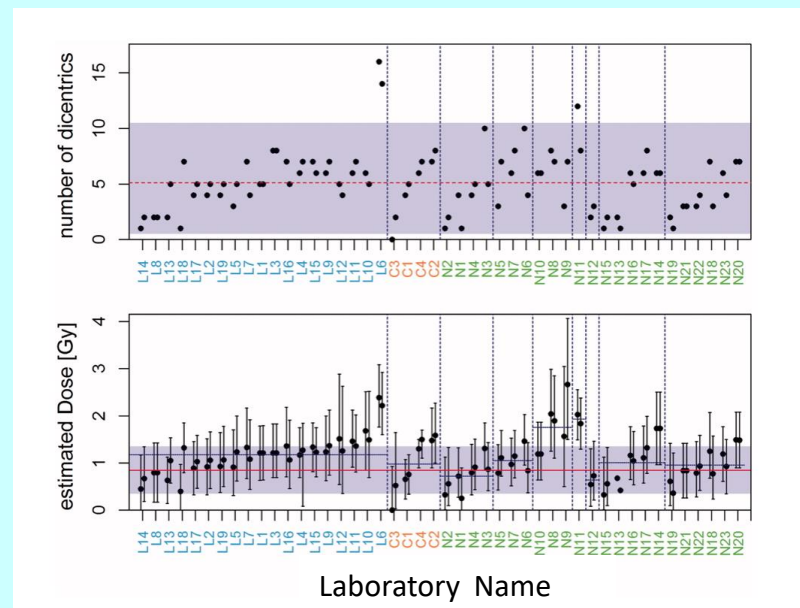


図1. 二動原体分析のラボ間比較 [U. Oestreicher ら(2017)より]。長年の訓練でもギムザ染色法による目視判定分析では、判定基準の統一化が難しい。ラボや人によって線量効果曲線は異なり(下図)、また同じ標本や画像を用いても、検出する染色体異常数や推定線量に違いが生じる(上図)。

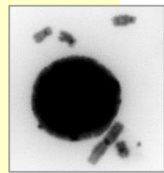
【今年度実施概要2 研究の流れ】

4~6月

AI導入試験
Phase 1 プレ試験

Faster R-CNN (畳み込みニューラルネットワークを利用した深層学習手法)

- ▶ 「COCOデータセット」(20万枚以上の一般的な画像)で形態を学習させる = 転移学習(少数データ対策)
- ▶ 少数のメタフェーズ画像で学習させる (66枚)
 - ・ 熟練者による教師データ作製(目視アノテーション)
 - ・ 教師データで深層学習→ 性能評価試験(染色体の「検出」と「分類」の性能)
- ▶ 教師データを増やして効果を確認 (95枚)

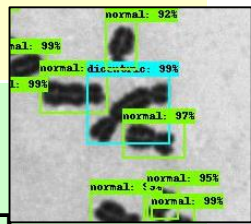


7月~

教師データの最適作成法を確立・データ増加

教師データ増加 (0~5.0 Gy照射標本画像のアノテーション)

- ▶ AIによる自動アノテーション → 熟練専門家による修正
 - ・ 準備画像(12月末まで) 14,688枚

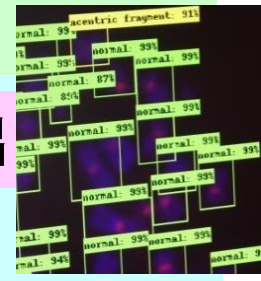


9月~

AI導入試験
Phase 1 本試験

AI画像判定モデルの向上・評価・課題抽出

- ▶ 再学習による染色体の検出および分類の性能向上
- ▶ 線量効果曲線の試作・性能評価



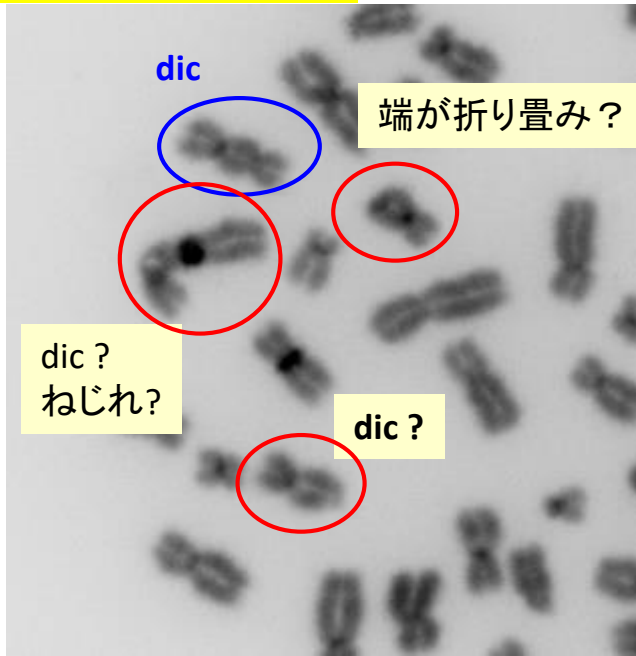
1月~

とりまとめ/2.0 Gy(800枚画像追加)線量推定試験・論文投稿準備/本格研究への準備

(注)一連の図は画像判別の進化をイメージしたもの

【今年度実施概要3 教師データ作成法】

単色染色 画像



PNA-FISH 画像

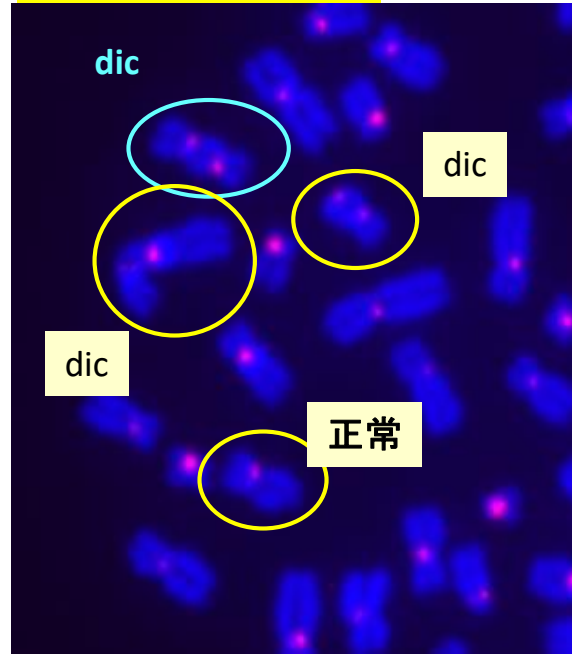


図2. メタフェーズ像(一部)。dic: 二動原体 (dicentric)

教師データ作成方法

■ **ギムザ染色**では、目視画像判定に時間を要し、染色体のねじれなどを正確に判定することはできない(図2)。アジアネットワークの調査経験もあり(右解説参照)、本研究では、PNA-FISHの併用(単色染色による染色体形態検出と蛍光標識による動原体検出)により、**正確な教師データ**を作製することとした。Phase 1 プレ試験において、AI 利用に最適な標本・画像作製条件を決定した。

■ 教師データ作成ではアノテーションソフトウェア Labellmg を用いた(図3)。各画像上の染色体に分類名をラベルすると、位置情報を伴うXML形式で保存される。

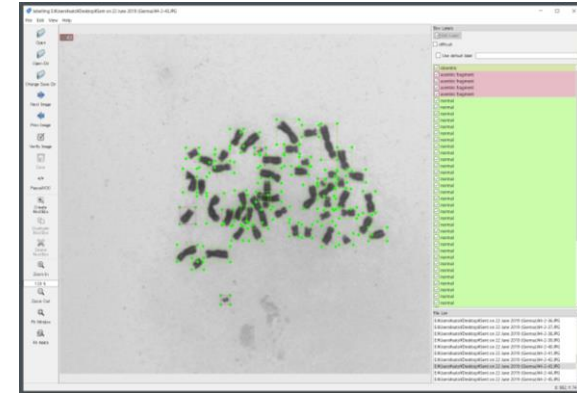


図3. Labellmg を用いたアノテーション実行イメージ。熟練者が染色体1個1個を矩形で囲み、分類名を与えることで教師データができる。

アジアネットワーク 二動原体分析比較調査 2019

* 量研機構が実施 (参加4カ国、11ラボ、32名、経験年数は2年~22年と多様)。本研究で作成した**単色染色画像**(3.0 Gy、200枚)で判定実施後、同一細胞の**PNA-FISH**画像を提供し判定、比較。(投稿論文準備中)

[結果]

PNA-FISH像では全ラボで採択細胞も判定も一致(推定線量3.0 Gy)。
→ **PNA-FISH**解析の精度の高さが証明された。

【今年度実施概要4 Phase 1・本試験の実施】

材料： 0～5.0 Gy照射の末梢血由来染色体画像 約15,000枚(内訳は表参照)

方法： AI導入本試験を実施し、AIによる画像判定支援の有効性を検証する。

1) 機械学習アルゴリズムを学習させるための正確なタグ(異常染色体・正常染色体など)を付したアノテーション済みのデータセットである画像1190枚(モデル学習の訓練用セット832枚、繰り返し検証用セット180枚、テスト用セット177枚)を用い、モデルの学習を行い、染色体の検出と分類の性能を調べる。

2) 1)のテスト用セットにより、線量効果関係を確認する。

3) アノテーションを行っていない残りの全画像のうち、染色体数42～50個の画像を用い、線量推定を試みる(最尤法とベイズ法)。

① 各線量からトリアージ・レベルの線量推定用のサンプルをとり(70枚×3セット)、残りのデータセットで線量効果曲線(式)を試作する。

② ①の線量効果曲線線に対して、2 Gy判別を試験する。(注)

表1. 画像の準備状況(上)および含まれる染色体の分類別個数(下)

線量 (Gy)	全画像枚数	アノテーション済み
0	1868	209
0.50	2260	0
1.0	2603	199
2.0	3630	217
3.0	2783	326
4.0	1432	127
5.0	112	112
合計	14688	1190

クラス分類	Total	⁶⁰ Co-ガンマ線照射における吸収線量					
		0 Gy	1.0 Gy	2.0 Gy	3.0 Gy	4.0 Gy	5.0 Gy
Normal	52717	9460	9080	9717	14337	5341	4782
Dicentric	598	2	15	68	197	145	171
Fragment	998	2	26	103	353	236	278
Others	137	16	5	14	51	17	34

(注)急性放射線症候群の前駆症状に対し医学的な介入が必要となる基準である2 Gy以上の放射線照射血液サンプルについて検証する。

【進捗1 染色体の検出・分類と線量効果関係】

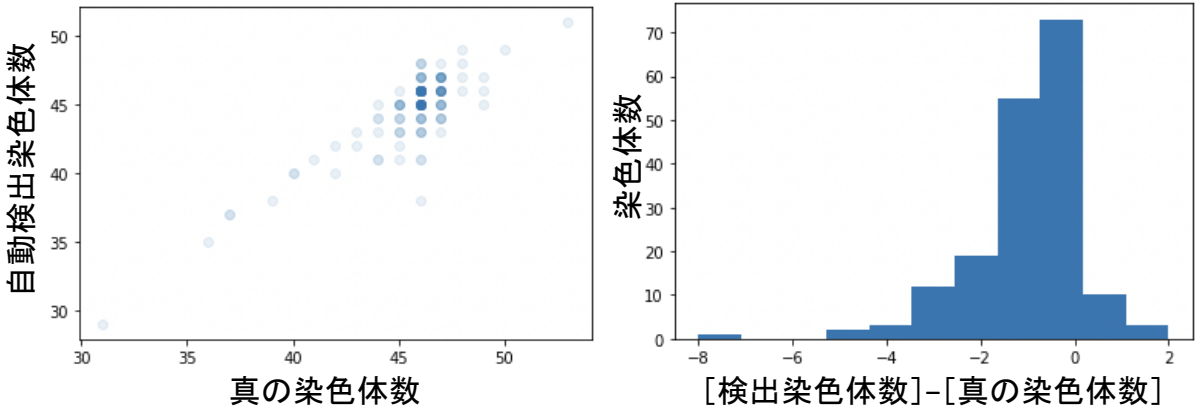


図4. 実際の染色体数とAI自動検出された染色体数(左)および自動検出された染色体数と実際の染色体数の差の分布(右)

→ 染色体検出力98%(人がパラメータ設定をする機械学習法では検出力40~50%)
 → 線量推定の検証では42~50染色体を含む画像を用いることとした。

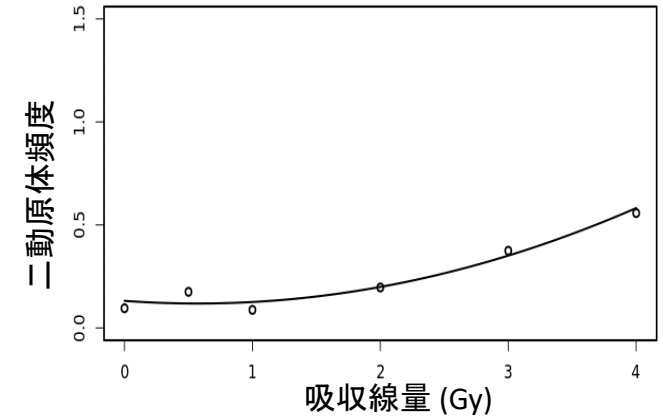


図5. 線量とAI自動検出による二動原体頻度

→ 線量効果関係が認められた。

参考:

感度 $Recall = \frac{True\ positives}{True\ positives + False\ negatives}$

精度 $Precision = \frac{True\ positives}{True\ positives + False\ positives}$

特異度 $Specificity = \frac{True\ negatives}{False\ positives + True\ negatives}$

	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	
① normal	7754	17	5	0	0	0	0	0	0	132	normal
② dicentric	36	36	0	0	0	0	0	0	0	5	Recall: 0.981
③ fragment	6	0	94	0	0	0	0	0	0	37	Precision: 0.989
④ centric ring	4	0	0	1	0	0	0	0	0	0	Specificity: 0.789
⑤ broken chromosome	2	0	1	0	0	0	0	0	0	6	Dicentric chromosome
⑥ tricentric	0	1	0	0	0	0	0	0	0	0	Recall: 0.468
⑦ iso-chromatid break	1	0	1	0	0	0	0	0	0	0	Precision: 0.667
⑧ other abnormalities	0	0	0	0	0	0	0	0	0	1	Specificity: 0.998
⑨ normal (partially broken)	0	0	0	0	0	0	0	0	0	1	Acentric fragment
⑩ not detected	35	0	3	1	0	0	0	0	0	0	Recall: 0.686
											Precision: 0.904
											Specificity: 0.999

Predicted (AIによる推論結果)

図6. テスト用データセットを用いたモデルの精度検証結果

→ 1月以降の開発でいっそうの精度向上を試みる。
 具体策としては、現解析で用いた標本では染色体正常細胞が多く含まれているという偏りがあるので、AI画像分析で使われている諸法などで染色体異常の画像数を増やす。

【進捗2 線量効果曲線試作試験】

表2. 二動原体染色体を指標としたAI画像判定によるテスト用画像のトリアージ線量推定試験結果

二動原体 (標準法) (注1)

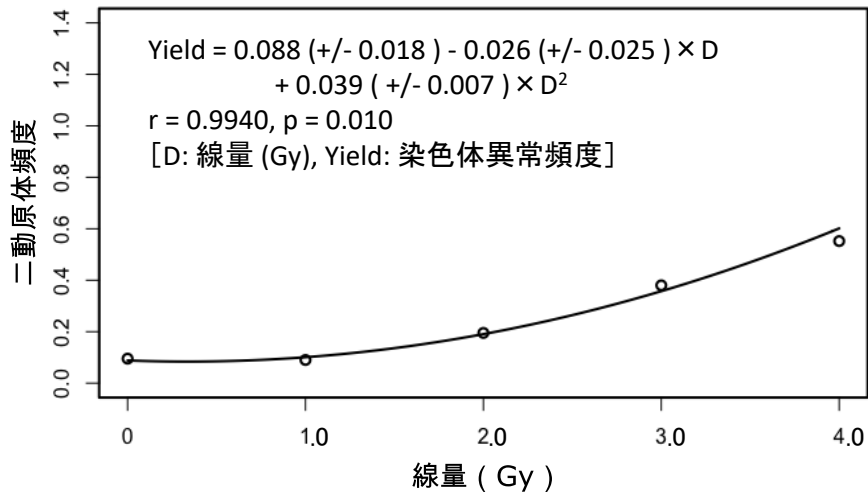


図7. AI自動画像判定により作成された線量効果曲線

→ 臨床判断で用いられる95%信頼限界上限値で、**2 Gyのトリアージ判定** (注2, 3) **が可能であることがわかった。**

* ベイズ法でも同様の結果を得た(データ省略)。

→ **十分な細胞数をとれば精度の高い線量推定も可能であると期待できる。**また、解析速度からいって、トリアージ用でも細胞数は100個以上としても良いといえる。

Dose (Gy)	Set	Dicentrics	Cells	Estimated Dose (Gy)	95% LCL	95% UCL
0	1	6	61	0.95	0.34	2.17
0	2	1	57	0.00	0.34	0.94
0	3	10	57	1.88	0.43	2.83
0	total (1-3)	17	175	0.92	0.34	1.70
0	4	126	1319	0.89	0.34	1.22
1	1	3	57	0.00	0.34	1.68
1	2	3	59	0.00	0.34	1.63
1	3	5	62	0.08	0.34	1.98
1	total (1-3)	11	178	0.00	0.34	1.17
1	4	183	2088	0.65	0.34	1.01
2	1	10	57	1.88	0.43	2.82
2	2	12	60	2.07	1.05	2.96
2	3	15	54	2.58	1.70	3.45
2	total (1-3)	37	171	2.19	1.67	2.69
2	4	553	2815	2.05	1.92	2.17
3	1	43	57	4.51	3.80	5.25
3	2	18	63	2.62	1.83	3.42
3	3	24	59	3.23	2.48	4.01
3	total (1-3)	85	179	3.52	3.10	3.95
3	4	770	2055	3.08	2.96	3.21
4	1	29	54	3.76	3.01	4.56
4	2	34	61	3.84	3.14	4.58
4	3	41	63	4.17	3.49	4.89
4	total (1-3)	104	178	3.94	3.53	4.36
4	4	576	1033	3.84	3.67	4.01

(注1) 国際標準法では染色体異常の出現がポアソン分布をとるものとして、最尤法を用いる。

(注2) 通常の被ばく事故時トリアージ・レベルの線量評価では、ISO推奨プロトコールに基づき、**目視で50メタフェーズの観察を行う。**

(注3) 深層学習を用いず人がパラメータ設定をする機械学習法では3 Gy以上しか分離できなかった。

【進捗2 線量効果曲線試作試験】

表3. 染色体断片を指標としたAI画像判定によるテスト用画像のトリアージ線量推定試験結果

染色体断片 (標準法)

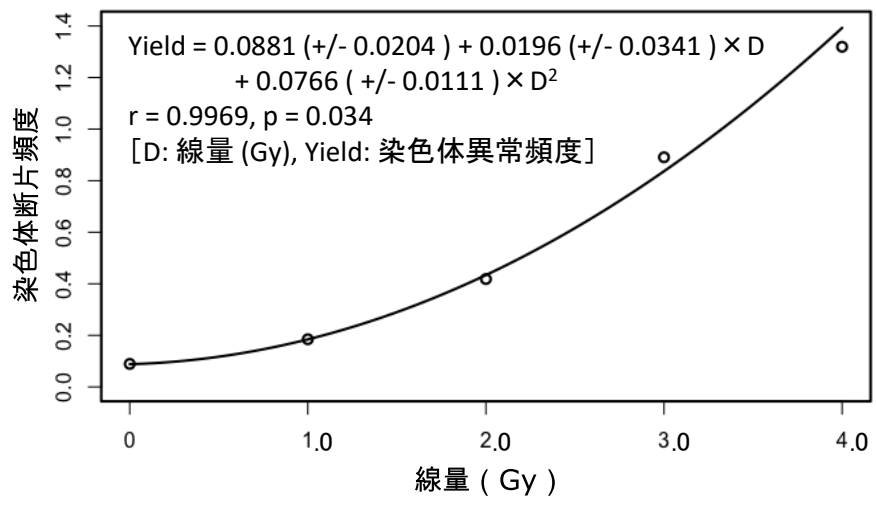


図8. AI自動画像判定により作成された線量効果曲線

Dose (Gy)	Set	Frg	Cells	Estimated Dose (Gy)	95% LCL	95% UCL
0	1	7	61	0.48	0.00	1.27
0	2	3	57	0.00	0.00	0.81
0	3	4	57	0.00	0.00	0.97
0	total	14	175	0.00	0.00	0.66
1	1	15	57	1.39	0.76	2.00
1	2	6	59	0.31	0.00	1.20
1	3	6	62	0.23	0.00	1.14
1	total	27	178	0.79	0.29	1.19
2	1	20	57	1.73	1.16	2.31
2	2	30	60	2.19	1.68	2.73
2	3	24	54	2.03	1.48	2.61
2	total	74	171	2.00	1.69	2.31
3	1	43	57	2.82	2.32	3.36
3	2	71	63	3.56	3.09	4.05
3	3	50	59	3.02	2.53	3.54
3	total	164	179	3.16	2.88	3.45
4	1	77	54	4.05	3.55	4.58
4	2	78	61	3.82	3.35	4.31
4	3	93	63	4.13	3.67	4.61
4	total	248	178	4.00	3.73	4.28

- 本研究で**染色体断片が線量評価の指標になる**ことが見出された。
- **2 Gyのトリアージ判定が可能**であることが示された。
 - * ベイズ法でも同様の結果を得た(データ省略)。
- 二動原体と染色体断片の情報を併用するの**もひとつの方法**となる。

【進捗3 考察・その他の情報】

■ AIによる画像分析で、トリアージ・レベルの線量推定が試作段階でも実現した。今後の本格開発により、0.5 Gyと1 Gy、1.5 Gyと2 Gy、といったより細かな判別を高度に効率よく実施することが期待できる。

■ 年度末まで試験セット数を増やすなどしてAIによる二動原体自動分析検証を重ね、この分野での先鞭をつけた成果の論文投稿を予定。今後、細胞数を増し、本フィージビリティ・スタディで得られた従来注目されていなかった情報(染色体断片生成頻度)も加えることにより、更に効率よく線量推定が可能となることもわかった。

■ 動原体が特定できるPNA-FISH画像を用いた教師データを増加することで、単色染色画像からの情報[形態、染色の濃淡、二動原体生成に伴う付加情報(染色体断片など)]により、二動原体判別がより効率化した。将来は単色染色画像の学習モデルのプロトタイプも改善されることが期待できる。

■ 通常のラボの画像解析用PCスペックでも、1検体(1000細胞)につき10分程度で可能。

■ PNA-FISHを利用することは容易であることがわかった。

* コスト: プローブ代1検体につき約4~8円、染色時間30分~1時間(ギムザ染色は15分)。国内の多くの染色体解析ラボ・検査会社で技術は確立されている。

AI を利用した染色体画像判別の性能向上(見込み)

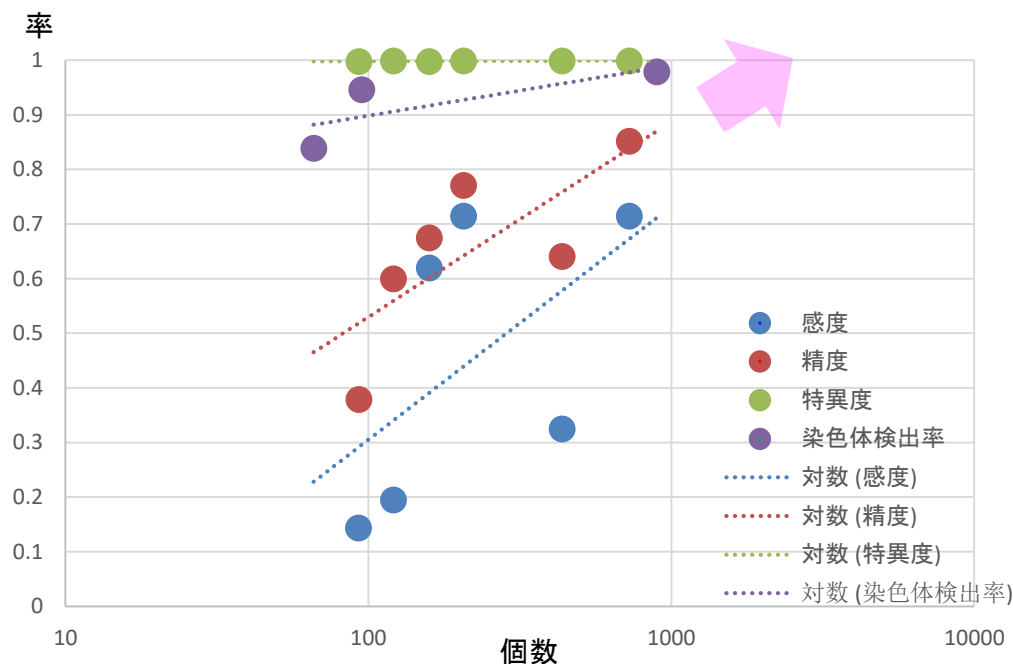


図9. テストごとの染色体検出率および検出された各種染色体異常の個数と正確さとの関係

染色体検査会社の AI による核型
分析正答率~100%(既に実用化)

医用画像研究実例
~2000画像

→ 多検体対応で
大きな力となる

【進捗4 今後の課題】

- ① 現段階では、正確な教師データが得られ、かつ、実際の運用上(コスト、技術をもつラボ数など)支障が無いことがわかったPNA-FISH画像を用いたAI自動画像判定技術開発を進めていくこととする。[AIの特性から、将来、教師データや訓練回数の増加によって単色染色(PNA-FISH画像のもつ対比染色像)のみでも判定が可能となることは期待できる。]
- ② 教師データの分類種間バランスを良くし、判別精度をさらに上げ、1 Gy判定やより細かな線量の区別を可能にする。
- ③ テストセット数を増やし、検証をより強く行う。
- ④ 画像品質差対策として、現段階では1000～3000枚の画像をAIにより自動選別(分析候補画像提示)する手法を提案するが、将来的には、AI自動分析ベーシック・モデルを開発し、ラボごとに一定枚数の画像を借り受けて調整する機能を持たせる方向で進める。モデルの向上・配布については量研機構で一括管理する方向で進める。
- ⑤ 現段階では線量推定計算で試用した標準法(最尤法)もベイズ法も差違が無いが、今後学習モデルの教師データ量ビッグ・データ化するとベイズ法の有用性が見えてくる可能性があるため、両方法に注視していく。
- ⑥ 以上の方向性を原案に、また平成31年度に確立した多検体全血培養法・長期保存細胞培養法等を交えて、運用ストラテジーを検討する。

【追記】 「進捗4 今後の課題」のうち、③および⑥について実施した。

③ 線量推定試験の追加試験

表4. 二動原体頻度による線量推定結果（2.0 Gy画像追加）

線量 (Gy)	セット番号	二動原体 (個)	メタフェーズ (個)	推定線量 (Gy)	95% LCL (Gy)	95% UCL (Gy)
2.0	4	13	58	2.24	1.30	3.12
2.0	5	13	55	2.33	1.38	3.22
2.0	6	8	48	1.80	0.34	2.85
2.0	7	10	64	1.71	0.34	2.63
2.0	8	16	58	2.57	1.72	3.41
2.0	9	15	57	2.49	1.62	3.35
2.0	10	18	55	2.85	2.03	3.69
2.0	11	14	57	2.38	1.48	3.25
2.0	12	10	59	1.83	0.34	2.77
2.0	13	9	54	1.80	0.34	2.79
2.0	14	9	61	1.62	0.34	2.59
2.0	15	12	60	2.07	1.05	2.96
2.0	16	9	59	1.67	0.34	2.65
2.0	セット4～6合算	34	161	2.15	1.61	2.68
2.0	セット7～9合算	41	179	2.28	1.78	2.76
2.0	セット10～12合算	42	171	2.38	1.89	2.87
2.0	セット13～15合算	30	175	1.84	1.25	2.34
2.0	セット1～16合算	193	916	2.15	1.93	2.37

(注) セット1～3については表2を参照。なお、真の値が2.0 Gy未満の検体の場合、線量の過大評価は被ばく患者候補の見逃しにはならないので、追加試験は進めていない。

■ 二動原体染色体を指標とした場合、95%信頼限界上限値はすべてで2.0 Gyを超え、見逃しは起こらなかった。試みに、3セット毎の合算データ(細胞数150以上)、およびセット1～16すべての合算データ(細胞数1000弱)での線量推定を行ったところ、精度が上がり、信頼限界の幅が減少する傾向がみられた。

■ AIの場合、1000細胞で10分程度を要するので、解析枚数を増やすことによる検査時間の増大影響は少なく、推定精度の向上が見込まれる。熟練者目視観察では1検体当たり500から1000個程度のメタフェーズを観察することで95%信頼限界上限値が± 0.15 Gyまでの線量計測が可能であることが知られている。AIの場合、1検体当たり染色体標本1枚を検査するという目安で画像取得する方法もあり得る。さらに、今後の研究では、細胞数をあらかじめ限定するのではなく、熟練専門家の目視判定によるトリアージの目標細胞数50個の場合と同等の、95%信頼限界上限値が± 0.5 Gyとなるレベルの細胞数までAIで判定させるといったような方法も考慮に入れる必要がある。

【追記】

③ 線量推定試験の追加試験(つづき)

表5. 染色体断片による線量推定結果 (2.0 Gy画像追加)

線量 (Gy)	セット番号	染色体断片 (個)	メタフェーズ (個)	推定線量 (Gy)	95% LCL (Gy)	95% UCL (Gy)
2.0	4	22	58	1.83	1.28	2.40
2.0	5	31	55	2.37	1.84	2.92
2.0	6	21	48	2.01	1.42	2.63
2.0	7	30	64	2.11	1.60	2.63
2.0	8	26	58	2.04	1.51	2.60
2.0	9	30	57	2.27	1.74	2.82
2.0	10	37	55	2.64	2.12	3.18
2.0	11	21	57	1.79	1.23	2.37
2.0	12	29	59	2.17	1.65	2.72
2.0	13	32	54	2.44	1.91	3.00
2.0	14	25	61	1.93	1.40	2.47
2.0	15	30	60	2.19	1.68	2.73
2.0	16	34	59	2.4	1.89	2.94
2.0	セット4～6合算	74	161	2.08	1.76	2.40
2.0	セット7～9合算	86	179	2.14	1.84	2.44
2.0	セット10～12合算	87	171	2.22	1.92	2.53
2.0	セット13～15合算	87	175	2.19	1.89	2.49
2.0	セット1～16合算	451	916	2.17	2.04	2.30

(注) セット1～3については表3を参照。

■ 染色体断片を指標とした場合も、95%信頼限界上限値はすべてで2.0 Gyを超え、見逃しは起こらなかった。3セット毎の合算データ(細胞数150以上)、およびセット1～16すべての合算データ(細胞数1000弱)での線量推定を行ったところ、精度が上がり、信頼限界の幅が減少する傾向がみられた。

■ 染色体異常の分類性能、線量効果曲線試作、線量推定試験において、染色体断片を指標とした場合、二動原体染色体よりも成績が良い傾向にある。これまで染色体線量評価では染色体断片は用いられることが無かったが、本研究により、その有用性が見出された。今後の研究開発では、AIを用いた線量推定の性能を上げるために、二動原体染色体との併用を検討していく。

【追記2】 「進捗4 今後の課題」の⑥(運用戦略)について検討を行った。

■ 原子力災害等被ばく事故における緊急時対応の戦略

実際の放射線被ばく事故における運用には、AI自動判定だけ多検体対応ができて迅速化しても十分とは言えない。事故の様々な場合を想定し、染色体線量評価のための血液検体処理から画像判定・線量評価までの戦略を準備しておく必要がある。

1) 血液検体の保存・培養

本研究では、長期凍結した末梢血リンパ球の培養法を確立した [Y. Suto et al., 2020]。これにより、例えば、職業上被ばくが想定される緊急作業員やファースト・レスポンドー、宇宙飛行士などについて事前に採血し分離したリンパ球を凍結保存しておけば、被ばく事故が起こった時に、この凍結保存細胞を用いてバックグラウンド値(被ばく事故前の染色体異常頻度)の調査を行えるので、より精度の高い線量評価を行うことができる。

また、長期凍結保存の完全培地が利用可能であることを示す研究を行った(投稿論文準備中)。完全培地を分注して凍結保存し備えることで、大規模放射線被ばく事故が起こったとき、検体間のクロスコンタミネーションを起こさず迅速に多検体の培養を行うことが可能である。

【追記2】

■ 原子力災害等被ばく事故における緊急時対応のストラテジー（つづき）

2) 細胞培養からから染色体画像取得まで

万一の大規模放射線事故・災害に備えて、染色体分析が行える専門要員の拡充は不可欠である。本研究事業における情報収集から、高度被ばく医療の基幹5センターだけでなく、国内の臨床染色体検査会社も担当可能になっていることが明らかになった。日本人類遺伝学会臨床細胞遺伝学認定士制度 (<http://cytogen.jp>) も活用し、放射線被ばく事故の状況・規模によって検査支援要請ができるようにしておくことも重要である。

3) 染色体画像データの取り扱いについて

現在、染色体画像は医用画像の扱いとなっており、個人情報の保護は必須である。このため、染色体画像データの電子的授受には制約が大きい。今後、画像データの安全管理・安全共有のためのシステム作りも必要である。

【研究評価委員会】

評価： A

総合コメント： AIを利用した染色体自動画像判定が短時間で実施できることを示すなど、重要な結果が得られており、研究は高く評価される。深層学習用の教師データを複数の専門家で検証するなど、次のステップに期待する。

【対応】

ラボ間の技術差や患者の状態・溶血などによって生じる多様な画質に対応できるようにする必要がある。また、実際的な標準化、利用者の確保・拡大には、信頼を得ることが重要である。それらのために、今後、以下のことを検討し進める。

- 1) 染色体線量評価の専門家の協力を得て、画像データの貸借または共有、AI自動染色体判別モデルの評価・検証を行い、信頼性を高める。
- 2) 国際誌や国際学会での成果発表、および国際標準化機構において標準化課題提案を目指す。これらにより、国内だけでなく各国の生物学的線量評価の主要な担当機関の専門家による評価・検証を得ることができる。

【平成31年度 成果公表】

■ 論文発表

1. Y. Suto, et al., Cytogenetic Examination of Human Peripheral Blood Lymphocytes Cryopreserved after Gamma Irradiation: A Pilot Study. *Cytologia*, 85: 71-77, 2020.
2. 弓至. 深層学習を用いた画像認識AIによる生物学的線量推定. 北海道大学大学院情報科学研究院 修士論文(2020年3月).

■ 学会発表

1. Y. Suto, Radiation cytogenetics in Japan: current state and perspectives. ICRR 2019, シンポジウム(マンチェスター、8月25日)[招待]
2. 弓至, et al., Chromosome dose evaluation by image recognition AI using deep learning. 日本進化学会第21回大会(札幌、8月7日～10日)
3. 弓至, et al., Abnormal chromosome detection by image recognition AI using deep learning. 日本バイオインフォマティクス学会2019年年会・第8回生命医薬情報学連合大会(東京、9月9日～11日)
4. Y. Suto, 演題未定. The 66th Annual Radiation Research Society Meeting(ハワイ、2020年10月18～21日を予定)[招待]

■ 取材(記事)

1. 共同通信社 臨界事故20年特集「被ばく線量 AIが推定」(年9月6日、西日本新聞ほかで掲載)

[参考](関連の業績)

1. 染色体画像等の図説提供 (Oxford社英語版1件、Springer社ドイツ語版1件、BBCドキュメンタリー英語版1件)
2. 国際標準化機構 プロトコール (Y. Suto、策定主要メンバー) [International Organization for Standardization (ISO)/TC85/SC2: ISO 20046(FISH分析)、ISO 19238(二動原体染色体分析)、ISO 21243(大規模事故における二動原体染色体分析)の出版・改訂]

【平成31年度 その他の研究会合・学会参加・情報収集】

- 1) 研究会合 3回 (個別会合1回)
- 2) 画像検討会(教師データ作成) 8回
- 3) 染色体線量評価・人工知能研究に関連した学会・会議
 - * 国際会議: ISO, Working Group 18(生物線量評価)(マンチェスター、8月)
ISO, 本会議およびWorking Group 25(放射線事故時の住民・作業員線量モニタリング)(岡山、11月)
 - * 国際会議: ARADOS 2019(北京、11月)(アジア線量評価ネットワーク会議。本研究のテストセットを用いて目視判定による Intercomparison Studyを主催。レビュー論文投稿予定)
 - * PHE訪問・講演実施(ジドコット、8月)
 - * 国内学会: 日本メディカルAI学会(東京、1月31日～2月1日参加)
 - * 第3回AI・人工知能EXPO(東京、4月)、AI関連企業ヒアリング、染色体検査会社ヒアリング

【自己評価】

1. 研究代表者(主任研究者)による自己評価

評価の視点	自己評価	コメント
評価時点までの研究の実施が研究計画に沿って行われているか	計画を上回る	AIによる画像判定支援の有効性を明らかにするために、0～5 Gyに相当する放射線照射血液から作成した染色体画像(約15,000枚)を用い、AI導入試験を実施した。深層学習により、現試作段階の学習モデルでも、これまで検査者の目視観察により、1検体(1,000細胞)当たり3日以上要していた画像判定が、AI自動画像判定10分間で良好な精度での評価が可能となった。さらに、急性放射線症候群の前駆症状に対し医学的な介入が必要となる基準である2 Gy以上の放射線照射血液サンプルについて、患者の見逃しなく判別が可能となった。本成果は原子力災害をはじめとする多数の被ばく者を生じる放射線事故に対し、トリアージに必要な判断材料として適用できる可能性を示唆するものである。AIにより画像判定の標準化が達成されることで、ラボや国の別なく協働で線量評価を行うことができるようになる。今後、本格開発研究で教師データの染色体画像を増やす等により、1 Gy以下の放射線照射血液の判別についても精度向上が期待される。
今年度の進捗や達成度を踏まえて、次年度の研究計画に変更が必要か	今年度で終了	平成31年度末に本研究(フィージビリティ・スタディ)は終了する。 本研究ではAIによる画像判定支援の有効性を示す重要な結果が得られており、今後、画像判定の標準化に向け発展させる。

2. 分担研究者による自己評価 (分担研究者 無し)

【放射線規制及び放射線防護分野への活用】

本研究事業(フィージビリティ・スタディ)で、以下の成果を得ることができた。

1) 基盤技術の構築

- * AI の一種である深層学習を導入した学習モデルの試作。
- * AI に適した標本作製法、画像作成法の確立。
- * 実際の AI 技術の運用で必要となる多検体対応の細胞培養法の確立。

2) AI 導入の有効性・実現可能性

- * 試作モデルで、二動原体染色体を指標とした線量推定試験において、約50個のメタフェーズ画像データから、2 Gy未満と2 Gy以上の被ばくが判別できた。
- * 新たに染色体断片も線量評価の良好な指標となることを見出した。二動原体染色体と併用することで、より正確な線量推定ができる可能性が高い。

3) 線量推定作業の高度の効率化

- ・1検体当たりの画像判定(1000メタフェーズ)に要する時間が実質30時間以上(目視)から10分(AI)へと短縮化できることがわかった。

以上の成果に基づいて研究開発することで、染色体画像判別の標準化・効率化が達成される。特に原子力災害をはじめとする多数の被ばく者を生じる放射線事故において、被ばく医療診断の支援(トリアージに有用な推定線量情報の提供)に大きく貢献できるものと期待できる。

【放射線規制及び放射線防護分野への活用】（補足）

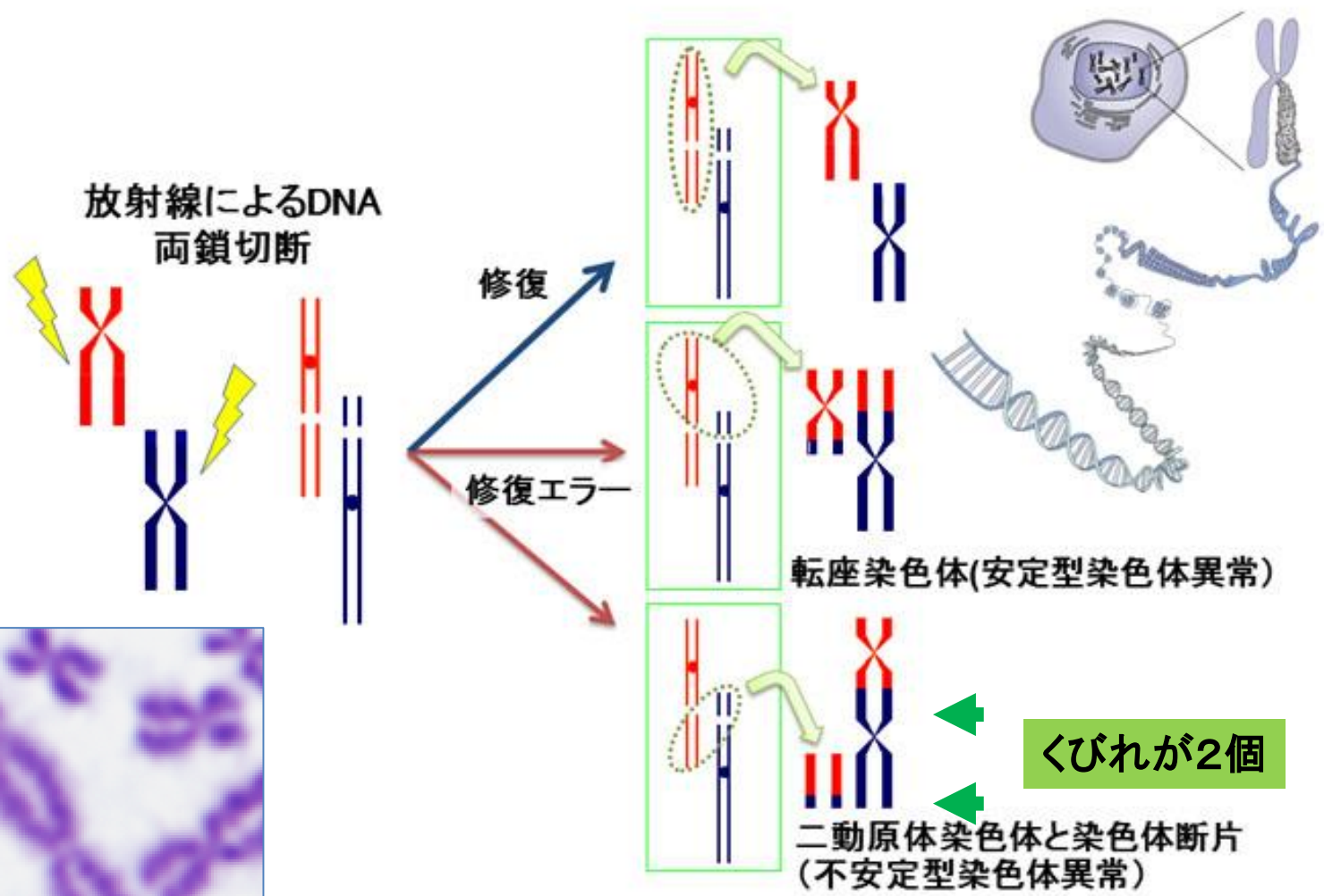
AI 自動染色体判定の応用(発展)

本研究事業は原子力災害における緊急時線量評価支援が主要な目的であり、二動原体染色体分析についてAI導入を試験した。本研究におけるプレ試験では、3-color FISH(染色体ペインティング)による交換型染色体異常の検出においてもAIが威力を発揮することが示唆された。

染色体ペインティング法は費用・時間がPNA-FISHより大きいため、緊急時対応ではPNA-FISH法に劣るが、染色体異常の検出効率は同様に高い。染色体番号も蛍光標識も多様な組み合わせが可能で、以下のことに有用な手法であるため、将来必要に応じてAIによる画像判別の応用を検討すると良い。

- 1) 1ヶ月～30年以上遡る過去の被ばくの線量評価も、現在になって採取する血液試料や長期凍結保存した染色体標本から可能である。また、長期間被ばくする作業従事者の累積線量の評価も可能である。
- 2) 一般的な臨床染色体検査だけでなく、放射線被ばくが原因で将来起こり得る疾患で特徴的な染色体異常を伴うものがあれば、早期発見スクリーニングや治療効果観察に役立つ可能性がある。

放射線被ばくによる染色体異常の形成



二動原体染色体像

[QST e-learning 教材を改変]

■ エルピクセル株式会社 (<https://lpixel.net/>)

→ 東京大学発。ライフサイエンスと画像解析を背景に独自の技術を研究・開発。オリンパス、富士フィルム、キャノンメディカル等から出資を受け複数の医療、医薬用システム開発

* 研究事例（一部抜粋。括弧内は共同研究先）

- ・Deep Learningを用いた脳動脈瘤検出アルゴリズム(大阪市立大学)
- ・生物画像自動分類ソフトウェア CARTA(東京理科大学、国立がん研究センター)

* 国家プロジェクトへの参加

- ・経済産業省 / 戦略的基盤技術高度化支援事業:
生体組織の立体構造情報と人工知能を活用する病理診断支援システム
- ・科学技術推進機構 / 戦略的創造研究推進事業(CREST):
栽培植物倍数体のマルチオミクス技術開発プロジェクト
- ・厚生労働省 / 革新的がん医療実用化研究事業:
人工知能技術を用いた大腸内視鏡検査における病変検出・診断支援技術の開発

* 特許

- ・第4688954号 能動学習型の生物医学画像自動分類ソフトウェアCARTA
- ・第4696278号 領域分割画像生成方法、領域分割画像生成装置及びコンピュータプログラム
- ・第6329651号 画像処理装置及び画像処理方法

* 論文

- ・2019年 IMACEL: A cloud-based bioimage analysis platform for morphological analysis and image classification. PLOS ONE DOI:10.1371/journal.pone.0212619 ほか
- ・2018年 Deep Learning for MR Angiography: Automated Detection of Cerebral Aneurysms. Radiology DOI: 10.1148/radiol.2018180901 ほか
(2017~13年分省略)