

Instytut Chemii Bioorganicznej
Polskiej Akademii Nauk
w Poznaniu

Zakład Inżynierii Genomowej

mgr inż. Mateusz Nowaczyk

**Identyfikacja procesów zaangażowanych w skracanie ciągu powtórzeń
CAG w wyniku indukcji dwuniciowych pęknięć DNA w genie *HTT***

Identification of processes involved in CAG repeat contraction following double strand break
induction in the *HTT* gene

Praca doktorska wykonana pod kierunkiem

prof. dr hab. Marty Olejniczak

Poznań, 2025

*Pragnę serdecznie podziękować
Pani promotor prof. dr hab. Marcie Olejniczak
za wszelkie wsparcie udzielane mi w czasie realizacji doktoratu.*

*Dziękuję również koleżankom i kolegom
z Zakładu Inżynierii Genomowej
oraz Zakładu Biotechnologii Medycznej
za dzielenie się swoim doświadczeniem,
dobrą współpracę i miłe towarzystwo.*

Pracę dedykuję moim Rodzicom.

Spis treści

1. Wstęp.....	12
1.1 Sekwencje mikrosatelitarne w genomie człowieka.....	12
1.2 Choroby związane z ekspansją sekwencji mikrosatelitarnych.....	13
1.3 Niestabilność powtórzeń CAG a naprawa DNA.....	18
1.4 Mechanizmy naprawy dwuniciowych pęknięć DNA.....	21
1.4.1 Wprowadzenie.....	21
1.4.2 Łączenie niehomologicznych końców (NHEJ).....	24
1.4.3 Rekombinacja homologiczna (HR).....	24
1.4.4 Alternatywne ścieżki naprawy dwuniciowych pęknięć DNA.....	26
1.5 Skracanie ciągów powtórzeń CAG w efekcie naprawy DNA.....	28
1.5.1 Narzędzia inżynierii genetycznej do indukcji dwuniciowych pęknięć DNA.....	28
1.5.1.1 Nukleazy palca cynkowego (ZFN).....	29
1.5.1.2 System TALEN.....	29
1.5.1.3 System CRISPR-Cas9.....	30
1.5.2 Wykorzystanie systemu CRISPR-Cas9 do skracania ciągów powtórzeń CAG.....	33
2. Cel pracy.....	36
3. Materiały.....	37
3.1 Linie komórkowe.....	37
3.2 Odczynniki do hodowli i transfekcji komórek.....	37
3.3 Oligonukleotydy.....	38
3.4 Przeciwciała.....	40
3.5 Chemiczne inhibitory białek oraz siRNA.....	40
3.6 Enzymy restrykcyjne.....	41
3.7 Składy buforów i żeli do elektroforezy.....	42
3.8. Pozostałe odczynniki i zestawy odczynników.....	45

3.9 Sprzęt.....	46
3.10 Wykorzystywane oprogramowanie	46
4. Metody.....	47
4.1 Hodowle komórkowe.....	47
4.2 System CRISPR-Cas9.....	47
4.2.1 Kompleksy RNP.....	47
4.2.2 Ekspresja systemu CRISPR-Cas9 z plazmidu.....	48
4.3 Transfekcja komórek metodą elektroporacji.....	48
4.4. Transfekcja komórek siRNA za pomocą lipofektaminy.....	49
4.5 Użycie inhibitorów chemicznych białek naprawczych.....	49
4.6. Transfekcja komórek plazmidem za pomocą polietylenoiminy.....	49
4.7 Analiza produktów naprawy DNA w <i>locus HTT</i> z wykorzystaniem metody NGS.....	50
4.8 Analiza dynamiki skracania ciągu powtórzeń CAG w punktach czasowych.....	51
4.9 Analiza resekcji końców DNA.....	51
4.10 Analiza udziału białek naprawy DNA w skracaniu ciągu powtórzeń CAG.....	54
4.11. Amplifikacja DNA.....	56
4.11.1. Zagnieżdżony PCR (nested) do analizy NGS.....	56
4.11.2. PCR do amplifikacji <i>locus HTT</i>	57
4.11.3 PCR do analiz resekcji końców DNA.....	57
4.11.4 Amplifikacja cDNA do analizy ekspresji genów.....	58
4.11.5 PCR do amplifikacji <i>locus off-target</i> dla gRNA1.....	59
4.11.6 PCR z gradientem temperatury przyłączania starterów.....	59
4.12. Odwrotna transkrypcja i analiza ekspresji genów.....	60
4.13 Elektroforetyczny rozdział produktów PCR w żelu agarozowym.....	61
4.14. Analiza efektywności edycji metodą T7E1.....	61

4.15 Elektroforetyczny rozdział białek w żelu poliakrylamidowym i analizie western blot....	62
5. Wyniki	63
5.1 Skracanie ciągów powtórzeń CAG w wyniku indukcji DSB w genie <i>HTT</i>	63
5.1.1 Przebieg skracania ciągu powtórzeń CAG w wyniku indukcji DSB z udziałem gRNA1	64
5.1.2 Przebieg skracania ciągu powtórzeń CAG w wyniku indukcji DSB z udziałem gRNA2	67
5.1.3 Przebieg skracania ciągu powtórzeń CAG w wyniku indukcji DSB z udziałem gRNA4	68
5.1.4 Wycięcie ciągu powtórzeń CAG w wyniku indukcji DSBs z udziałem gRNA1 oraz gRNA4.....	69
5.2 Analiza dynamiki procesu skracania ciągu powtórzeń CAG w wyniku indukcji DSB w genie <i>HTT</i>	70
5.3 Optymalizacja metody określania zasięgu i intensywności resekcji końców DNA w komórkach ludzkich.....	72
5.4 Analiza resekcji końców DNA w czasie naprawy DSBs w regionach powtórzeń CAG w różnych modelach komórkowych.....	76
5.4.1 Analiza resekcji końców DNA w czasie naprawy DSB po edycji z udziałem gRNA1 w linii komórkowej HEK293T_41CAG.....	76
5.4.2 Analiza resekcji końców DNA w czasie naprawy DSB po edycji z udziałem gRNA2 w linii komórkowej HEK293T_41CAG.	80
5.4.3 Analiza resekcji końców DNA w czasie naprawy DSB po edycji z udziałem gRNA2 w linii komórkowej fibroblastów HD.....	83
5.5 Analiza udziału białek naprawczych w skracaniu ciągu powtórzeń CAG w genie <i>HTT</i> ...	86

5.5.1 Analiza udziału białek naprawczych w skracaniu ciągu powtórzeń CAG po indukcji DSB przy użyciu gRNA1.....	87
5.5.2 Analiza udziału białek naprawczych w skracaniu ciągu powtórzeń CAG po indukcji DSB przy użyciu gRNA2.....	95
5.6 Wpływ POLθ na wyniki naprawy DSBs w regionie z powtórzeniami CAG w genie <i>HTT</i>	105
6. Dyskusja.....	108
7. Wnioski.....	120
8. Spis rysunków i tabel prezentowanych w pracy.....	121
8.1 Rysunki.....	121
8.2 Tabele.....	124
9. Literatura.....	126

Dorobek naukowy autora

1) Computational Tools and Resources Supporting CRISPR-Cas Experiments.

Cells (2020) Śledziński P., Nowaczyk M., Olejniczak M.

2) Paving the way towards precise and safe CRISPR genome editing.

Biotechnology Advances (2021) Śledziński P., Dąbrowska M., Nowaczyk M., Olejniczak M.

3) CRISPR/Cas9-induced double-strand breaks in the huntingtin locus lead to CAG repeat contraction through DNA end resection and homology-mediated repair.

BMC Biology (2024) Śledziński P⁺, Nowaczyk M.⁺, Śmiełowska M. I., Olejniczak M.

⁺ Paweł Śledziński oraz Mateusz Nowaczyk mają wspólne pierwszoautorstwo.

Streszczenie

Ekspansje powtórzeń CAG są przyczyną wielu neurodegeneracyjnych, dziedzicznych i obecnie nieuleczalnych chorób człowieka (choroby poliQ). Jednym z takich schorzeń, jest choroba Huntingtona (HD). Nowatorską strategią terapii dedykowanej tej chorobie jest kontrolowane skracanie ciągu powtórzeń, za pomocą narzędzi inżynierii genetycznej, takich jak CRISPR-Cas9. Indukcja dwuniciowych pęknięć DNA (DSBs) w obrębie ciągów powtórzeń CAG prowadzi do ich niestabilności, zarówno wydłużania, jak i skracania ciągu, lecz mechanizmy stojące za obserwowaną niestabilnością nie są do końca poznane. Wcześniejsze badania sugerują zaangażowanie w ten proces niektórych ścieżek naprawy DNA. Brak jest jednak badań dotyczących mechanizmów naprawy DSBs w obrębie powtórzeń CAG w endogennym *locus* w ludzkim modelu komórkowym.

Celem niniejszej pracy było dokładniejsze poznanie mechanizmów naprawy DSBs w regionach powtórzeń CAG w endogennym *locus* w komórkach ludzkich. Wyniki niniejszej pracy ukazały złożoność naprawy DSBs w regionach z powtórzeniami CAG. Pozwoliły na identyfikację czynników i mechanizmów zaangażowanych w ten proces na różnych jego etapach oraz przyczyniły się do stworzenia modeli naprawy DNA w regionach z powtórzeniami CAG u ludzi. W niniejszej pracy wykazano udział resekcji końców DNA w naprawie DSBs w regionach z powtórzeniami CAG i przedstawiono przesłanki sugerujące zależność intensywności resekcji a w konsekwencji wyboru mechanizmu naprawy DNA i jej efektów od długości ciągu. Stwierdzono silny wpływ umiejscowienia DSB względem powtórzeń CAG oraz sekwencji flankujących ciąg na wybór mechanizmu naprawy DNA. Wykazano, że naprawa DSB indukowanego na początku ciągu powtórzeń CAG angażuje mechanizm NHEJ oraz polimerazę DNA theta (POL θ), prowadząc do wycięcia całego ciągu powtórzeń. Stwierdzono, że indukcja DSB wewnątrz ciągu, skutkuje jego skracaniem z zachowaniem ramki odczytu, co odbywa się z udziałem mechanizmu NHEJ oraz wykorzystującego dalekie resekcje mechanizmu angażującego białka rekombinacji homologicznej takie jak RAD51.

W efekcie praca ta przyczyniła się do lepszego poznania zjawisk odpowiedzialnych za niestabilność ciągów powtórzeń CAG. Poszerzenie wiedzy w tej dziedzinie może przyczynić się do opracowania skutecznej i bezpiecznej terapii chorób poliQ wykorzystujących technologie edycji genomu, takie jak CRISPR-Cas9.

Abstract

Expansions of CAG repeats are the cause of many neurodegenerative, hereditary, and currently incurable human diseases (polyQ diseases). One such condition is Huntington's disease (HD). An innovative therapeutic strategy dedicated to this disease is the controlled shortening of the repeat tract using genetic engineering tools such as CRISPR-Cas9. Induction of double-strand breaks (DSBs) within CAG repeat tracts leads to their instability, resulting in both expansions and contractions of the tract. However, the mechanisms underlying this observed instability are not fully understood. Previous studies suggest the involvement of certain DNA repair pathways in this process. Nevertheless, there is a lack of studies on DSBs repair mechanisms within CAG repeats in an endogenous locus in a human cell model.

The aim of this study was to better understand the mechanisms of DSBs repair in CAG repeat regions at an endogenous *locus* in human cells. The results of this work revealed the complexity of DSBs repair in regions containing CAG repeats. They enabled the identification of factors and mechanisms involved in this process at various stages and contributed to the development of models of DNA repair in regions with CAG repeats in humans. This study demonstrated the involvement of DNA end resection in DSBs repair within CAG repeat regions and presented evidence suggesting a correlation between the extent of resection (and consequently, the choice of DNA repair mechanism and its outcome) and the length of the repeat tract. A strong influence of the DSB position relative to the CAG repeats and the flanking sequences on the choice of DNA repair pathway was observed. It was shown that repair of a DSB induced at the beginning of the CAG repeat tract involves the NHEJ mechanism and DNA polymerase theta (POL θ), which participates in the deletion of the entire repeat tract. It was found that induction of a DSB within the repeat tract results in its contraction while preserving the reading frame, a process involving the NHEJ pathway and a mechanism utilizing extensive resection and homologous recombination proteins such as RAD51.

As a result, this work contributed to a better understanding of the phenomena responsible for the instability of CAG repeat tracts. Expanding knowledge in this field may aid in the development of effective and safe therapies for polyQ diseases using genome editing technologies such as CRISPR-Cas9.

Wykaz skrótów

alt-EJ (ang. *alternative end joining*)- alternatywne łączenie końców

CDK (ang. *cyclin-dependent kinase*)- kinaza zależna od cyklin

CRISPR-Cas9 (ang. *clustered regularly interspaced short palindromic repeats- CRISPR associated 9*)

DM1 (ang. *Myotonic dystrophy type 1*)- dystrofia miotoniczna typu 1

DSB (ang. *double strand break*)- dwuniciowe pęknięcie DNA

dsDNA (ang. *double-stranded DNA*)- dwuniciowy DNA

en-ChIP-MS (ang. *engineered DNA-binding molecule-mediated chromatin immunoprecipitation- mass spectrometry*)

HD (ang. *Huntington's disease*)- choroba Huntingtona

HR (ang. *homologous recombination*)- rekombinacja homologiczna

MMEJ (ang. *microhomology-mediated end joining*)- łączenie końców za pośrednictwem mikrohomologii

MSNs (ang. *medium spiny neurons*)- średnie neurony kolczaste

NGS (ang. *next generation sequencing*)- sekwencjonowanie następnej generacji

NHEJ (ang. *non-homologous end joining*)-łączenie niehomologicznych końców

ORF (ang. *open reading frame*)- otwarta ramka odczytu

PAM (ang. *protospacer adjacent motif*)

PCR (ang. *polymerase chain reaction*)- reakcja łańcuchowa polimerazy

poliQ- poliglutaminowe (choroby)

RNP- rybonukleoproteina

SSA (ang. *single strand annealing*)- łączenie pojedynczej nici

ssDNA (ang. *single-stranded DNA*)- jednoniciowy DNA

STR (ang. *short tandem repeats*)- krótkie powtórzenia tandemowe

TALEN (ang. *transcription activator-like effector nuclease*)

TINS (ang. *templated insertions*)

TMEJ (ang. *DNA polymerase theta-mediated end joining*)- łączenie końców zależne od polimerazy theta

UTR (ang. *Untranslated Rregion*)- region nieulegający translacji

ZFN (ang. *zinc finger nuclease*)- nukleaza palca cynkowego

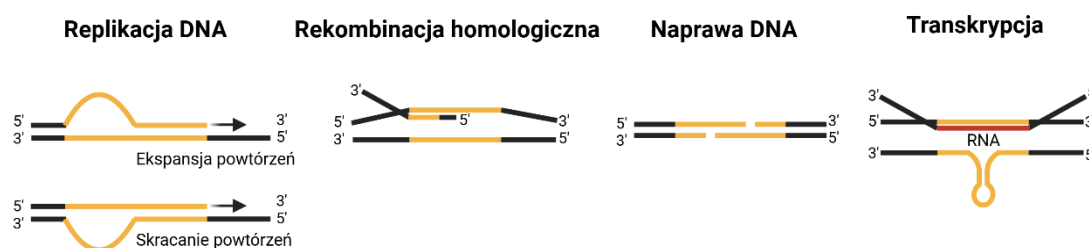
1. Wstęp

1.1 Sekwencje mikrosatelitarne w genomie człowieka

Sekwencje mikrosatelitarne (STR, ang. *Short Tandem Repeats*- krótkie powtórzenia tandemowe) to powtarzające się sekwencje składające się z 2-6 par zasad, występujące powszechnie w genomach eukariotycznych.¹ W ludzkim genomie występuje ponad 1,5 miliona STR i stanowią one około 3% jego zawartości.² Funkcje sekwencji mikrosatelitarnych przez wiele lat pozostawały nieznane. Uznawano, że nie odgrywają one żadnej roli i należą do tzw. śmieciowego DNA- *junk DNA*. Obecnie twierdzi się, że umiejscowione w obszarach pozagenowych mikrosatelity mają udział w regulacji ekspresji genów, poprzez wpływ na zmianę metylacji DNA, upakowanie chromatyny czy alternatywny splicing.^{1,3,4} W przypadku występowania STR w sekwencjach kodujących ale również w regionach 3' i 5' niepodlegających translacji (UTR, ang. *untranslated region*), jak i regionach regulatorowych, ekspansja powtórzeń powyżej pewnej długości, może mieć negatywne skutki w postaci rozwoju szeregu chorób.¹

Sekwencje mikrosatelitarne to elementy wysoce polimorficzne. Szacuje się, że na pojedynczego człowieka przypada około 75 mutacji *de novo* powstałych w obrębie STR.⁵ Duża zmienność długości sekwencji mikrosatelitarnych (liczby powtórzeń) w populacjach wynika z ich wysokiej niestabilności, za którą odpowiada zjawisko nieprawidłowego sparowania nici DNA w regionach powtórzeń, zachodzące podczas takich procesów jak: replikacja DNA, rekombinacja homologiczna czy też naprawa DNA (rysunek 1.1).⁶ Zjawisko tak wysokiej polimorficzności tych sekwencji (zmienność 100-10000 razy większa od zmienności wynikającej z mutacji punktowych) między osobnikami jednego gatunku, pozwoliło wykorzystać STR do badań populacyjnych, tworzenia map genetycznych⁷ opracowania testów na ojcostwo czy identyfikacji sprawców przestępstw.⁸

Zazwyczaj obserwuje się niewielkie zmiany liczby krótkich tandemowych powtórzeń, które mają nikły lub żaden wpływ na fenotyp danego osobnika. Natomiast duże ekspansje mikrosatelit, szczególnie tych umiejscowionych w regionach promotorowych, eksonach i intronach genów, jak również 5' i 3' UTR mogą być przyczyną ponad 50 różnych chorób u ludzi.²



Rysunek 1.1 Procesy, w których dochodzi do ekspansji lub skracania ciągów powtórzeń mikrosatelitarnych. Powtórzenia zaznaczono na żółto. (na podstawie Verbiest i wsp. 2022⁶)

1.2 Choroby związane z ekspansją sekwencji mikrosatelitarnych

Jak wspomniano w poprzednim rozdziale, ekspansje sekwencji mikrosatelitarnych mogą być przyczyną wielu chorób genetycznych, takich jak choroba Huntingtona (HD), dystrofia miotoniczna typu 1 (DM1), zespół łamliwego chromosomu X (FXS) czy różne typy ataksji⁹ (tabela 1.2). Mechanizmy patogenezy zależą od lokalizacji STR w obrębie genu oraz od charakteru i funkcji tego genu. Mechanizmy te można podzielić na: mechanizmy utraty funkcji (LOF, ang. *loss of function*) lub nabycia funkcji (toksycznej) (GOF, ang. *gain of function*). Mechanizmy LOF obejmują hipermetylację DNA lub histonów, czego efektem jest wyciszenie genu, jak również negatywny wpływ na proces transkrypcji danego genu prowadzący do degradacji mRNA.^{10,11} Przykładem ekspansji powtórzeń prowadzącej do wyciszenia ekspresji genu jest ekspansja powtórzeń CGG powyżej 200 w regionie promotora genu *FMR1*, obserwowana w FXS. W przypadku tak dużej ekspansji dochodzi do oddziaływania między powtórzeniami obecnymi w mRNA i komplementarnymi do nich powtórzeniami w części 5' UTR genu *FMR1*. Oddziaływanie prowadzi do zmian epigenetycznych w postaci metylacji promotora i w konsekwencji wyciszenia genu. Brak ekspresji białka FMRP, produktu *FMR1*, które jest kluczowe dla regulacji translacji w synapsach neuronów, skutkuje nieprawidłowym funkcjonowaniem połączeń neuronalnych, czego efektem jest upośledzenie intelektualne osoby obciążonej tym schorzeniem.^{12,13}

Większość znanych chorób związanych z ekspansją powtórzeń rozwija się według mechanizmów GOF, które obejmują: toksyczność RNA, nieprawidłowy splicing, związaną z powtórzeniami i niezależną od kodonu start translację (RAN-translacja, ang. *Repeat Associated Non-AUG translation*), zwiększoną aktywność promotora i agregację toksycznych

białek. Co więcej, mechanizmy te zwykle występują razem, przyczyniając się do wywołania poszczególnych chorób.¹⁰

Tabela 1.2 Zestawienie przykładowych chorób związanych z ekspansją powtórzeń mikrosatelitarnych (na podstawie Chintalaphani i wsp. 2021¹⁰)

Nazwa choroby	Rodzaj powtórzeń	Umiejscowienie ciągu powtórzeń	Patologiczna liczba powtórzeń
Dystrofia miotoniczna typu 1 (DM1)	CTG	3' UTR <i>DMPK</i>	50-10000
Dystrofia miotoniczna typu 2 (DM2)	CCTG	1 intron <i>ZNF9</i>	49-93
Zanik jądra zębatego, jądra czerwienego, gałki bladej i jądra niskowzgórzowego (DRPLA)	CAG	5 ekson <i>ATN1</i>	49-93
Niepełnosprawność intelektualna (FRAXE)	CCG	5' UTR <i>AFF2</i>	>200
Otępienie czołowo-skroniowe (FTD) i stwardnienie zanikowe boczne (ALS)	GGGGCC	5' UTR <i>C9ORF72</i>	24-4000
Choroba Huntingtona (HD)	CAG	1 ekson <i>HTT</i>	36-250
Rdzeniowo-opuszkowy zanik mięśni (SBMA)	CAG	1 ekson <i>AR</i>	38-68
Ataksja rdzeniowo-mózdkowa typu 1 (SCA1)	CAG	8 ekson <i>ATXN1</i>	39-91
Ataksja rdzeniowo-mózdkowa typu 3 (SCA3)	CAG	10 ekson <i>ATXN3</i>	53-87
Ataksja rdzeniowo-mózdkowa typu 7 (SCA7)	CAG	1 ekson <i>ATXN7</i>	34-460
Ataksja rdzeniowo-mózdkowa typu 10 (SCA10)	ATTCT	9 intron <i>ATXN10</i>	280-4500
Ataksja rdzeniowo-mózdkowa typu 36 (SCA36)	GGCCTG	1 intron <i>NOP56</i>	650-2500

W przypadku stwardnienia zanikowego bocznego (ALS) i otępienia czołowo-skroniowego (FTD) za obydwie choroby odpowiada ekspansja powtórzeń GGGGCC w genie *C9ORF72*, która prowadzi do zaburzeń zgodnie z mechanizmem GOF. Osoby chore zazwyczaj charakteryzuje liczba powtórzeń od 250 do 2000, choć znane są przypadki pojawienia się symptomów chorobowych u osób z liczbą powtórzeń wynoszącą około 20.¹⁴ W przypadku ALS/FTD patogeneza odbywa się na poziomie zarówno RNA, jak i białek. Toksyczne formy RNA zawierające ciągi powtórzeń GGGGCC tworzą w neuronach, astrocytach, mikrogleju i oligodendrocytach tzw. skupienia RNA (*RNA foci*), które wiążą białka zaangażowane w obróbkę RNA, prowadząc do zaburzeń w jego transporcie i translacji. RAN-translacja, z kolei, prowadzi do powstawania toksycznych białek DPR (ang. *dipeptide repeat proteins*), które agregują w komórkach nerwowych i prowadzą do ich degeneracji.^{15–17}

Kolejną chorobą wywoływaną przez ekspansje powtórzeń mikrosatelitarnych jest DM1, za której rozwój odpowiada przekroczenie określonej liczby powtórzeń CTG w rejonie 3' UTR genu *DMPK*. Przebieg choroby jest zależny od liczby powtórzeń. Osoby o słabych objawach mają od 50 do 150 powtórzeń CTG, zaś pełna penetracja choroby pojawia się u osób z liczbą powtórzeń mieszczącą się w przedziale od 100 do 1000. W przypadku DM1 obserwowane jest zjawisko antycypacji polegające na nasileniu objawów choroby i wcześniejszym jej ujawnianiu się w kolejnych pokoleniach.¹⁸ Proces ten zostanie szerzej opisany w dalszej części pracy. Transkrypty genu *DMPK* zawierające długie ciągi powtórzeń CTG ulegają akumulacji w jądrze komórkowym, wiążąc ze sobą regulatory transkrypcji i splicingu, w tym m.in. białka MBNL1 i 2, kluczowe w prawidłowym rozwoju mięśni i układu nerwowego. Konsekwencją tego są zmiany w splicingu setek genów, obserwowane u chorych na DM1, których jednym z efektów jest osłabienie mięśni.¹⁹

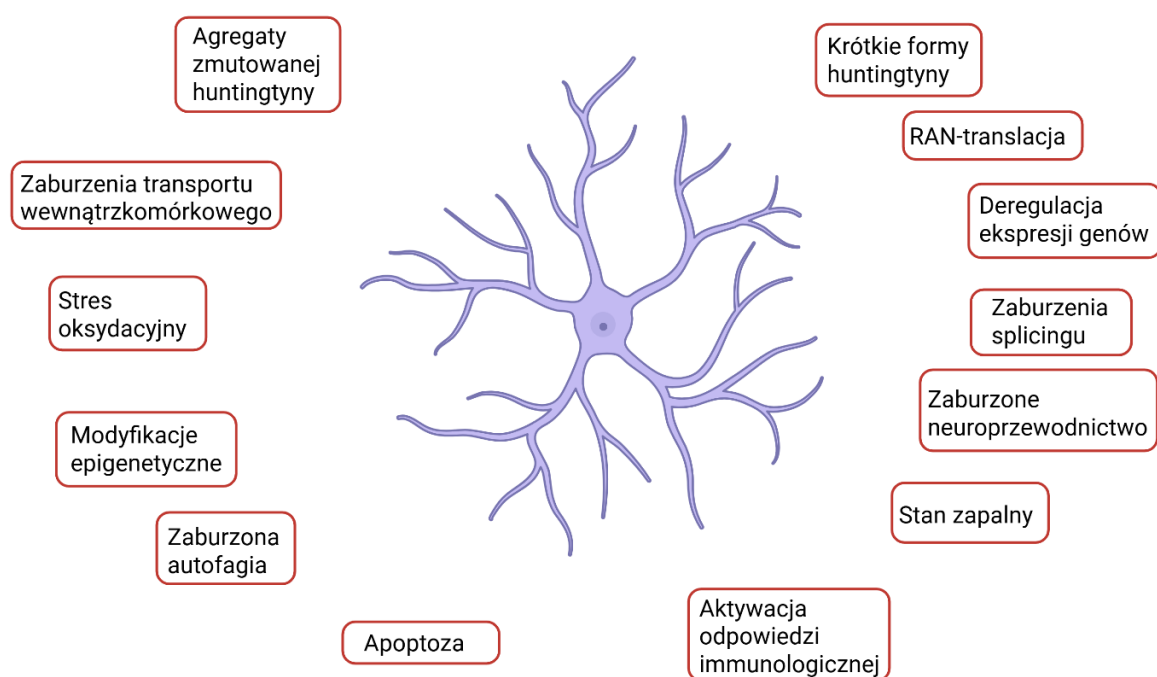
Szczególnym przypadkiem chorób wywołanych ekspansją powtórzeń mikrosatelitarnych są choroby związane z powtórzeniami CAG, określane mianem chorób poliglutaminowych (poliQ). Jest to związane z faktem, że w sekwencji kodującej poszczególnych genów odpowiadających za daną chorobę, występuje ciąg powtórzeń CAG, a każde takie powtórzenie koduje jedną glutaminę. Do najpowszechniej znanych chorób poliQ należą ataksje rdzeniowo-mózdkowe np. ataksja rdzeniowo-mózdkowa typu 3 (SCA3) wywoływana, ekspansją powtórzeń CAG w eksonie 10 genu *ATXN3* czy HD, za której rozwój odpowiada ekspansja powtórzeń CAG w eksonie 1 genu *HTT*. Choroby poliQ są chorobami neurodegeneracyjnymi, mającymi charakter postępujący z wiekiem, na które obecnie nie ma skutecznej terapii. W przypadku większości chorób poliQ maksymalna niepatologiczna liczba powtórzeń CAG

mieści się między 30 a 40 powtórzeniami. Poszczególne choroby charakteryzują się różnym poziomem niestabilności długości ciągów, która dodatkowo jest różna w różnych tkankach organizmu (niestabilność somatyczna). W przypadku wszystkich chorób poliQ obecność dłuższego ciągu powtórzeń skutkuje silniejszymi objawami i wcześniejszym ich ujawnieniem.²⁰

W HD obecność wydłużonego ciągu glutamin w zmutowanej huntingtynie (mHTT) powoduje nieprawidłowe jej fałdowanie i tworzenie się agregatów w neuronach, których nagromadzenie powszechnie uważane jest za jedną z przyczyn rozwoju choroby. Agregaty mHTT zaburzają transport mitochondriów w obrębie neuronów, uniemożliwiając ich dotarcie do miejsc o dużym zapotrzebowaniu energetycznym. Jednocześnie nagromadzenie mitochondriów w innych częściach komórki przyczynia się do lokalnego wzrostu stężenia toksycznych reaktywnych form tlenu w tych obszarach.²¹ Ponadto obecność agregatów mHTT może indukować odpowiedź immunologiczną wobec komórek nerwowych, przyspieszając neurodegenerację.²² Udział w patogenezie mają również krótkie, toksyczne formy huntingtyny zawierające ekson 1 z wydłużonymi ciągami glutamin (HTT_{exon1}). Powstają one w wyniku translacji krótszego transkryptu, będącego efektem zaburzonego splicingu pre-mRNA, który powstaje z wykorzystaniem alternatywnego miejsca poliadenylacji znajdującego się w intronie 1 genu *HTT*.^{23–25} Badania wykazały, że zaburzenia splicingu mogą wynikać z samej obecności zmutowanego mRNA *HTT*, które sekwestruje czynniki obróbki potranskrypcyjnej pre-mRNA takie jak MBNL1.^{26,27} Duży udział w wytworzeniu, wspomnianych krótkich form huntingtyny mają kaspazy, szczególnie kaspaza 6, która ma swój udział w degeneracji neuronów w HD.^{28,29} Kolejnym produktem zmutowanego genu *HTT* są białka powstałe w wyniku RAN-translacji, również mogące tworzyć agregaty, nasilając w ten sposób stres komórkowy i aktywując szlak odpowiedzi na nieprawidłowe białka (UPR, ang. *unfolded protein response*), mogący skutkować apoptozą neuronów.³⁰ mHTT działa również na poziomie epigenetycznym, indukując zmiany w metylacji DNA, czy prowadząc do zaburzeń w modyfikacji histonów.²² Wykazano, że w neuronach HD dochodzi do zmniejszonej ekspresji genów zaangażowanych w funkcje mitochondrialne, neuroprzewodnictwo i plastyczność synaptyczną.³¹ Ponadto mHTT wydaje się odpowiedzialne również za deregulację autofagii w komórce, co może skutkować nagromadzeniem toksycznych białek w cytoplazmie.³² Kolejnym zjawiskiem wywołanym przez mHTT, przyczyniającym się do neurodegeneracji, jest zaburzenie sygnalizacji wapniowej w średnich neuronach kolczastych (MSNs- ang. *medium spiny neurons*) prądkowia, skutkujące nadmiernym napływem Ca^{2+} do komórek, prowadzącym do utraty przez

nie kolców dendrytycznych.³³ W opozycji do przytoczonych powyżej wyników badań znajduje się hipoteza o protekcyjnej funkcji agregatów mHTT. Odbywać by się to miało poprzez zmniejszenie ilości biologicznie dostępnych, rozpuszczalnych i toksycznych form tego białka, w sytuacji gdy tworzy ono agregaty.³⁴ Hipotezę tę wspierają badania przeprowadzone na mysich modelach HD, w których nie obserwowano neurodegeneracji, pomimo obecności dużej ilości agregatów zmutowanej huntingtyny³⁵ lub stwierdzano neurodegenerację, jeszcze zanim agregaty zostały wytworzone³⁶.

Jak wynika z powyższego, patogeneza chorób wywołanych ekspansją powtórzeń mikrosatelitarnych, w tym chorób poliQ napędzana jest przez wiele zjawisk, które w synergiczny sposób prowadzą do rozwoju opisywanych chorób (rysunek 1.2).



Rysunek 1.2 Czynniki wpływające na patogenezę choroby Huntingtona (na podstawie Jiang i wsp. 2023³²).

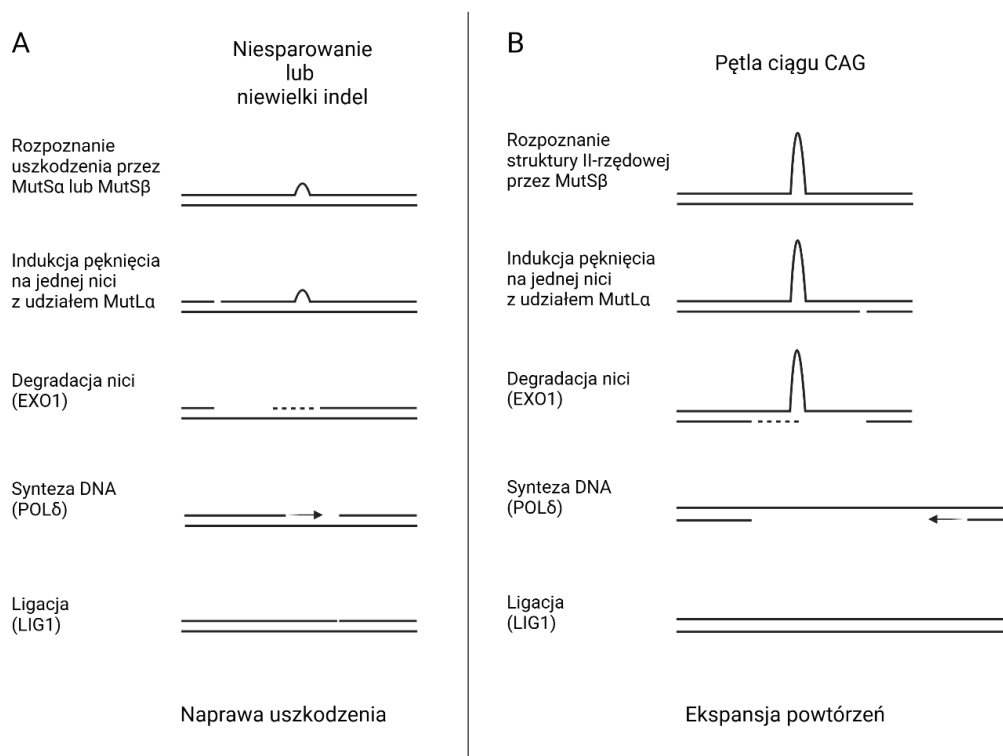
1.3 Niestabilność powtórzeń CAG a naprawa DNA

Zjawiskiem, które odróżnia choroby wywoływane niestabilnością powtórzeń od innych chorób, jest zjawisko antycypacji, polegające na wzroście liczby powtórzeń i nasilaniu się objawów w kolejnych pokoleniach. Jest to związane z niestabilnością germinálną, zachodzącą w komórkach linii płciowej- plemnikach i oocytach. Powtórzenia, które uległy ekspansji w komórkach rozrodczych, będą dziedziczone przez potomstwo. Niektóre spośród tych chorób wykazują wyraźne różnice w niestabilności germinálnej w zależności od płci. Na przykład w HD ekspansja powtórzeń CAG jest bardziej prawdopodobna w linii ojcowskiej, natomiast w zespole łamliwego chromosomu X niestabilność powtórzeń CGG częściej występuje w linii matczynej.³⁷ Odziedziczona liczba powtórzeń, które dodatkowo mogą ulec ekspansji w trakcie rozwoju embrionalnego³⁸, nie jest jedyną przyczyną patogenezy opisywanych chorób. W przypadku wielu z nich, oprócz niestabilności germinálnej, obserwuje się również niestabilność somatyczną. Powoduje ona stopniowe zwiększanie liczby powtórzeń w określonych tkankach organizmu, prowadząc do mozaicyzmu somatycznego, w którym różne komórki mogą zawierać różną liczbę powtórzeń CAG. W przypadku HD tkanką, w której dochodzi do najsilniejszych ekspansji, jest prądkowie, obszar, który jednocześnie ulega najsilniejszej degeneracji. W mózgach osób chorych, w MSNs prądkowia obserwowano ekspansje wynoszące do tysiąca powtórzeń CAG.^{37,39,40} Obecnie zjawisko ekspansji somatycznej w MSNs uważa się za kluczowe w patogenezie HD.⁴¹ Fakt, że w przypadku chorób neurodegeneracyjnych, takich jak HD dochodzi do bardzo dużych ekspansji powtórzeń CAG w neuronach- komórkach dzielących się, w których procesy replikacji DNA czy rekombinacji są zahamowane, sugeruje, że to mechanizmy naprawy DNA mogą być głównie zaangażowane w ekspansje tych powtórzeń w komórkach nerwowych.⁴²

Zidentyfikowano kilka czynników mających duży wpływ na promowanie ekspansji somatycznych, które są intensywnie badane, również pod względem rozwoju potencjalnej terapii chorób. Za mechanizm szczególnie zaangażowany w ekspansję powtórzeń CAG w HD uważa się MMR (ang. *mismatch repair*)⁴³⁻⁴⁶. Mechanizm ten odpowiedzialny jest za korygowanie błędów, powstałych podczas replikacji DNA, takich jak niewłaściwie sparowane zasady azotowe, krótkie insercje i delecje. W rozpoznaniu nieprawidłowo sparowanych nukleotydów biorą udział kompleksy MutSa (heterodimer MSH2-MSH6) i MutSβ (heterodimer MSH2-MSH3). Jeden z tych kompleksów razem z kompleksem MutLa (składającym się z MLH1-PMS2) oraz PCNA tworzy kompleks inicjujący naprawę. PCNA aktywuje MutLa, który następnie nacina świeżo zsyntetyzowaną nić DNA w pewnej odległości

po stronie 5' od niesparowania. W kolejnym kroku EXO1 dokonuje degradacji naciętej nici w kierunku uszkodzenia. Następnie na odcinku jednoniciowym POL δ syntetyzuje ponownie drugą nić, wiązanie fosfodiesterowe zaś zostaje odtworzone przez Ligazę DNA I (rysunek 1.3 A)⁴⁷

Rolą mechanizmów naprawy DNA jest utrzymywanie stabilności genomu, jednakże w przypadku obecności wydłużonych ciągów powtórzeń CAG, które mogą tworzyć struktury drugorzędowe, takie jak pętle czy struktury typu spinki (*hairpin*), MMR może działać nieprawidłowo, przyczyniając się do ekspansji somatycznych, a w konsekwencji patogenezy chorób takich jak HD.⁴⁸ Badania wykazały, że kompleks MutS β wykazuje duże powinowactwo do struktur drugorzędowych tworzonych przez DNA, w przeciwieństwie do MutS α , które ma większą tendencję do wiązania się do niesparowań zasad lub niewielkich insercji i delecji.^{47,49} Po związaniu się MutS β z wytworzoną przez powtórzenia strukturą drugorzędową, zrekrutowany kompleks MutL α nacina nić komplementarną do tej, na której doszło do wypętlenia. Następnie nić zawierająca pętlę zostaje użyta jako matryca przez polimerazę, co skutkuje wydłużeniem ciągu powtórzeń (rysunek 1.3 B).⁴¹



Rysunek 1.3 A) Mechanizm naprawy DNA na drodze MMR. B) Udział MMR w ekspansji powtórzeń CAG (na podstawie Lahue 2020⁴¹).

W wielu badaniach z wykorzystaniem mysich modeli HD, jak również DM1 wykazano wpływ kompleksu MutS β na ekspansję powtórzeń CAG/CTG. Za szczególnie istotne w promowaniu ekspansji somatycznych uznano białko MSH3, będące elementem kompleksu MutS β , a konkretnie jego aktywność ATP-azową.^{46,50–55} Ponadto w dwóch niezależnych badaniach pośmiertnych mózgów ludzi chorych na HD wykazano podwyższoną ekspresję białek MSH2 i MSH3 w MSNs, które posiadały największą liczbę powtórzeń CAG spośród badanych komórek nerwowych, a które to komórki są najbardziej podatne na degenerację w HD.^{56,57}

Obecnie obniżanie ekspresji MSH3 jest w centrum zainteresowania badaczy chcących zahamować rozwój ekspansji somatycznych i rozwijających terapię HD^{58,59} W niedawno opublikowanych pracach, stosując interferencję RNA, wykazano, że obniżenie ekspresji MSH3 skutecznie zahamowało ekspansję powtórzeń CAG w prądkowiu w mysich modelach HD. W niektórych przypadkach zaobserwowano również spadek ilości agregatów zmutowanej huntingtyny.^{59,60} Zastosowanie antysensownego oligonukleotydu (ASO) do wyciszenia ekspresji MSH3 w ludzkich neuronach, wytworzonych z komórek pobranych od pacjenta posiadającego 125 powtórzeń CAG w genie *HTT*, również przyniosło spektakularne rezultaty w postaci zahamowania ekspansji powtórzeń.⁶¹

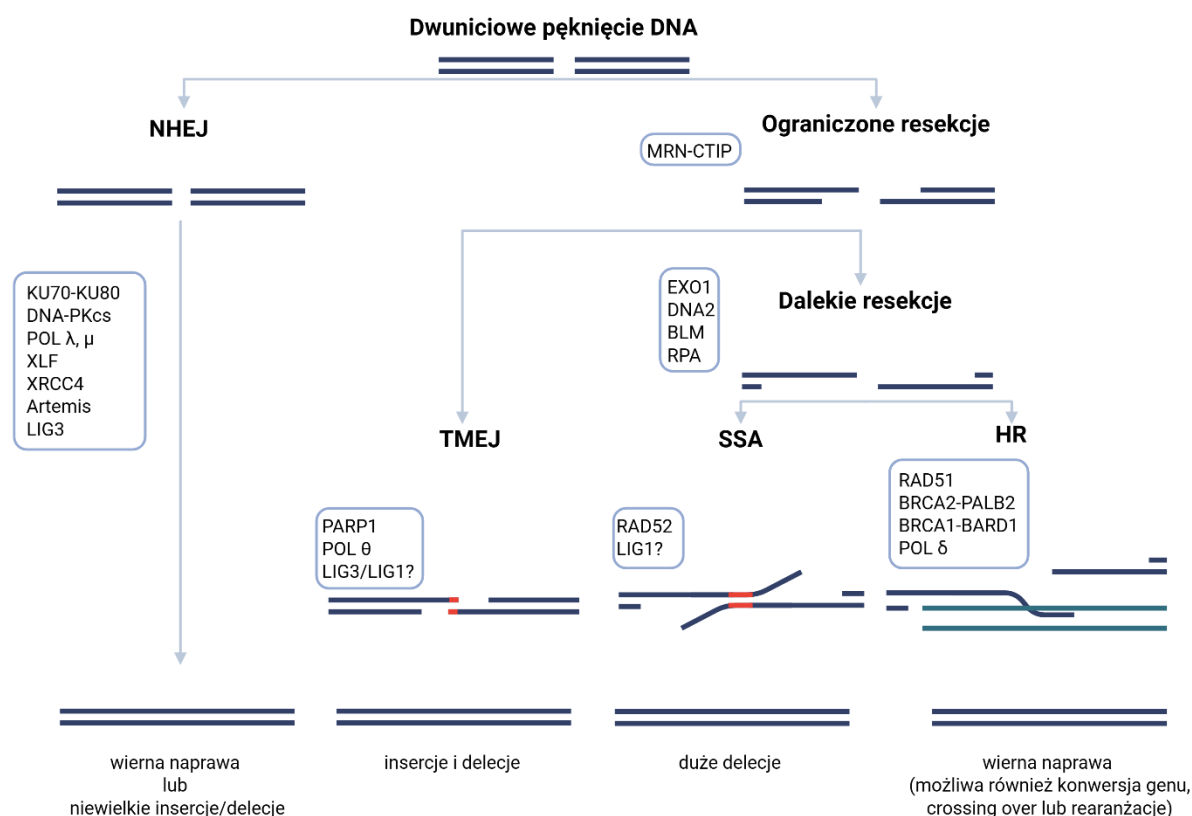
Kolejnym czynnikiem biorącym udział w naprawie DNA, a którego wpływ na rozwój HD został potwierdzony licznymi badaniami, jest nukleaza FAN1. W przeciwieństwie do mechanizmu MMR, skutkiem działania FAN1 jest stabilizacja długości ciągu powtórzeń. Badania *in vitro* wykazały, że FAN1 jest zdolna do wycinania wypętlonych powtórzeń CAG/CTG.⁶² W badaniach na modelach mysich oraz w ludzkich liniach komórkowych wykazano, że FAN1 poprzez oddziaływanie z białkiem MLH1, jednym ze składowych kompleksu MutL α , działa antagonistycznie wobec mechanizmu MMR promującego ekspansję powtórzeń CAG. Stabilizujące działanie FAN1 opiera się na dwóch mechanizmach: po pierwsze poprzez blokowanie rekrutacji MLH1 przez MSH3, co zapobiega tworzeniu funkcjonalnego kompleksu MMR sprzyjającego ekspansjom. Po drugie FAN1 działając jako nukleaza, wycina niewielkie wypętlenia CAG/CTG stabilizując w ten sposób ciąg powtórzeń.^{63–68}

W chwili obecnej nadal poszukiwane są kolejne czynniki mogące wpływać na ekspansje somatyczne powtórzeń CAG, które pozwoliłyby opracować nowe strategie terapeutyczne.^{58,68,69}

1.4 Mechanizmy naprawy dwuniciowych pęknięć DNA

1.4.1 Wprowadzenie

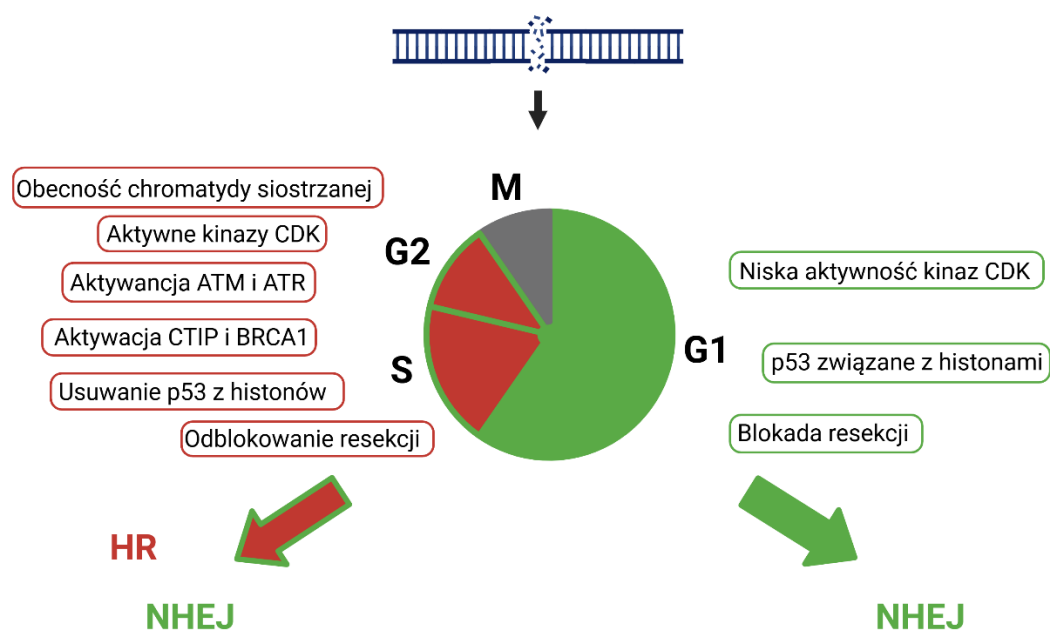
Dwuniciowe pęknięcia DNA (DSBs, ang. *double strand breaks*) stanowią jedno z najbardziej toksycznych zjawisk zachodzących w komórce, mogące skutkować powstaniem mutacji, rearanżacją chromosomów, utratą części materiału genetycznego lub transformacją nowotworową komórki.⁷⁰ DSBs mogą być wywołane czynnikami mutagennymi, takimi jak oddziaływanie na DNA promieniowania jonizującego lub reaktywnych form tlenu. Endogenne dwuniciowe pęknięcia nici DNA pojawiają się również podczas naturalnych procesów zachodzących w komórce, takich jak replikacja DNA, mejoza czy transkrypcja.⁷¹ Ze względu na wysoką toksyczność dwuniciowych pęknięć, odpowiedź komórkowa na tego typu uszkodzenie DNA jest natychmiastowa i angażuje szereg czynników, mających uchronić cały organizm przed negatywnymi skutkami tego zjawiska.⁷² Wyróżnia się dwie główne ścieżki naprawy DSBs w komórce eukariotycznej: łączenie niehomologicznych końców (NHEJ, ang. *non-homologous end joining*), które jest naprawą relatywnie szybką i odbywającą się bez intensywnego procesowania końców DNA oraz rekombinację homologiczną (HR, ang. *homologous recombination*), w której końce DNA ulegają znacznym resekcom. Resekcja jest procesem polegającym na degradacji jednej nici z dwuniciowych końców DNA w celu odsłonięcia fragmentów wykazujących homologię, sprzyjającą łączeniu obydwu końców. Naprawa HR trwa dłużej niż NHEJ ale uznawana jest za naprawę wolną od błędów, w przeciwieństwie do tej drugiej. Poza wymienionymi „kanonicznymi” mechanizmami naprawy DSBs występują jeszcze dwie alternatywne ścieżki, obie wykorzystujące zjawisko resekcji, choć o różnym jej zasięgu i obie związane z wysokim prawdopodobieństwem mutagenyzy. Poszczególne ścieżki naprawy zostaną szczegółowo przedstawione poniżej (rysunek 1.4.1.1).^{72–75}



Rysunek 1.4.1.1 Ścieżki naprawy dwuniciowych pęknięć DNA. W ramach przedstawiono białka uczestniczące w danym procesie, na czerwono zaznaczono regiony homologiczne DNA (na podstawie Chang i wsp. 2017 oraz Scully i wsp. 2019^{73,74}).

Na wybór sposobu naprawy DSBs kluczowy wpływ ma kilka czynników. Jednym z nich jest faza cyklu, w której znajduje się komórka w momencie pojawienia się uszkodzenia. NHEJ uważane jest za domyślny sposób naprawy DSBs w komórce, a wykorzystanie tego mechanizmu możliwe jest w przeciągu całego cyklu jej życia. HR ograniczona jest do faz S i G2, w których materiał genetyczny jest zreplikowany, w związku z czym, dostępna jest, konieczna w tym mechanizmie naprawy, matryca DNA w postaci chromatydy siostrzanej. W fazie G1 białko p53 wiąże się do histonów, w pobliżu DSB, zabezpieczając końce DNA przed resekcją.^{74,76} Ponadto niska aktywność kinaz zależnych od cyklin (CDK) w tej fazie cyklu komórkowego powoduje supresję enzymów odpowiedzialnych za resekcje DNA. W fazie S i G2 gdy aktywność CDK wzrasta, następuje aktywacja białek punktu kontrolnego odpowiedzi na uszkodzenia DNA: ATM i ATR jak również białka stymulującego resekcję CTIP. Aktywacja tych czynników pozwala na usunięcie p53 z histonów w pobliżu końców DNA, umożliwiając tym samym dostęp do nich białkom odpowiedzialnym za resekcje (rysunek

1.4.1.2).^{77–82} Kolejnym czynnikiem wpływającym na wybór mechanizmu naprawy DSB przez komórkę jest rodzaj końców DNA powstałych w wyniku pęknięcia. Końce tępe lub z jedynie niewielkimi wystającymi jednoniciowymi nawisami (ang. *overhang*) są łatwo wiązane przez kluczowe białka NHEJ jakimi są KU70 i KU80, w przeciwieństwie do końców z długimi fragmentami jednoniciowymi. Obecność tych ostatnich zwiększa prawdopodobieństwo skierowania naprawy ku mechanizmom wykorzystującym resekcję DNA.^{74,83} Ważnym aspektem, o którym warto wspomnieć w kontekście wyboru ścieżek naprawy DSBs, jest kwestia stanu chromatyny w obszarze, w którym dochodzi do jej uszkodzenia. Badania wskazują na udział potranslacyjnych modyfikacji histonów i ich pośredniczenie w rekrutacji białek należących do różnych ścieżek naprawy. Postulowana jest nawet koncepcja istnienia kodu histonowego naprawy DNA.^{84–87} Biorąc pod uwagę różne czynniki mogące wpływać na wybór mechanizmu naprawy DSB, warto podkreślić, że fundamentalną kwestią pozostaje wpływ cyklu komórkowego. Białka uczestniczące w mechanizmach naprawy wykorzystujących resekcję w zdrowej komórce w fazie G1 pozostają nieaktywne lub przegrywają konkurencję o dostęp do miejsca pęknięcia z białkami promującymi NHEJ, niezależnie od innych czynników.^{74,82} W kolejnym rozdziale poszczególne ścieżki naprawy dwuniciowych pęknięć DNA zostały przedstawione w sposób bardziej szczegółowy.



Rysunek 1.4.1.2 Zależność wyboru ścieżki naprawy dwuniciowego pęknięcia DNA od fazy cyklu komórkowego (na podstawie Hustedt i Durocher 2017⁸²).

1.4.2 Łączenie niehomologicznych końców (NHEJ)

Naprawa dwuniciowego pęknięcia na drodze NHEJ polega na zbliżeniu oraz ligacji dwóch końców DNA bez lub z minimalną resekcją nici DNA. W przypadku tej ścieżki naprawy, resekcja odbywa się w bardzo niewielkim zakresie i obejmuje do 4 nukleotydów.^{81,88} NHEJ rozpoczyna się od związania przez heterodimery KU70-KU80 obydwu końców DNA uwolnionych po powstaniu dwuniciowego pęknięcia. Zabezpieczają one owe końce przed degradacją. Białka KU-70-KU80 następnie rekrutują katalityczną podjednostkę zależnej od DNA kinazy białkowej (DNA-PKcs), która stabilizuje i utrzymuje w bliskości końce DNA, jednocześnie rekrutując kolejne białka naprawcze- kompleks ligazy IV (LIG4) z XRCC4-XLF, którego zadaniem jest ligacja obydwu końców.^{73,89} W sytuacji, gdy na końcach DNA obecne są wystające jednoniciowe fragmenty 3' lub 5' mogące tworzyć bardziej złożone struktury jak struktura *hairpin*, zaangażowana zostaje nukleaza Artemis, aktywowana przez DNA-PKcs, dysponująca aktywnością 5'→3' egzonukleazową wobec jednoniciowego DNA (ssDNA) oraz (w kompleksie z DNA-PKcs) aktywnością endonukleazową, pozwalającą na usuwanie struktur typu *hairpin* lub *overhang*. Usunięcie struktur drugorzędowych z końców DNA oraz ograniczone resekcje pozwalają na ujawnienie fragmentów mikrohomologicznych, które służą do parowania nici na zasadzie komplementarności par zasad.^{83,88,90,91} Ewentualne brakujące nukleotydy zostają dołączone do końców 3' DNA przez polimerazy z rodziny X (λ i μ). Odtworzenie wiązania fosfodiesterowego następuje dzięki aktywności kompleksu LIG IV-XRCC4-XLF.⁹²⁻⁹⁴ NHEJ uważana jest za naprawę obciążoną ryzykiem wprowadzania mutacji w miejscach łączenia nici DNA. Charakterystycznymi dla tego mechanizmu naprawy zmianami są niewielkie insercje lub delecje (*indels*), długości od jednego do kilku par zasad, mogące prowadzić do przesunięcia ramki odczytu (*frame shift*), jeśli do pęknięcia doszłoby w sekwencji kodującej genu.

1.4.3 Rekombinacja homologiczna (HR)

Drugą z kanonicznych ścieżek naprawy dwuniciowego pęknięcia DNA jest HR. Jak zostało to już wspomniane powyżej, jest to proces, który zachodzi tylko w specyficznych warunkach panujących w komórce, jedynie w obecności zreplikowanego DNA, czyli w fazie cyklu komórkowego S lub G2. Chromatyda siostrzana stanowi w tym mechanizmie matrycę do naprawy uszkodzonej nici DNA, w efekcie czego, jest to naprawa wierna.⁷⁴ Po rozpoznaniu

dwuniciowego pęknięcia DNA przez kompleks MRN, na który składają się białka MRE11, RAD50 oraz NBS1, dochodzi do fosforylacji kinazy ATM, która w ten sposób ulega aktywacji i poprzez fosforylację pobliskich histonów i innych białek zaangażowanych w kontrolę cyklu komórkowego i naprawę DNA uruchamia cały proces.⁹⁵ Z kompleksem MRN współpracuje kolejne białko- CTIP, które stymuluje aktywność endonukleazową MRE11.⁹⁶ Kompleks MRN-CTIP inicjuje pęknięcie nici DNA w pewnym oddaleniu od dwuniciowego pęknięcia, wynoszącym do około 300 par zasad. Od tego miejsca MRN-CTIP dokonuje degradacji nici DNA w kierunku 3'-5' pozostawiając długie jednoniciowe fragmenty DNA z wolnymi 3' końcami. Białka KU70-KU80, które mogą w tym momencie okupować nic DNA są usuwane przez kompleks MRN-CTIP.^{74,97} Resekcje nici 5' są następnie kontynuowane przez egzonukleazę 1 (EXO1) degradującą nic 5' oraz helikazę zespołu Blooma (BLM) i endonukleazę DNA 2, które stymulowane przez CTIP odpowiadają za oddzielenie jednej nici DNA od drugiej, tworząc substrat dla EXO1 w postaci ssDNA z 5'-końcem.⁹⁸⁻¹⁰⁰ Odsłonięte w wyniku resekcji długie jednoniciowe fragmenty DNA (mogące obejmować tysiące nukleotydów) są pokrywane kompleksami białka RPA na które składają się: RPA 1, 2 i 3, chroniącymi nic ssDNA przed jej degradacją lub tworzeniem nieprawidłowych struktur drugorzędowych.¹⁰¹ Następnie kompleksy RPA z pomocą BRCA2 i PALB2 są zastępowane przez białka RAD51, które tworzą długie spiralne filamenty na jednoniciowym DNA. Jedna z takich nici z udziałem BRCA1-BARD1 poszukuje homologii między końcem 3' jednoniciowego DNA a dwuniciowym DNA (dsDNA) chromatydy siostrzanej. Następnie koniec 3' dokonuje inwazji do dsDNA, tworzy pętlę D oraz służy jako starter do syntezy DNA przez polimerazę δ .¹⁰²⁻¹⁰⁸

Naprawa DNA na drodze HR może zostać zakończona na kilka różnych sposobów. W przypadku gdy synteza nowej nici na matrycy drugiego końca DNA zostanie przerwana na wczesnym etapie, wydłużana nic 3' która współtworzyła pętlę D, paruje się z nicią drugiego końca DNA powstałego po pęknięciu, który do tej pory pozostawał bierny, a synteza dokańczana jest już na wyjściowym chromosomie. Nie dochodzi tu do konwersji genu, a ścieżka ta nosi nazwę parowania nici zależnego od syntezy (SDSA, ang. *synthesis-dependent strand annealing*). Alternatywną ścieżką jest klasyczny HR zachodzący z wytworzeniem podwójnej struktury Hollidaya (dHJ), co może prowadzić do wymiany fragmentów materiału genetycznego między chromatydami siostrzanymi, czyli *crossing-over* lub jeśli struktura dHJ zostanie przerwana, może zakończyć się bez *crossing-over*. Trzecią opcją jest długodystansowa konwersja genu (LTGC, ang. *long-tract gene conversion*) zachodząca w przypadku

nieprzerwanej syntezy DNA na matrycy chromatydy siostrzanej, która może zostać zakończona naprawą indukowaną pęknięciem (BIR, ang. *break induced replication*). Skutkuje to asymetryczną wymianą informacji genetycznej między chromatydami, może również prowadzić do rearanżacji chromosomowych.^{70,74,109}

1.4.4 Alternatywne ścieżki naprawy dwuniciowych pęknięć DNA

Istnieją alternatywne dla dwóch powyżej opisanych ścieżki naprawy DSBs, obie wykorzystujące zjawisko resekcji, choć o różnym jej zasięgu. Ścieżka o nazwie MMEJ (ang. *microhomology-mediated end joining*- łączenie końców za pośrednictwem mikrohomologii) inaczej nazywana alt-EJ (ang. *alternative end joining*- alternatywne łączenie końców) wymaga resekcji o niewielkim zasięgu, w których uczestniczy kompleks MRN-CTIP. Z kolei ścieżka SSA (ang. *single strand annealing*- łączenie pojedynczej nici) wymaga resekcji o dużym zasięgu sięgających nawet kilku tysięcy nukleotydów, w których uczestniczą EXO1, DNA2 i BLM. Obie ścieżki wykorzystują zjawisko łączenia się nici DNA w miejscach występowania homologii, którego skutkiem są delecje fragmentów DNA zawartych pomiędzy obszarami homologicznymi.^{73–75,110}

W specyficznym rodzaju mechanizmu MMEJ zwanym TMEJ (ang. *DNA polymerase theta-mediated end joining*) jednoniciowe 3' końce DNA łączą się w obrębie regionów mikrohomologicznych, które mogą obejmować od 2 do 20 nukleotydów.¹¹¹ Chociaż TMEJ jest poznany słabiej w porównaniu do poprzednio opisywanych ścieżek naprawy, to zidentyfikowano kilka białek uczestniczących w tym mechanizmie, jak i kilka czynników działających przeciwnie, inhibujących TMEJ i promujących inne mechanizmy.^{75,112} Co do znaczenia niektórych czynników istnieją sprzeczne doniesienia. Jest tak w przypadku polimerazy PARP1 opisywanej z jednej strony jako białko stabilizujące dwa końce DNA i umożliwiające ich łączenie, konkurujące z białkami KU70 i KU80 o wiązanie do końców DNA i sprzyjające naprawie z udziałem polimerazy theta (POLθ)^{113,114}. W innej pracy z kolei PARP1 jest opisywana jako czynnik odpowiadający za rekrutację białek KU w miejsce pęknięcia DNA i hamującego resekcyjne¹¹⁵. Z pewnością jednak PARP1 jest jednym z pierwszych białek, uczestniczących w sygnalizacji uszkodzenia DNA i rekrutujących inne białka naprawcze w miejsce pęknięcia. Dokonuje ono tzw. PARylacji, czyli przyłączenia zarówno do siebie, jak i innych białek cząsteczek poli ADP-rybozy (PAR) tworzących szkielet, do którego mogą wiązać się białka naprawcze (zawierające domenę wiążącą PAR) jednocześnie ulegające

wówczas aktywacji. PARylacji ulegają również histony znajdujące się w pobliżu dwuniciowego pęknięcia. Skutkiem tej modyfikacji dochodzi do obszarowego rozluźnienia chromatyny, która staje się bardziej podatna na działanie czynników naprawczych.^{116–118}

Kluczowym białkiem w TMEJ jest POLθ, która odpowiada za łączenie jednoniciowych fragmentów DNA w miejscach występowania mikrohomologii, jednocześnie usuwając z nici białka RPA.^{119–122} Po połączeniu dwóch końców DNA w regionach mikrohomologicznych POLθ może dokonać degradacji niewielkich fragmentów z 3' końca DNA i rozpocząć syntezę DNA na matrycy drugiej nici wydłużając końce 3'.¹²³ Ewentualne powstałe w wyniku połączenia jednoniciowych 3'końców DNA fragmenty (znajdujące się pomiędzy regionami homologicznymi) tworzące struktury typu flap są usuwane przez XPF-ERCC1 lub FEN1^{124–127}. W następnym kroku ligaza DNA III (LIG3) lub ligaza DNA I (LIG1) odtwarzają wiązanie fosfodiesterowe, łącząc obie nici DNA.¹²⁸

Naprawa TMEJ charakteryzuje się niską wiernością i tendencją do wprowadzania mutacji w miejscu naprawy. Mutacje te wynikają zarówno z faktu usuwania fragmentów DNA znajdujących się między regionami mikrohomologicznymi, co skutkuje delecjami, jak również mogą być związane z aktywnością samej POLθ, mającej tendencję do wprowadzania insercji. Charakterystycznymi rezultatami naprawy TMEJ są małe insercje powstałe przy wykorzystaniu pobliskiego DNA jako matrycy (TINS ang. *templated insertions*). Mogą być one efektem niskiej procesywności POLθ, która przerywa elongację nici DNA i oddysocjowuje od nici matrycowej, aby za chwilę znów przyłączyć się do niej w innym miejscu i rozpocząć kolejną rundę wydłużania. POLθ może również używać tej samej nici zarówno jako matrycy, jak i jako nici wydłużanej, wprowadzając tym samym insercje w orientacji odwrotnej (*reverse complement* TINS)^{121,122,129}. Co więcej, POLθ wykazuje również aktywność terminalnej transferazy, a więc ma zdolność dobudowywania nukleotydów do końca 3' niezależnie od matrycy DNA, aby uzyskać homologię pomiędzy dwiema nićmi, które mają zostać połączone, co również może być przyczyną insercji.¹³⁰ Ponadto TMEJ uważane jest za jeden z procesów prowadzących do translokacji oraz innych rearanżacji chromosomowych.^{103,131,132} W tym miejscu warto również nadmienić, że POLθ jest polimerazą nieobecną u drożdży, które niezwykle często wykorzystywane są jako model do badania mechanizmów naprawy DNA.

Kolejną, alternatywną ścieżką, którą może podążać naprawa dwuniciowego pęknięcia DNA, jest wspomniane już SSA. Ścieżka ta, w przeciwieństwie do TMEJ, wymaga resekcji nici 5' sięgających nawet ponad 1000 nukleotydów, powstałych przy współudziale EXO1, DNA2

i BLM. Długie jednoniciowe fragmenty DNA z wolnymi 3' końcami zostają pokryte białkami RPA, które następnie ulegają wymianie na białka RAD52, te zaś, odpowiadają za poszukiwanie regionów homologicznych znacznie dłuższych niż w przypadku TMEJ (>20 nt) i za parowanie nici DNA w tych obszarach.^{75,133–135} Podobnie jak w przypadku TMEJ niesparowane 3' końce obejmujące obszary między regionami homologicznymi są usuwane przez XPF-ERCC1.¹³⁶ Udział polimeraz i ligaz, które uzupełniałyby brakujące nukleotydy i ostatecznie połączyły obie nici, nie jest do końca jasny.^{74,75,135}

Naprawa dwuniciowych pęknięć DNA na drodze SSA jest wysoce mutagenna. Jednym z możliwych niekorzystnych efektów jest utrata części materiału genetycznego- fragmentu DNA znajdującego się pomiędzy sekwencjami homologicznymi, które uległy połączeniu. Kolejnym, jest możliwość parowania jednoniciowego fragmentu DNA pokrytego RAD52 z regionami homologicznymi znajdującymi się na innych chromosomach, czego skutkiem będą rearanżacje genomowe.^{137,138}

1.5 Skracanie ciągów powtórzeń CAG w efekcie naprawy DNA

1.5.1 Narzędzia inżynierii genetycznej do indukcji dwuniciowych pęknięć DNA

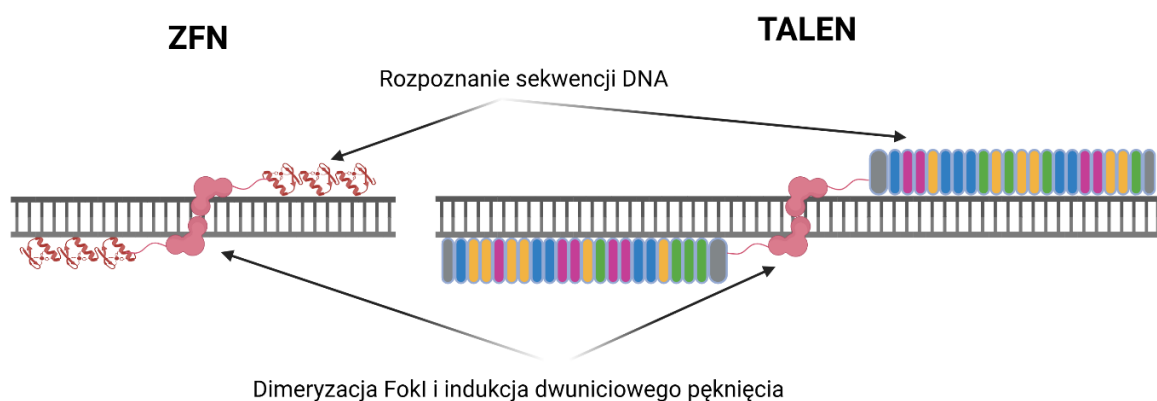
Rozwój biologii molekularnej i wytworzenie nowych narzędzi inżynierii genetycznej pozwoliły na poszukiwanie nowych rozwiązań walki z chorobami genetycznymi. Wśród tych metod pojawiły się również takie, które umożliwiają indukowanie dwuniciowych pęknięć w konkretnym miejscu genomu. Dzięki tym kierowanym nukleazom możliwe jest wprowadzanie zmian w sekwencjach odpowiedzialnych za wywoływanie chorób, co zostało już wykorzystane w leczeniu anemii sierpowatej i talasemii beta.¹³⁹ W przypadku chorób wywołanych ekspansją powtórzeń mikrosatelitarnych nadzieję na przyszłą terapię genową mogą budzić obserwacje skracania ciągów powtórzeń CAG w wyniku indukcji w ich obrębie dwuniciowych pęknięć DNA.¹⁴⁰ Poniżej przedstawiono dotychczas wykorzystywane w badaniach nad skracaniem ciągów powtórzeń (głównie CAG) narzędzia inżynierii genetycznej.

1.5.1.1 Nukleazy palca cynkowego (ZFN)

Jednym z narzędzi inżynierii genetycznej, służących do indukcji DSBs są nukleazy palca cynkowego (ZFNs, ang. *zinc finger nucleases*). Na ZFN składają się dwa elementy: palec cynkowy, w liczbie kilku, ustawione szeregowo jeden obok drugiego oraz nukleaza FokI. Palec cynkowy składa się z około 30 aminokwasów tworzących strukturę β - β - α , która wiąże jeden jon cynku. Fragmenty palca cynkowego wchodzą w bezpośrednią interakcję z zasadami azotowymi w łańcuchu DNA, a każdy taki palec zdolny jest do wiązania 3 par zasad. O tym jakie to pary zasad, decyduje sekwencja aminokwasów w oddziałującym z DNA obszarze palca. Połączenie kilku domen palca cynkowego pozwala rozpoznawać dłuższe sekwencje (będące wielokrotnością 3 nukleotydów) a poprzez odpowiednią kombinację różnych palców cynkowych, można zaprojektować ZFN rozpoznający niemal dowolną sekwencję DNA. Drugi element ZFN, nukleaza FokI, jest zdolna do cięcia DNA jedynie, gdy znajduje się w formie dimeru. Dlatego stosuje się dwie cząsteczki ZFN, każda złożona z kilku palców cynkowych i jednej domeny FokI. Po rozpoznaniu odpowiedniej sekwencji DNA przez palce cynkowe monomery FokI znajdują się blisko siebie, ulegają dimeryzacji i hydrolizują DNA w oczekiwanym miejscu.¹⁴¹

1.5.1.2 System TALEN

Drugi z popularnych systemów edycji genomu, TALEN (ang. *transcription activator-like effector nuclease*), również wykorzystuje nukleazę FokI oraz fakt konieczności jej dimeryzacji dla uaktywnienia jej funkcji nukleazowej (rysunek 1.5.1.2). Za rozpoznanie specyficznej sekwencji DNA odpowiada w tym przypadku białko TALE składające się z centralnej domeny rozpoznającej sekwencję nukleotydów i wiążącej DNA oraz domeny aktywującej transkrypcję genu. Domena wiążąca DNA składa się z monomerów, z których każdy wiąże jeden nukleotyd w sekwencji docelowej. Monomery zbudowane są z tandemowo powtórzonych 34 reszt aminokwasowych, z których 12 i 13 są wysoce zmienne (RVD- *repeat variable diresidue*) i odpowiadają za specyficzność rozpoznawania nukleotydów. Cięcie DNA jak w przypadku ZFN następuje po rozpoznaniu docelowych sekwencji przez dwa elementy TALEN i dimeryzacji FokI.¹⁴²



Rysunek 1.5.1.2 Schemat działania systemów ZFN i TALEN.

1.5.1.3 System CRISPR-Cas9

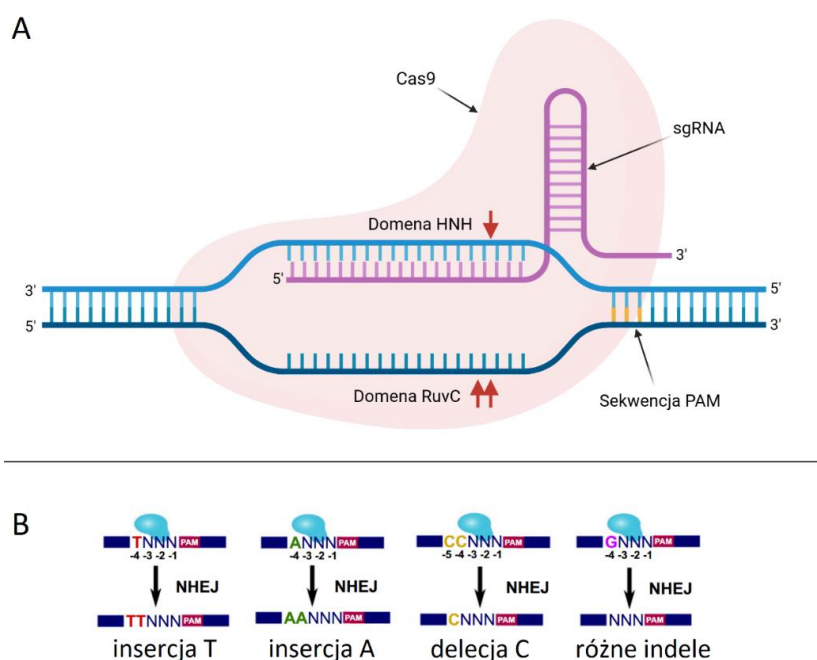
Najnowszym narzędziem inżynierii genetycznej, za pomocą którego dokonywać można indukcji dwuniciowych pęknięć DNA w konkretnym, zaplanowanym miejscu, jest system CRISPR-Cas9 (ang. *Clustered regularly interspaced short palindromic repeats- CRISPR associated 9*). Wywodzi się on z systemu odpornościowego bakterii, w którym obcy materiał genetyczny jest cięty przez nukleazę Cas9 kierowaną krótkim fragmentem CRISPR-RNA (crRNA). Kluczowym dla aktywności nukleazy jest obecność krótkiej sekwencji PAM (ang. *protospacer adjacent motif*) w sekwencji kwasu nukleinowego, który ma być cięty. Wśród bakterii występuje duża różnorodność elementów systemu CRISPR-Cas. Wyróżnia się dwie klasy systemu. Do klasy pierwszej należą systemy, w których za degradację obcego materiału genetycznego odpowiedzialny jest kompleks białek, do klasy drugiej zaś takie, w których wykorzystywana jest pojedyncza nukleaza. Drugim podziałem, jest podział na typy, których jest sześć. Poszczególne typy różnią się między sobą ze względu na obróbkę crRNA, rozpoznawanie sekwencji docelowej, czy rodzaj kwasu nukleinowego, który ma być cięty przez nukleazę (RNA lub DNA).¹⁴³ Do najczęściej wykorzystywanych w inżynierii genetycznej systemów należy system CRISPR-Cas9, przedstawiciel typu II i klasy drugiej. Najpowszechniej stosowane Cas9 pochodzące ze *Streptococcus pyogenes* (SpCas9), posłuży jako przykład do opisu systemu.^{144,145}

Białko SpCas9 składa się z 1368 aminokwasów i jest białkiem wielodomenowym. Zawiera m.in. trzy domeny helikazowe oraz dwie domeny nukleazowe. Do specyficznego rozpoznawania określonej sekwencji DNA wykorzystuje ono dwie cząsteczki RNA: crRNA

oraz tracr-RNA (ang. *trans-activating CRISPR RNA*) razem tworzące cząsteczkę gRNA (ang. *guide RNA*). Cas9 dopiero w następstwie związania gRNA przechodzi zmianę konformacyjną, dzięki której zyskuje zdolność do skanowania dwuniciowego DNA w poszukiwaniu sekwencji PAM. W przypadku SpCas9 PAM składa się z trzech nukleotydów ułożonych w kolejności: 5'-NGG-3', gdzie N oznacza dowolny nukleotyd. Rozpoznanie PAM skutkuje kolejną zmianą konformacyjną Cas9, uruchamiającą aktywność nukleazową kompleksu. W następstwie rozpoznania PAM dochodzi do separacji nici DNA i poszukiwania komplementarności pomiędzy sekwencją rozpoznającą miejsce docelowe (ang. *target recognition site*) crRNA a DNA.^{146–151} Występowanie komplementarności par zasad na odcinku zaledwie 5 nukleotydów jest wystarczające do czasowego związania kompleksu Cas9-gRNA z danym miejscem w DNA, lecz nie jest wystarczające do indukcji dwuniciowego pęknięcia.¹⁵² Dopiero w sytuacji, gdy parowanie nici obejmie co najmniej 17 nukleotydów, domena HNH tnie nie zawierającą sekwencję komplementarną do sekwencji crRNA. W wyniku czego, aktywacji ulega domena RuvC, która następnie tnie drugą nić DNA.^{153,154}

Sekwencja około 10-13 nukleotydów od strony 5' sekwencji PAM, w przypadku których dochodzi do parowania między DNA a crRNA, jest kluczowa dla efektywnego cięcia. Obecność w tym regionie niesparowań łatwo może doprowadzić do przerwania dalszego parowania nici i dysocjacji kompleksu Cas9-gRNA od DNA. Niesparowania w dalszym regionie są lepiej tolerowane i mogą skutkować indukcją dwuniciowego pęknięcia, nawet pomimo niepełnej komplementarności między DNA a fragmentem crRNA. Takie cięcie określa się mianem *off-target*.^{155–160} Dość powszechny jest pogląd, że Cas9 tnie, pozostawiając tzw. tępe końce DNA, za trzecim nukleotydem licząc od 5' końca PAM, czyli w pozycji -3. Wiele prac jednak wskazuje na to, że domena RuvC może ciąć nić DNA również w pozycji -4 od PAM tworząc w ten sposób 1 nukleotydowy *overhang* na jednej nici (non-target).^{161–164} Może to wyjaśniać obecność niewielkich insercji w miejscu naprawy DNA często obserwowanych w przypadku edycji z użyciem Cas9. Badania wykazały, niebagatelny wpływ sekwencji DNA znajdującej się w pobliżu miejsca pęknięcia, a szczególnie nukleotydu w pozycji -4 na rezultat naprawy takiej sekwencji. Obecność tyminy lub adeniny w pozycji -4 najczęściej skutkuje insercją odpowiednio tyminy lub adeniny w wyniku naprawy, podczas gdy obecność cytozyny w tej pozycji częściej generuje jej delecję. Obecność guaniny zaś nie generuje specyficznych rezultatów, w efekcie obserwuje się różnego rodzaju delecje lub insercje.^{165–170} Schemat działania kompleksu Cas9-gRNA oraz zależność wyniku naprawy DSB od nukleotydu w pozycji -4 przedstawiono na rysunku 1.5.1.3.

W kontekście systemu CRISPR-Cas9 należy wspomnieć o pewnych modyfikacjach, poszerzających możliwości tego narzędzia inżynierii genetycznej. Dość wcześnie po poznaniu systemu CRISPR-Cas9 zastosowano zamiast dwóch cząsteczek crRNA i tracrRNA pojedynczą cząsteczkę sgRNA (ang. *single-guide RNA*), która umożliwiła ekspresję systemu w komórce, bez konieczności wcześniejszej hybrydyzacji dwóch cząsteczek RNA.¹⁴⁵ Wiele prac poświęcono modyfikacjom zarówno długości, jak i struktury gRNA w celu poprawy specyficzności systemu, lecz często jej poprawa skutkowała obniżeniem efektywności edycji.^{170–176} Kolejnym elementem systemu, który podlegał modyfikacjom jest białko Cas9. Dzięki mutacjom w odpowiednich lokalizacjach w sekwencji kodującej białko uzyskano nukleazy o zmienionych właściwościach. Jednym z najważniejszych jest często wykorzystywane laboratoryjnie białko nikaza Cas9 (nCas9), występujące w dwóch wariantach. W każdym wariancie jedna domena nukleazowa jest inaktywowana, w związku z czym nCas9 generuje cięcie tylko jednej nici dwuniciowego DNA.¹⁷⁷ Kolejną modyfikacją jest dCas9 (ang. *deactivated Cas9*), w którym obie domeny nukleazowe są nieaktywne, zachowana jednak jest zdolność do specyficznego rozpoznawania sekwencji DNA. dCas9 wykorzystywane jest do różnych celów m.in. w połączeniu z czynnikami transkrypcyjnymi do wpływania na ekspresję genów.¹⁷⁸ Podobnie jak w przypadku gRNA modyfikacje Cas9 miały głównie na celu uzyskać nukleazę bardziej specyficzną. W niektórych pracach starano się poszerzyć zakres rozpoznawanych przez nukleazę sekwencji PAM.^{170,179–181}



Rysunek 1.5.1.3 A) Schemat działania kompleksu Cas9-sgRNA. Czerwonymi strzałkami zaznaczono potencjalne miejsce cięcia dwuniciowego DNA. B) Zależność wyniku naprawy od nukleotydu w pozycji -4 (na podstawie Sledzinski i wsp. 2024¹⁸²).

1.5.2 Wykorzystanie systemu CRISPR-Cas9 do skracania ciągów powtórzeń CAG

Obserwowana w komórkach drożdżowych naturalna skłonność długich ciągów powtórzeń trójnukleotydowych do ich pękania, skierowała uwagę badaczy na wpływ, jaki mogą mieć DSBs na niestabilność powtórzeń.¹⁸³ W związku z tym w wielu pracach koncentrowano się badaniu niestabilności ciągów CAG, wykorzystując do indukcji DSBs, narzędzia inżynierii genetycznej, opisywane w poprzednim rozdziale. Skracanie ciągu, bądź jego całkowite wycięcie, zaobserwowano w komórkach drożdży, modelach mysich, jak również w komórkach ludzkich. W jednym z pierwszych badań indukowano DSBs w obrębie powtórzeń CAG w systemach reporterowych, w drożdżach analizując wybrane ścieżki naprawy DNA.^{184,185} Wykazano w nich, że dwuniciowe pęknięcia w obrębie powtórzeń CAG prowadzą do ich niestabilności, skutkując wydłużeniami lub skróceniami ciągu. Zwrócono uwagę na obecność resekcji oraz udział białek takich jak Mre11 czy Rad50 w naprawie tego typu uszkodzeń DNA. W kolejnej pracy, w której wykorzystano model drożdżowy i system TALEN zaobserwowano bardzo wydajne skracanie ciągu powtórzeń CAG w wyniku indukcji DSB w jego obrębie.¹⁸⁶ W kolejnych dwóch pracach, w których szczególny nacisk położono na badanie mechanizmów naprawy, prowadzących do skracania, bądź wycinania całego ciągu CTG/ CAG w drożdżach, wykorzystano zarówno system TALEN, jak i CRISPR-Cas9.^{187,188} W sekwencję genu *SUP4* wprowadzono powtórzenia CTG/CAG pochodzące od pacjenta chorego na DM1, a następnie indukowano w ich obrębie DSB. W pracach tych zaobserwowano dalekie resekcje DNA zachodzące w czasie naprawy, sięgające od kilkuset do kilku tysięcy nukleotydów. Stwierdzono, że do efektywnych resekcji końców DNA wymagane jest białko Sae2 będące homologiem ludzkiego białka CTIP, w modelu ludzkim również stymulującego resekcje. Stwierdzono ponadto udział w naprawie Rad50 elementu kompleksu MRX, odpowiednika ludzkiego kompleksu MRN, także zaangażowanego w resekcje końców DNA.¹⁸⁹ Kolejnym białkiem, którego udział był konieczny w naprawie dwuniciowego pęknięcia w powtórzeniach CTG/CAG jest Rad52. Białko to, kluczowe w mechanizmie naprawy SSA, działając na jednoniciowych odcinkach DNA, odsłoniętych w wyniku resekcji, odpowiedzialne jest za łączenie końców DNA w miejscach występowania między nimi homologii. Wykazano również pewne różnice w rezultatach naprawy DNA, zależne od systemu użytego do indukcji DSBs. W wyniku użycia systemu TALEN dochodziło głównie do skróceń powtórzeń CTG/CAG,

podczas gdy zastosowanie Cas9 skutkowało rozległymi delecjami wokół miejsca indukowanego pęknięcia. Ponadto dla naprawy DNA, w której do indukcji pęknięcia użyto Cas9, w niektórych przypadkach niezbędna okazała się ligaza IV (DNL4), co wskazuje na udział NHEJ w naprawie. W kolejnym badaniu wykorzystującym model drożdżowy badano efektywność różnych nukleaz CRISPR-Cas (SpCas9, SaCas9, Cas12a, nCas9) w skracaniu powtórzeń trójnukleotydowych, nie koncentrowano się jednak na mechanizmach naprawy DNA.¹⁹⁰

Jedną z pierwszych prac, w której autorzy analizowali efekty indukcji DSBs w obrębie powtórzeń CAG w komórkach ssaczych, jest praca Mittelmana i wsp. z 2009.¹⁹¹ W pracy tej wykorzystując system ZFN, indukowano DSB w ciągu powtórzeń CAG umieszczonym w plazmidzie w komórkach ludzkich i chemicznych. Zaobserwowano, że dłuższe ciągi powtórzeń CAG mają większą tendencję do skracania się podczas naprawy DNA w porównaniu do ciągów krótszych. Ponadto badając udział w naprawie białka RAD51- jednego z kluczowych białek ścieżki HR, stwierdzono spadek liczby skróceń ciągu w przypadku ekspresji zmutowanego i niefunkcjonalnego RAD51 w komórkach chemicznych. W kolejnej pracy, wykorzystującej system reporterowy z GFP w ludzkich komórkach, stwierdzono skracanie ciągu powtórzeń CAG po indukcji jednoniciowego pęknięcia DNA za pomocą nCas9.¹⁹²

Warto podkreślić, iż w wymienionych wyżej pracach, zarówno dotyczących systemu drożdżowego, jaki i modelu ludzkiego, nie uwzględniono endogennego kontekstu, jako że stosowano systemy reporterowe, w których ciągi CAG/CTG były sztucznie wprowadzone w sekwencje różnych genów lub umieszczone w plazmidach. Skracania ciągów powtórzeń trójnukleotydowych w endogennym *locus* badano w pracy van Agtmaal i wsp. z 2017.¹⁹³ W tej pracy zaobserwowano zarówno skracanie, jak i wydłużanie ciągu 2600 powtórzeń CTG w genie *DMPK* po indukcji DSBs przy użyciu systemu CRISPR-Cas9 w komórkach ludzkich oraz modelu mysim. Co ciekawe indukcja DSBs w regionach flankujących powtórzenia CTG skutkowała dużymi delecjami obejmującymi cały region z powtórzeniami. W kolejnej pracy zastosowano potencjalnie bezpieczniejszy pod względem użycia w terapii system nCas9 z dwiema różnymi cząsteczkami sgRNA, zaprojektowanymi w taki sposób by wprowadzone pęknięcia flankowały ciąg powtórzeń CAG w genie *HTT*.¹⁹⁴ Fragment pomiędzy jednoniciowymi pęknięciami ulegał wycięciu z wysoką efektywnością i precyzją. Jednocześnie potwierdzono wysoką specyficzność systemu, badając poziom edycji w *loci*, które mogły stanowić potencjalne miejsca off-target dla użytych sgRNA. W najnowszych pracach do

indukowania skrótów ciągu CAG/CTG również wykorzystano nCas9. W jednej z nich badacze, wykorzystując ludzki model komórkowy, potwierdzili wycinanie ciągu powtórzeń CAG/CTG w genie reporterowym z użyciem nCas9¹⁹⁵ Wykazali, że jest to proces zależny od transkrypcji, a nie od replikacji, co oznacza, że może zachodzić w komórkach dzielących się. W kolejnej pracy stwierdzono skrócenie powtórzeń CAG/CTG w uzyskanych od pacjentów z HD oraz DM1 neuronach i astrocytach, jak również w modelu mysim po dostarczeniu do prądkowia nCas9 i sgRNA za pomocą wektora AAV (ang. *adeno-associate virus*).¹⁹⁶ W kolejnym badaniu, o którym warto wspomnieć, badacze z udziałem zmodyfikowanego SpCas9 (SpCas9-NG) rozpoznającego inny PAM (NGN zamiast NGG) doprowadzili do skrótów powtórzeń CAG w genie *HTT* w embrionalnych komórkach mysich. Następnie z edytowanych komórek, wyprowadzono neurony, charakteryzujące się prawidłową morfologią i funkcjonalnością.¹⁹⁷ Wyniki tych prac stanowią ważny krok w rozwoju terapii chorób poliQ opartej na skracaniu ciągów powtórzeń.

Pomimo dużej liczby przedstawionych powyżej prac, koncentrujących się na zagadnieniu skracania ciągów powtórzeń CAG, jedynie w części z nich badacze skupiali się na identyfikacji zaangażowanych w ten proces mechanizmów naprawy DNA. Z kolei w innych przedstawionych badaniach, nie analizowano udziału poszczególnych białek, lecz indukowano pęknięcia DNA w endogennie występujących powtórzeniach, co stanowi dużą wartość poznawczą. Istnieją również prace, w których nie dokonywano uszkodzeń DNA, koncentrując się na samoistnie powstających uszkodzeniach lub uszkodzeniach powstałych w wyniku działania czynników chemicznych (zatem w warunkach niepozwalających na kontrolę miejsca indukcji pęknięcia DNA). W pracach tych koncentrowano się na niestabilności powtórzeń i analizowano czynniki zaangażowane w ten proces, często stosując wyciszenie ekspresji poszczególnych genów.^{51,198–208} Znane są również prace, w których analizowano jedynie eksperymenty *in vitro* z udziałem ciągów CAG oraz wybranych białek, potwierdzając udział niektórych z nich w naprawie DSBs indukowanych w obrębie ciągów.^{209–211}

Pośród tej dużej liczby prac, przedstawionej powyżej, brak jest takich, które koncentrowałyby się na badaniu mechanizmów naprawy dwuniciowych pęknięć DNA, indukowanych w powtórzeniach CAG w endogennym *locus* w ludzkim modelu komórkowym. Niniejsza praca koncentruje się właśnie na tym brakującym aspekcie.

2. Cel pracy

Ekspansje powtórzeń mikrosatelitarnych są przyczyną wielu nieuleczalnych chorób człowieka. Jednym z takich schorzeń, jest choroba Huntingtona, za którą odpowiada ekspansja powtórzeń CAG w eksonie 1 genu *HTT*. Nowatorską strategią terapii dedykowanej tej chorobie jest kontrolowane skracanie ciągu powtórzeń, za pomocą narzędzi inżynierii genetycznej. Indukcja pęknięć DNA w obrębie ciągów powtórzeń CAG prowadzi do ich niestabilności, zarówno wydłużania, jak i skracania ciągu, lecz mechanizmy stojące za obserwowaną niestabilnością nie są do końca poznane. Wcześniejsze badania prowadzone na modelach drożdżowych, ale także w komórkach ludzkich, wykorzystujące systemy reporterowe, sugerują zaangażowanie w ten proces niektórych ścieżek naprawy DNA. Brak jest jednak badań dotyczących mechanizmów naprawy DSBs w obrębie powtórzeń CAG w endogennym *locus* w ludzkim modelu komórkowym.

W związku z tym, celem ogólnym tej pracy było dokładniejsze poznanie mechanizmów naprawy dwuniciowych pęknięć DNA w regionach powtórzeń CAG w endogennym *locus* w komórkach ludzkich. Jako model wybrano zmodyfikowane komórki HEK293T z ekspansją powtórzeń CAG w *locus HTT*.

Powyższy cel został osiągnięty przez realizację następujących zadań szczegółowych:

- 1) scharakteryzowanie produktów naprawy dwuniciowych pęknięć DNA w regionie powtórzeń CAG w genie *HTT* w komórkach ludzkich,
- 2) określenie udziału i zasięgu resekcji końców DNA w czasie naprawy dwuniciowych pęknięć DNA w regionach powtórzeń CAG w komórkach ludzkich,
- 3) określenie wpływu miejsca indukcji dwuniciowego pęknięcia DNA względem ciągu powtórzeń CAG oraz sekwencji flankujących ciąg na przebieg i rezultat naprawy,
- 4) identyfikację białek zaangażowanych w skracanie ciągu powtórzeń CAG.

3. Materiały

3.1 Linie komórkowe

W pracy wykorzystywano dwie linie komórkowe:

- 1) linia HEK293T posiadająca 41 powtórzeń CAG na obu allelech *HTT* (HEK293T_41CAG).²¹²
- 2) fibroblasty GM04208 pozyskane z Coriell Cell Repositories (Camden, NJ, USA) posiadające 21 oraz 44 powtórzenia CAG w genie *HTT*.

3.2 Odczynniki do hodowli i transfekcji komórek

Tabela 3.2. Odczynniki używane w pracy do hodowli i transfekcji komórek.

Odczynnik	Producent
Dulbecco's Modified Eagle Medium (DMEM)	Gibco
Minimum Essential Media (MEM)	Gibco
Opti-MEM	Gibco
Surowica krwi bydlęcej (FBS)	Eurx
DPBS	Gibco
Trypsyna 0,5% (10X)	Gibco
L-glutamina 200mM (100X)	Gibco
Antybiotyk-antymykotyk (100X)	Gibco
Polyethylenimine, Linear, MW 25000, Transfection Grade (PEI) 1 mg/ml	Polysciences
Lipofectamine 2000	Thermo Fisher
Trypan Blue Solution 0,4%	Gibco
Cas9	VBCF Protein Technologies
Alt-R™ CRISPR-Cas9 crRNA	Integrated DNA Technologies
Alt-R™ CRISPR-Cas9 tracrRNA	Integrated DNA Technologies
Electrolytic Buffer E i E2	Invitrogen
Plazmid PX458 (#48138)	Addgene

3.3 Oligonukleotydy

Wszystkie oligonukleotydy z wyjątkiem crRNA były syntetyzowane w Pracowni Sekwencjonowania i Syntezy DNA Instytutu Biochemii i Biofizyki PAN. Starter IG36 użyty do analizy skróceń ciągu CAG w elektroforezie kapilarnej wyznakowany na końcu 5' fluoresceiną był syntetyzowany przez FutureSynthesis sp. z o.o. Oligonukleotydy crRNA były syntetyzowane przez Integrated DNA Technologies.

Tabela 3.3 Zestawienie oligonukleotydów wykorzystywanych w pracy.

Nazwa oligonukleotydu	Sekwencja w orientacji 5'-3'	Opis
EXO1_F	ACTAAGCTACGCTGGGCAATATG	Startery do amplifikacji cDNA genu <i>EXO1</i>
EXO1_R	ATGACTTCTTGAATGGGCAGG	
POLQ_F	GCCTTGGTCGCTGCCTGAA	Startery do amplifikacji cDNA genu <i>POLQ</i>
POLQ_R	AGGAAGTCCCCAGTTTGCCA	
GAPDH_F	GAAGGTGAAGGTCGGAGTC	Startery do amplifikacji cDNA genu <i>GAPDH</i>
GAPDH_R	GAAGATGGTGATGGGATTTC	
IG36	CCGCTCAGGTTCTGCTTTTA	Startery do amplifikacji regionu z powtórzeniami CAG w genie <i>HTT</i>
IG37	GGCTGAGGCAGCAGCGGCTG	
IG126	AGTTGTCAAAGTGCTACCCTTC	Startery do amplifikacji regionu A (analiza resekcji)
IG127	GGTGTGCTAACATTCCCATGTC	
IG124	AAGCTCCTTACCAGCGGTGC	Startery do amplifikacji regionu B (analiza resekcji)
IG125	CCCGTCTCATAGGATGGCTG	
IG50	GGGGTCACACTTGGGGTCCTCA	Startery do amplifikacji regionu C (analiza resekcji)
IG51	GCCAGAGCCATACTCACCCGGA	
IG68	CACACACAGCTTCGCCTCAC	Startery do amplifikacji regionu D (analiza resekcji)
IG69	GTTCTGCCTCACACAGCAAGG	
IG46	CTGCACCGACCGTGAGTTTGGG	Startery do amplifikacji regionu E (analiza resekcji)
IG47	TGGGTCACTCTGTCTCTGCGGG	
IG48	CCAACACGTTGCTGATGGGGAGG	Startery do amplifikacji regionu F (analiza resekcji)
IG49	GGCACATCTGAGATGCCACGG	
IG128	CAGTAGCTTGCGTTATCAGGTT	Startery do amplifikacji regionu G (analiza resekcji)
IG129	TGTCTTCTGATAAGCTCTTGCTTG	
IG130	CAGCCATTGGTGAACCTCGTGC	Startery do amplifikacji regionu H (analiza resekcji)
IG131	GTTATACTCCATGTTGCGGGC	

IG94	GTTCTCCGGATGAGTCCTGC	Startery do amplifikacji regionu C' (analiza resekcji)
IG95	ACCCTGGTCCTCGTTGTTTC	
IG104	GGGTTCTGGAGGAAAGCTCT	Startery do amplifikacji regionu D' (analiza resekcji)
IG105	TCCCTCTCCCCACCTCTGA	
IG110	GTTGCCACTGGACAGTTGAT	Startery do amplifikacji regionu E' (analiza resekcji)
IG111	GCCTCCAAGTCCACTTGCC	
IG100	CCCTGGTCTGCAGAACTGA	Startery do amplifikacji regionu F' (analiza resekcji)
IG101	GACAGAAGAAGCTTCAAGCCC	
IG112	GCTCGCCTGATGGATTGACT	Startery do amplifikacji <i>locus</i> off-target dla gRNA1
IG113	CGCCACAAGAGCATTAGCAA	
IG52	AGGGCTGTCAATCATGCTGG	Startery do amplifikacji regionu d1 (analiza resekcji)
IG53	AAAAGCAGAACCTGAGCGGC	
IG64	CCGCAGGCTAGGGCTGTCAA	Startery do amplifikacji regionu d2 (analiza resekcji)
IG65	CAGCTTTTCCAGGGTCGCCAT	
IG66	TCGCCTCACCCCATTACAGTC	Startery do amplifikacji regionu d3 (analiza resekcji)
IG67	TTGACAGCCCTAGCCTGCGG	
IG96	AGAGGAGGGACGCGGAG	Startery do amplifikacji regionu d'1 (analiza resekcji)
IG97	GCCCTCGCGGGTGAAC	
IG114	AAAGGTTGCTTCGCAGGCTG	Startery do amplifikacji regionu d'2 (analiza resekcji)
IG115	CCCTCTCCCCACCTCTGA	
IG108	CGGTGTTTGTCCCTGGTCT	Startery do amplifikacji regionu f'1 (analiza resekcji)
IG109	TTCAAACCCAATTCTCGTGC	
crRNA1	CUGCUGCUGCUGCUGGGA	crRNA do wytworzenia gRNA1
crRNA2	AGCAGCAGCAGCAGCAGCAG	crRNA do wytworzenia gRNA2
crRNA4	GCUUCCUCAGCCGCCGCCGC	crRNA do wytworzenia gRNA4
sgRNA1_S	CACCGCTGCTGCTGCTGCTGGA	Oligonukleotydy wykorzystane do stworzenia konstruktów kodujących sgRNA1 w plazmidzie
sgRNA1_A	AAACTCCAGCAGCAGCAGCAGCAGC	
sgRNA2_S	CACCGAGCAGCAGCAGCAGCAGCAG	Oligonukleotydy wykorzystane do stworzenia konstruktów kodujących sgRNA2 w plazmidzie
sgRNA2_A	AAACCTGCTGCTGCTGCTGCTGCTC	

3.4 Przeciwciała

Tabela 3.4. Przeciwciała pierwszo i drugorzędowe wykorzystane w pracy.

Nazwa przeciwciała	Producent
Królicze poliklonalne przeciwciało anti-MRE11 (NB100-142)	Novus Biologicals
Królicze monoklonalne przeciwciało anti-Artemis (D708V)	Cell Signaling Technology
Królicze monoklonalne przeciwciało anti-RAD51 (D4B10)	Cell Signaling Technology
Królicze poliklonalne przeciwciało anti-β-actin (#4967)	Cell Signaling Technology
Ośle przeciwciało anti-Mouse IgG (715-035-150)	Jackson ImmunoResearch
Ośle przeciwciało anti-Rabbit IgG (711-035-152)	Jackson ImmunoResearch

3.5 Chemiczne inhibitory białek oraz siRNA

Tabela 3.5.1. Inhibitory białek naprawy DNA wykorzystane w pracy.

Białko naprawcze	Inhibitor	Użyte stężenie	Producent
CTIP	Triapine (3AP)	0.8 μM	Sigma Aldrich
KU70/80	STL127705 (Compound L)	7 μM	MedChemExpress
MRE11	Mirin	25 μM	Sigma Aldrich
POLθ	Novobiocin (NVB)	25 μM	Selleck Chemicals

Tabela 3.5.2. siRNA wykorzystane w pracy.

Gen	siRNA ID	Użyte stężenie	Producent
<i>DCLRE1C</i> (Artemis)	s532106	100 nM	Life Technologies
<i>EXO1</i>	s17502	100 nM	Life Technologies
<i>MRE11</i>	s8959	50 nM	Life Technologies
<i>POLQ</i>	122556	100 nM	Life Technologies
<i>RAD51</i>	s531928	100 nM	Life Technologies

3.6 Enzymy restrykcyjne

Tabela 3.6. Enzymy restrykcyjne wykorzystane w pracy.

Nazwa enzymu	Producent
AluI	New England Biolabs
AvaII	New England Biolabs
BstNI	New England Biolabs
Bsu36I	New England Biolabs
BpmI (GsuI)	ThermoFisher Scientific
CviKI-1	New England Biolabs
Eco0109I	ThermoFisher Scientific
Esp3I	ThermoFisher Scientific
HaeIII	New England Biolabs
HindIII	ThermoFisher Scientific
HpyAV	New England Biolabs
HpyCH4V	New England Biolabs
MluI	ThermoFisher Scientific
MnII	New England Biolabs
MspA1I	New England Biolabs
NlaI	New England Biolabs
PstI	New England Biolabs
RsaI	New England Biolabs
SfcI	New England Biolabs
SmaI	ThermoFisher Scientific

3.7 Składy buforów i żeli do elektroforezy

-Bufor TBE do elektroforezy kwasów nukleinowych

Skład na 1000 ml (X10):

-Tris base	108 g
-Kwas borowy	55 g
-EDTA 0,5 M (pH 8.0)	40 ml
-H ₂ O	do 1000 ml

-Bufor RIPA 0,2% SDS do lizy komórek

Skład na 50 ml:

-Tris HCl 1 M (pH 8)	2,5 ml
-NaCl	0,438 g
-EDTA pH8 0,5 M	200,75 µl
-Deoksycholan sodu	0,25 g
-Triton X-100	0,5 ml
-SDS 20%	0,5 ml
-H ₂ O	do 50 ml

-Obciążacz do elektroforezy białek

Skład na 10 ml (X4):

-SDS	0,8 g
-β-merkaptetanol	0,6 µl
-Glicerol	3,8 ml
-Tris 0,5M (6,8 pH)	4,8 ml
-Błękit bromofenolowy	10 ng
-H ₂ O	do 10 ml

-Bufor Tris-Glycine SDS do elektroforezy białek

Skład na 1000ml (X10):

-Tris base	30,3 g
-Glicyna	144,0 g
-SDS	10,0 g
-H ₂ O	do 1000 ml

-Bufor do transferu białek

Skład na 1000 ml:

-Towbin (X10)	100 ml
-Metanol	200 ml
-SDS roztwór 10%	1 ml
-H ₂ O	do 1000 ml

-Bufor Towbin

Skład na 1000 ml (X10):

-Tris base (0,25 M)	30,28 g
-Glicyna (1,92 M)	144 g
-uzupełnić wodą	do 1000 ml

-Bufor PBS

Skład na 1000 ml (X10):

-NaCl	80 g
-KCl	2 g
-Na ₂ HPO ₄ *7 H ₂ O	11,5 g
-KH ₂ PO ₄	1,9 g
-H ₂ O	do 1000 ml

-Bufor TBS

Skład na 500 ml (X10):

-Tris base	12,114 g
-NaCl	43,83 g
-H ₂ O	do 500 ml

-Żel poliakrylamidowy 12% do elektroforezy białek:

Skład żelu rozdzielającego (12%, 5 ml):

-H ₂ O	1,6 ml
-Akrylamid/bis-akrylamid (29:1) 30%	2 ml
-Tris 8,8pH (1,5 M)	1,3 ml
-SDS 10%	50 µl
-APS 10%	50 µl
-TEMED	2 µl

Skład żelu zagęszczającego (5%, 2 ml):

-H ₂ O	1,4 ml
-Akrylamid/bis-akrylamid (29:1) 30%	330 µl
-Tris 6,8pH (1 M)	250 µl
-SDS 10%	20 µl
-APS 10%	20 µl
-TEMED	2 µl

-Żel agarozowy 1,5% do rozdzielania kwasów nukleinowych

Skład na 100ml

-Bufor TBE (X10)	10 ml
-Agaroza	1,5 g
-Bromek etydyny	5 µl
-H ₂ O	do 100 ml

Tabela 3.7 Składniki buforów i żeli do elektroforezy wykorzystywanych w pracy.

Odzywnnik	Producent
Agaroza	Eurx
Akrylamid/bis-akrylamid (29:1) 30%	Bioshop
APS	Bioshop
BSA	Cell Signaling Technology
EDTA 0,5 M (pH 8.0)	Bioshop
SDS	Bioshop
TEMED	Bioshop
Tween-20	Bioshop
Bromek etydyny 10 mg/10 ml	Sigma-Aldrich

3.8. Pozostałe odczynniki i zestawy odczynników

-Marker wielkości do rozdzielania białek:

PageRuler Plus Prestained Protein Ladder (Thermo Fisher)

-Marker wielkości do rozdzielania kwasów nukleinowych:

Perfect 100-1000 bp DNA Ladder (EURx)

-Nitrocelulozowa membrana do transferu białek:

Amersham Protran 0,45 µm (Cytiva)

-Substrat chemiluminescencyjny do western blot:

Westar Supernova (Cyagen)

-Mieszanina trójfosforanów deoksyrybonukleotydów:

10 mM Ultrapure dNTPs mix (EURx)

-Endonukleaza faga T7:

T7 Endonuclease I (New England Biolabs)

Startery heksamerytowe:

Random Primers (Promega)

-BLOCK-it Fluorescent Oligo (Thermo Fisher)

Polimerazy:

-Q5 High-Fidelity DNA Polymerase (New England Biolabs)

-Herculase II Fusion DNA Polymerase (Agilent)

-GoTaq G2 Flexi DNA Polymerase (Promega)

-SsoAdvanced Universal SYBR Green Supermix (BioRad)

Tabela 3.8. Zestawy odczynników wykorzystywane w pracy.

Zestaw	Producent
Zestaw do izolacji DNA	Genomic DNA Isolation Kit (Norgen, Biotek Corp.)
Zestaw do izolacji DNA	QuickExtract DNA Extraction Solution (Lucigen)
Zestaw do izolacji RNA	Direct-zol RNA Microprep (Zymo Research)
Zestaw do odwrotnej transkrypcji	SuperScript III Reverse Transcriptase (Thermo Fisher)
Zestaw do oczyszczania produktu PCR	Gene JET PCR Purification Kit (Thermo Fisher)

3.9 Sprzęt

-Automatyczny licznik komórek TC20 (Biorad)

-Elektroporator Neon Transfection System (Invitrogen)

-Spektrofotometr DS-11 (DeNovix)

-Termocykler T100 Thermal Cycler (BioRad) oraz (do qPCR) CFX Connect Real Time System (BioRad)

-Aparat do wizualizacji żeli i membran G-BOX (Syngene)

3.10 Wykorzystywane oprogramowanie

Do analizy wyników rozdziałów produktów PCR w elektroforezie kapilarnej wykorzystano Peak Scanner Software V1.0 (Applied Biosystems).

Analizy statystyczne wykonano w programie GraphPad Prism 8.

Rysunki wykonano w programie Biorender.

Do analizy ekspresji genów wykorzystano CFX Maestro (Biorad).

Do analizy rozdziału DNA w żelach agarozowych wykorzystywano Gel-Pro Analyzer.

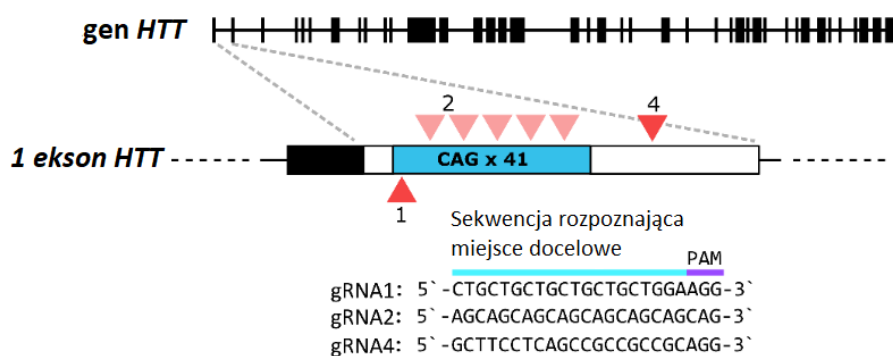
4. Metody

4.1 Hodowle komórkowe

Komórki hodowano w inkubatorze, w stałej temperaturze 37 °C w atmosferze z 5% zawartością CO₂. Do hodowli komórek HEK293T stosowano medium DMEM z 10% FBS, 1% antybiotyku-antymykotyku oraz 1% L-glutaminy. Do hodowli fibroblastów stosowano medium MEM z 10% FBS oraz 1% antybiotyku-antymykotyku oraz 1% L-glutaminy.

4.2 System CRISPR-Cas9

W poszczególnych eksperymentach korzystano z jednego z dwóch systemów: rybonukleoproteiny (RNP) składającej się z białka Cas9 w kompleksie z odpowiednim gRNA (gRNA1, gRNA2 lub gRNA4) lub plazmidu z sekwencją kodującą Cas9, białko zielonej fluorescencji (GFP) oraz wklonowaną sekwencją kodującą odpowiedni sgRNA (sgRNA1, sgRNA2 lub sgRNA4). Rysunek 4.2 przedstawia sekwencje i miejsca rozpoznawane przez poszczególne gRNA w *locus HTT*.



Rysunek 4.2 Schemat przedstawiający ludzki gen *HTT* oraz miejsca rozpoznawane przez poszczególne gRNA wykorzystane w pracy. Potencjalne miejsca cięcia dla każdego gRNA zaznaczono czerwonymi trójkątami (na podstawie Śledziński i in 2024¹⁸²).

4.2.1 Kompleksy RNP

Kompleksy RNP wytworzono w następujący sposób: odpowiedni crRNA i tracrRNA zmieszano w ilościach równomolowych do uzyskania końcowego stężenia dupleksu 44 μM. Mieszaninę denaturowano przez 5 min w 95 °C a następnie schładzano do temperatury pokojowej przez 20 min. Następnie do uzyskanego gRNA dodawano Cas9 w proporcji 22 pmol

do 18 pmol i inkubowano 20 min w temperaturze pokojowej przez 20 min. Tak wytworzony kompleks RNP wprowadzano do komórek metodą elektroporacji.

4.2.2 Ekspresja systemu CRISPR-Cas9 z plazmidu

W eksperymentach w których wyciszano ekspresję genów kodujących białka naprawcze za pomocą siRNA oraz w których stosowano inhibitory chemiczne białek naprawczych wykorzystywano plazmid PX458 (Addgene #48138) kodujący Cas9 wraz z wyklonowanymi odpowiednim sgRNA wykorzystywany w poprzednich pracach^{194,213}. Transfekcje komórek plazmidem wykonywano przy użyciu polietylenoiminy (PEI).

4.3 Transfekcja komórek metodą elektroporacji

W celu wprowadzenia do komórek kompleksów RNP stosowano Neon Transfection System. Komórki zawieszone w PBS, wraz z wytworzonymi kompleksami RNP wprowadzano do odpowiedniej końcówki Neon Tip i poddawano elektroporacji, po czym wysiewano na szalkę z właściwym dla danych komórek medium. Do analizy produktów naprawy metodą NGS elektroporacji poddawano milion komórek w przypadku każdej próbki. W przypadku analizy resekcji końców DNA oraz analizy dynamiki skróceń w punktach czasowych elektroporacji poddawano po 200 tysięcy komórek na każdą próbkę. Szczegółowe warunki elektroporacji dla różnych typów i liczby komórek przedstawiono w tabeli 4.3

Tabela 4.3. Szczegółowe warunki elektroporacji dla różnych typów i liczby komórek

Typ komórek	Liczba komórek jednorazowo poddawanych elektroporacji	Wykorzystane elementy zestawu Neon Transfection System	Warunki elektroporacji
HEK293T_41CAG	200 tysięcy	Neon Tip 10 µl, bufor E	1150 V, 20 ms, 2 impulsy
HEK293T_41CAG	1 milion	Neon Tip 100 µl, bufor E2	1150 V, 20 ms, 2 impulsy
Fibroblasty GM04208	1 milion	Neon Tip 100 µl, bufor E2	1350 V, 30 ms, 1 impuls

4.4. Transfekcja komórek siRNA za pomocą lipofektaminy

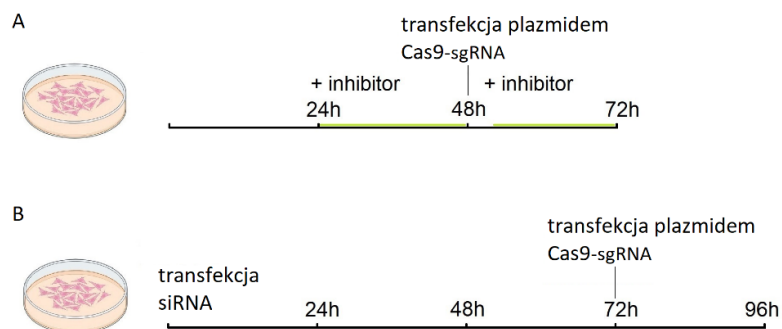
W celu wprowadzenia do komórek siRNA stosowano odwróconą transfekcję. Do pustego dołka płytki 6-dołkowej dodano 500 μ l Opti-MEM oraz odpowiednie siRNA w ilości właściwej do uzyskania końcowego stężenia 50 lub 100 nM (tabela 3.5.2). W kolejnym kroku dodano 5 μ l lipofektaminy i inkubowano przez 20 min w temperaturze pokojowej. Po inkubacji, do dołka dodawano 200 tys. komórek zawieszonych w 2 ml DMEM bez antybiotyku. Medium zmieniono po 24 h.

4.5 Użycie inhibitorów chemicznych białek naprawczych

Dzień przed podaniem inhibitora komórki były wysiewane w liczbie 300 tysięcy na dołek płytki 6-dołkowej. Po 24 h od pasażu w każdym dołku wymieniano medium na nowe, zawierające inhibitor w odpowiednim stężeniu (tabela 3.5.1). Po 24 h komórki transfekowano plazmidem kodującym Cas9 i odpowiednie sgRNA (rozdział 4.6). Następnie, 4 h od transfekcji plazmidem medium ponownie wymieniano na nowe zawierające inhibitor w stężeniu jak uprzednio.

4.6. Transfekcja komórek plazmidem za pomocą polietylenoiminy

Komórki, były poddawane transfekcji plazmidem po 72 h od transfekcji siRNA lub po 24 h od momentu podania inhibitora. Dla każdego dołka, w którym znajdowały się komórki, mające być poddane transfekcji, wytwarzano dwie mieszaniny: A oraz B. Mieszanina A składała się ze 125 μ l DMEM oraz 2,5 μ g plazmidu. Mieszanina B ze 125 μ l DMEM oraz 7,5 μ l PEI o stężeniu 1 mg/ml. Mieszaniny A i B mieszano razem i inkubowano w temperaturze pokojowej przez 20 minut. Następnie w każdym dołku wymieniono medium na 500 μ l DMEM a następnie wkroplono mieszaninę A+B. Medium zmieniono po 24 h. Procedurę przedstawiono na rysunku 4.6.



Rysunek 4.6 Schemat procedury użycia inhibitorów chemicznych (A) lub siRNA (B) oraz transfekcji komórek HEK293T_41CAG plazmidem kodującym Cas9 i sgRNA.

4.7 Analiza produktów naprawy DNA w *locus HTT* z wykorzystaniem metody NGS

Komórki HEK293T_41CAG zostały zebrane w celu izolacji genomowego DNA 24 h po elektroporacji kompleksami RNP z gRNA1, gRNA2, gRNA4 lub łącznie gRNA1 i gRNA4. Izolacja DNA odbywała się przy użyciu zestawu Genomic DNA Isolation Kit zgodnie z zaleceniami producenta. Stężenie DNA zostało zmierzone spektrofotometrem. Ze względu na ograniczenia długości produktów PCR, które mogły być sekwencjonowane metodą NGS Illumina NovaSeq 6000 do 250 par zasad, konieczne było wykonanie dwóch reakcji PCR, gdzie matrycą w drugiej reakcji był produkt PCR z pierwszej (nested PCR). W obu reakcjach użyto Herculanse II Fusion DNA Polymerase. W pierwszej reakcji wykorzystano startery IG36 i IG37 a długość produktu nieedytowanego wynosiła 377 pz. Produkty PCR z pierwszej reakcji zostały oczyszczone przy użyciu zestawu GeneJET PCR Purification Kit, zgodnie z zaleceniami producenta. Produkty po dziesięciokrotnym rozcieńczeniu zostały poddane drugiej amplifikacji ze starterami IG30 i IG31. Długość nieskróconego produktu wynosiła 242 pz. Przygotowanie biblioteki oraz sekwencjonowanie zostało wykonane przez Novogene z wykorzystaniem zestawu NEBNext Ultra II DNA Library Prep Kit for Illumina. Analiza bioinformatyczna została wykonana przez ideas4biology sp. z o. o. Dalsza analiza i opracowanie wyników zostało wykonane wspólnie ze współpracownikiem z Zakładu Inżynierii Genomowej Instytutu Chemii Bioorganicznej PAN Pawłem Śledzińskim. W eksperymencie, w którym analizowano wyniki NGS prób z wyciszeniem genu *POLQ* oraz bez jego wyciszenia, transfekcję siRNA przeprowadzono tak jak to opisano w rozdziale 4.4. Zbiór komórek do izolacji DNA przeprowadzono po 96 godzinach od transfekcji siRNA i 24h od transfekcji plazmidem (rysunek 4.6 B)

4.8 Analiza dynamiki skracania ciągu powtórzeń CAG w punktach czasowych

Komórki HEK293T_41CAG poddane elektroporacji kompleksem RNP zawierającym gRNA1 lub gRNA2 zebrano w celu izolacji genomowego DNA w czasie 1 h, 4 h oraz 24 h od transfekcji. Do izolacji DNA użyto zestawu Genomic DNA Isolation Kit zgodnie z zaleceniami producenta. Stężenie DNA zostało zmierzone spektrofotometrem. Uzyskany DNA poddano amplifikacji z użyciem Q5 High-Fidelity DNA Polymerase oraz starterów IG36 (wyznakowanym na końcu 5' fluoresceiną) i IG37. Produkty PCR były następnie oczyszczane przy użyciu zestawu GeneJET PCR Purification Kit, zgodnie z instrukcją dostarczoną przez producenta. Elektroforeza kapilarna została wykonywana w Laboratorium Techniki Biologii Molekularnej, Uniwersytetu Adama Mickiewicza na urządzeniu ABI Prism 3130xl (Applied Biosystems) z wykorzystaniem polimera POP-7 i filtra G5. W przypadku każdej próbki objętość 1 mikrolitra była rozcieńczana w 9 μ l Hi-Di formamidu (Applied Biosystems) zawierającego 0,25 μ l standardu GeneScan 600 LIZ. Przed rozdzieleniem próbki były denaturowane w 95 °C przez 5 min a następnie schładzane do 10 °C. Wyniki były analizowane przy użyciu programu Peak Scanner Software V1.0.

4.9 Analiza resekcji końców DNA

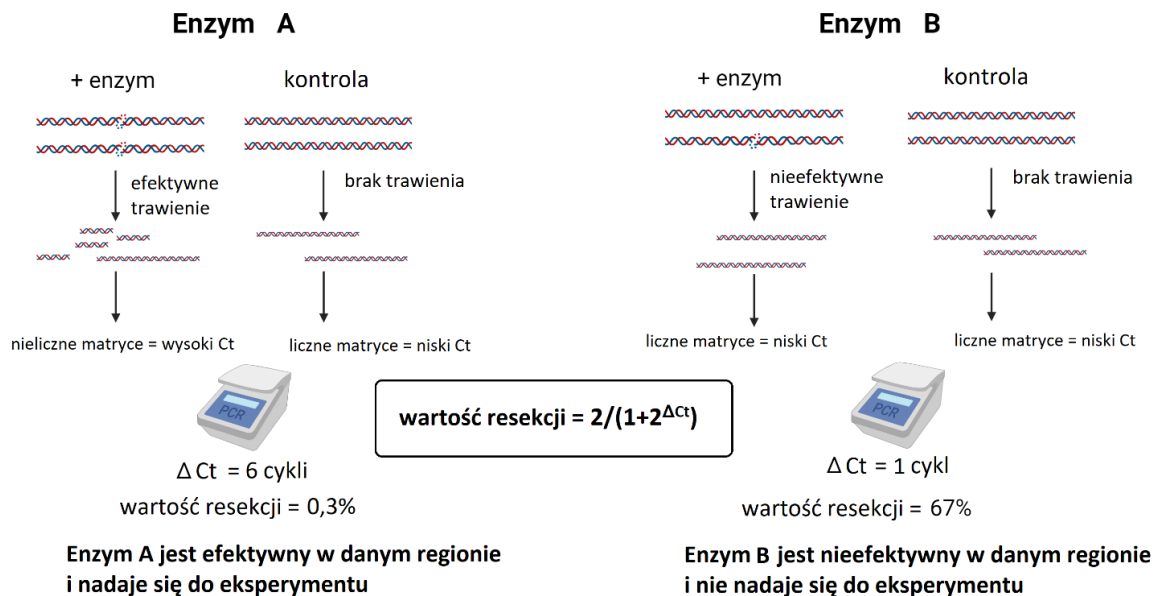
Metoda pozwalająca na określenie zasięgu i intensywności resekcji końców DNA została wcześniej wykorzystana w pracach dotyczących naprawy DNA w komórkach drożdży.^{187,214} W niniejszej pracy została ona zaadoptowana do badania resekcji w komórkach ludzkich. Metoda ta wykorzystuje fakt trawienia przez enzymy restrykcyjne dwuniciowego DNA i ich nieaktywności wobec jednoniciowego DNA. Degradacja dwuniciowego DNA, będąca efektem trawienia enzymami restrykcyjnymi skutkuje obecnością niewielkiej liczby matryc podczas amplifikacji w qPCR, co wiąże się z wysoką wartością parametru Ct. Obecność jednoniciowego DNA w regionie, który następnie podlega amplifikacji, a został uprzednio traktowany odpowiednim enzymem restrykcyjnym, skutkuje zachowaniem dużej liczby matryc w qPCR i niską wartością Ct. Dla każdego regionu, w którym badano intensywność resekcji podzielono materiał DNA uzyskany po edycji na dwie próbki: pierwszą, inkubowaną z enzymem oraz drugą, inkubowaną bez enzymu. Jeżeli region pozostałby dwuniciowy, enzym dokonałby degradacji DNA, a różnica w wartości Ct dla próby z enzymem i bez enzymu (Δ Ct) byłaby wysoka, co po podstawieniu do wzoru:

$$\text{wartość resekcji} = 2/(1+2^{\Delta C_t})$$

$$\Delta C_t = C_{t_{\text{enzym}}} - C_{t_{\text{bez enzymu}}}$$

dałoby niską wartość resekcji. Jeśli zaś, dany region podlegałby resekcji, ssDNA nie uległby degradacji, w związku z czym amplifikacja zasłabłaby efektywnie i wartość ΔC_t byłaby niska, co po podstawieniu do wzoru dałoby wyższą wartość resekcji.

Kluczowe w tym eksperymencie jest zweryfikowanie efektywności trawienia DNA w każdym badanym regionie przez dany enzym restrykcyjny, tak by uniknąć wyników fałszywie pozytywnych. W związku z tym, przed eksperymentem, w którym określano intensywność i zakres resekcji DNA w komórkach po indukcji dwuniciowych pęknięć, wcześniej dokonano określenia efektywności poszczególnych enzymów w danych regionach wykorzystując DNA izolowany z komórek nie poddanych edycji. Dokładniej proces ten został zaprezentowany na rysunku 4.9. Tabela 4.9 przedstawia enzymy restrykcyjne, które były testowane dla poszczególnych regionów.



Rysunek 4.9 Schemat selekcji enzymów restrykcyjnych na podstawie ich efektywności trawienia DNA. Wzór do obliczania wartości resekcji zaczerpnięty z pracy Zierhut i Diffey 2008.²¹⁵

Docelowy eksperyment został wykonany w następujący sposób. Komórki HEK293T_41CAG lub fibroblasty GM04208 poddane elektroporacji z kompleksem RNP zawierającym gRNA1 lub gRNA2 zebrano w celu izolacji genomowego DNA w czasie: 1 h, 6 h, 24 h oraz 48 h od transfekcji. Do izolacji DNA użyto zestawu Genomic DNA Isolation Kit zgodnie z zaleceniami producenta. Stężenie DNA zostało zmierzone spektrofotometrem. Każda próbka została podzielona na część inkubowaną z enzymami restrykcyjnymi oraz część inkubowaną bez enzymów przez 16 h w odpowiedniej dla każdego enzymu temperaturze. Po inkubacji DNA został poddany amplifikacji w qPCR przy użyciu SsoAdvanced Universal SYBR Green Supermix ze starterami odpowiednimi dla każdego badanego regionu. Każda próbka amplifikowana była w trzech powtórzeniach technicznych. Wartości parametru Ct dla amplifikacji każdego regionu w próbce traktowanej enzymem i nietraktowanej zostały wykorzystane do obliczenia wartości resekcji w danym regionie za pomocą wyżej przedstawionego wzoru. Eksperyment został wykonany w trzech powtórzeniach biologicznych. Istotność statystyczna wyników została określona za pomocą jednoczynnikowej analizy wariancji (one-way ANOVA).

Tabela 4.9 Enzymy restrykcyjne testowane w poszczególnych regionach.

HEK293T_41CAG	
Region	Testowane enzymy
Regiony a	NlaI, CviKI-1
Regiony b	Bsu36I, HpyCH4V
Regiony c	AvaII
Regiony d	AvaII, AluI, BstNI, HaeIII
Regiony e	AvaII
Regiony f	AvaII, RsaI
Regiony g	Esp3I, HpyCH4V
Regiony h	MspAI, CviKI-1
Regiony c'	MnII, SfcI
Regiony d'	PstI, BpmI, Eco0109I, PstI, SfcI
Regiony e'	MluI
Regiony f'	AluI, PstI, HindIII, BstNI SfcI

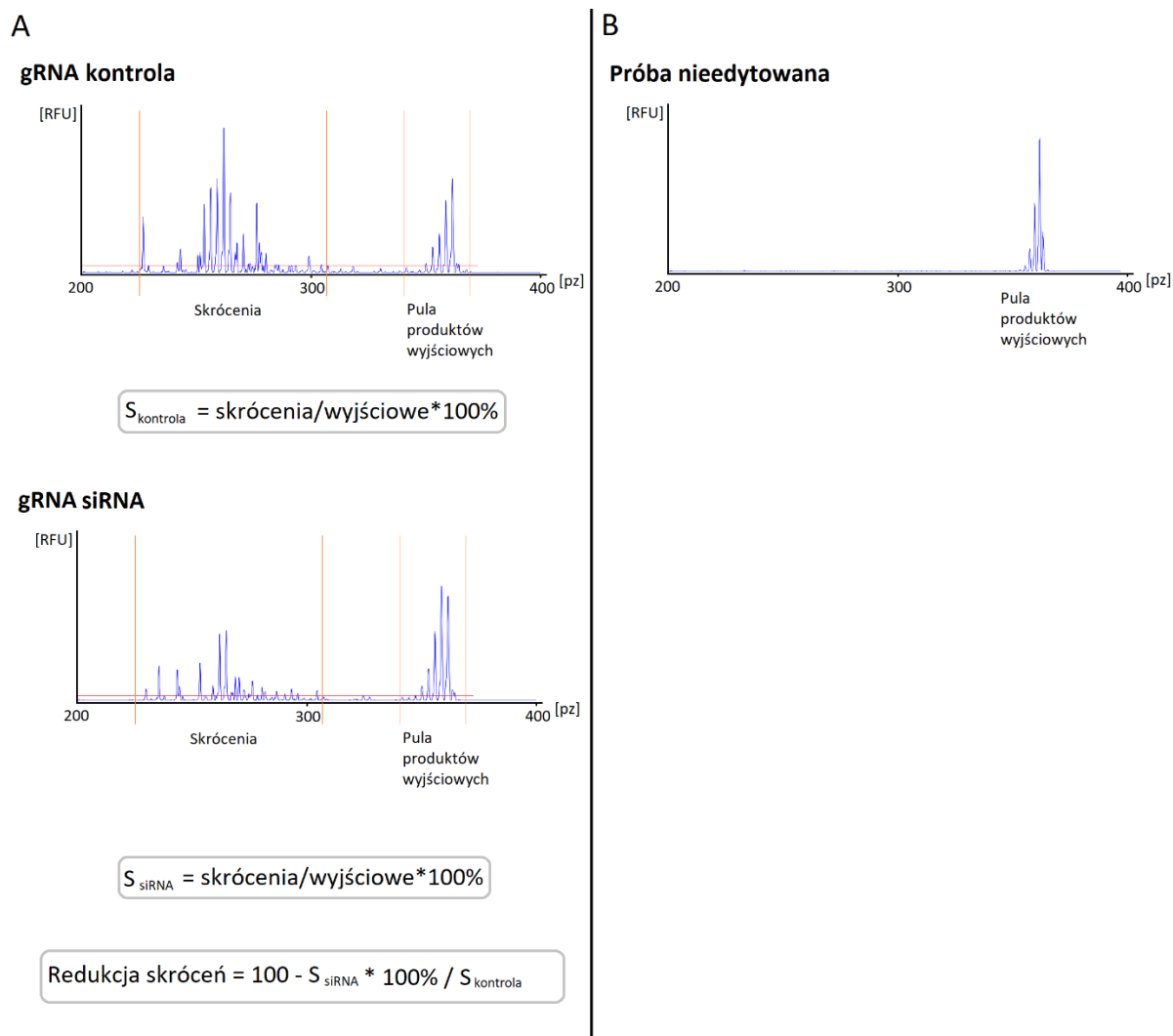
Fibroblasty GM04208	
Region	Testowane enzymy
Regiony a	CviKI-1
Regiony b	HpyCH4V
Regiony c	AvaII
Regiony d	HaeIII, SmaI, MspA1I, HpyAV
Regiony e	AvaII
Regiony f	RsaI
Regiony g	HpyCH4V
Regiony h	CviKI-1

4.10 Analiza udziału białek naprawy DNA w skracaniu ciągu powtórzeń CAG

Komórki HEK293T_41CAG poddane transfekcji z plazmidem kodującym Cas9, sgRNA1 lub sgRNA2 oraz GFP zebrano 24 h po transfekcji w celu wysortowania 100 tysięcy komórek z ekspresją GFP. Jako kontrolę do ustawienia bramkowania użyto komórek HEK293T_41CAG niepoddanych transfekcji plazmidem. Sortowanie komórek wykonano na urządzeniu FACSaria w Pracowni Analiz Pojedynczych Komórek w Instytucie Chemii Bioorganicznej PAN. Z wysortowanych komórek z ekspresją GFP wyizolowano DNA przy użyciu QuickExtract DNA Extraction Solution, zgodnie z zaleceniami producenta. Następnie DNA był amplifikowany przy użyciu Q5 High-Fidelity DNA Polymerase i starterów IG36 (wyznakowanym na końcu 5' fluoresceiną) oraz IG37. Produkty PCR były następnie oczyszczane przy użyciu zestawu Gene JET PCR Purification Kit, zgodnie z instrukcją dostarczoną przez producenta. Elektroforeza kapilarna została wykonana tak, jak to opisano w rozdziale 4.8.

Wyniki rozdziału elektroforetycznego były analizowane przy użyciu programu Peak Scanner Software V1.0. Elektroforegram każdej próbki przedstawia piki odpowiadające produktom PCR o określonej wielkości (liczbie par zasad). Każdy pik określany jest takimi parametrami jak: rozmiar (size), wysokość (height) i pole powierzchni pod pikiem (area in BP). W celu ograniczenia wpływu tła, wszystkie piki, których wysokość była niższa od 5% wysokości najwyższego piku dla danej próbki zostały pominięte w analizie. Wartości area in BP zostały następnie zsumowane dla każdej próbki, podobnie wartości area in BP pików odpowiadających produktom PCR zawierającym skrócenia ciągu CAG. Dla każdej próbki traktowanej inhibitorem lub siRNA wyliczono stosunek produktów PCR skróconych do puli produktów wyjściowych (obserwowanych w rozdziale produktów amplifikacji tego *locus* bez edycji)

i odniesiono do takiego samego stosunku uzyskanego w próbie kontrolnej, nietraktowanej inhibitorem lub siRNA (rysunek 4.10). Eksperyment został wykonany w trzech powtórzeniach biologicznych dla każdego siRNA oraz inhibitora. Istotność statystyczna wyników została określona za pomocą testu t-Studenta dla prób niezależnych.



Rysunek 4.10 A) Schemat przedstawiający sposób obliczania redukcji skróceń uzyskanych po edycji z udziałem sgRNA w próbach po zastosowaniu siRNA względem kontroli. W identyczny sposób wyliczano redukcję skróceń dla prób traktowanych inhibitorami chemicznymi. Czerwone linie oznaczają punkt odcięcia (5% wysokości najwyższego piku dla danej próbki). Pomarańczowe linie wyznaczają granice, w których zawierają się piki reprezentujące skrócenia ciągu CAG oraz produkty wyjściowe, obserwowane w rozdziale produktów amplifikacji tego *locus* bez edycji (B).

4.11. Amplifikacja DNA

4.11.1. Zagnieżdżony PCR (nested) do analizy NGS

Do analizy NGS amplifikowano fragment zawierający ciąg CAG w genie *HTT* przy użyciu Herculase II Fusion DNA Polymerase oraz starterów: do pierwszej reakcji IG36 i IG37, do drugiej reakcji IG30 i IG31. Programy i skład reakcji przedstawiono poniżej.

I reakcja

Program:

98 °C 3 min

98 °C 20 s

61,1 °C 20 s x35

72 °C 20 s

72 °C 3 min

I reakcja

Skład reakcji:

H ₂ O	6,3 µl
5× Herculase II Reaction Buffer	2 µl
Starter IG36	0,25 µl
Starter IG37	0,25 µl
dNTP	0,1 µl
Herculase II Fusion DNA Polymerase	0,1 µl
gDNA(100 ng)	1 µl

II reakcja

Program:

98 °C 3 min

98 °C 20 s

56 °C 20 s x30

72 °C 12 s

72 °C 3 min

II reakcja

Skład reakcji:

H ₂ O	6,3 µl
5× Herculase II Reaction Buffer	2 µl
Starter IG30	0,25 µl
Starter IG31	0,25 µl
dNTP	0,1 µl
Herculase II Fusion DNA Polymerase	0,1 µl
Produkt PCR z I reakcji rozcieńczony dziesięciokrotnie	1 µl

4.11.2. PCR do amplifikacji *locus HTT*

Do analiz, w których produkty PCR rozdzielano w elektroforezie kapilarnej wykorzystywano Q5 High-Fidelity DNA Polymerase oraz starter IG36 wyznakowany na 5' końcu fluoresceiną i starter IG37. W innych przypadkach amplifikowano ten region z użyciem starterów niewyznakowanych. Program i skład reakcji przedstawiono poniżej.

Program:

98 °C	3 min
98 °C	10 s
68 °C	30 s x35
72 °C	5 s
72 °C	2 min

Skład reakcji:

H ₂ O	4,7 µl
5X Q5 Reaction Buffer	2 µl
High GC enhancer	1 µl
Starter IG36	0,5 µl
Starter IG37	0,5 µl
dNTP	0,2 µl
Q5 High-Fidelity DNA Polymerase	0,1 µl
gDNA (50 ng)	1 µl

4.11.3 PCR do analiz resekcji końców DNA

Do amplifikacji poszczególnych regionów, które miały być badane pod względem obecności resekcji DNA wykorzystywano SsoAdvanced Universal SYBR Green Supermix. Ze względu na pojemność płytki 96-dółkowej, wykorzystywanej w eksperymencie, jednorazowo, na pojedynczej płytce przeprowadzano analizę czterech regionów dla czterech czasów. Ze względu na różne optymalne temperatury topnienia starterów (T_m) wykorzystywanych do amplifikacji różnych miejsc restrykcyjnych, wykorzystywano trzy programy z różnymi, optymalnymi T_m dla danych starterów. Dla każdego regionu kontrolę stanowiły reakcje (3 powtórzenia biologiczne) na matrycy nieedytowanego DNA, które zachodziły w tych samych warunkach co reakcje z DNA izolowanym z komórek poddanych edycji. Skład reakcji był w każdym przypadku taki sam. Programy i składy reakcji przedstawiono poniżej.

Program PCR dla regionów A, B, G, H:

98 °C 3min
 95 °C 15s
 67,7°C 30 s x45
 72 °C 30 s

Skład reakcji:

H₂O 1,5 µl
 SsoAdvanced Universal SYBR 5 µl
 Green Supermix
 Starter F 0,5 µl
 Starter R 0,5 µl
 gDNA (50 ng) 2,5 µl

Program PCR dla regionów C, D, E, F:

98 °C 3min
 95 °C 15s
 66,2°C 30 s x45
 72 °C 30 s

Program PCR dla regionów C', D', E', F':

98 °C 3min
 95 °C 15s
 67,5°C 30 s x45
 72 °C 30 s

4.11.4 Amplifikacja cDNA do analizy ekspresji genów

Do amplifikacji cDNA przy analizie ekspresji genów stosowano SsoAdvanced Universal SYBR Green Supermix z odpowiednimi starterami przedstawionymi w tabeli 3.3. Program i skład reakcji przedstawiono poniżej.

Program:

95 °C 30 s
 95 °C 15 s
 60°C 20 s x4

Skład reakcji:

H₂O do 10 µl
 SsoAdvanced Universal SYBR 5 µl
 Green Supermix
 Starter F 0,5 µl
 Starter R 0,5 µl
 cDNA 100 ng

4.11.5 PCR do amplifikacji *locus off-target* dla gRNA1

Do amplifikacji *locus off-target* dla gRNA1 stosowano GoTaq G2 Flexi DNA Polymerase, ze starterami IG112 oraz IG113. Program i skład reakcji przedstawiono poniżej.

Program:

95 °C 3 min
95 °C 30 s
61 °C 30 s x 35
72 °C 20 s
72 °C 5 min

Skład reakcji:

H ₂ O	5,3 µl
5X GoTaq Flexi Buffer	2 µl
MgCl ₂	0,6 µl
Starter IG112	0,4 µl
Starter IG113	0,4 µl
dNTP	0,2 µl
GoTaq G2 Flexi DNA Polymerase	0,1 µl
gDNA (100 ng)	1 ul

4.11.6 PCR z gradientem temperatury przyłączania starterów

PCR z gradientem temperatury przyłączania starterów wykonywany był przy użyciu GoTaq G2 Flexi DNA Polymerase. Program i skład reakcji przedstawiono poniżej.

Program:

95 °C 3 min
95 °C 30 s
gradient °C 30 s x 30
72 °C 30 s
72 °C 5 min

Skład reakcji:

H ₂ O	5,3 µl
5X GoTaq Flexi Buffer	2 µl
MgCl ₂	0,6 µl
Starter F	0,4 µl
Starter R	0,4 µl
dNTP	0,2 µl
GoTaq G2 Flexi DNA Polymerase	0,1 µl
gDNA (100 ng)	1 ul

4.12. Odwrotna transkrypcja i analiza ekspresji genów

Aby potwierdzić obniżenie ekspresji genów kodujących białka naprawcze w komórkach po transfekcji siRNA wykonano analizę ekspresji tych genów, stosując jako gen referencyjny *GAPDH*. Jako kontrolę stosowano komórki transfekowane BLOCK-it zamiast siRNA. Komórki HEK293T_41CAG w liczbie 500 tysięcy zebrano 96 h po transfekcji siRNA do izolacji RNA. Izolację wykonano z użyciem zestawu Direct-zol RNA Microprep, zgodnie z zaleceniami producenta. Stężenie RNA zostało zmierzone spektrofotometrem a następnie wykonano odwrotną transkrypcję (RT) z użyciem zestawu SuperScript III w następujących krokach:

Zmieszano:

RNA	500 ng
dNTP	1 µl
Heksamer Primers	1 µl
H ₂ O	do 13 µl

Następnie mieszaninę inkubowano przez 5 minut w 65 °C i wstawiono na lód na 1 minutę.

Po inkubacji do mieszaniny dodano:

First Strand Buffer	4 µl
DTT	1 µl
RNase OUT	1 µl
SuperScript III	1 µl

Następnie mieszaninę wstawiono do reakcji odwrotnej transkrypcji przedstawionej poniżej.

Program RT:

25 °C 5min

55 °C 60min

70 °C 15min

Stężenie cDNA zmierzono a następnie wykorzystano do reakcji opisanej w rozdziale 4.11.4.

Analizę ekspresji genów wykonano w trzech powtórzeniach biologicznych dla każdego badanego genu w programie CFX Maestro. Istotność statystyczna wyników została określona za pomocą testu t-Studenta.

4.13 Elektroforetyczny rozdział produktów PCR w żelu agarozowym

Produkty PCR rozdzielano w 1,5 % żelu agarozowym pod napięciem 120V przez czas od 30 do 60 minut.

4.14. Analiza efektywności edycji metodą T7E1

W celu potwierdzenia edycji *locus* off-target dla gRNA1 w komórkach HEK293T_41CAG, w których analizowano resekcje końców DNA wykonano analizę efektywności edycji tego *locus* metodą wykorzystującą endonukleazę faga T7.

Z komórek poddanych transfekcji kompleksem RNP wyizolowano DNA 24 h po transfekcji przy użyciu zestawu Genomic DNA Isolation Kit zgodnie z zaleceniami producenta. DNA posłużył jako matryca do amplifikacji badanego *locus*. Po amplifikacji opisanej w rozdziale 4.11.5 produkty PCR oczyszczono przy użyciu zestawu Gene JET PCR Purification Kit. W kolejnym kroku, do oczyszczonego produktu PCR w ilości 200 ng dodano 2 µl NEBuffer oraz uzupełniono wodą do 19 ul. Następnie przeprowadzono annealing produktów PCR według programu opisanego poniżej:

Program annealingu produktów PCR:

95 °C	5 min
95 °C - 85 °C	obniżanie temperatury o 2 °C na sekundę
85 °C - 25 °C	obniżanie temperatury o 0,1 °C na sekundę
4 °C	

Po annealingu dodano 1 µl endonukleazy T7 i inkubowano próbkę przez 30 minut w 37 °C. Po inkubacji, do próbki dodano 0,35 µl EDTA 0,5 M. Produkt reakcji rozdzielono w żelu agarozowym, a następnie analizowano w programie Gel-Pro Analyzer uzyskując liczbową wartość intensywności prążków. Efektywność edycji obliczono, dzieląc wartość intensywności prążków powstałych po inkubacji z T7E1 przez wartość intensywności wszystkich prążków i wyrażono w procentach.

4.15 Elektroforetyczny rozdział białek w żelu poliakrylamidowym i analizie western blot

Komórki HEK293T_41CAG zebrano 96 h po transfekcji siRNA do lizy w buforze RIPA 0,2% SDS. Jako kontroli użyto takich samych komórek nietransfekowanych siRNA. Stężenie białka zostało zmierzone spektrofotometrem. Do każdej kieszonki żelu 12% nałożono próbkę zawierającą 30 µg białka i 5 µl obciążacza do białek. W razie potrzeby uzupełniono próbkę buforem RIPA do 20 µl. Przed nałożeniem na żel próbki denaturowano przez 5 min w 95 °C . Rozdział był prowadzony w buforze x1 Tris-Glycine SDS Running Buffer w warunkach:

-20 min 60V

-90 min 125 V

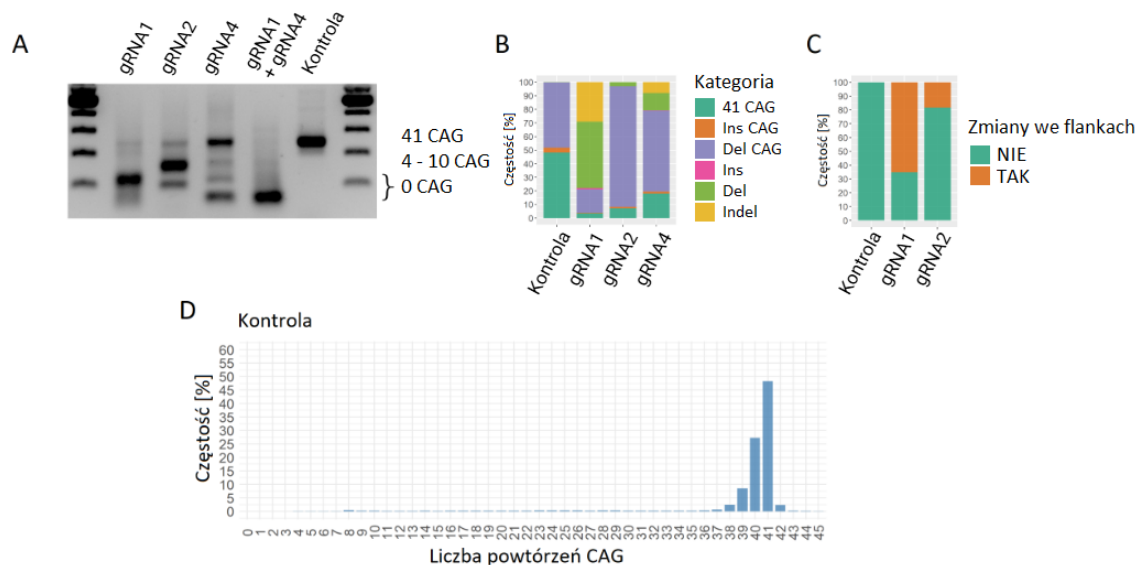
Następnie wykonano transfer białek na membranę nitrocelulozową w buforze do transferu przez 90 minut pod napięciem 25 V w 4 °C. Membranę blokowano w 3-5% roztworze BSA w buforze PBS lub TBS z 0,1% Tween-20. Inkubacje z przeciwciałami pierwszorzędowymi prowadzono przez noc w 4 °C. Inkubacje z przeciwciałami II-rzędowymi prowadzono przez 1 h w temperaturze pokojowej. Do wizualizacji białek wykorzystano urządzenie G-BOX.

5. Wyniki

W celu identyfikacji procesów zaangażowanych w skracanie ciągów powtórzeń CAG w wyniku indukcji DSBs, przeprowadzono szereg eksperymentów, w których wykorzystano ludzki model komórkowy HD. Były to komórki HEK293T z 41 powtórzeniami CAG na obu allelach genu *HTT* (HEK293T_41CAG)²¹². Aby zbadać wpływ miejsca indukcji DSB oraz sekwencji go flankujących na naprawę DNA do eksperymentów użyto trzech gRNA indukujących DSBs w różnych miejscach względem ciągu powtórzeń CAG w genie *HTT*. Wykorzystano: gRNA1 indukujący DSB na początku ciągu powtórzeń, gRNA2 tnący DNA w obrębie ciągu powtórzeń, oraz gRNA4 indukujący DSB poniżej ciągu powtórzeń (rysunek 4.2).

5.1 Skracanie ciągów powtórzeń CAG w wyniku indukcji DSB w genie *HTT*

W celu scharakteryzowania produktów naprawy DSBs indukowanych w regionach z powtórzeniami CAG, wykorzystując model HD (komórki HEK293T_41CAG) wykonano sekwencjonowanie następnej generacji (NGS, targeted sequencing) produktów PCR, powstałych po amplifikacji regionu z powtórzeniami CAG w genie *HTT*, edytowanego z udziałem różnych gRNA. Wyniki efektów naprawy różniły się w zależności od użytego gRNA (rysunek 5.1 A-C) W wynikach NGS z nieedytowanej próby kontrolnej zaobserwowano wysoką częstość występowania wariantów z liczbą powtórzeń CAG różniącą się w niewielkim stopniu od wyjściowych 41 powtórzeń, co jest prawdopodobnie wynikiem błędów polimerazy powstałych w czasie amplifikacji ciągu powtórzeń (rysunek 5.1 B i D). Warianty zaobserwowane w próbach po edycji systemem CRISPR-Cas9 zostały przyporządkowane do sześciu kategorii: „41 CAG” oznacza odczyty identyczne z sekwencją wyjściową; „Ins CAG” oraz „Del CAG” oznaczają warianty, w których doszło do odpowiednio zwiększenia bądź zmniejszenia liczby powtórzeń CAG, bez zmian w postaci innych insercji bądź delecji w sekwencjach flankujących edytowane miejsce; kategorie „Ins” i „Del” oznaczają odpowiednio insercje bądź delecje, których skutkiem było przesunięcie otwartej ramki odczytu (ORF); w kategorii „Indel” zawierają się inne złożone zmiany, których efektem również było przesunięcie ORF.

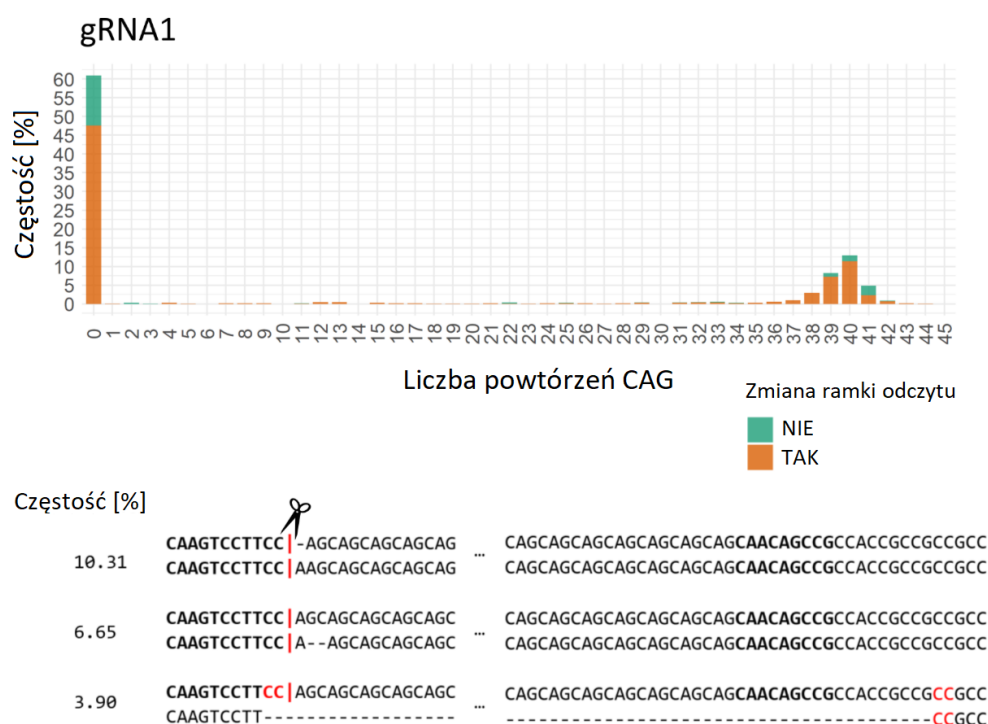


Rysunek 5.1. Efekty naprawy DSB w regionie z powtórzeniami CAG w genie *HTT* w komórkach HEK293T_41CAG. A) Obraz żelu agarozowego z rozdzielonymi produktami amplifikacji regionu powtórzeń CAG edytowanego kompleksami RNP z poszczególnymi gRNA oraz nieedytowanego (kontrola). B) Udział poszczególnych kategorii wyników naprawy DNA po edycji z udziałem poszczególnych gRNA. C) Udział zmian w sekwencjach flankujących ciąg CAG po edycji z udziałem gRNA1 oraz gRNA2. D) Częstość występowania wariantów z określoną liczbą powtórzeń CAG w próbce kontrolnej (nieedytowanej). Rysunek na podstawie Śledziński i wsp. 2024¹⁸².

5.1.1 Przebieg skracania ciągu powtórzeń CAG w wyniku indukcji DSB z udziałem gRNA1

W wyniku indukcji DSB przy użyciu gRNA1 (które rozpoznaje miejsce na początku ciągu powtórzeń CAG, indukując DSB za pierwszą cytozyną pierwszego powtórzenia w ciągu) locus *HTT* zostało edytowane z wysoką wydajnością, co potwierdzają wyniki NGS, wskazujące, że niezmienną liczbą powtórzeń CAG była obecna jedynie w przypadku 4,9% wszystkich odczytów (rysunek 5.1 B). W przypadku 79% odczytów, stwierdzono przesunięcia ORF. Różnych rozmiarów delecje zostały zaobserwowane w 66% odczytów: w przypadku 49% doszło do przesunięcia ORF, w pozostałych przypadkach stanowiących 17% wszystkich odczytów, doszło do delecji, w których ORF została zachowana. W 65% przypadków doszło do zmian w sekwencjach flankujących ciąg powtórzeń CAG (rysunek 5.1 C), zaś wycięcie całego ciągu zostało zaobserwowane w przypadku 61% odczytów. W większości przypadków dochodziło do wycięcia nie tylko samego ciągu, ale również sekwencji flankujących z obu stron ciągu. Najczęściej występującym wariantem (10,3% odczytów) była insercja pojedynczej

adeniny występująca w miejscu indukcji dwuniciowego pęknięcia (rysunek 5.1.1.1). Ponadto, w wynikach sekwencjonowania ujawniono charakterystyczne dla naprawy z udziałem POL θ małe insercje, powstałe przy wykorzystaniu pobliskiego DNA jako matrycy (TINS). Ich przykłady przedstawiono na rysunku 5.1.1.2. Ponadto, zaobserwowano, że znaczna część wariantów charakteryzowała się obecnością krótkich sekwencji mikrohomologicznych (MH) w miejscu ligacji końców DNA (rysunek 5.1.1.3). Wykorzystanie takich regionów do łączenia końców DNA może być wykorzystywane przez POL θ .



Rysunek 5.1.1.1 Częstość występowania wariantów z określoną liczbą powtórzeń CAG w próbce edytowanej z udziałem gRNA1. Poniżej trzy występujące z największą częstością warianty. Na czerwono zaznaczono regiony mikrohomologiczne mające udział w delecjach ciągu powtórzeń.

5.1.3 Przebieg skracania ciągu powtórzeń CAG w wyniku indukcji DSB z udziałem gRNA4

gRNA4

Y-axis: Częstość [%]

X-axis: Liczba powtórzeń CAG

Legend: Zmiana ramki odczytu
 ■ NIE (green)
 ■ TAK (orange)

Liczba powtórzeń CAG	NIE [%]	TAK [%]
0	0	5
1	0	0
2	0	0
3	0	0
4	0	0
5	0	0
6	2	0
7	2	0
8	8	0
9	7	0
10	0	0
11	2	0
12	1	0
13	0	0
14	0	0
15	0	0
16	0	0
17	0	0
18	0	0
19	0	0
20	1	0
21	0	0
22	0	0
23	0	0
24	0	0
25	0	0
26	0	0
27	0	0
28	1	0
29	0	0
30	0	0
31	0	0
32	0	0
33	0	0
34	0	0
35	0	0
36	0	0
37	0	0
38	2	0
39	5	2
40	14	4
41	13	9
42	0	0
43	1	0
44	0	0
45	0	0

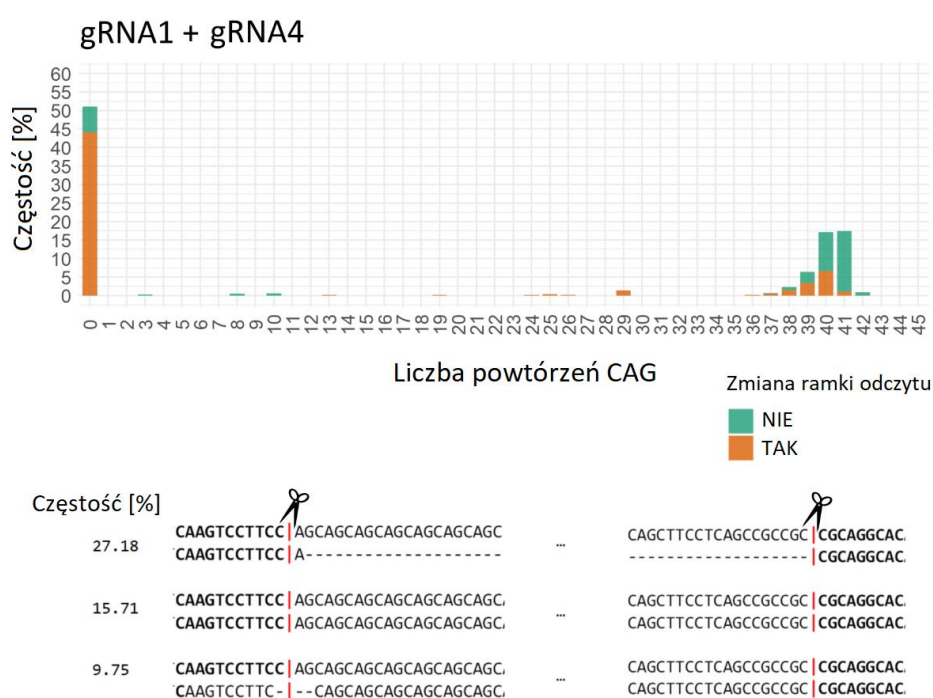
Częstość [%]

11.58	CAAGTCTCTCCAGCAGCAG CAAGTCTCTCCAGCAGCAG	... GCCGCCGCCGCCGCCGCCCTCTCAGTTCTCTCAGCCGCCGC	CGCAGGCAC CGCAGGCAC
11.23	CAAGTCTCTCCAGCAGCAG CAAGTCTCTC---CAGCAG	... GCCGCCGCCGCCGCCGCCCTCTCAGTTCTCTCAGCCGCCGC	CGCAGGCAC CGCAGGCAC
6.91	CAAGTCTCTCCAGCAGCAG CAAGTCTCTCCAGCA---	... GCCGCCGCCGCCGCCGCCCTCTCAGTTCTCTCAGCCGCCGC ---GCCGCCGCCGCCGCCGCCCTCTCAGTTCTCTCAGCCGCCGC	CGCAGGCAC CGCAGGCAC

Rysunek 5.1.3 Częstość występowania wariantów z określoną liczbą powtórzeń CAG w próbie edytowanej z udziałem gRNA4. Poniżej trzy występujące z największą częstością warianty. Na czerwono zaznaczono regiony mikrohologiczne mające udział w delecjach ciągu powtórzeń.

5.1.4 Wycięcie ciągu powtórzeń CAG w wyniku indukcji DSBs z udziałem gRNA1 oraz gRNA4

W przypadku indukcji dwóch dwuniciowych pęknięć DNA flankujących ciąg powtórzeń wycięcie całego ciągu zostało zaobserwowane w przypadku 51% odczytów (rysunek 5.1 B). Najczęstszym wariantem było wycięcie regionu dokładnie pomiędzy dwoma miejscami indukcji DSBs z zachowaną adeniną znajdującą się w pozycji -4 od PAM (27,2%) (rysunek 5.1.4), co jest charakterystyczne dla naprawy DSB indukowanego z udziałem gRNA1 (porównaj rozdział 5.1.1). Przesunięcie ramki odczytu wystąpiło w przypadku 60% odczytów.

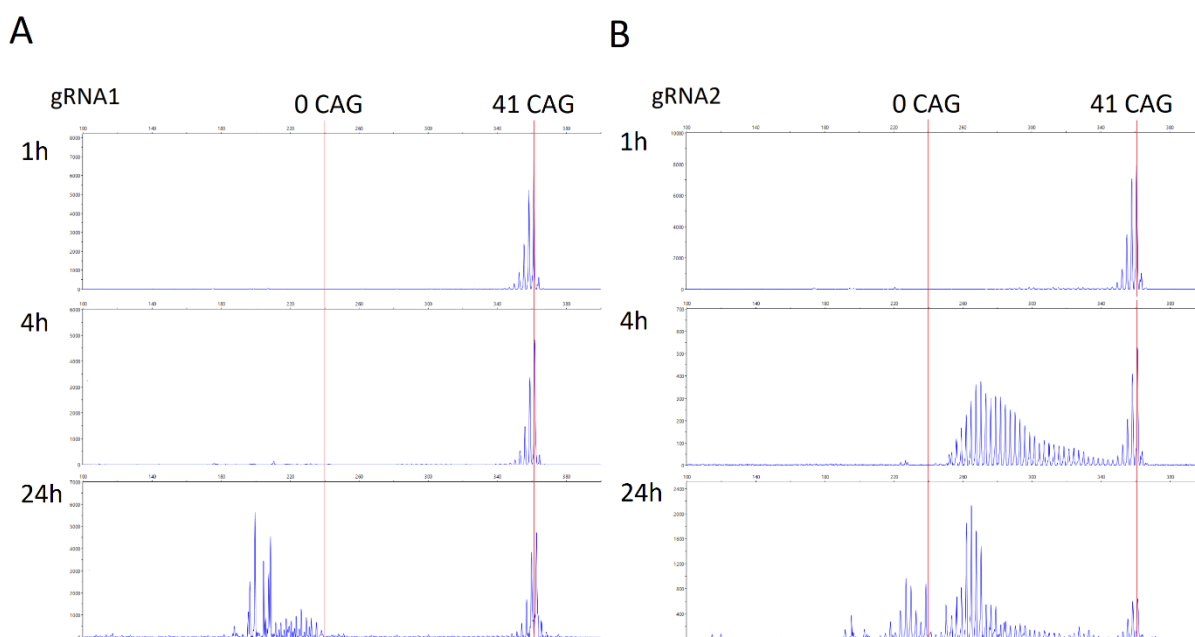


Rysunek 5.1.4 Częstość występowania wariantów z określoną liczbą powtórzeń CAG w próbie edytowanej jednocześnie z udziałem gRNA1 i gRNA4. Poniżej trzy występujące z największą częstością warianty.

5.2 Analiza dynamiki procesu skracania ciągu powtórzeń CAG w wyniku indukcji DSB w genie *HTT*

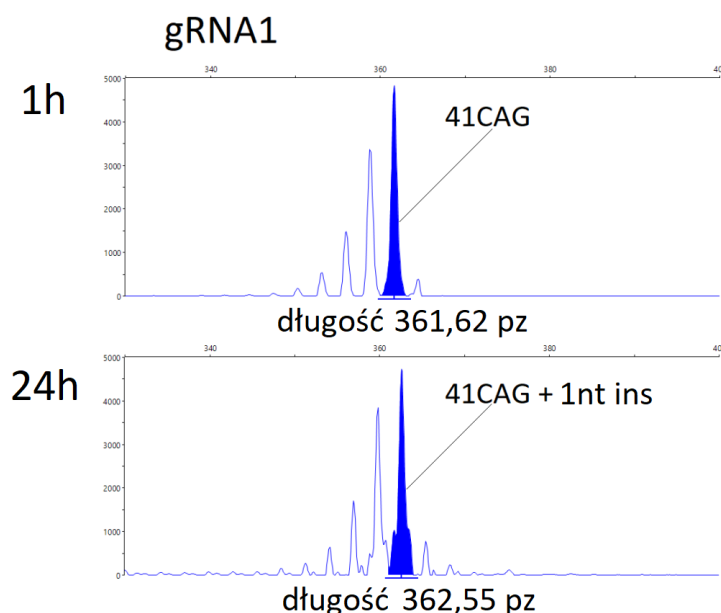
W celu lepszego poznania dynamiki naprawy DSBs powstałych w obrębie powtórzeń CAG dokonano analizy produktów PCR po amplifikacji regionu z powtórzeniami w genie *HTT* w komórkach HEK293T_41CAG, z których izolowano DNA w punktach czasowych (1 h, 4 h, 24 h) po transfekcji RNP. Produkty PCR zostały rozdzielone w elektroforezie kapilarnej dostarczając informacji o przebiegu naprawy w czasie.

W przypadku edycji z udziałem gRNA1 krótsze produkty PCR mogące odpowiadać skróceniom ciągu CAG zostały zaobserwowane dopiero po 24 h od transfekcji (rysunek 5.2.1 A), podczas gdy po edycji z udziałem gRNA2 tego typu produkty PCR zostały zaobserwowane już po 4 godzinach od transfekcji (rysunek 5.2.1 B). Co więcej, skrócone przy użyciu gRNA2 produkty naprawy obserwowane po 4h ulegały dalszym skróceniom, co zostało zaobserwowane po 24 godzinach od transfekcji. Zauważalny jest zanik w czasie produktów o długości między 270 a 330 par zasad oraz silny wzrost udziału produktów o długości między 250 a 260 par zasad, jak również pojawienie się nieobecnych w próbie 4 h produktów o długości między 190 a 250 par zasad. Porównując wyniki po edycji z udziałem gRNA2 po 4 h i 24 h od transfekcji RNP, widoczny jest również spadek wysokości pików odpowiadających produktom wyjściowym w czasie, co wskazuje na ich dalsze procesowanie przez mechanizmy naprawy między 4 a 24 godziną po transfekcji.



Rysunek 5.2.1 Rozdział produktów PCR uzyskanych po amplifikacji regionu z powtórzeniami CAG w genie *HTT* na matrycy DNA izolowanego w opisanych punktach czasowych z komórek HEK293T_41CAG po transfekcji RNP z gRNA1 (A) lub gRNA2 (B). Czerwoną linią zaznaczono wielkość produktu PCR z 41 powtórzeniami CAG oraz bez powtórzeń.

Kolejnym zjawiskiem wykrytym w tej analizie była insercja pojedynczego nukleotydu, zaobserwowana w wynikach po edycji z udziałem gRNA1 i nieobserwowana w wynikach po edycji z udziałem gRNA2 (rysunek 5.2.2). Odpowiada to występującym z dużą częstością w wynikach z NGS wariantom z insercją pojedynczej adeniny po edycji z udziałem gRNA1.

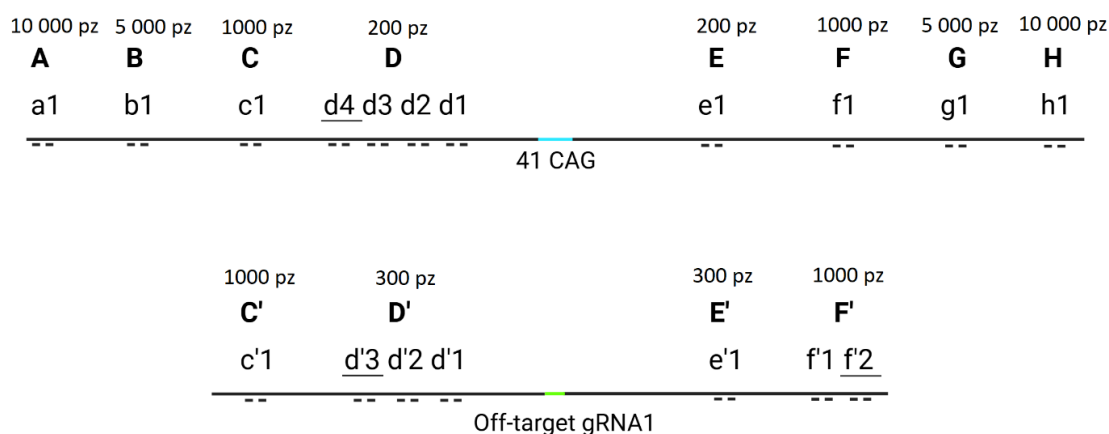


Rysunek 5.2.2 Jednonukleotydowa insercja obserwowana po edycji z udziałem gRNA1 po 24 godzinach od transfekcji komórek.

5.3 Optymalizacja metody określania zasięgu i intensywności resekcji końców DNA w komórkach ludzkich

W celu określenia użyteczności poszczególnych enzymów restrykcyjnych w analizie resekcji końców DNA w komórkach ludzkich, przetestowano poszczególne enzymy inkubując z nimi nieedytowany DNA izolowany z komórek HEK293T_41CAG i fibroblastów. Proces selekcji enzymów restrykcyjnych na podstawie ich efektywności w danych regionach DNA został przedstawiony na rysunku 4.9.

Zaprojektowano po kilka par starterów do amplifikacji regionów znajdujących się w pewnych odległościach od miejsca, które miało podlegać edycji. Były to region powtórzeń CAG w genie *HTT* oraz miejsce off-target dla gRNA1 (chr15:47717511+47717530; hg38). Dla ciągu powtórzeń w genie *HTT* były to miejsca znajdujące się około 10000 pz (regiony a), około 5000 pz (regiony b), około 1000 pz (regiony c) oraz około 200 pz (regiony d) powyżej oraz poniżej (odpowiednio regiony: h, g, f oraz e) tego *locus*. Dla miejsca off-target gRNA1 były to miejsca znajdujące się około 1000 pz (regiony c') i około 300 pz (regiony d') powyżej oraz poniżej (odpowiednio regiony: f' i e') tego *locus*. (rysunek 5.3.1). Dla każdej pary starterów wykonano PCR z gradientem temperatury przyłączania starterów (T_m) (rozdział 4.11.6) i na podstawie oceny produktów PCR po rozdiale elektroforetycznym w żelu agarozowym dobrano optymalne warunki reakcji.



Rysunek 5.3.1 Schematyczne przedstawienie lokalizacji starterów względem edytowanych *loci*: regionu z powtórzeniami CAG w genie *HTT* oraz *locus* off-target dla gRNA1, które zawiera sześć powtórzeń CAG. Małymi literami oznaczono regiony definiowane przez poszczególne startery (tabela 5.3). Oznaczenie wielkimi literami nadano regionom, które ostatecznie zostały wybrane do eksperymentu (podkreślone tam gdzie testowano więcej niż jedną parę starterów).

Po wyborze par starterów, efektywnie amplifikujących dany region i wykorzystujących tę samą T_m , za pomocą programu NEBcutter2.0 wybrano enzymy restrykcyjne rozpoznające sekwencje zawierające się pomiędzy miejscami wiązania starterów, mających służyć do amplifikacji określonego regionu. Enzymy restrykcyjne, które były testowane dla poszczególnych regionów przedstawiono w tabeli 4.9.

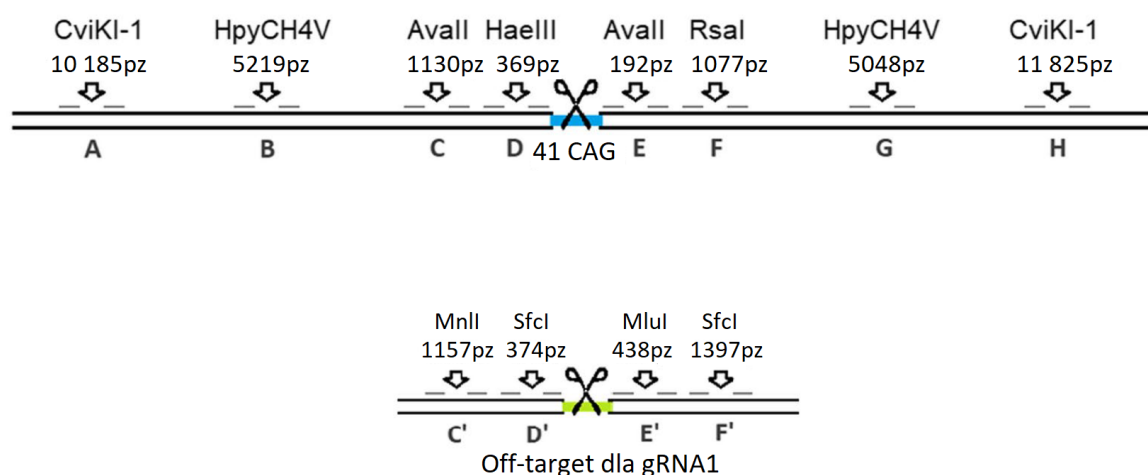
Przetestowano efektywność trawienia poszczególnych enzymów restrykcyjnych w określonych regionach genomu, poprzez porównanie parametrów C_t uzyskanych po amplifikacji DNA izolowanego z komórek nieedytowanych, inkubowanego z danym enzymem i próby nietraktowanej enzymem. Schemat wyboru enzymu efektywnego w danym regionie i przydatnego do analizy przedstawiono na rysunku 4.9. Tabela 5.3 zawiera zestawienie wartości ΔC_t i poziomu resekcji uzyskanych podczas testowania poszczególnych enzymów w określonych regionach. Przyjęto, że w przypadku nieedytowanego DNA wartość resekcji może wynosić maksymalnie 10% co odpowiada różnicy ΔC_t wynoszącej około 4 cykle. Regiony, które zostały ostatecznie wybrane do dalszych eksperymentów, opisano odpowiednimi wielkimi literami (rysunek 5.3.1). Rysunek 5.3.2 przedstawia wybrane do eksperymentu miejsca flankujące region z powtórzeniami CAG w genie *HTT* oraz miejsce off-target dla gRNA1 wraz z zastosowanymi enzymami restrykcyjnymi i odległościami miejsca cięcia enzymu od odpowiedniego miejsca edycji.

Tabela 5.3. Zestawienie wartości ΔC_t i poziomu resekcji uzyskanych podczas testowania poszczególnych enzymów w określonych regionach w komórkach HEK293T_41CAG i w fibroblastach.

HEK293T_41CAG				
Region	Enzym	Użyte startery	ΔC_t	Poziom resekcji
Region a1 (A)	NlaI	IG126, IG 127	4,4	9%
Region a1 (A)	CviKI-1	IG126, IG 127	6,7	0,2%
Region b1 (B)	Bsn3	IG124, IG125	1,33	57%
Region b1 (B)	HpyCH4V	IG124, IG125	5,3	5%
Region c1 (C)	Ava II	IG50, IG51	5,26	5,1%
Region d1	Ava II	IG52, IG53	0,85	71,4%
Region d2	Ava II	IG64, IG65	2,82	24,8%
Region d1	AluI	IG52, IG53	2,46	30,8%

Region d2	AluI	IG64, IG65	0,71	75,9%
Region d1	BstNI	IG52, IG53	1,61	49,4%
Region d2	BstNI	IG64, IG65	5,72	3,7%
Region d3	BstNI	IG66, IG67	3,35	17,9%
Region d4 (D)	HaeIII	IG68, IG69	5,02	6,0%
Region e1 (E)	Ava II	IG46, IG47	5,55	4,2%
Region f1 (F)	RsaI	IG48, IG49	4,67	7,6%
Region g1 (G)	Esp3I	IG128, IG129	0,25	91%
Region g1 (G)	HpyCH4V	IG128, IG129	5,9	3,3%
Region h1 (H)	MspAI	IG130, IG131	-0,08	1,03
Region h1 (H)	CviKI-1	IG130, IG131	4,42	8,9%
Region c'1 (C')	MnII	IG94, IG95	5,19	5,3%
Region c'1 (C')	Sfcl	IG94, IG95	4,19	10,4%
Region d'3 (D')	PstI	IG104, IG105	3,18	19,9%
Region d'3 (D')	BpmI	IG104, IG105	3,56	15,6%
Region d'1	BpmI	IG96, IG97	2,99	22,4%
Region d'1	Eco0109I	IG96, IG97	4,40	9,0%
Region d'2	PstI	IG114, IG115	2,55	29%
Region d'3 (D')	Sfcl	IG104, IG105	5,59	4,1%
Region e'1 (E')	MluI	IG110, IG111	5,33	4,9%
Region f'1	AluI	IG108, IG109	3,63	14,9%
Region f'1	PstI	IG108, IG109	0,94	68,5%
Region f'1	HindIII	IG108, IG109	3,02	21,9%
Region f'2 (F')	BstNI	IG100, IG101	3,28	18,7%
Region f'2 (F')	PstI	IG100, IG101	3,99	11,8%
Region f'2 (F')	AluI	IG100, IG101	0,74	74,9%
Region f'2 (F')	HindIII	IG100, IG101	3,53	15,9%
Region f'2 (F')	Sfcl	IG100, IG101	4,7	7,4%
Fibroblasty GM04208				
Region	Enzym	Użyte startery	dCt	Wartość resekcji
Region A	CviKI-1	126,127	7,9	0,8%
Region B	HpyCH4V	124,125	4,75	7,2%

Region C	AvaII	IG50, IG51	4,31	9,6%
Region D	HaeIII	IG68, IG69	6,59	2%
Region D	SmaI	IG68, IG69	4,86	6,7%
Region D	MspA1I	IG68, IG69	0,14	95%
Region D	HpyAV	IG68, IG69	6,52	2,2%
Region E	AvaII	IG46, IG47	4,55	8,2%
Region F	RsaI	IG48, IG49	4,8	6,9%
Region G	HpyCH4V	IG128, IG129	5,74	3,7%
Region H	CviKI-1	IG130, IG131	6,59	2,1%



Rysunek 5.3.2 Przedstawienie miejsc flankujących region z powtórzeniami CAG w genie *HTT* oraz miejsce off-target dla gRNA1 wraz z użytymi enzymami restrykcyjnymi i odległościami miejsca cięcia enzymu od odpowiedniego miejsca edycji (w przypadku *locus HTT* jest to odległość od miejsca cięcia enzymu do początku lub końca ciągu CAG, odpowiednio dla regionów znajdujących się powyżej oraz poniżej edytowanego miejsca). W regionie D w przypadku fibroblastów zastosowano enzym HpyAV zamiast HaeIII. Wyjaśnienie w tekście.

Ze względu na to, że wyniki resekcji w fibroblastach w regionie D po zastosowaniu enzymu HaeIII wskazywały na brak resekcji w tym regionie (wartości resekcji z trzech powtórzeń biologicznych: 3,1%; 0,1%; 2%) pomimo, że w regionach sąsiednich resekcje zostały wykryte, postanowiono przetestować inny enzym, mający miejsce restrykcyjne w tym samym regionie. Zastosowanie HpyAV ujawniło obecność resekcji w regionie D w fibroblastach.

Wynik uzyskany przy użyciu enzymu HaeIII wynikał z wysokiej wartości Ct w próbie traktowanej enzymem, co świadczyłoby o obecności dwuniciowego DNA w regionie D, który został efektywnie strawiony przez enzym. Obecność dwuniciowego odcinka w tym miejscu mogła wynikać z utworzenia struktury drugorzędowej przez nić DNA, która uległa resekcji. Następnie dwuniciowy fragment zawierający miejsce restrykcyjne HaeIII został rozpoznany przez ten enzym i uległ przecięciu.

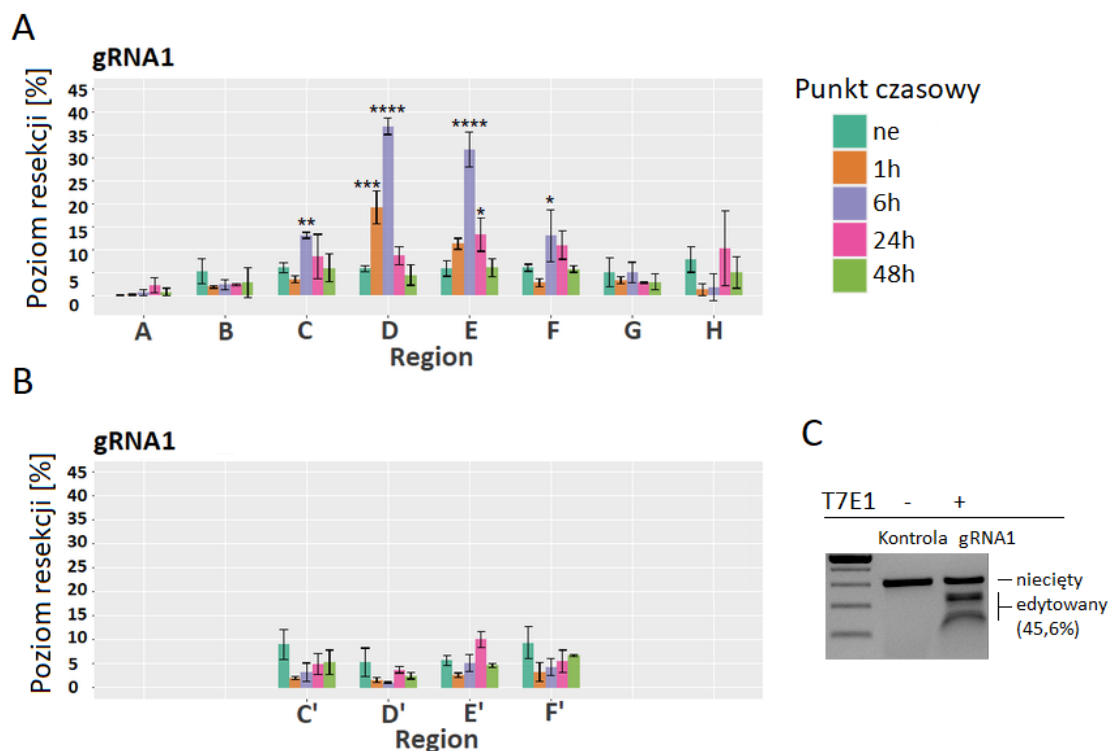
5.4 Analiza resekcji końców DNA w czasie naprawy DSBs w regionach powtórzeń CAG w różnych modelach komórkowych

Aby odpowiedzieć na pytanie czy podczas naprawy DSB w regionach mikrosatelitarnych w komórkach ludzkich dochodzi do resekcji, przeprowadzono eksperymenty na modelu HEK293T_41CAG oraz fibroblastach posiadających 21 oraz 44 powtórzenia CAG w genie *HTT*. Aby określić wpływ długości ciągu powtórzeń na intensywność i zasięg resekcji, wyniki dla *locus HTT* porównano z wynikami dla *locus off-target* dla gRNA1, zawierającego jedynie 6 powtórzeń CAG.

5.4.1 Analiza resekcji końców DNA w czasie naprawy DSB po edycji z udziałem gRNA1 w linii komórkowej HEK293T_41CAG

Badanie resekcji końców DNA zachodzących w czasie naprawy DSB po edycji regionu z powtórzeniami CAG w genie *HTT* z udziałem gRNA1 w komórkach HEK293T_41CAG ujawniło, że do resekcji dochodzi w regionach bliskich edytowanemu miejscu (rysunek 5.4.1 A). Maksymalny wykryty zasięg resekcji w tym przypadku sięgał 1000 nukleotydów powyżej i poniżej miejsca cięcia (regiony C i F). Jednak z największą intensywnością resekcje zachodziły w obszarach najbliższych, edytowanemu miejscu (regiony D i E) i osiągały tu wartość około 35% po 6 h od transfekcji RNP. Dla większości regionów (poza regionem E) wartość resekcji nie była istotnie różna od wartości dla nieedytowanej kontroli już po 24 h od transfekcji, co może świadczyć o zakończeniu procesu naprawy DNA w przeciągu tego czasu. Jest to zbieżne z wynikami analizy dynamiki naprawy DNA dla gRNA1 (rysunek 5.2.1 A). Uzyskane wyniki poziomu resekcji końców DNA w różnych punktach czasowych i regionach flankujących ciąg powtórzeń CAG zostały przedstawione w tabeli 5.4.1.1.

W przypadku badania resekcji w *locus* off-target dla gRNA1, które zawiera sześć powtórzeń CAG, resekcje nie zostały wykryte w żadnym czasie od transfekcji komórek, pomimo stwierdzenia efektywnej edycji tego miejsca (rysunek 5.4.1 B i C). Wyniki poziomu resekcji w różnych punktach czasowych i regionach flankujących region off-target dla gRNA1 zostały przedstawione w tabeli 5.4.1.2.



Rysunek 5.4.1 Poziom i zasięg resekcji końców DNA w czasie po edycji z udziałem gRNA1 w *locus* *HTT* (A) oraz w miejscu off-target dla gRNA1 (B) w komórkach HEK293T_41CAG. C) Potwierdzenie edycji *locus* off-target dla gRNA1 metodą T7E1. Wyniki przedstawiają średnią z trzech powtórzeń biologicznych wraz z odchyleniem standardowym. Istotność statystyczna wyników została określona za pomocą jednoczynnikowej analizy wariancji (one-way ANOVA) * $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$; **** $p \leq 0,0001$.

Tabela 5.4.1.1 Wyniki poziomu resekcji w różnych punktach czasowych od transfekcji RNP z gRNA1 oraz kontroli (nietransfekowanej RNP) w regionach flankujących region z powtórzeniami CAG w genie *HTT* w komórkach HEK293T_41CAG dla trzech powtórzeń biologicznych eksperymentu wraz z wartością średnią. Wyniki posłużyły do sporządzenia wykresu z rysunku 5.4.1.A.

Region A	1	2	3	Średnia
1 h	0,2%	0,1%	0,3%	0,2%
6 h	0,2%	0,4%	1,5%	0,7%
24 h	0,4%	3,1%	3,4%	2,3%
48 h	0,2%	0,4%	1,8%	0,8%
Kontrola	0,2%	0,2%	0,0%	0,1%
Region B	1	2	3	Średnia
1 h	1,7%	2,3%	1,8%	1,9%
6 h	3,2%	1,2%	2,8%	2,4%
24 h	2,2%	2,6%	2,5%	2,4%
48 h	0,9%	1,1%	6,6%	2,8%
Kontrola	5,0%	8,1%	2,8%	5,3%
Region C	1	2	3	Średnia
1 h	3,3%	4,5%	2,9%	3,6%
6 h	13,7%	13,4%	12,4%	13,2%
24 h	8,0%	5,8%	11,6%	8,5%
48 h	4,2%	9,5%	4,3%	6,0%
Kontrola	7,4%	5,3%	5,6%	6,1%
Region D	1	2	3	Średnia
1 h	21,5%	15,0%	21,1%	19,2%
6 h	37,8%	38,1%	34,9%	36,9%
24 h	10,8%	6,9%	8,3%	8,7%
48 h	2,0%	6,0%	5,6%	4,5%
Kontrola	6,0%	5,1%	6,5%	5,9%
Region E	1	2	3	Średnia
1 h	12,6%	10,3%	11,0%	11,3%
6 h	35,2%	27,7%	32,5%	31,8%
24 h	10,0%	12,7%	17,3%	13,3%
48 h	4,6%	8,4%	5,4%	6,1%
Kontrola	7,5%	4,2%	6,0%	5,9%

Region F	1	2	3	Średnia
1 h	3,6%	1,9%	3,0%	2,8%
6 h	6,5%	16,8%	15,8%	13,0%
24 h	7,9%	11,0%	14,1%	11,0%
48 h	5,1%	6,5%	5,8%	5,8%
Kontrola	5,9%	6,9%	5,4%	6,1%
Region G	1	2	3	Średnia
1 h	3,1%	4,2%	2,8%	3,4%
6 h	7,4%	4,9%	2,9%	5,1%
24 h	2,8%	2,8%	3,0%	2,9%
48 h	1,0%	3,7%	4,3%	3,0%
Kontrola	3,3%	3,1%	8,7%	5,0%
Region H	1	2	3	Średnia
1 h	0,2%	2,7%	1,0%	1,3%
6 h	0,1%	0,2%	5,3%	1,8%
24 h	1,1%	16,4%	13,5%	10,3%
48 h	2,5%	8,9%	3,7%	5,0%
Kontrola	8,3%	4,9%	10,4%	7,9%

Tabela 5.4.1.2 Wyniki poziomej resekcji w różnych punktach czasowych od transfekcji RNP z gRNA1 oraz kontroli (nietransfekowanej RNP) w regionach flankujących *locus* off-target dla gRNA1 w komórkach HEK293T_41CAG dla trzech powtórzeń biologicznych eksperymentu wraz z wartością średnią. Wyniki posłużyły do sporządzenia wykresu z rysunku 5.4.1.B.

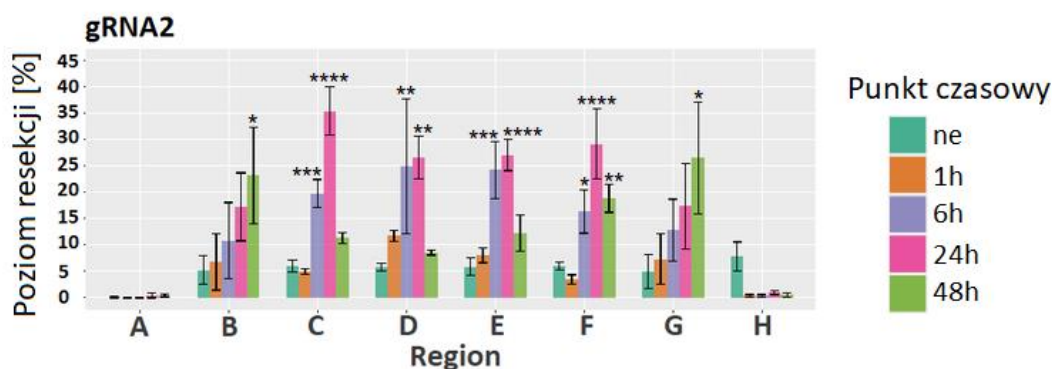
Region C'	1	2	3	Średnia
1 h	2,1%	2,4%	1,6%	2,1%
6 h	2,2%	1,7%	6,0%	3,3%
24 h	2,1%	7,4%	5,6%	5,0%
48 h	2,1%	8,3%	5,8%	5,4%
Kontrola	9,1%	12,8%	5,3%	9,1%
Region D'	1	2	3	Średnia
1 h	1,6%	1,0%	2,3%	1,6%
6 h	1,1%	1,0%	1,4%	1,1%

24 h	4,5%	4,0%	2,9%	3,8%
48 h	1,6%	3,0%	3,1%	2,6%
Kontrola	2,7%	4,0%	9,5%	5,4%
Region E'	1	2	3	Średnia
1 h	3,0%	3,0%	2,0%	2,7%
6 h	4,0%	4,0%	7,7%	5,2%
24 h	9,0%	9,0%	12,5%	10,2%
48 h	5,0%	5,0%	4,1%	4,7%
Kontrola	5,3%	4,8%	7,2%	5,8%
Region F'	1	2	3	Średnia
1 h	6,0%	1,7%	2,3%	3,4%
6 h	6,8%	3,9%	2,5%	4,4%
24 h	5,0%	3,1%	8,7%	5,6%
48 h	6,7%	7,1%	6,7%	6,8%
Kontrola	7,4%	6,9%	14,1%	9,5%

5.4.2 Analiza resekcji końców DNA w czasie naprawy DSB po edycji z udziałem gRNA2 w linii komórkowej HEK293T_41CAG

Analiza resekcji w komórkach HEK293T_41CAG w *locus HTT* po edycji z udziałem gRNA2 ujawniła znacznie dalszy zasięg resekcji (rysunek 5.4.2). Regiony jednoniciowego DNA zostały w tym przypadku wykryte w odległości 5000 pz od edytowanego miejsca (regiony B i G). Nie zostały zaś wykryte w regionach dystalnych, oddalonych o ok. 10000 pz od ciągu powtórzeń CAG (regiony A i H). Największa intensywność resekcji została wykryta 24 h po transfekcji i osiągnęła wartość ok. 35% w regionie C oddalonym o ok. 1000 pz od edytowanego *locus*. Wyniki wskazują na inną dynamikę naprawy w porównaniu do naprawy po edycji z udziałem gRNA1 jako, że po edycji z udziałem gRNA2 najintensywniej resekcje zachodziły 24 h po transfekcji i były wykrywane nawet po 48 h. Świadczy to tym, że naprawa nie została jeszcze ukończona w tym czasie. Dłuższy czas naprawy DNA po edycji z udziałem gRNA2 w porównaniu do gRNA1 może wynikać z następujących po pierwotnym cięciu DNA kolejnych cięć indukowanych przez kompleks RNP, zachodzących aż do skrócenia ciągu CAG do długości, która nie będzie wystarczająca do związania ponownie kompleksu Cas9-gRNA2 lub do momentu degradacji kompleksu w komórce. Wynik ten jest zgodny z analizą dynamiki naprawy DNA po edycji z udziałem gRNA2 (rysunek 5.2.1 B). Uzyskane wyniki poziomu

resekcji końców DNA w różnych punktach czasowych i regionach flankujących ciąg powtórzeń CAG w genie *HTT* zostały przedstawione w tabeli 5.4.2.



Rysunek 5.4.2 Poziom i zasięg resekcji końców DNA w czasie po edycji z udziałem gRNA2 w *locus HTT* w komórkach HEK293T_41CAG. Wyniki przedstawiają średnią z trzech powtórzeń biologicznych wraz z odchyleniem standardowym. Istotność statystyczna wyników została określona za pomocą jednoczynnikowej analizy wariancji (one-way ANOVA) * $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$; **** $p \leq 0,0001$.

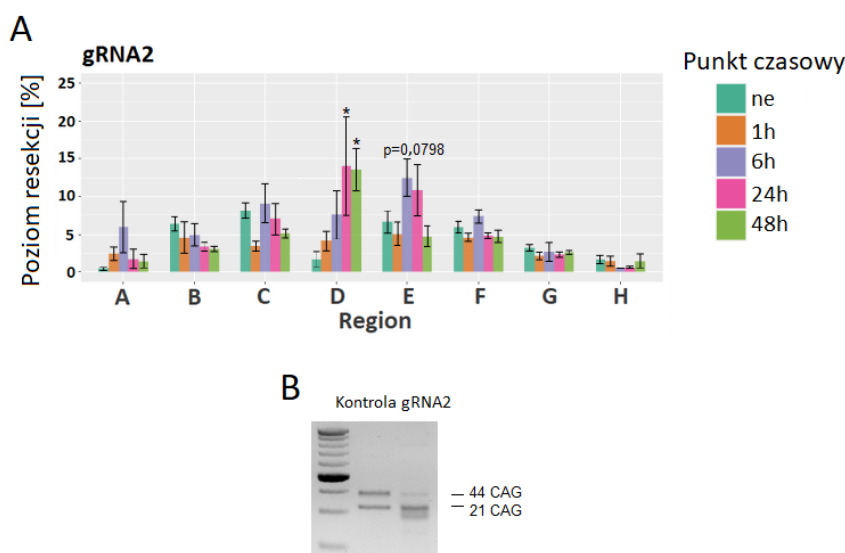
Tabela 5.4.2 Wyniki poziomu resekcji w różnych punktach czasowych od transfekcji RNP z gRNA2 oraz kontroli (nietransfekowanej RNP) w regionach flankujących region z powtórzeniami CAG w genie *HTT* w komórkach HEK293T_41CAG dla trzech powtórzeń biologicznych eksperymentu wraz z wartością średnią. Wyniki posłużyły do sporządzenia wykresu z rysunku 5.4.2.

Region A	1	2	3	Średnia
1 h	0,1%	0,1%	0,0%	0,1%
6 h	0,1%	0,0%	0,1%	0,1%
24 h	0,3%	0,2%	1,1%	0,5%
48 h	0,7%	0,4%	0,3%	0,5%
Kontrola	0,2%	0,2%	0,0%	0,1%
Region B	1	2	3	Średnia
1 h	13,1%	3,7%	3,8%	6,9%
6 h	5,6%	19,2%	8,0%	10,9%
24 h	16,8%	24,0%	11,1%	17,3%
48 h	28,8%	28,3%	12,6%	23,2%
Kontrola	5,0%	8,1%	2,8%	5,3%

Region C	1	2	3	Średnia
1 h	5,1%	5,4%	4,4%	5%
6 h	22,8%	19,3%	17,5%	20%
24 h	33,7%	32,1%	40,7%	36%
48 h	12,0%	12,0%	10,2%	11%
Kontrola	7,4%	5,3%	5,6%	6%
Region D	1	2	3	Średnia
1 h	10,7%	12,8%	11,8%	12%
6 h	14,8%	39,4%	20,8%	25%
24 h	28,5%	22,0%	29,5%	27%
48 h	8,2%	8,5%	9,1%	9%
Kontrola	6,0%	5,1%	6,5%	6%
Region E	1	2	3	Średnia
1 h	9,6%	6,8%	7,9%	8%
6 h	28,0%	18,1%	26,9%	24%
24 h	30,4%	24,7%	26,2%	27%
48 h	11,2%	16,2%	9,5%	12%
Kontrola	7,5%	4,2%	6,0%	6%
Region F	1	2	3	Średnia
1 h	2,7%	4,4%	3,5%	4%
6 h	13,7%	14,5%	21,2%	16%
24 h	21,8%	31,2%	34,8%	29%
48 h	20,9%	20,1%	15,9%	19%
Kontrola	5,9%	6,9%	5,4%	6%
Region G	1	2	3	Średnia
1 h	4,7%	12,9%	4,5%	7%
6 h	10,0%	19,7%	9,1%	13%
24 h	12,9%	26,8%	12,5%	17%
48 h	32,8%	32,6%	14,3%	27%
Kontrola	3,3%	3,1%	8,7%	5%
Region H	1	2	3	Średnia
1 h	0,1%	0,4%	0,6%	0%
6 h	0,3%	0,7%	0,1%	0%
24 h	0,7%	1,4%	1,0%	1%
48 h	0,1%	0,8%	0,7%	1%
Kontrola	8,3%	4,9%	10,4%	8%

5.4.3 Analiza resekcji końców DNA w czasie naprawy DSB po edycji z udziałem gRNA2 w linii komórkowej fibroblastów HD

Analiza resekcji została wykonana również w fibroblastach posiadających 21 oraz 44 powtórzenia CAG w genie *HTT*. Aby określić poziom resekcji użyto kompleksu RNP z gRNA2. W przypadku edycji fibroblastów intensywność resekcji osiągała poziom o około połowę niższy w porównaniu do resekcji w komórkach HEK293T_41CAG tj. około 15% w regionie znajdującym się w odległości ok. 350 pz od miejsca cięcia (region D) (rysunek 5.4.3 A) Relatywnie wysoki poziom resekcji został zaobserwowany w tym regionie również po 48 h od transfekcji podobnie jak w komórkach HEK293T_41CAG, jednak nie odpowiadał mu równie wysoki poziom resekcji w symetrycznie oddalonym od edytowanego miejsca regionie E. W bardziej oddalonych od edytowanego *locus* regionach nie wykryto istotnych statystycznie różnic w poziomach resekcji pomiędzy próbkami po edycji i nieedytowanymi. Niższy zakres resekcji może wynikać z faktu heterozygotyczności użytej linii komórkowej posiadającej również allel z krótkim ciągiem powtórzeń, którego edycja mogła zachodzić z niską efektywnością (rysunek 5.4.3 B). Uzyskane wyniki poziomu resekcji końców DNA w różnych punktach czasowych i regionach flankujących ciąg powtórzeń CAG w genie *HTT* w fibroblastach zostały przedstawione w tabeli 5.4.3.



Rysunek 5.4.3 Poziom i zasięg resekcji końców DNA w czasie po edycji z udziałem gRNA2 w regionie z powtórzeniami CAG w genie *HTT* w fibroblastach (A). B) Produkty PCR powstałe po amplifikacji *locus HTT* na matrycy DNA z fibroblastów nieedytowanych (kontrola) i edytowanych z udziałem gRNA2. Wyniki przedstawiają średnią z trzech powtórzeń

biologicznych wraz z odchyleniem standardowym. Istotność statystyczna wyników została określona za pomocą jednoczynnikowej analizy wariancji (one-way ANOVA) $*p \leq 0,05$.

Tabela 5.4.3 Wyniki poziomu resekcji w różnych punktach czasowych od transfekcji RNP z gRNA2 oraz kontroli (nietransfekowanej RNP) w regionach flankujących region z powtórzeniami CAG w genie *HTT* w fibroblastach dla trzech powtórzeń biologicznych eksperymentu wraz z wartością średnią. Wyniki posłużyły do sporządzenia wykresu z rysunku 5.4.3.

Region A	1	2	3	Średnia
1 h	3,6%	2,2%	1,6%	2,5%
6 h	1,4%	9,5%	7,1%	6,0%
24 h	0,5%	2,0%	3,1%	1,8%
48 h	1,2%	0,5%	2,7%	1,5%
Kontrola	0,8%	0,1%	0,5%	0,5%
Region B	1	2	3	Średnia
1 h	4,4%	7,2%	2,2%	4,6%
6 h	4,1%	7,1%	3,7%	4,9%
24 h	2,8%	4,2%	3,1%	3,4%
48 h	3,2%	3,4%	2,6%	3,1%
Kontrola	7,2%	6,9%	5,2%	6,4%
Region C	1	2	3	Średnia
1 h	3,9%	3,8%	2,5%	3,4%
6 h	11,3%	10,5%	5,5%	9,1%
24 h	9,1%	7,9%	4,2%	7,1%
48 h	5,5%	5,6%	4,4%	5,2%
Kontrola	9,6%	7,3%	7,6%	8,2%
Region D	1	2	3	Średnia
1 h	5,6%	2,4%	4,3%	4,1%
6 h	7,5%	3,7%	11,5%	7,6%
24 h	23,2%	8,1%	10,9%	14,0%
48 h	17,3%	13,2%	10,3%	13,6%
Kontrola	0,3%	2,6%	2,3%	1,7%
Region E	1	2	3	Średnia
1 h	5,5%	6,7%	3,0%	5,1%
6 h	13,3%	15,0%	9,2%	12,5%

24 h	13,7%	12,7%	6,1%	10,8%
48 h	4,9%	6,4%	3,0%	4,8%
Kontrola	4,8%	7,0%	8,2%	6,7%
Region F	1	2	3	Średnia
1 h	4,1%	5,5%	4,3%	4,6%
6 h	7,1%	8,6%	6,5%	7,4%
24 h	4,4%	5,1%	5,2%	4,9%
48 h	3,7%	4,7%	5,8%	4,7%
Kontrola	5,1%	6,9%	5,9%	6,0%
Region G	1	2	3	Średnia
1 h	2,4%	2,6%	1,5%	2,2%
6 h	4,4%	2,0%	1,7%	2,7%
24 h	2,1%	2,8%	2,1%	2,3%
48 h	2,6%	3,0%	2,4%	2,6%
Kontrola	3,7%	3,1%	2,7%	3,2%
Region H	1	2	3	Średnia
1 h	0,7%	2,3%	1,5%	1,5%
6 h	0,6%	0,5%	0,5%	0,5%
24 h	0,5%	0,9%	0,6%	0,7%
48 h	2,8%	0,7%	0,9%	1,5%
Kontrola	2,1%	2,1%	1,0%	1,7%

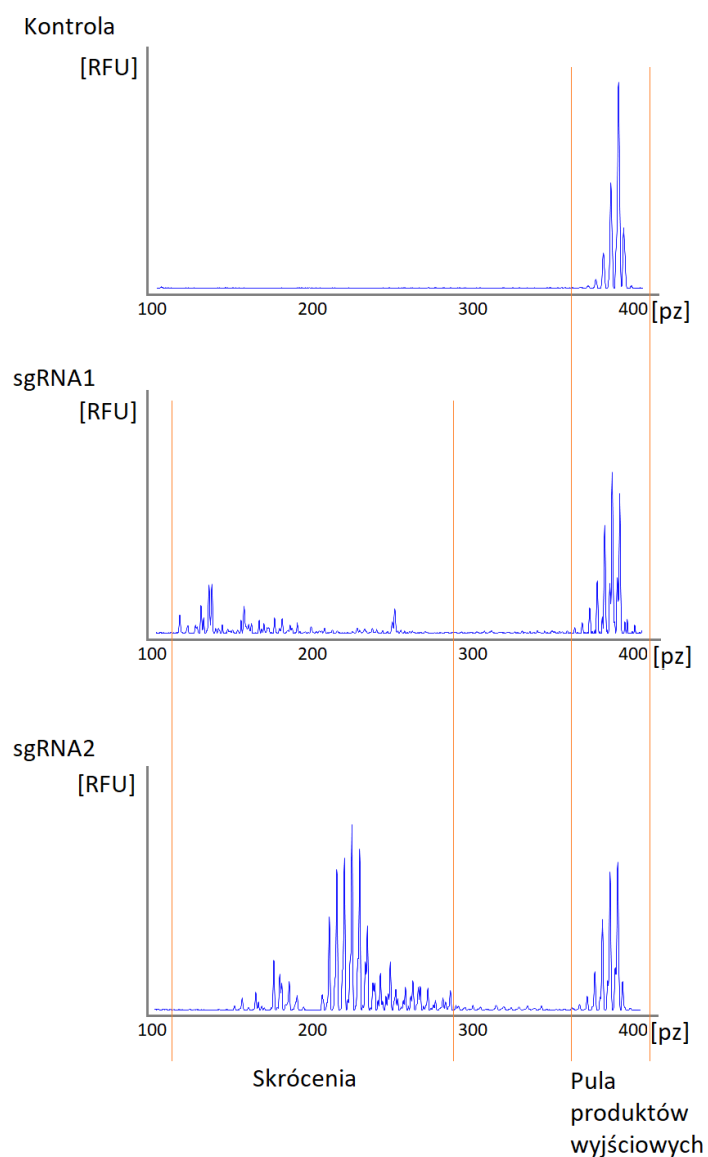
W wyniku analizy resekcji końców DNA w czasie naprawy DSBs w regionach z powtórzeniami CAG ujawniono różnice w ich zasięgu i intensywności, jak również przebiegu w czasie wynikające z różnicy w miejscu indukcji DSB. Większe zakresy resekcji oraz wyższą jej intensywność stwierdzono w komórkach HEK293T_41CAG w porównaniu do linii fibroblastów. Ponadto stwierdzono wpływ długości ciągu powtórzeń CAG na intensywność resekcji, wskazujące, że dłuższe ciągi powtórzeń promują mechanizmy naprawy wykorzystujące resekcję końców DNA.

5.5 Analiza udziału białek naprawczych w skracaniu ciągu powtórzeń CAG w genie *HTT*

W celu zbadania udziału poszczególnych białek przypisanych do różnych ścieżek naprawy DNA, w skracaniu ciągu powtórzeń CAG w genie *HTT* przeprowadzono eksperymenty wykorzystujące inhibitory chemiczne wybranych białek oraz wyciszanie ekspresji poszczególnych genów kodujących białka naprawcze za pomocą interferencji RNA w komórkach HEK293T_41CAG. Wyboru białek dokonano na podstawie literatury, w której sugerowano ich udział w skracaniu ciągów powtórzeń CAG/CTG w komórkach drożdży lub ludzkich^{187,191,214} oraz na podstawie własnych badań, w których potwierdzono zaangażowanie w naprawę regionów z powtórzeniami CAG mechanizmów wykorzystujących resekcje końców DNA (rozdział 5.4). Białka, których udział w skracaniu ciągu CAG w genie *HTT* był badany przedstawiono w tabeli 5.5. Produkty PCR uzyskane po amplifikacji edytowanego *locus* były rozdzielane w elektroforezie kapilarnej, co pozwoliło na ilościową analizę skróceń ciągu powtórzeń (rysunek 5.5). Metoda obliczania redukcji udziału produktów skróconych względem kontroli została przedstawiona w rozdziale 4.10.

Tabela 5.5 Białka naprawy DNA, dla których wykonano badanie ich udziału w skracaniu ciągu powtórzeń CAG w genie *HTT*.

Nazwa białka	Mechanizm
MRE11	Resekcje
CTIP	Resekcje
KU70/KU80	NHEJ
Artemis	NHEJ
POLθ	TMEJ
EXO1	Resekcje

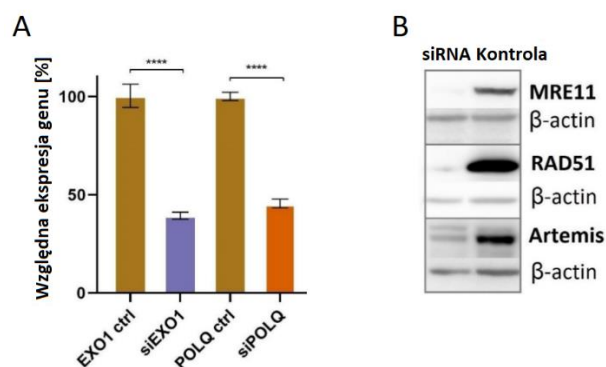


Rysunek 5.5 Elektroforegram przedstawiający rozdział produktów PCR po amplifikacji *locus HTT* w komórkach HEK293T_41CAG nieedytowanego oraz po edycji z udziałem sgRNA1 i sgRNA2.

5.5.1 Analiza udziału białek naprawczych w skracaniu ciągu powtórzeń CAG po indukcji DSB przy użyciu gRNA1

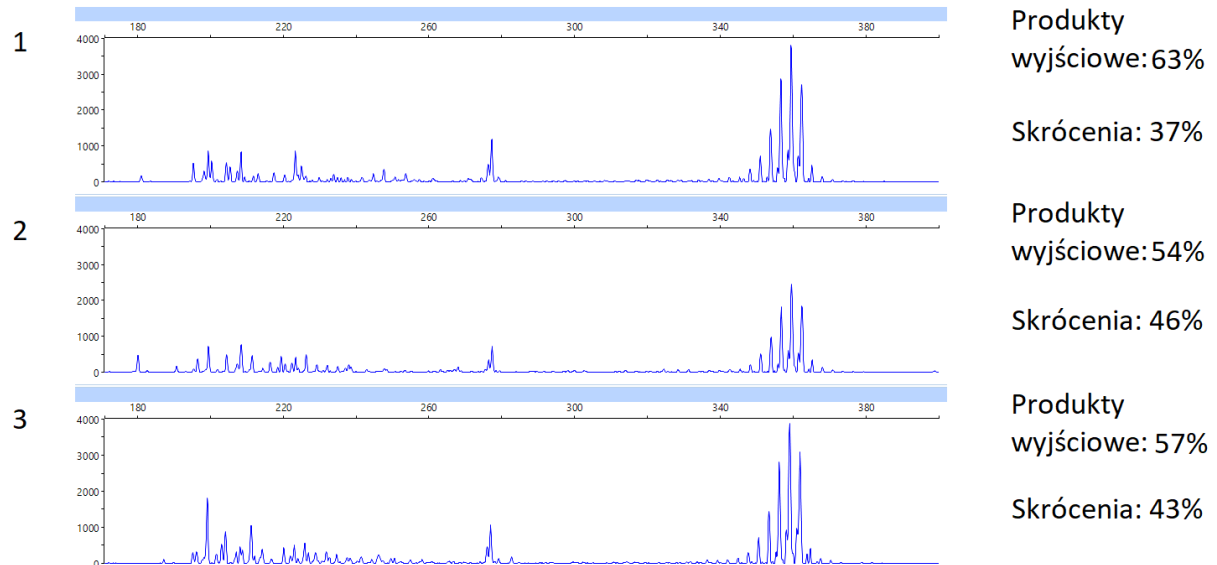
W przypadku edycji z udziałem gRNA1 zarówno inhibicja chemiczna POLθ- kluczowego białka w naprawie TMEJ, jak i wyciszenie kodującego to białko genu *POLQ* do poziomu ok. 45% (rysunek 5.5.1.1 A) spowodowały spadek udziału produktów skróconych w całej puli produktów PCR w porównaniu do kontroli o ok. 30% (rysunek 5.5.1.8). Podobny wynik został

uzyskany w przypadku wyciszenia do poziomu ok. 39% genu kodującego nukleazę EXO1, biorącą udział w resekcjach końców DNA. Po zastosowania inhibitora białek KU70 i KU80 (białka zaangażowane w NHEJ) uzyskano spadek udziału produktów skróconych względem próby kontrolnej wynoszący 18%. W przypadku pozostałych białek nie stwierdzono istotnych statystycznie różnic między kontrolą a próbą traktowaną, pomimo efektywnego wyciszenia poszczególnych genów (rysunek 5.5.1.1 B). Elektroforegramy dla poszczególnych powtórzeń biologicznych eksperymentu z udziałem sgRNA1, wraz z procentowym udziałem produktów skróconych i wyjściowych przedstawiono na poniższych rysunkach (5.5.1.2-5.5.1.7). Wykresy podsumowujące, wraz z obliczonym procentowym obniżeniem udziału produktów skróconych względem odpowiedniej kontroli przedstawiono na rysunku 5.5.1.8.

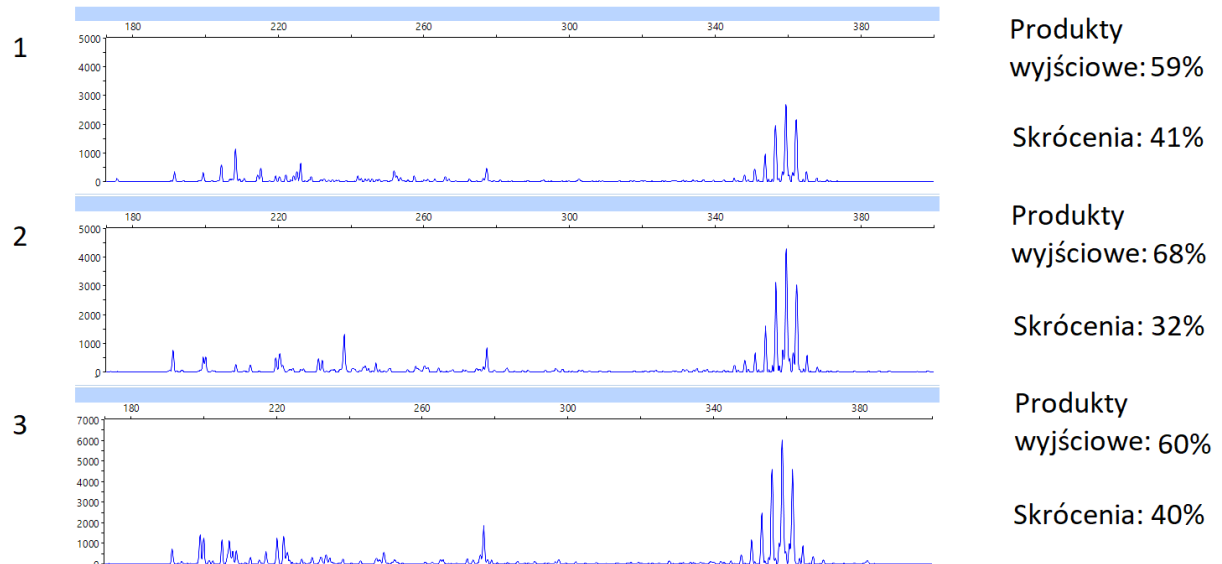


Rysunek 5.5.1.1 Efekt wyciszenia genów kodujących poszczególne białka naprawy DNA przy użyciu siRNA określony na poziomie mRNA (A) lub białka (B). Wyciszenie genu na poziomie mRNA zostało wykonane ze względu na brak specyficznych przeciwciał pozwalających na detekcję białek EXO1 i POL θ metodą western blot. Wyniki względnej ekspresji genów są średnią z trzech powtórzeń biologicznych. Istotność statystyczna wyników została określona za pomocą testu t-Studenta, **** $p \leq 0,0001$.

Kontrola

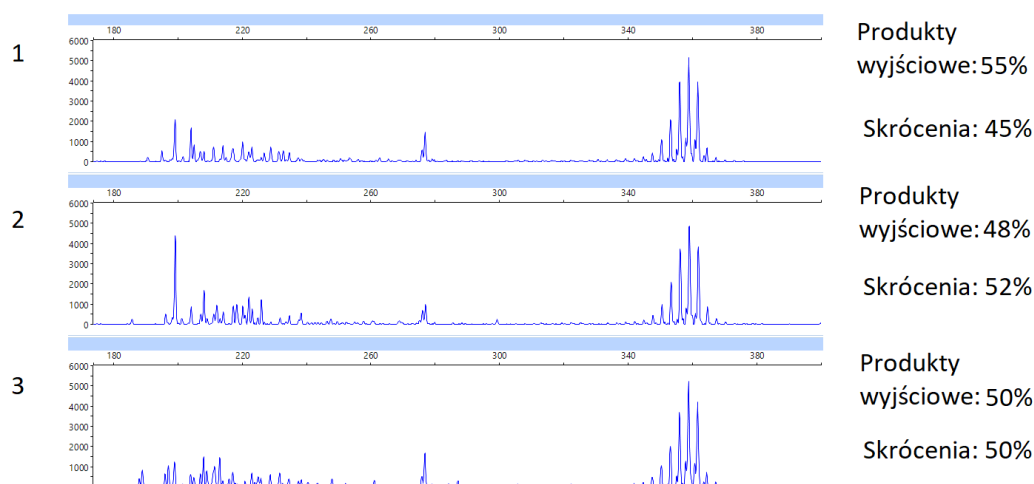


Inhibicja CTIP

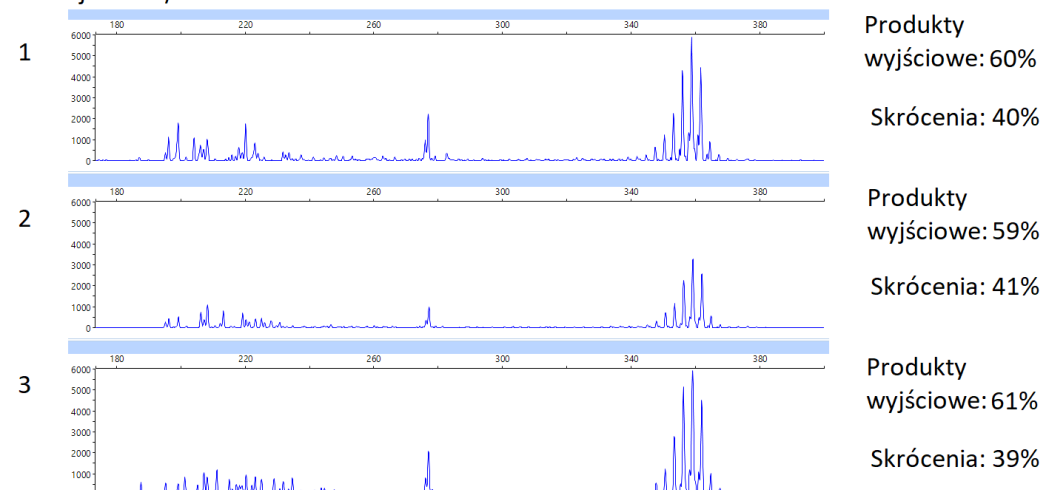


Rysunek 5.5.1.2 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT* w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA1, traktowanych inhibitorem CTIP oraz bez udziału inhibitora (kontrola).

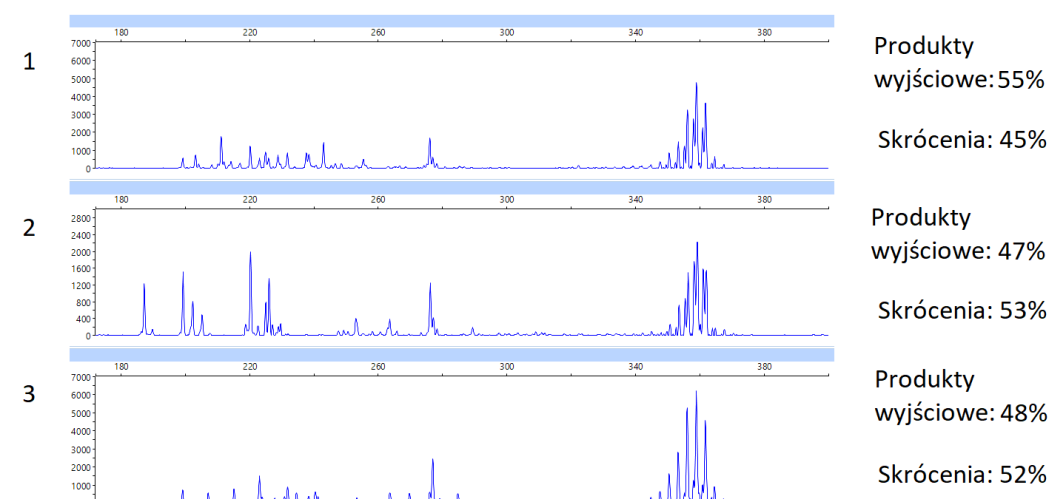
Kontrola



Inhibicja KU70/KU80

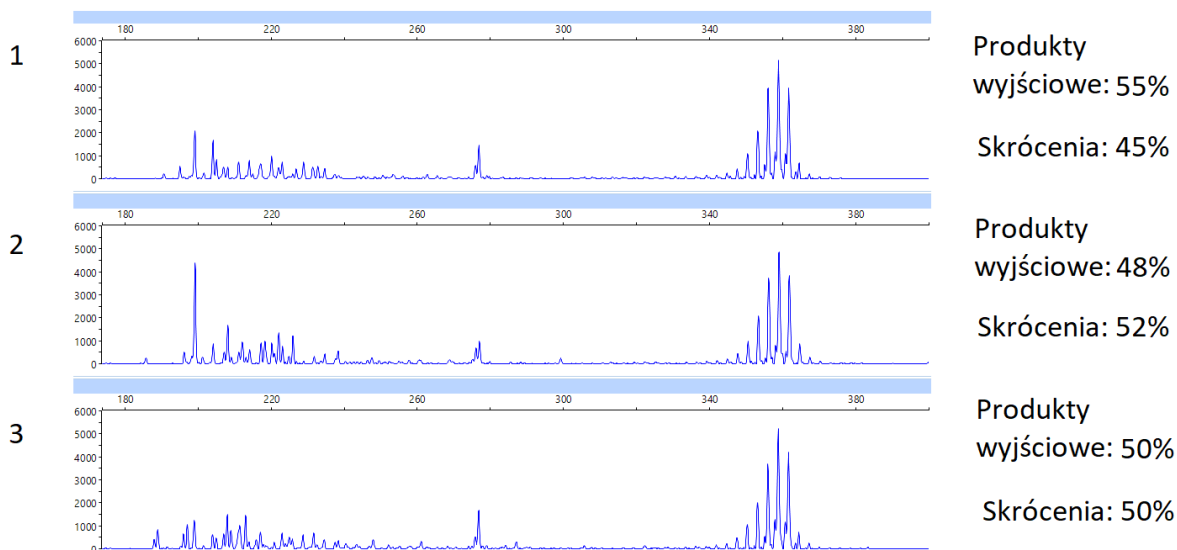


Inhibicja MRE11

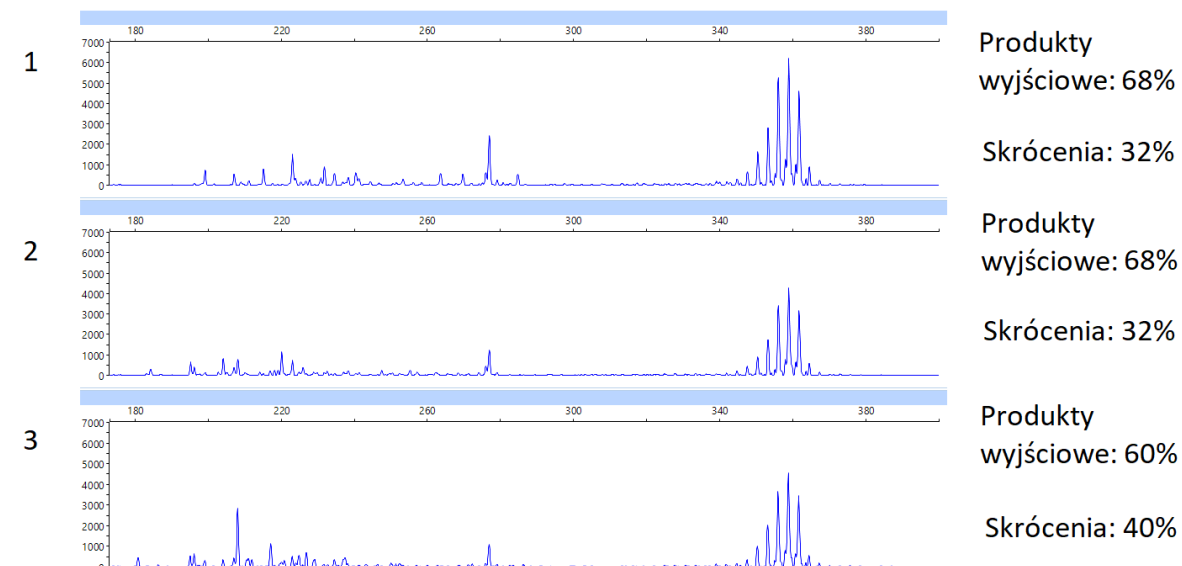


Rysunek 5.5.1.3 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT*, w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA1, traktowanych poszczególnymi inhibitorami oraz bez udziału inhibitorów (kontrola).

Kontrola

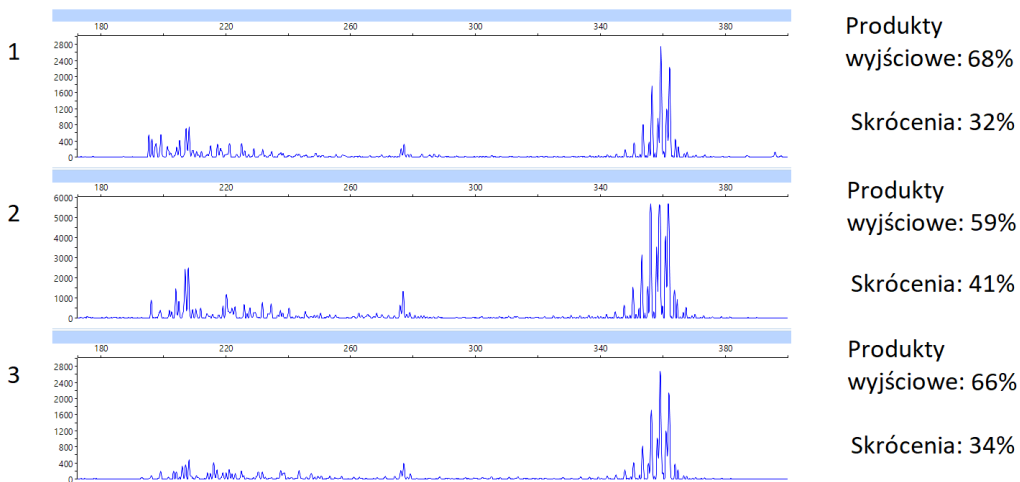


Inhibicja POLθ

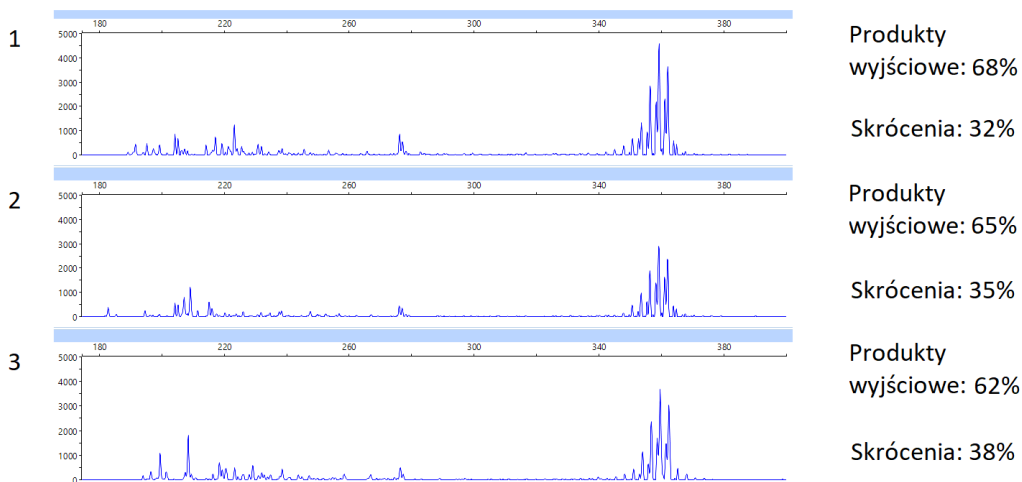


Rysunek 5.5.1.4 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT*, w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA1, traktowanych inhibitorem POLθ oraz bez udziału inhibitora (kontrola).

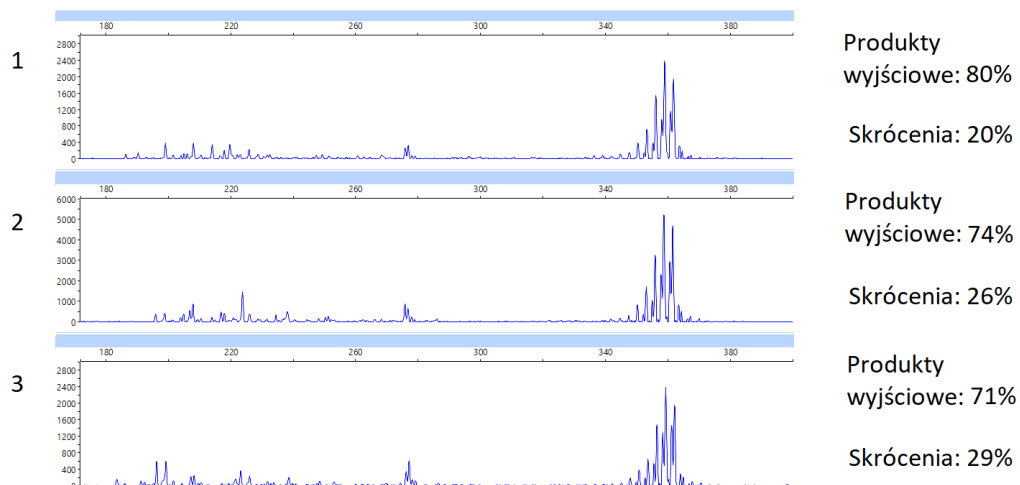
Kontrola



siDCLRE1C (Artemis)

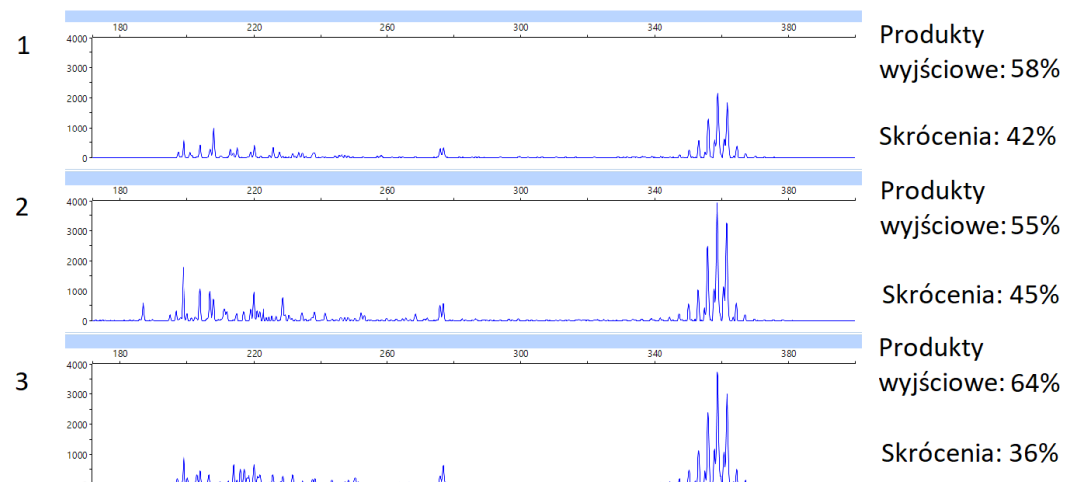


siEXO1

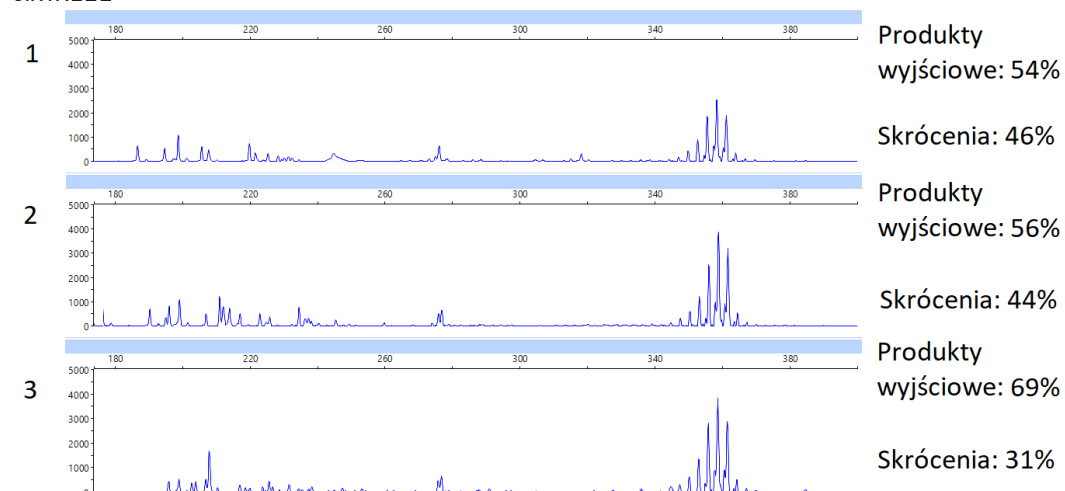


Rysunek 5.5.1.5 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT*, w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA1, traktowanych siRNA oraz bez udziału siRNA (kontrola).

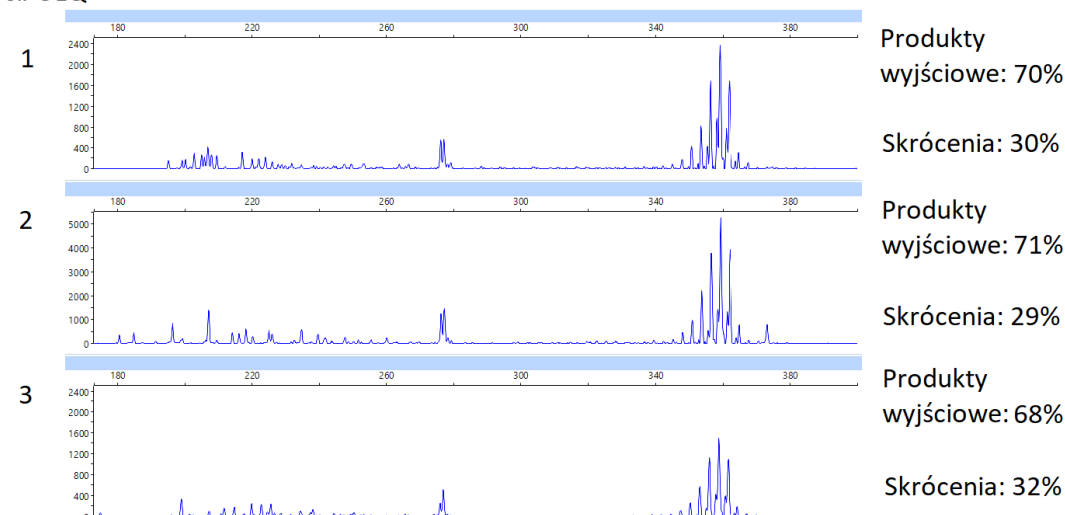
Kontrola



siMRE11

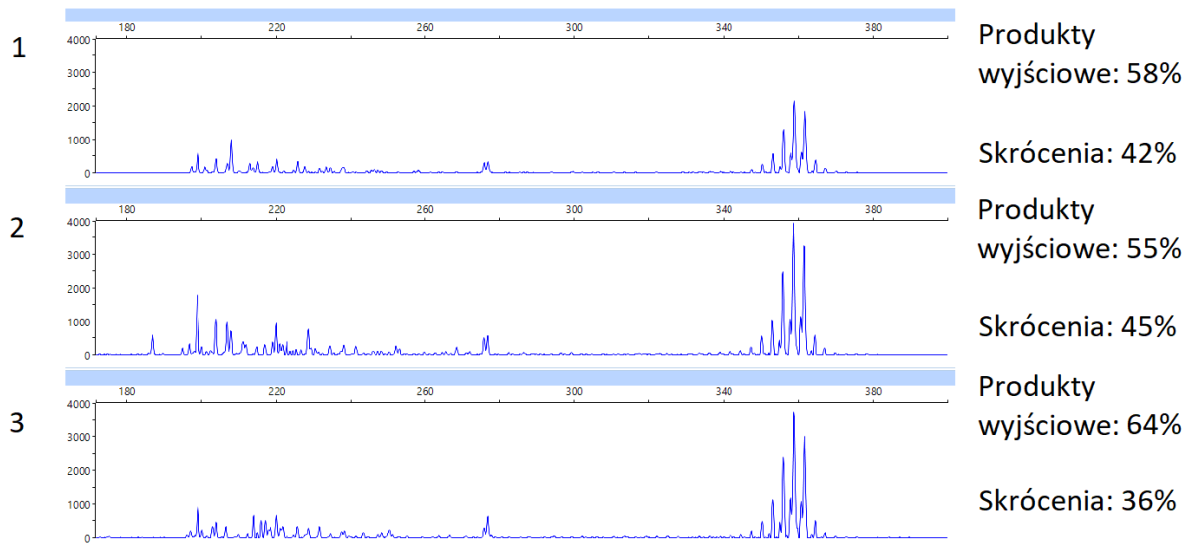


siPOLQ

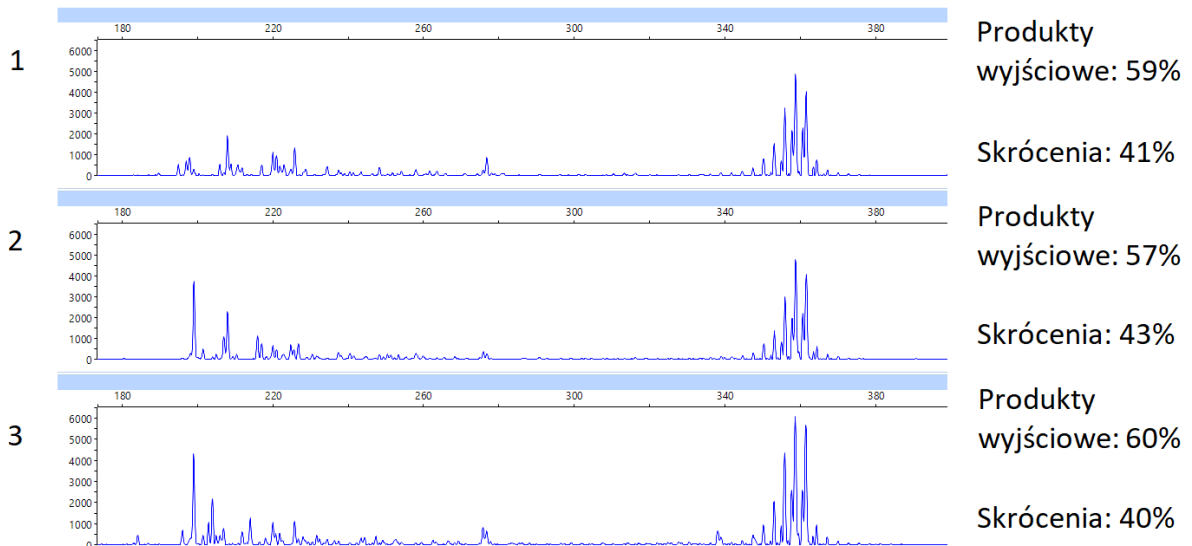


Rysunek 5.5.1.6 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT*, w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA1, traktowanych siRNA oraz bez udziału siRNA (kontrola).

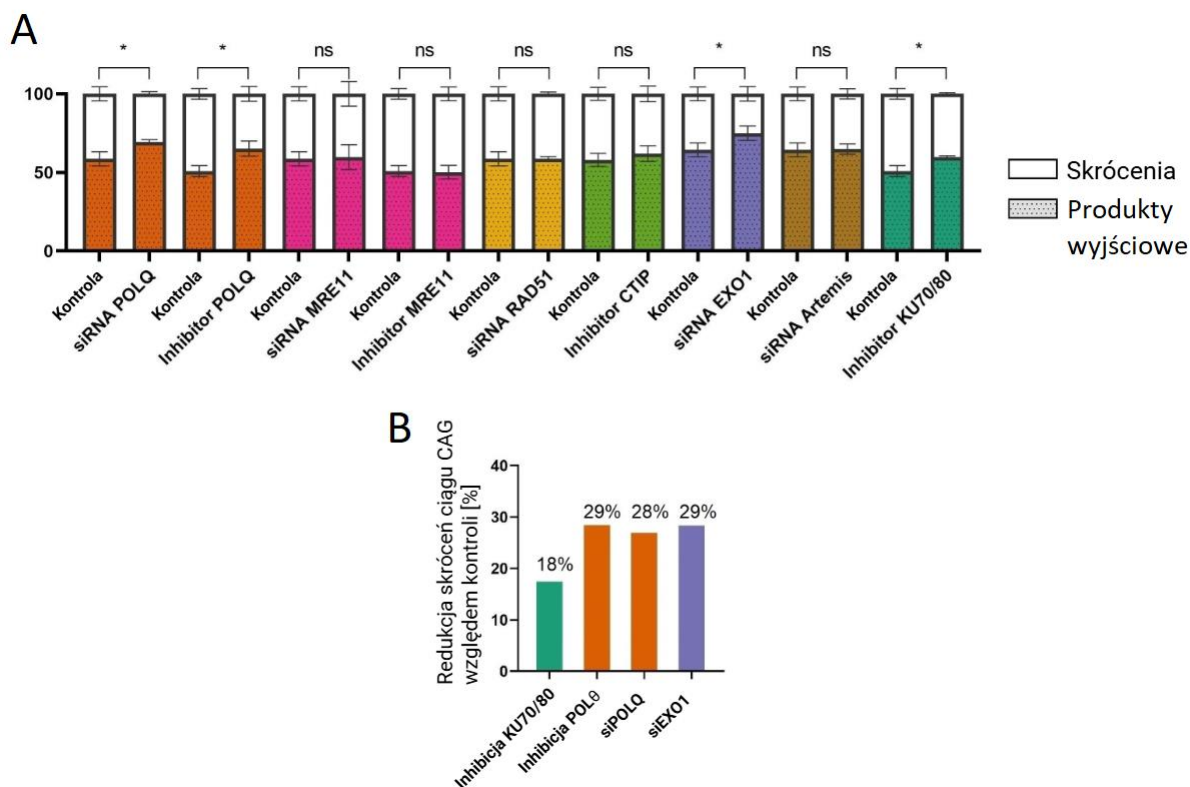
Kontrola



siRAD51



Rysunek 5.5.1.7 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT*, w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA1, traktowanych siRNA oraz bez udziału siRNA (kontrola).



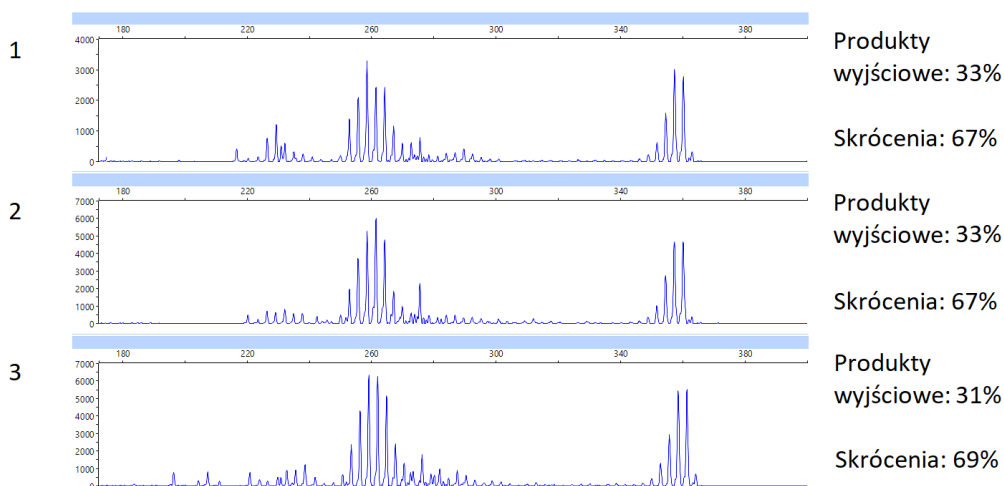
Rysunek 5.5.1.8 Udział produktów zawierających skrócone ciągi powtórzeń CAG w puli produktów po amplifikacji regionu powtórzeń w genie *HTT*, w komórkach HEK293T_41CAG po edycji z udziałem sgRNA1. A) Średnie wyniki procentowego udziału produktów PCR ze skróceniami ciągu powtórzeń oraz produktów wyjściowych w próbach traktowanych inhibitorem lub siRNA oraz w próbach kontrolnych. B) Redukcja skróceń ciągu CAG po zastosowaniu inhibitora lub siRNA po odniesieniu do kontroli, wyliczona ze statystycznie istotnych wyników przedstawionych na wykresie powyżej. Wyniki przedstawiają średnią z trzech powtórzeń biologicznych z odchyleniem standardowym. Istotność statystyczna wyników została określona za pomocą testu t-Studenta, * $p \leq 0.05$.

5.5.2 Analiza udziału białek naprawczych w skracaniu ciągu powtórzeń CAG po indukcji DSB przy użyciu gRNA2

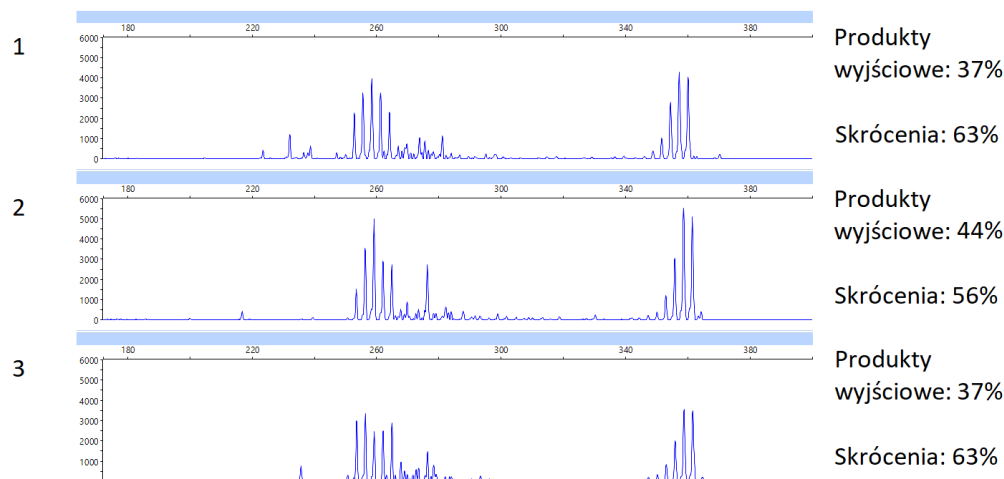
Po edycji z udziałem sgRNA2, podobnie jak w przypadku edycji z udziałem sgRNA1, zidentyfikowano udział w skróceniach ciągu CAG białek KU70 i KU80. Spadek udziału produktów skróconych względem kontroli po inhibicji KU70/KU80 wyniósł 19% (rysunek 5.5.2.7). Ponadto stwierdzono, że relatywnie duży wpływ na skrócenia ma białko MRE11 (23% redukcja skróceń względem kontroli w przypadku siRNA i 6% w przypadku użycia inhibitora). Chemiczna inhibicja, bądź wyciszenie ekspresji genu za pomocą siRNA, wykazały udział

w skracaniu ciągu powtórzeń białka CTIP, współdziałającego z MRE11 (redukcja skróceń o 11%) oraz biorącego udział w HR białka RAD51 (redukcja skróceń o 10%). Elektroforegramy dla poszczególnych powtórzeń biologicznych eksperymentu z udziałem sgRNA2, wraz z procentowym udziałem produktów skróconych i wyjściowych przedstawiono na poniższych rysunkach (5.5.2.1-5.5.2.6). Wykresy podsumowujące, wraz z obliczonym procentowym obniżeniem udziału produktów skróconych względem odpowiedniej kontroli przedstawiono na rysunku 5.5.2.7.

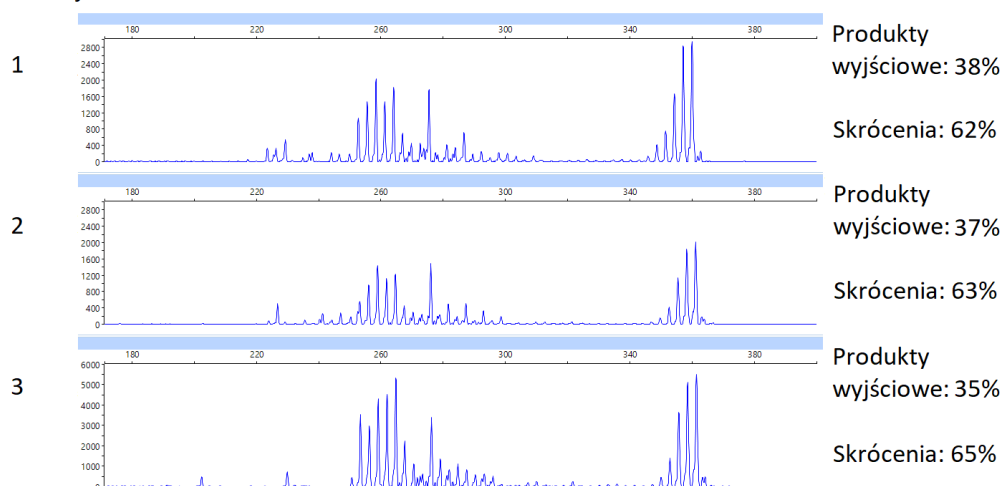
Kontrola



Inhibicja CTIP

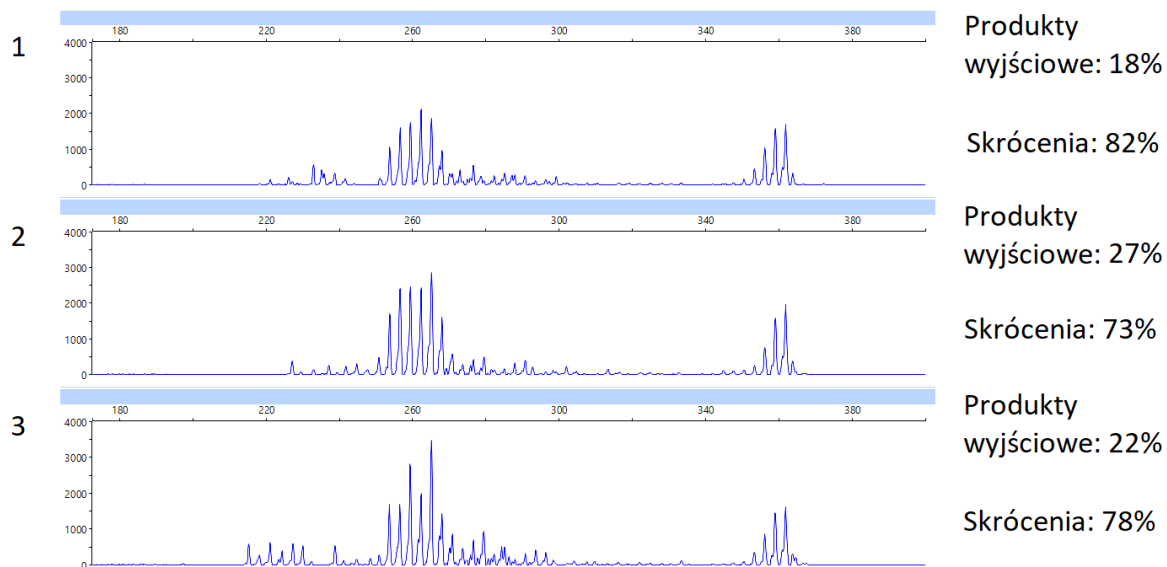


Inhibicja MRE11

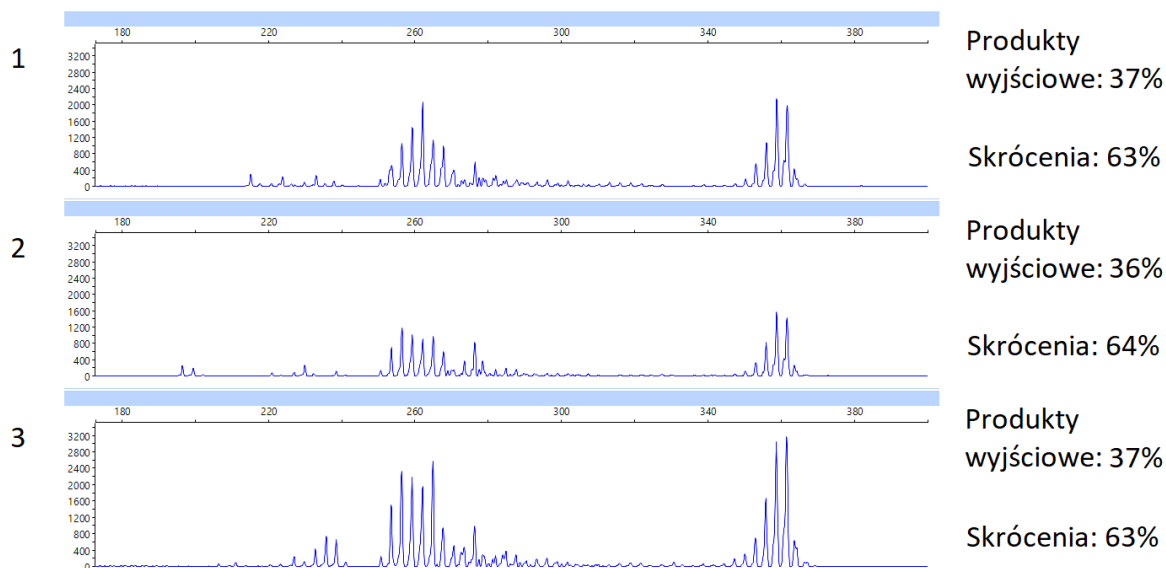


Rysunek 5.5.2.1 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT*, w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA2, traktowanych poszczególnymi inhibitorami oraz bez udziału inhibitorów (kontrola).

Kontrola

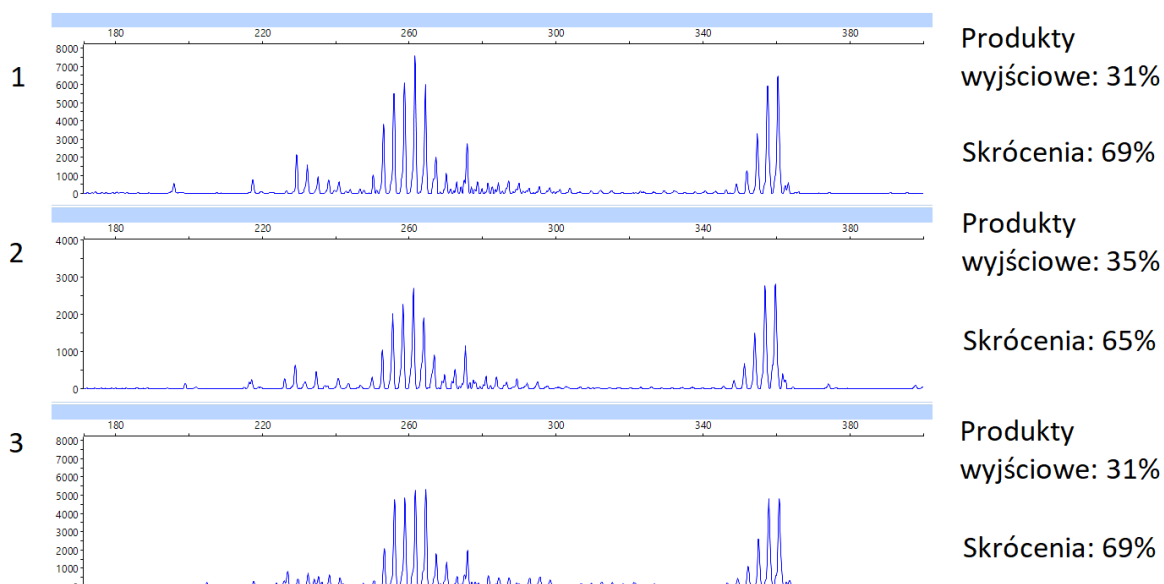


Inhibicja KU70/KU80

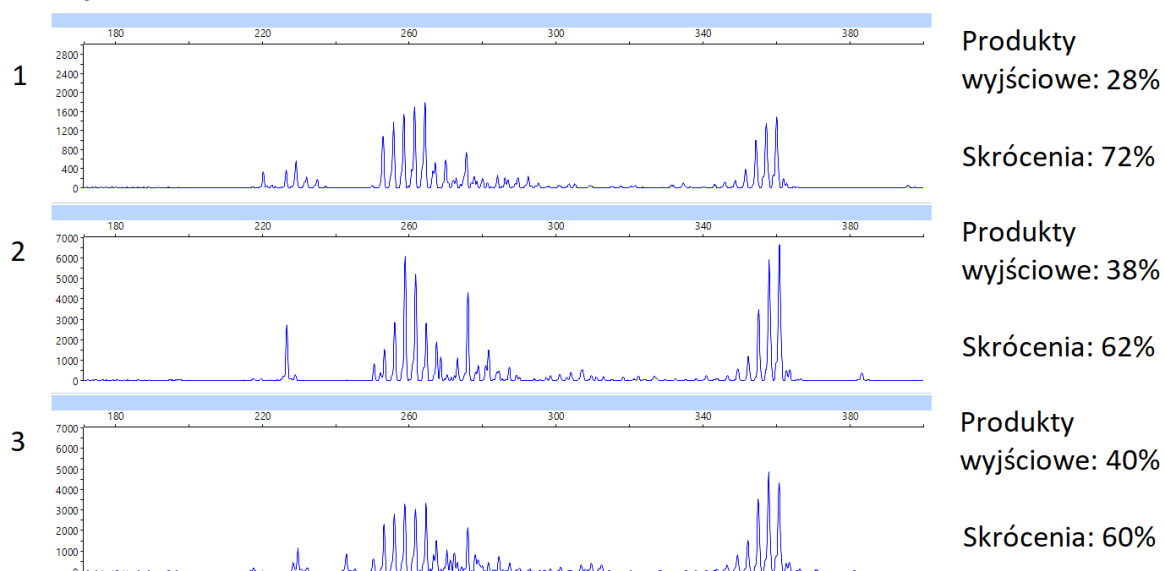


Rysunek 5.5.2.2 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT*, w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA2, traktowanych inhibitorem białek KU70 i KU80 oraz bez udziału inhibitora (kontrola).

Kontrola

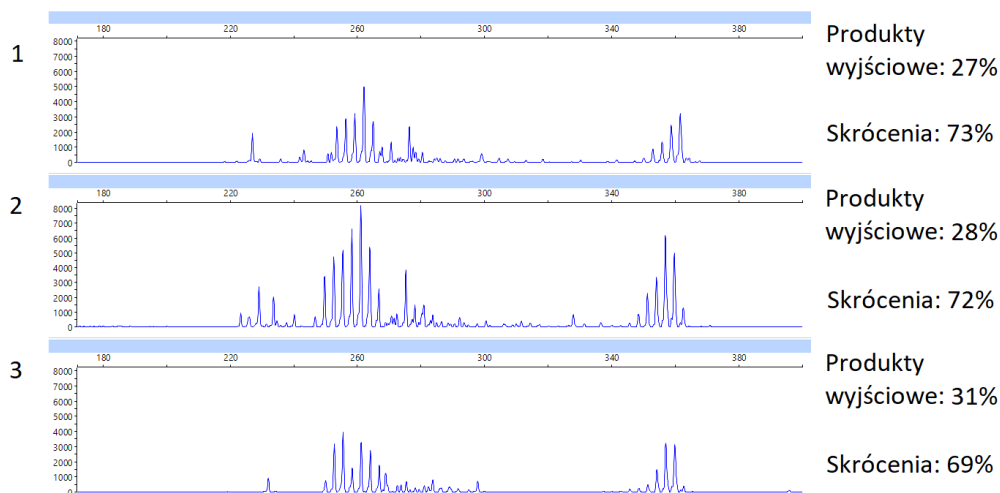


Inhibicja POLθ

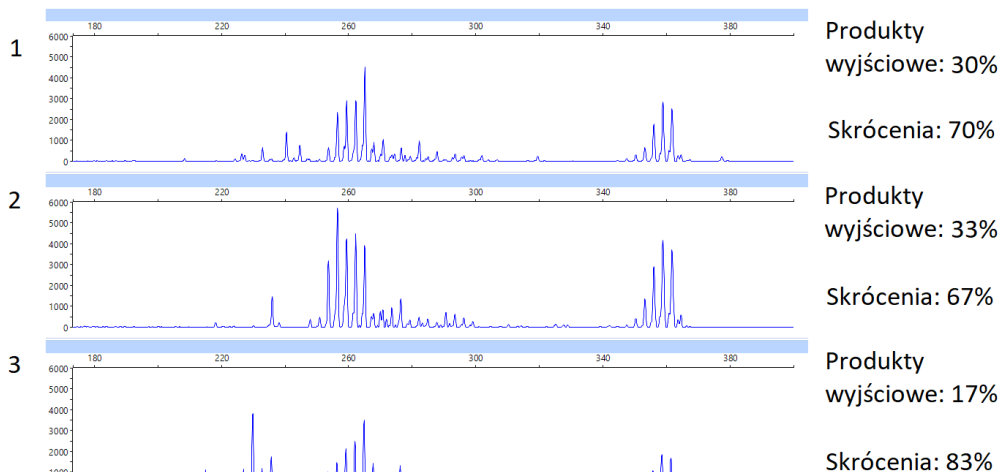


Rysunek 5.5.2.3 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT*, w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA2, traktowanych inhibitorem POLθ oraz bez udziału inhibitora (kontrola).

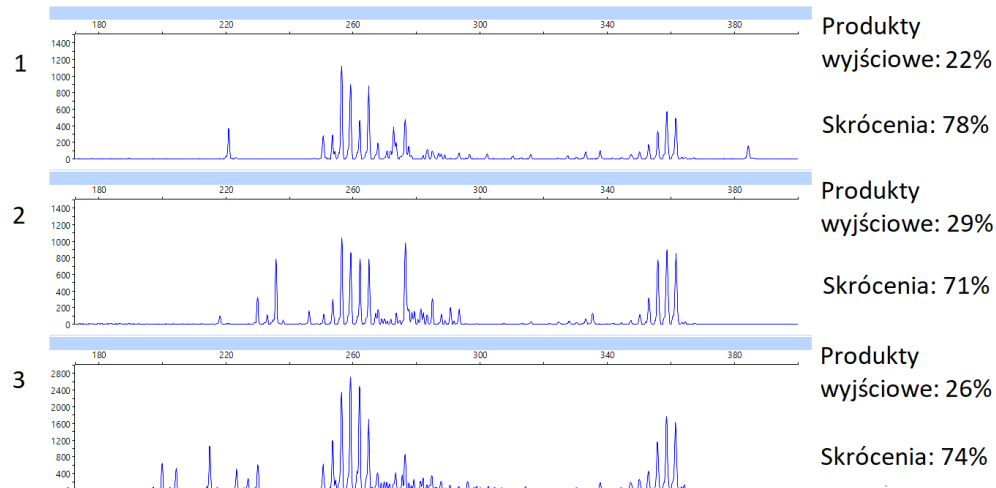
Kontrola



siDCLRE1C (Artemis)

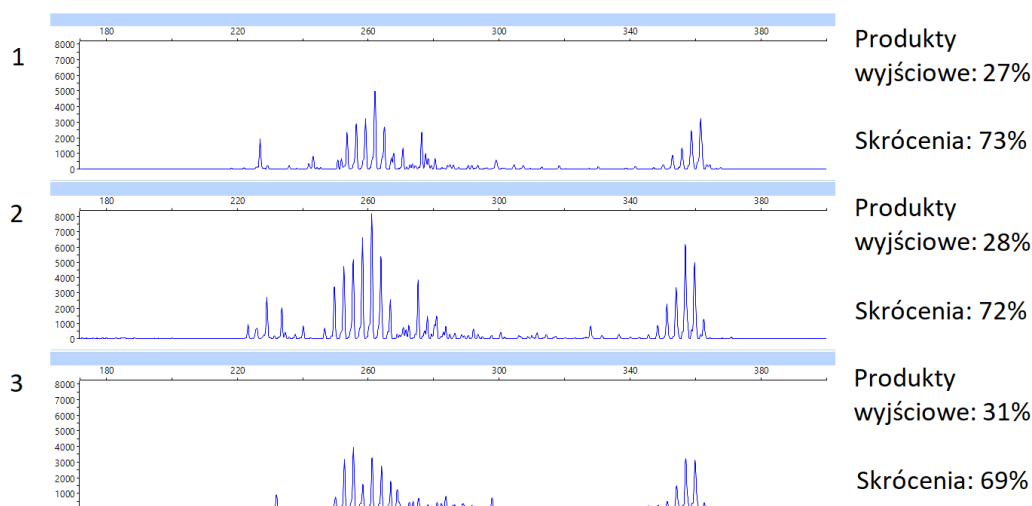


siEXO1

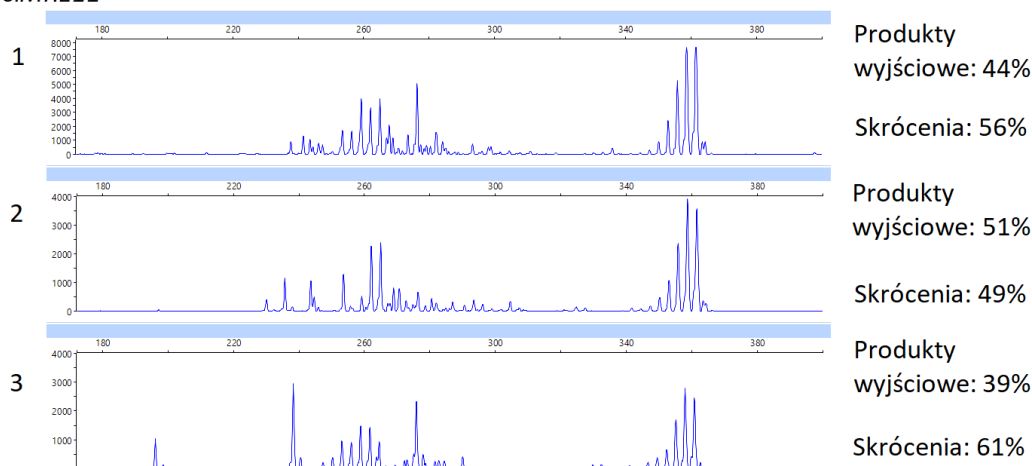


Rysunek 5.5.2.4 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT*, w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA2, traktowanych siRNA oraz bez udziału siRNA (kontrola).

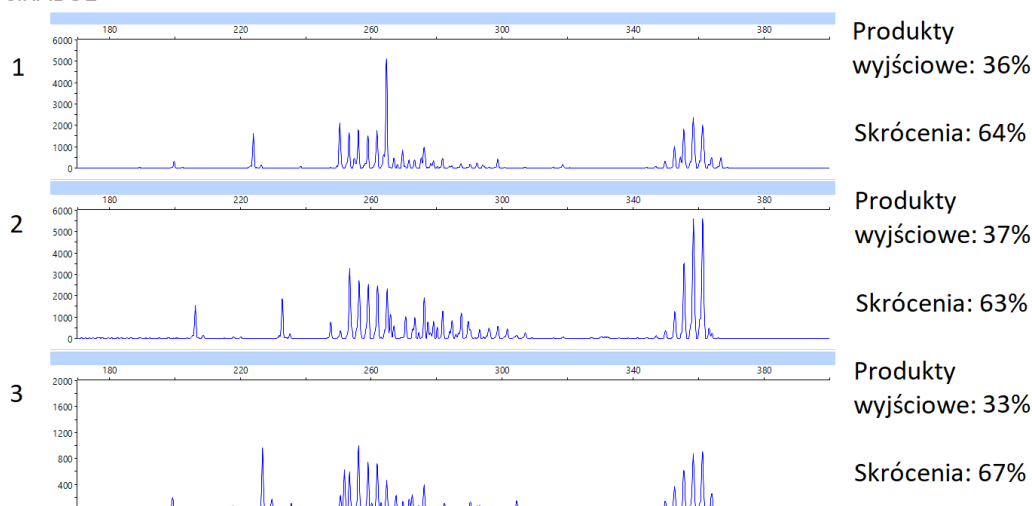
Kontrola



siMRE11

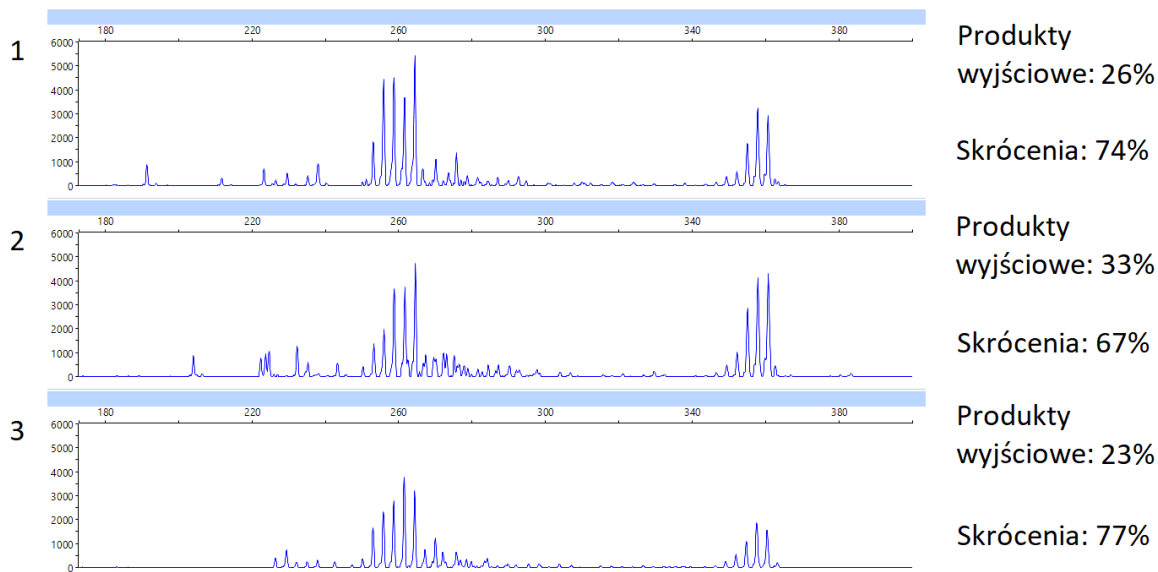


siRAD51

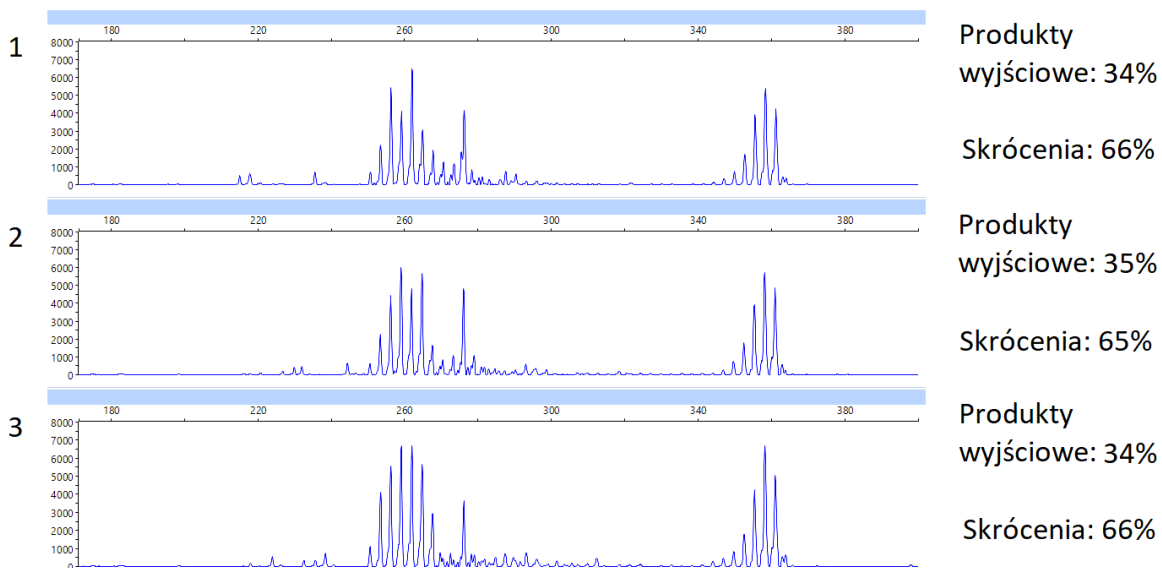


Rysunek 5.5.2.5 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT*, w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA2, traktowanych siRNA oraz bez udziału siRNA (kontrola).

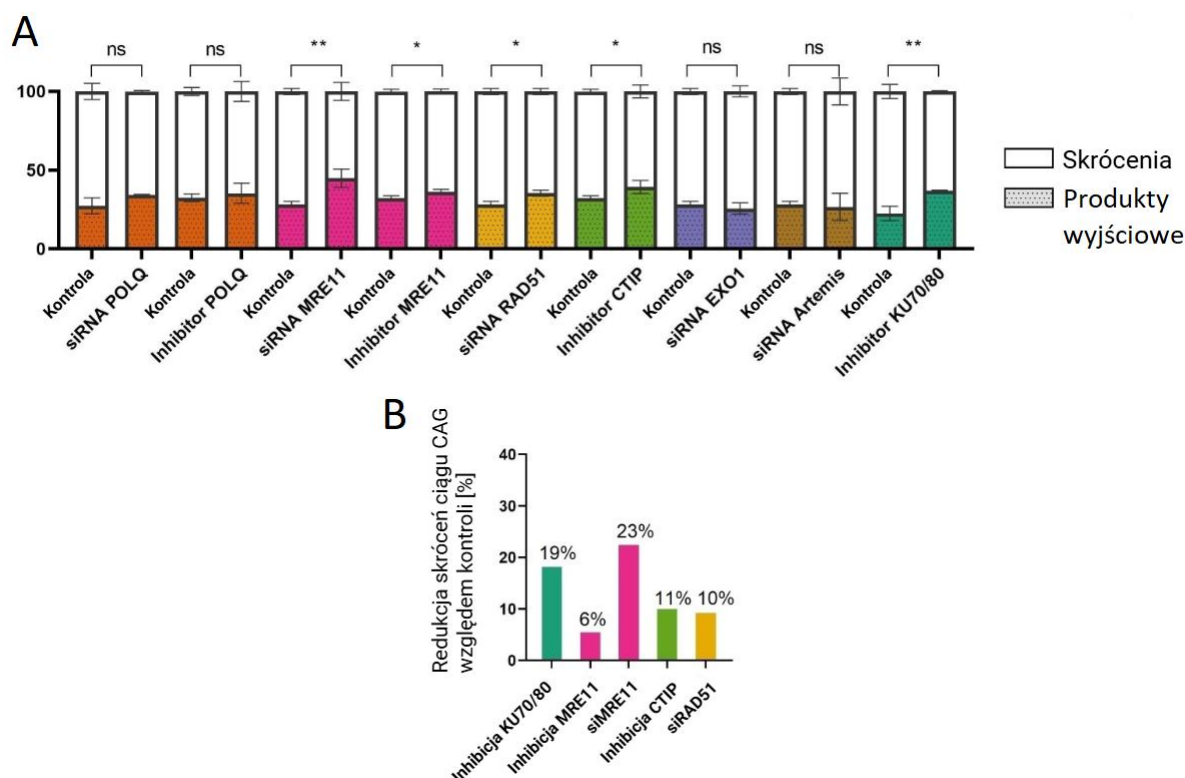
Kontrola



siPOLQ

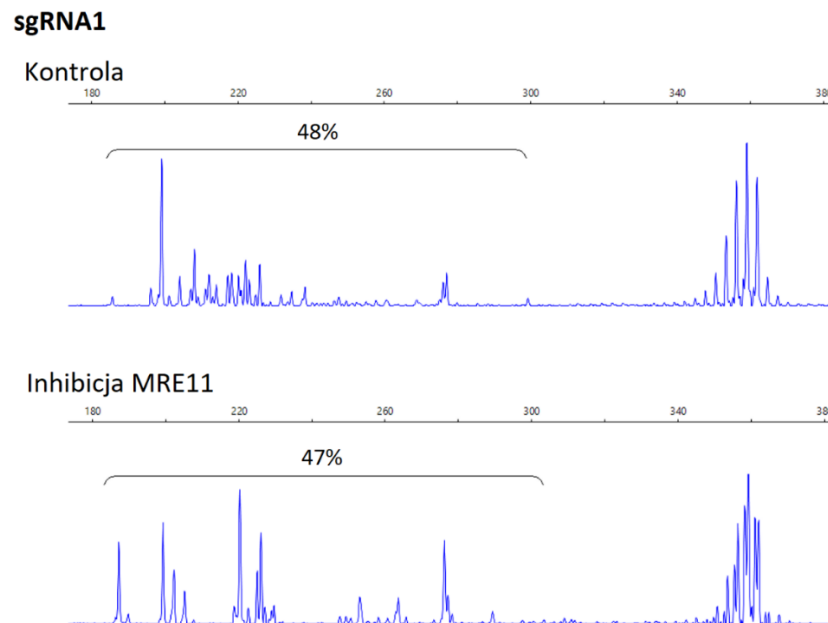


Rysunek 5.5.2.6 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT*, w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA2, traktowanych siRNA oraz bez udziału siRNA (kontrola).



Rysunek 5.5.2.7 Udział produktów zawierających skrócone ciągi powtórzeń CAG w puli produktów po amplifikacji regionu powtórzeń w genie *HTT* w komórkach HEK293T_41CAG po edycji z udziałem sgRNA2. A) Średnie wyniki procentowego udziału produktów PCR ze skróceniami ciągu powtórzeń oraz produktów wyjściowych w próbach traktowanych inhibitorem lub siRNA oraz w próbach kontrolnych. B) Redukcja skróceń ciągu CAG po zastosowaniu inhibitora lub siRNA po odniesieniu do kontroli, wyliczona ze statystycznie istotnych wyników przedstawionych na wykresie powyżej. Wyniki przedstawiają średnią z trzech powtórzeń biologicznych z odchyleniem standardowym. Istotność statystyczna wyników została określona za pomocą testu t-Studenta, * $p \leq 0.05$; ** $p \leq 0.01$.

Zarówno po edycji z udziałem sgRNA1, jak i sgRNA2 w przypadku inhibicji niektórych białek lub wyciszenia genów je kodujących nie zaobserwowano statystycznie istotnych różnic w udziale produktów skróconych i wyjściowych pomiędzy próbami traktowanymi a kontrolą. Warto tu podkreślić, że użyta w tym eksperymencie metoda służy do analizy ilościowej powstałych produktów. Określana jest w niej ich długość a nie sekwencja, w związku z tym jakościowe zmiany, widoczne gdy dokonać porównania wizualnego poszczególnych elektroforegramów, nie miały wpływu na ostateczny wynik eksperymentu. Przykład różnic w rozkładzie i liczbie pików w elektroforegramach dla próby traktowanej inhibitorem białka MRE11 i kontroli, dla których analiza ilościowa nie wykazała istotnych różnic, przedstawiono na rysunku 5.5.2.8

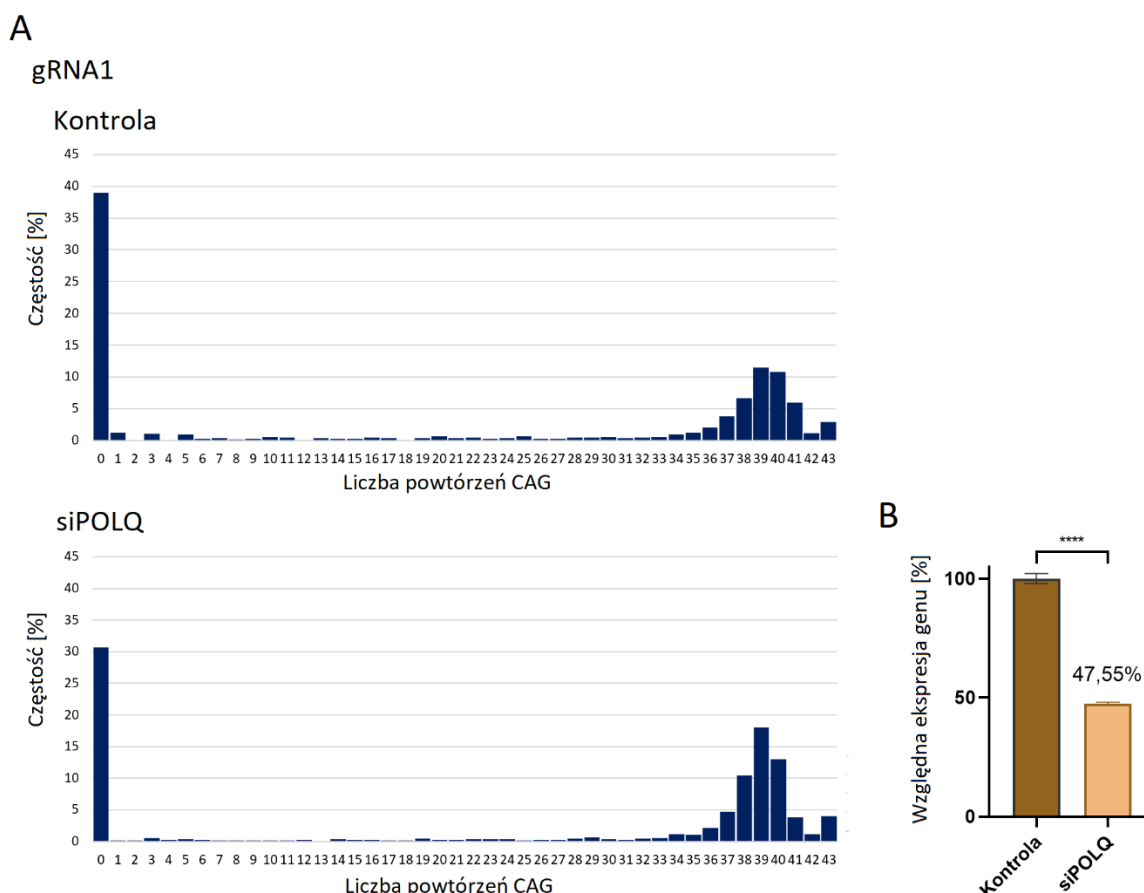


Rysunek 5.5.2.8 Porównanie rozkładu pików na elektroforegramach z rozdziałem produktów PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT* dla prób edytowanych sgRNA1: kontrolnej i traktowanej inhibitorem MRE11.

5.6 Wpływ POLθ na wyniki naprawy DSBs w regionie z powtórzeniami CAG w genie *HTT*

Aby określić precyzyjniej wpływ POLθ na wyniki naprawy DSBs indukowanych w obrębie powtórzeń CAG w genie *HTT* przeprowadzono eksperyment, w którym dokonano porównania wyników sekwencjonowania (NGS) produktów PCR powstałych w wyniku amplifikacji edytowanego z udziałem gRNA1 regionu z powtórzeniami CAG w genie *HTT* w komórkach HEK293T_41CAG po wyciszeniu genu *POLQ* oraz bez jego wyciszenia.

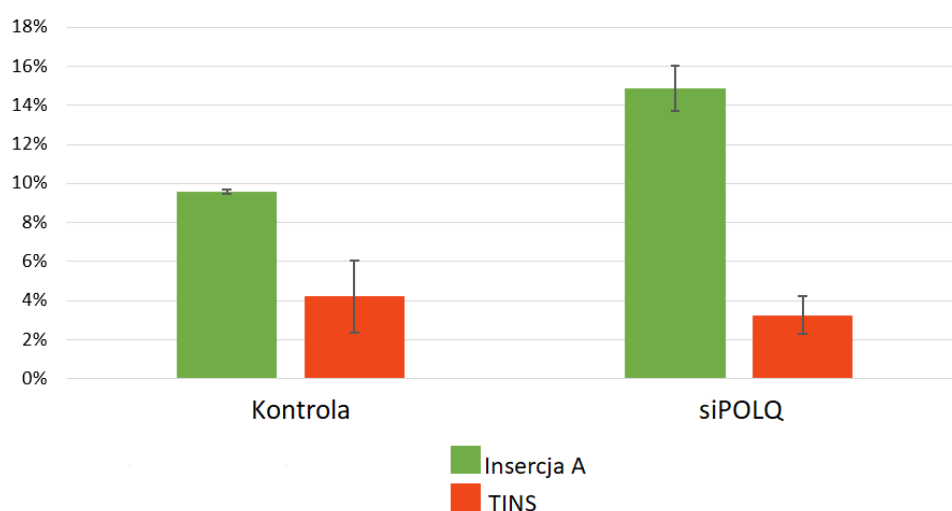
We wszystkich próbach po edycji z udziałem gRNA1 zaobserwowano wycięcia oraz skrócenia ciągu powtórzeń analogicznie do wyników prezentowanych w rozdziale 5.1.2. W próbach kontrolnych warianty całkowicie pozbawione ciągu CAG stanowiły ok. 39%, zaś w próbach traktowanych siRNA ok. 31% (rysunek 5.6.1 A). Najczęściej występującym wariantem zarówno w próbach kontrolnych, jak i traktowanych siRNA była sekwencja z ciągiem 41 powtórzeń CAG i insercją adeniny znajdującą się bezpośrednio za cytozyną z pierwszego powtórzenia CAG, czyli w miejscu indukcji DSB. Częstość tego wariantu w próbach kontrolnych była niższa w porównaniu do prób traktowanych siRNA i wynosiła ok. 5% odczytów, podczas gdy w próbach traktowanych siRNA było to ok. 11,5% odczytów. Łączna liczba odczytów z insercjami adeniny bez uwzględnienia innych zmian była również niższa w kontrolach i wynosiła ok. 9,5% odczytów, podczas gdy dla prób traktowanych siRNA było to ok. 15% odczytów (rysunek 5.6.2). Większy udział wariantów z insercją adeniny w pozycji -4 od PAM, charakterystyczną dla naprawy z udziałem NHEJ, w warunkach wyciszenia ekspresji *POLQ*, może świadczyć o większym udziale tej ścieżki naprawy w przypadku zablokowania ścieżki alternatywnej, jaką jest TMEJ.



Rysunek 5.6.1 A) Rozkład częstości występowania wariantów z określoną liczbą powtórzeń CAG w produktach PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT* w komórkach HEK293T_41CAG edytowanych z udziałem gRNA1 z wyciszeniem ekspresji *POLQ* oraz bez wyciszenia (kontrola). Wyniki przedstawiają średnią z dwóch powtórzeń biologicznych. B) Względna ekspresja genu *POLQ* po zastosowaniu siRNA. Wynik przedstawiają średnią z trzech powtórzeń biologicznych oraz odchylenie standardowe. Istotność statystyczna wyników została określona za pomocą testu t-Studenta, **** $p \leq 0,0001$.

Ponieważ *POLθ* odpowiedzialna jest za wprowadzanie charakterystycznych insercji, dla których matrycą jest DNA, znajdujący się w pobliżu miejsca, w którym doszło do DSB (TINS), szczegółowo przeanalizowano wyniki naprawy pod względem wprowadzonych insercji o długości od 3 do 30 par zasad, znajdujących się w odległości do 50 par zasad od miejsca cięcia, gdyż tak najczęściej definiowane są TINS¹²¹. Nie stwierdzono jednak wyraźnych różnic w liczbie odczytów zawierających TINS pomiędzy próbkami traktowanymi siRNA a próbkami kontrolnymi (rysunek 5.6.2). W przypadku wariantów z delecją całego ciągu powtórzeń CAG TINS występowały w próbkach kontrolnych w ok. 1,5% wszystkich odczytów. W przypadku próbek traktowanych siRNA występowały w 2,7% odczytów. W przypadku

wariantów, w których doszło do skrócenia ale nie wycięcia całego ciągu powtórzeń, TINS zaobserwowano w próbach kontrolnych w 1,4% odczytów, a w próbach traktowanych siRNA w przypadku 0,6% odczytów. W wariantach z wyjściową liczbą powtórzeń CAG lub większą TINS zidentyfikowano tylko w jednej z prób kontrolnych i stanowiły one 2,7% wszystkich odczytów. Sumując wszystkie TINS zidentyfikowane w analizowanych próbach określono ich częstość na 4,25% w próbach kontrolnych oraz 3,25% w próbach traktowanych siRNA.



Rysunek 5.6.2 Procentowy udział poszczególnych typów insercji w wynikach sekwencjonowania regionu z powtórzeniami CAG w genie *HTT* w komórkach HEK293T_41CAG po edycji z udziałem gRNA1 z wyciszeniem ekspresji *POLQ* oraz bez wyciszenia (kontrola). Wynik przedstawia średnią z dwóch powtórzeń biologicznych z odchyleniem standardowym.

6. Dyskusja

Choroby poliQ takie jak HD stanowią grupę dziedzicznych, neurodegeneracyjnych schorzeń o wspólnym mechanizmie molekularnym, związanym z ekspansją powtórzeń CAG w obrębie poszczególnych genów. Dotychczas opracowywane terapie HD, znajdujące się na etapie badań klinicznych, koncentrują się na obniżeniu ekspresji zmutowanego transkryptu *HTT*, nie wpływając na pierwotną przyczynę choroby, jaką są wspomniane ekspansje powtórzeń w DNA pacjentów^{216,217}. Rozwój narzędzi inżynierii genetycznej takich jak technologia CRISPR-Cas9 przyniósł nadzieję na opracowanie skutecznych terapii chorób, o jasno zdefiniowanym podłożu genetycznym, takich jak HD. Jedną z nowatorskich strategii terapii chorób poliQ jest skracanie ciągu powtórzeń w wyniku indukcji w jego obrębie pęknięć DNA. Liczne badania koncentrujące się na naprawie DSBs w regionach z powtórzeniami CAG/CTG miały na celu udzielenie odpowiedzi na pytanie, jakie czynniki i mechanizmy są odpowiedzialne za to zjawisko. Większość z tych badań wykorzystywała modele drożdżowe, niebędące idealnym modelem, w kontekście potencjalnej przyszłej terapii ludzkich chorób. Warto tu choćby wspomnieć o braku u drożdży genu *POLQ* kodującego ważne białko związane z naprawą DNA POLθ. Gen ten występuje zaś u ludzi i jak wykazano w niniejszej pracy, POLθ ma swój udział w skracaniu ciągu powtórzeń CAG w genie *HTT* w procesie naprawy DNA. Liczne prace badające skracanie ciągów powtórzeń CAG/CTG w komórkach ludzkich wykorzystywały systemy reporterowe, sztucznie wprowadzone do ludzkiego genomu, bądź plazmidy zawierające ciągi powtórzeń, będące całkiem odrębną cząsteczką DNA. Badania te pomimo swych ograniczeń i braku uwzględnienia endogennego kontekstu przyniosły jednak wiele interesujących odkryć i pozwoliły na sformułowanie hipotez, które następnie mogą podlegać weryfikacji w ludzkich modelach w endogennych *loci* z ekspansją powtórzeń. Niniejsza praca miała na celu identyfikację czynników i procesów zaangażowanych w skracanie ciągu powtórzeń CAG w genie *HTT* w komórkach ludzkich, których poznanie jest kluczowe do potencjalnego przyszłego wykorzystania tego zjawiska w terapii HD i innych chorób poliQ. Praca ta jest częścią większego projektu finansowanego przez Narodowe Centrum Nauki noszącego tytuł „Badanie mechanizmów naprawy dwuniciowych pęknięć DNA w regionach mikrosatelitarnych z wykorzystaniem systemu CRISPR/Cas9”. Wyniki uzyskane w ramach tego projektu przyczyniły się do powstania pracy „CRISPR/Cas9-induced double-strand breaks in the huntingtin locus lead to CAG repeat contraction through DNA end resection and homology-mediated repair” opublikowanej w 2024 roku w czasopiśmie BMC Biology. Wyniki zawarte we wspomnianej publikacji będą również dyskutowane w niniejszym rozdziale.

W wyniku indukcji dwuniciowych pęknięć w obrębie powtórzeń CAG w genie *HTT* zaobserwowano znaczne różnice w wynikach naprawy, zależne od miejsca pęknięcia. Różnice dotyczyły nie tylko końcowego efektu naprawy DNA, ale również przebiegu naprawy w czasie oraz zaangażowania poszczególnych procesów i czynników. Naprawa DSB indukowanego na początku ciągu powtórzeń (edycja z udziałem gRNA1) skutkowała w większości przypadków przesunięciem ramki odczytu i jednonukleotydowymi insercjami, wprowadzonymi zgodnie z „regułą -4”.^{161–164} Zgodnie z tą regułą nukleotyd znajdujący się w pozycji -4 od sekwencji PAM ma kluczowe znaczenie na wynik naprawy, a obecność w tym miejscu adeniny skutkuje zwykle jej duplikacją, co zostało zaobserwowane w wynikach z NGS (rysunek 5.1.1.1) i rozdzieleniu produktów PCR w elektroforezie kapilarnej (rysunek 5.2.2 B). W przypadku naprawy DSB po edycji z udziałem gRNA1 dominującym zdarzeniem było wycięcie całego ciągu powtórzeń CAG, zwykle razem z sekwencjami flankującymi ciąg. Znane modele naprawy DSBs sugerują, że duże delecje będące skutkami naprawy DNA zachodzą z udziałem resekcji końców DNA, których następstwem jest łączenie się jego jednoniciowych fragmentów w miejscach występowania między nimi homologii. W ten sposób fragment znajdujący się pomiędzy homologicznymi regionami ulega delecji. Zaobserwowano, że w czasie naprawy DSB indukowanych z udziałem gRNA1 w komórkach HEK293T_41CAG dochodziło do resekcji końców DNA w ograniczonym zakresie (rysunek 5.4.1 A). Wykryte regiony jednoniciowe wynosiły około 1100 nukleotydów po obu stronach od DSB, a największy ich udział obserwowany był w rejonach oddalonych o ok. 200 par zasad od DSB po 6 godzinach od wprowadzenia kompleksu Cas9-gRNA1 do komórek. Po 24h od transfekcji poziom resekcji był zbliżony do wartości wyjściowej we wszystkich regionach, co może świadczyć o zakończeniu naprawy w tym czasie. Jest to zbieżne z wynikami uzyskanymi z elektroforetycznego rozdzielania produktów PCR po amplifikacji regionu z powtórzeniami CAG w genie *HTT* po edycji z udziałem gRNA1, które ujawniły obecność skróceń ciągu powtórzeń CAG po 24 godzinach od transfekcji komórek kompleksem RNP (żadne wyraźne skrócenia nie zostały zaobserwowane po 4 h od transfekcji; rysunek 5.2.1 A) Wyniki z eksperymentu ChIP-qPCR (immunoprecypitacji chromatyny przy użyciu specyficznych dla danego białka przeciwciał, po której następuje amplifikacja badanego *locus*) prezentowane we wspomnianej wcześniej publikacji potwierdzają obecność białka inicjującego resekcję końców DNA MRE11 w regionach po stronie 3' i 5' od edytowanego miejsca w czasie od 12 do 48 godzin po transfekcji plazmidem kodującym Cas9 i sgRNA1.¹⁸² Eksperyment, w którym analizowano wpływ poszczególnych białek naprawczych na udział produktów skróconych po edycji z udziałem sgRNA1, nie potwierdził zaangażowania MRE11 w skracanie ciągu powtórzeń

w tym przypadku (rysunek 5.5.1.8). Pomimo, że zaobserwowano pewne zmiany jakościowe w obrazie pików, porównując elektroforegramy przedstawiające rozdział produktów PCR amplifikowanych po edycji regionu z powtórzeniami CAG w genie *HTT* z udziałem sgRNA1 w próbach po wyciszeniu lub inhibicji MRE11 i kontrolach, to efekt ten nie był zauważalny w analizie ilościowej (rysunek 5.5.2.8). Wyniki z tego eksperymentu pozwoliły jednak określić udział innych białek w skracaniu ciągu powtórzeń CAG w badanym *locus*. Jednym z nich była zaangażowana w resekcie DNA nukleaza EXO1 (rysunek 5.5.1.8). Wyniki z ChIP-qPCR¹⁸² ujawniły również wzbogacenie w sekwencji znajdującej się w pobliżu edytowanego ciągu powtórzeń CAG w genie *HTT* białka FEN1, nukleazy odpowiedzialnej za usuwanie struktur typu 5' flap oraz utworzonych przez 5' ssDNA struktur drugorzędowych. Dwukrotne wzrost wzbogacenia FEN1 został zaobserwowany po 12 i 24 godzinach od transfekcji komórek HEK293T_41CAG plazmidem kodującym sgRNA1 i Cas9. Co więcej, ten sam eksperyment wykonany na modelu HEK293T, w którym na obu allelach *HTT* obecne były 83 powtórzenia CAG, ujawnił czterokrotne wzbogacenie FEN1 w sekwencjach flankujących ciąg powtórzeń. Sugeruje to wzrost liczby tworzonych struktur rozpoznawanych przez FEN1 podczas naprawy DSBs indukowanych w dłuższych ciągach powtórzeń CAG. Łączenie dwóch fragmentów DNA, które uległy resekcji może nastąpić z udziałem mechanizmu, w który zaangażowana jest POLθ (TMEJ). Jej udział w skracaniu ciągu powtórzeń CAG po edycji z udziałem sgRNA1 został potwierdzony zarówno po użyciu inhibitora POLθ, jak i siRNA skierowanego na gen *POLQ* (rysunek 5.5.1.8). Wpływ POLθ na efekt naprawy DSB indukowanego na początku ciągu powtórzeń CAG starano się określić precyzyjniej w eksperymencie, w którym porównano sekwencje DNA po naprawie DSB w warunkach wyciszenia genu *POLQ* z sekwencjami uzyskanymi z prób, w których gen *POLQ* nie został wyciszony (rozdział 5.6) Charakterystyczne dla udziału POLθ w naprawie DNA insercje (TINS) stwierdzono zarówno w próbach kontrolnych, jak i traktowanych siRNA, a ich udział nie był znacząco różny (rysunek 5.6.2). Mogło być to spowodowane niewystarczającym dla uzyskania spodziewanego efektu wyciszeniem *POLQ* (rysunek 5.6.1 B) lub udziałem innych czynników. Jednakże, w owym eksperymencie zaobserwowano obecność większej liczby wariantów z insercją A w warunkach wyciszenia *POLQ* (produktem naprawy DNA charakterystycznym dla NHEJ), w porównaniu do prób bez wyciszenia *POLQ*. Może to świadczyć o wzroście udziału NHEJ w naprawie DNA w przypadku braku POLθ i zablokowaniu naprawy na drodze TMEJ. Pomimo braku potwierdzenia zależności tworzenia TINS od obecności POLθ, należy zauważyć, że wyniki z tego eksperymentu w zakresie dotyczącym spadku procentowego udziału skróceń ciągu powtórzeń CAG w warunkach wyciszenia *POLQ* są zbieżne z wynikami z analizy

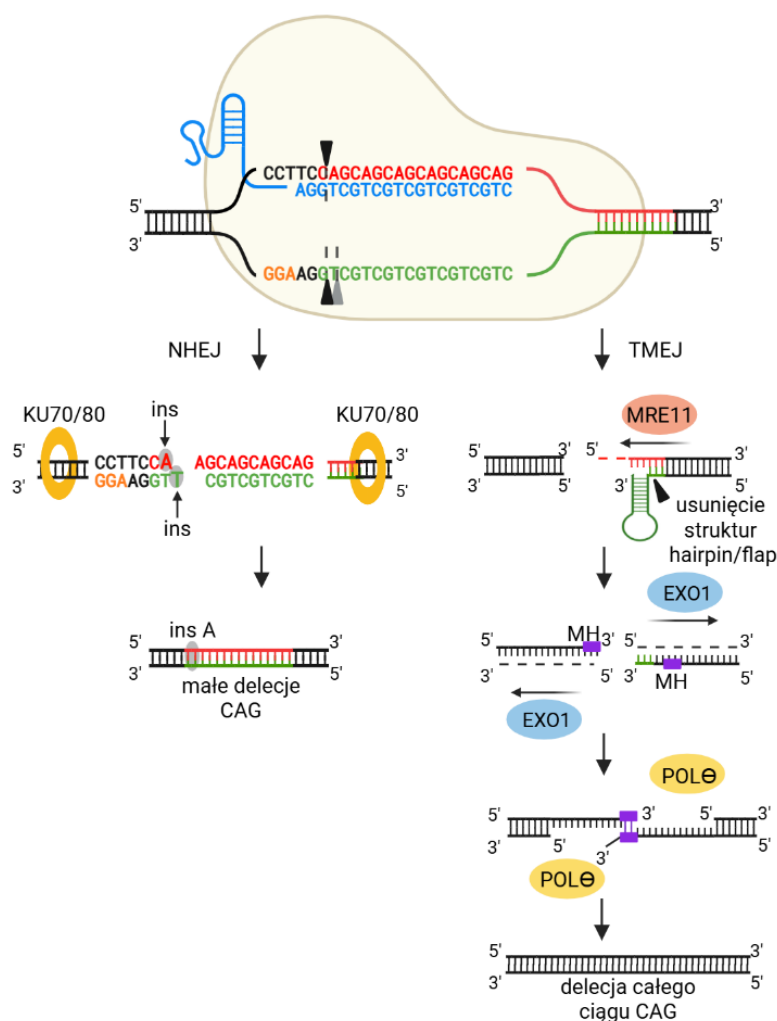
przeprowadzonej po rozdzieleniu produktów PCR w kapilarze, gdzie udział produktów skróconych w przypadku edycji po sgRNA1 wynosił średnio dla prób traktowanych siRNA 30% a dla kontroli 41% (rysunek 5.5.1.8). Chociaż udział POLθ w naprawie DSBs indukowanych przez Cas9 w komórkach ludzkich jest zjawiskiem znanym^{218,219}, a ostatnio potwierdzono wzrost ekspresji genu *POLQ* w neuronach, do których dostarczono kompleks RNP (Cas9-sgRNA)²²⁰, to w niniejszej pracy po raz pierwszy zaprezentowano udział tej polimerazy w skracaniu ciągu powtórzeń CAG w trakcie naprawy DNA. Jednakże, aby jednoznacznie określić udział tej polimerazy w tworzeniu zidentyfikowanych insercji, należałoby w kolejnym eksperymencie użyć linii komórkowej bez ekspresji *POLQ* oraz wykonać eksperyment na większej liczbie prób. Białkami, których udział w skracaniu ciągu powtórzeń CAG po edycji z udziałem sgRNA1 został zaobserwowany w niniejszej pracy, są białka KU70 i KU80 działające w ścieżce naprawy NHEJ (rysunek 5.5.1.8). Znanym efektem naprawy DNA w skutek NHEJ są niewielkie delecje lub insercje obserwowane w miejscu, w którym doszło do DSBs. Wspomniane już wyniki z NGS oraz z rozdzielania elektroforetycznego produktów PCR w kapilarze, uwidaczniające insercję adeniny w miejscu indukcji dwuniciowego pęknięcia DNA, potwierdzają udział mechanizmu NHEJ w tym przypadku. Intrygującym pozostaje zaobserwowany wpływ inhibicji białek KU70 i KU80 na spadek udziału produktów skróconych po edycji z udziałem sgRNA1, jako że duże delecje nie należą do wyników charakterystycznych dla naprawy DNA z udziałem NHEJ. Udział nukleazy Artemis, działającej po związaniu końców DNA przez KU70/80, która potencjalnie mogłaby generować niewielkie resekcje lub usuwać struktury drugorzędowe i odpowiadać za delecję ciągu powtórzeń nie został tu potwierdzony, pomimo efektywnego wyciszenia ekspresji genu kodującego tę nukleazę (rysunek 5.5.1.1 B i 5.5.1.8). Hipotezą mogącą wyjaśnić udział KU70 i KU80 w skracaniu ciągu powtórzeń po edycji sgRNA1 jest proponowany w niektórych pracach udział KU70/80 w rekrutacji do końców dwuniciowego pęknięcia białka PARP1, które następnie angażuje w naprawę odpowiadającą za delecje POLθ.¹¹⁵

W przeciwieństwie do edycji z udziałem gRNA1 indukcja DSB w obrębie ciągu powtórzeń CAG w genie *HTT* (z udziałem gRNA2) skutkowałą głównie skróceniami ciągu z zachowaniem ramki odczytu (rysunek 5.1.2). Jak zostało już wcześniej wspomniane, sekwencja fragmentu gRNA2 rozpoznającego miejsce docelowe, pozwala na wiązanie się kompleksu RNP do sekwencji komplementarnej do powtórzeń CAG. Ze względu na długość ciągu wynoszącą 41 powtórzeń CAG w genie *HTT*, czyli 123 pary zasad, teoretycznie do ciągu mogło przyłączyć się maksymalnie 6 kompleksów Cas9-gRNA2, z których każdy mógł indukować DSB.

Wyniki z rozdziału produktów PCR w elektroforezie kapilarnej uwidoczniły pojawienie się pierwszych skróceń już po 4 godzinach od transfekcji komórek kompleksem RNP (rysunek 5.2.1 B). Zaobserwowane produkty różniły się od siebie trzema lub wielokrotnością trzech par zasad, co sugeruje, że końce DNA łączyły się, w regionach homologicznych, które stanowiły same powtórzenia CAG. Ze względu na to, że skrócony ciąg powtórzeń CAG nadal mógł być rozpoznawany przez kompleks RNP zawierający gRNA2, ciąg ten mógł podlegać ponownej edycji, aż do skrócenia, do długości niebędącej już rozpoznawaną przez kompleks. Dalsze skracanie sekwencji potwierdza porównanie elektroforegramów prezentujące rozdział produktów PCR po edycji z udziałem gRNA2, uzyskanych po amplifikacji DNA izolowanego z komórek zebranych po 4 i 24 godzinach od transfekcji (rysunek 5.2.1 B) Tempo naprawy wynoszące 4 godziny od wprowadzenia kompleksu Cas9-gRNA2 sugeruje zaangażowanie szybkiego mechanizmu naprawy DNA, za jaki uchodzi NHEJ. Jego udział w skracaniu ciągu powtórzeń CAG po edycji z udziałem sgRNA2 został potwierdzony w eksperymencie, w którym badano wpływ inhibicji białek KU70 i KU80 na skracanie ciągu powtórzeń (rysunek 5.5.2.7). Wyniki z analizy resekcji po edycji regionu z ciągiem powtórzeń CAG w genie *HTT* z udziałem gRNA2 sugerują, że NHEJ nie jest jedynym mechanizmem naprawy zaangażowanym w tym przypadku. Po edycji z udziałem gRNA2 zaobserwowano rozległe resekcje końców DNA obejmujące nawet 5000 nukleotydów (rysunek 5.4.2). Z najwyższą intensywnością resekcje występowały w regionach oddalonych o około 1100 par zasad od miejsca indukcji DSB, szczególnie po stronie 5' od ciągu powtórzeń. W przeciwieństwie do edycji z udziałem gRNA1, po edycji z udziałem gRNA2 maksymalna intensywność resekcji została zaobserwowana po 24 godzinach od transfekcji komórek i pozostawała relatywnie wysoka po 48 godzinach szczególnie w regionach oddalonych o około 5000 par zasad. Wskazuje to na udział mechanizmu naprawy wykorzystującego duże resekcje końców DNA i trwającego znacznie dłużej od naprawy NHEJ. Podobnie jak w przypadku analizy po edycji komórek z udziałem sgRNA1, również po edycji z udziałem sgRNA2 stwierdzono wzbogacenie MRE11 w sekwencjach flankujących ciąg powtórzeń CAG między 12 a 48 godziną od transfekcji komórek plazmidem.¹⁸² Udział białka MRE11 w skracaniu ciągu powtórzeń CAG po edycji z udziałem sgRNA2, jak również współdziałającego z nim białka CTIP został potwierdzony w eksperymencie wykorzystującym inhibitory i siRNA (rysunek 5.5.2.7). Dodatkowo wyniki z CHIP-qPCR ujawniły wzbogacenie FEN1 w edytowanym *locus* z ciągiem powtórzeń CAG, lecz w późniejszych punktach czasowych, w porównaniu do edycji z udziałem sgRNA1¹⁸². Około trzykrotne wzbogacenie FEN1 stwierdzono po 24 i 48 godzinach po transfekcji plazmidem w komórkach HEK293T_41CAG. Podobnie jak w przypadku edycji

z udziałem sgRNA1, w przypadku edycji z udziałem gRNA2 zaobserwowano silniejsze, bo nawet pięciokrotne wzbogacenie sekwencji flankującej ciąg powtórzeń w białko FEN1 w komórkach o dłuższym ciągu powtórzeń CAG (w tym przypadku wykryto wzbogacenie nawet w punkcie czasowym 72 h). Kolejnym białkiem zaangażowanym w mechanizm naprawy DNA wykorzystujący resekcję, jest uczestniczące w HR białko RAD51. Stwierdzono, że wyciszenie genu kodującego to białko znacząco obniżyło udział produktów skróconych po edycji z udziałem sgRNA2 (rysunek 5.5.2.7). Udział HR w naprawie DSB indukowanego przy użyciu gRNA2 wspierają również wyniki z en-ChIP-MS (ang. *engineered DNA-binding molecule-mediated chromatin immunoprecipitation- mass spectrometry*) prezentowane we wspomnianej publikacji¹⁸². Zaobserwowano w nich większe wzbogacenie w białka PCNA i SMC1A regionu z powtórzeniami CAG w genie *HTT* w komórkach HEK293T_41CAG w próbkach edytowanych z udziałem sgRNA2, w porównaniu do prób edytowanych z udziałem sgRNA1 (odpowiednio o 60% i 250%)¹⁸². PCNA bierze udział w HR i zwiększa procesywność polimerazy DNA δ , również uczestniczącej w tej ścieżce naprawy²²¹, SMC1A jest zaś elementem kompleksu rekombinacyjnego i współdziała z BRCA1.²²²

Omówione powyżej wyniki poszczególnych eksperymentów pozwoliły na identyfikację białek uczestniczących w skracaniu powtórzeń CAG w wyniku indukcji w ich obrębie DSBs w komórkach ludzkich oraz na stworzenie modeli przedstawiających procesy prowadzące do tych skróceń i zależność wyboru mechanizmu naprawy DNA od miejsca indukcji DSB. Modele te zostały przedstawione na rysunkach 6.1 i 6.2



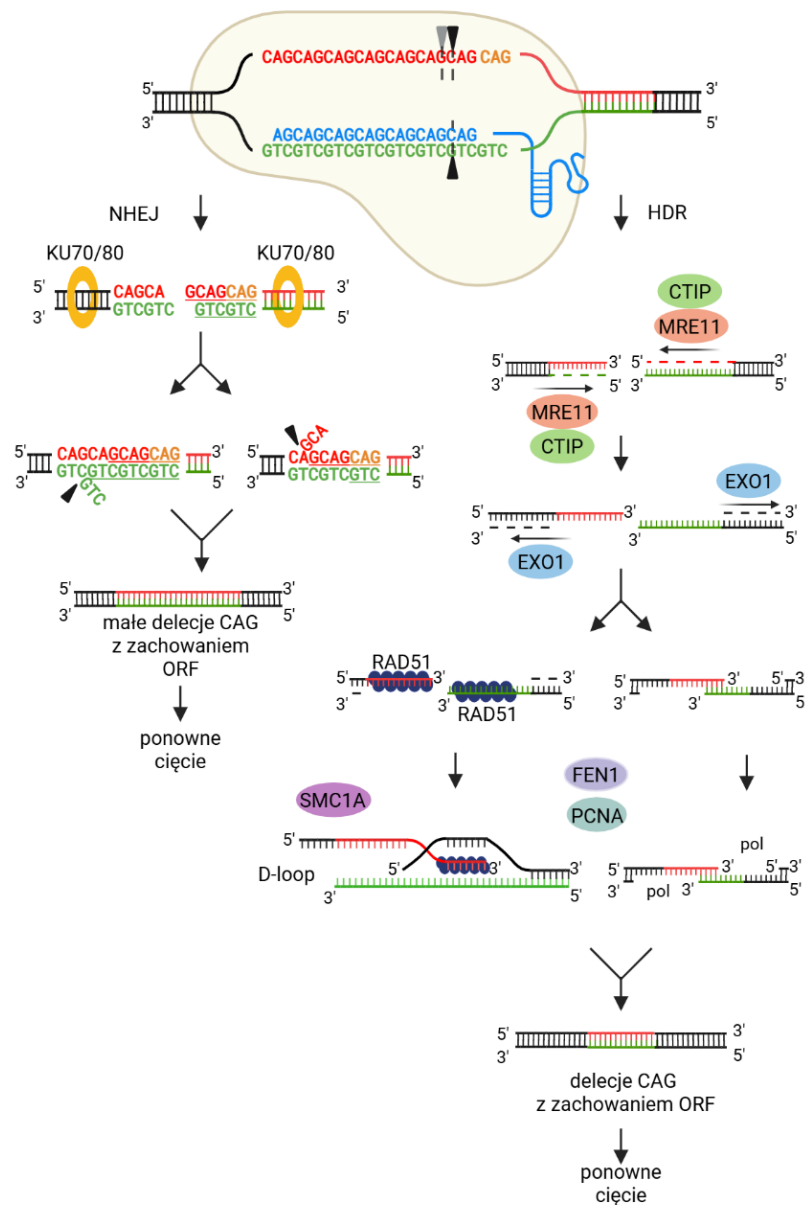
Rysunek 6.1 Model naprawy DSB indukowanego na początku ciągu powtórzeń CAG w komórkach ludzkich. Kolorem czerwonym oznaczono ciąg powtórzeń CAG, zielonym komplementarny do niego ciąg powtórzeń GTC, a pomarańczowym sekwencję PAM. Czarne trójkąty oznaczają miejsce indukcji pęknięcia nici DNA. Trójkąt szary oznacza miejsce możliwej indukcji pęknięcia na jednej z nici dsDNA przez Cas9, skutkującego insercją adeniny. Kolorem fioletowym zaznaczono miejsca występowania mikrohomologii (MH). Strzałki poziome wskazują kierunek resekcji końców DNA. (Na podstawie Sledzinski i wsp. 2024¹⁸²)

Efekt naprawy DSB w regionie znajdującym się w pewnym oddaleniu, poniżej ciągu powtórzeń CAG w genie *HTT* (edycja z udziałem gRNA4) również skutkowało skracaniem ciągu powtórzeń lub jego całkowitym wycięciem, co jest bardzo ciekawym zjawiskiem, wymagającym wyjaśnienia (rysunek 5.1.3). Ze względu na najniższą efektywność gRNA4 w porównaniu do pozostałych użytych gRNA, aspekt ten nie był rozwijany w niniejszej pracy. Jednoczesne wykorzystanie gRNA1 i gRNA4 do wycięcia całego ciągu powtórzeń CAG w genie *HTT*, w sposób spodziewany doprowadziło do delecji całego fragmentu znajdującego

się pomiędzy dwuniciowymi pęknięciami (rysunek 5.1.4). Ze względu na bardziej interesujące zjawiska skracania ciągu powtórzeń CAG w wyniku indukcji tylko jednego DSB, również ten aspekt nie był dalej rozwijany.

Jak przedstawiono powyżej, miejsce indukcji DSB ma kluczowe znaczenie w wyborze mechanizmu naprawy i drastycznie wpływa na jej końcowy efekt. Wyniki uzyskane w niniejszej pracy pozwalają na wysunięcie wniosków dotyczących wpływu długości ciągu powtórzeń CAG i sekwencji go flankującej na wybór mechanizmu naprawy. Wyniki z analizy resekcji w *locus* będącym off-targetem dla gRNA1 wykazały, że indukcja DSB w miejscu występowania jedynie sześciu powtórzeń CAG nie prowadzi do uruchomienia ścieżek naprawy DNA wykorzystujących resekcyjne (rysunek 5.4.1 B). Nie sposób tu odpowiedzieć na pytanie, w jakim zakresie wpływ na ten wynik ma sama długość ciągu, a w jakim sekwencje flankujące ten krótki ciąg. Nie bez znaczenia może być w tym przypadku stan chromatyny, który również ma udział w wyborze ścieżek naprawy DNA.

Rozstrzygnięcie powyższych wątpliwości wymagałoby przeprowadzenia dalszych badań, w których jedyną zmienną byłaby długość ciągu powtórzeń CAG, badaniu zaś poddane by było jedno *locus*, tak, by zniwelować ewentualny wpływ sekwencji flankujących i stanu chromatyny na wynik eksperymentu. W przypadku analizy resekcji w innym modelu komórkowym – fibroblastach z liczbą powtórzeń CAG w genie *HTT* wynoszącą 21 i 44 powtórzenia stwierdzono występowanie resekcji jedynie w regionach najbliższych edytowanemu miejscu (około 350 par zasad od DSB) po edycji z udziałem gRNA2, który w komórkach HEK293T_41CAG indukował resekcyjne o bardzo dużym zasięgu (rysunek 5.4.3 A). Niższy o około 50% poziom resekcji w przypadku fibroblastów w porównaniu do HEK293T_41CAG, może tłumaczyć nieefektywna edycja allelu zawierającego krótszy ciąg powtórzeń (rysunek 5.4.3 B), jednak nie tłumaczy to tak drastycznej różnicy w zasięgu resekcji. Zdaje się być to zjawiskiem komórkowo specyficznym. Niemniej udział resekcji końców DNA, choć o niewielkim zakresie, w czasie naprawy DSB indukowanego w obrębie ciągu powtórzeń CAG o patologicznej długości w ludzkich fibroblastach został potwierdzony.



Rysunek 6.2 Model naprawy DSB indukowanego wewnątrz ciągu powtórzeń CAG w komórkach ludzkich. Kolorem czerwonym oznaczono ciąg powtórzeń CAG, zielonym komplementarny do niego ciąg powtórzeń GTC, a pomarańczowym sekwencję PAM. Czarne trójkąty oznaczają miejsce indukcji pęknięcia nici DNA. Trójkąt szary oznacza miejsce możliwej indukcji pęknięcia na jednej z nici dsDNA przez Cas9. Strzałki poziome wskazują kierunek resekcji końców DNA. (Na podstawie Sledzinski i wsp. 2024¹⁸²)

Konfrontując wyniki uzyskane w niniejszej pracy z rezultatami innych publikacji, koncentrujących się na badaniu naprawy DSBs w obrębie powtórzeń CAG/CTG, zarówno w komórkach ludzkich, jak i drożdżowych, zauważalne są pewne zbieżności, ale również i różnice. Przedstawione tutaj wyniki eksperymentów, podobnie jak rezultaty badań w modelach drożdżowych, potwierdzają zaangażowanie w naprawę DSBs, indukowanych w obrębie ciągu powtórzeń CAG/CTG, białek uczestniczących w mechanizmach naprawy wykorzystujących resekcje końców DNA. W badaniach, w których do indukcji DSB wykorzystano system TALEN, wykazano zależność delekcji ciągu powtórzeń CTG/CAG od Mre11 i Rad50, składników kompleksu MRX, będącego homologiem ludzkiego kompleksu MRN oraz białka Sae2 homologa ludzkiego CTIP.¹⁸⁷ W pracy, w której indukcję DSB przeprowadzono z wykorzystaniem systemu CRISPR-Cas9, również potwierdzono udział Sae2 w dużych delekcjach obejmujących ciąg powtórzeń, a ponadto zaobserwowano zależność mniejszych delekcji od obecności drożdżowej ligazy IV, występującej także w komórkach ludzkich, charakterystycznej dla NHEJ.²¹⁴ W tej samej pracy stwierdzono zależność dużych delekcji od Rad52, które to białko, w komórkach drożdży uczestniczy w rekombinacji homologicznej i wspomaga wiązanie Rad51 do jednoniciowych końców DNA, choć udział samego Rad51 w indukowaniu delekcji ciągu nie został potwierdzony w komórkach drożdży.¹⁸⁷ Zaangażowanie RAD51 w wycinanie ciągu powtórzeń CAG, w efekcie naprawy DSB indukowanego w jego obrębie, zaobserwowane zostało zaś w pracy, w której wykorzystano system reporterowy zawierający wspomniany ciąg powtórzeń w komórkach ludzkich, co jest zbieżne z wynikami prezentowanymi w niniejszej pracy.¹⁹¹ Ciekawym aspektem prezentowanym w powyższych publikacjach dotyczących badań resekcji końców DNA w komórkach drożdży, jest asymetria resekcji występująca w warunkach braku w komórce Sae2 i Rad50. Resekcje po stronie obejmującej większą część ciągu CTG/CAG (78 powtórzeń) były silnie zahamowane w przypadku braku Sae2 i Rad50, podczas gdy drugi koniec DNA zawierający jedynie dwa powtórzenia ulegał w tych warunkach resekcji, chociaż z mniejszą intensywnością.²¹⁴ Jest to kolejna przesłanka sugerująca niebagatelny wpływ umiejscowienia DSB względem ciągu powtórzeń CAG/CTG na przebieg naprawy DNA.

Rozważając, czy stosowane w niniejszej pracy gRNA, celujące w ciąg powtórzeń CAG, mogłyby być użyteczne w terapii HD, należy wziąć pod uwagę kilka kwestii, takich jak efektywność indukcji DSB przez użytą cząsteczkę gRNA, jej specyficzność oraz uzyskany efekt naprawy, który ma decydujący wpływ na terapeutyczną skuteczność i bezpieczeństwo jej użycia. Zważywszy na te czynniki, wykorzystana w niniejszej pracy

cząsteczka gRNA2 zdaje się być najlepszym kandydatem do wykorzystania w ewentualnej przyszłej terapii, jako że indukuje DSB w obrębie ciągu powtórzeń CAG z dużą wydajnością, prowadząc w większości do skracania ciągu z zachowaniem ramki odczytu. Należy zauważyć, że cechą edycji genomu z wykorzystaniem technologii CRISPR-Cas9 jest tendencja do indukcji DSBs również w miejscach innych niż docelowe (off-target), co może skutkować wprowadzaniem niepożądanych zmian genetycznych, łącznie z rearanżacjami chromosomowymi. Jest to szczególnie prawdopodobne w przypadku stosowania gRNA2, który rozpoznaje wszelkie ciągi CAG obecne w genomie. Niemniej jak wykazano w eksperymencie na fibroblastach, krótsze ciągi powtórzeń CAG są edytowane ze znacznie niższą efektywnością w porównaniu do ciągów z ekspansją powtórzeń. Kluczowe w tym aspekcie wydaje się być dokładniejsze przeanalizowanie skutków edycji w potencjalnych miejscach off-target dla cząsteczki gRNA2. W końcu, aby określić realną, terapeutyczną użyteczność tej cząsteczki wymagane byłoby przeprowadzenie dalszych eksperymentów na komórkach neuronalnych, których terapia miałaby dotyczyć.

Podsumowując, w niniejszej pracy dokonano identyfikacji czynników i procesów odpowiedzialnych za skracanie ciągu powtórzeń CAG w endogennym *locus* w ludzkich komórkach. Uzyskane wyniki przyczyniły się do stworzenia modeli naprawy DNA w regionach z powtórzeniami CAG u ludzi. Zebrane wyniki pozwalają stwierdzić, że naprawa DNA w regionach powtórzeń CAG jest złożona i angażuje kilka mechanizmów na różnych jej etapach. Wykazano udział resekcji końców DNA w naprawie DSBs w regionach z powtórzeniami CAG i przedstawiono przesłanki sugerujące zależność jej intensywności a w konsekwencji wyboru mechanizmu naprawy DNA i jej efektów od długości ciągu, choć potwierdzenie tej zależności wymaga przeprowadzenia dalszych badań. Stwierdzono silny wpływ umiejscowienia DSB względem powtórzeń CAG oraz sekwencji flankujących ciąg na wybór mechanizmu naprawy DNA. Wykazano, że naprawa DSB indukowanego na początku ciągu powtórzeń angażuje mechanizm NHEJ, co potwierdza obecność insercji adeniny w naprawionej sekwencji, w miejscu indukcji cięcia. Jednocześnie naprawa DSB ulokowanego w tym miejscu może skutkować wycięciem całego ciągu, w co zaangażowana jest POL θ . Z kolei indukcja DSB wewnątrz ciągu, skutkuje jego skracaniem z zachowaniem ramki odczytu, co dobywa się z udziałem mechanizmu NHEJ oraz wykorzystującego dalekie resekcje mechanizmu angażującego białka rekombinacji homologicznej takie jak RAD51. W efekcie praca ta przyczyniła się do lepszego poznania zjawisk odpowiedzialnych za niestabilność ciągów powtórzeń CAG, która jest przyczyną wielu nieuleczalnych chorób. Poszerzenie

wiedzy w tej dziedzinie może przyczynić się do opracowania skutecznej i bezpiecznej terapii chorób poliQ wykorzystujących technologie edycji genomu takie jak CRISPR-Cas9. Należy też zaznaczyć, że badania podjęte w niniejszej pracy wymagają pogłębienia i rozszerzenia. Różnice zaobserwowane w naprawie DNA w komórkach HEK293T_41CAG i fibroblastach wskazują, że pogłębienia wymagają kwestie związane z komórkowo specyficzną odpowiedzią na indukcję DSBs w obrębie ciągów powtórzeń CAG. Przeprowadzenie eksperymentów na komórkach neuronalnych, w których mechanizmy naprawy DNA oparte o rekombinację homologiczną są ograniczone, ze względu na to, że są to komórki nieulegające podziałom, a które przede wszystkim ulegają degeneracji związanej z chorobami poliQ, jest kluczowe dla realistycznej oceny skuteczności i bezpieczeństwa strategii terapeutycznych opartych na indukcji DSBs w obrębie powtórzeń CAG. Tylko w tym kontekście komórkowym możliwe jest wierne odwzorowanie naturalnych warunków patologicznych oraz odpowiedzi neuronów na modyfikacje genomowe.

7. Wnioski

- 1) Naprawa DSBs w regionach powtórzeń CAG w komórkach ludzkich jest procesem złożonym, angażującym różne białka i mechanizmy naprawy DNA.
- 2) Naprawa DSBs zlokalizowanych w powtórzeniach CAG odbywa się przy zaangażowaniu mechanizmów wykorzystujących resekcje końców DNA.
- 3) Lokalizacja DSB względem ciągu powtórzeń CAG oraz sekwencje flankujące ciąg silnie wpływają na wybór mechanizmu naprawy DNA.
- 4) Indukcja DSB na początku ciągu powtórzeń CAG skutkuje wycięciem całego ciągu, angażując mechanizm naprawy zależny od POL θ (TMEJ).
- 5) Indukcja DSB wewnątrz ciągu powtórzeń CAG skutkuje skracaniem ciągu z zachowaniem ramki odczytu, wykorzystując mechanizm NHEJ, jak również zależną od dużych resekcji końców DNA naprawę z udziałem białek charakterystycznych dla rekombinacji homologicznej.

8. Spis rysunków i tabel prezentowanych w pracy

8.1 Rysunki

Rysunek 1.1 Procesy, w których dochodzi do ekspansji lub skracania ciągów powtórzeń mikrosatelitarnych.....	13
Rysunek 1.2 Czynniki wpływające na patogenezę choroby Huntingtona.....	17
Rysunek 1.3 Mechanizm naprawy DNA na drodze MMR i udział MMR w ekspansji powtórzeń CAG.....	19
Rysunek 1.4.1.1 Ścieżki naprawy dwuniciowych pęknięć DNA.....	22
Rysunek 1.4.1.2 Zależność wyboru ścieżki naprawy dwuniciowego pęknięcia DNA od fazy cyklu komórkowego.....	23
Rysunek 1.5.1.2 Schemat działania systemów ZNF i TALEN.....	30
Rysunek 1.5.1.3 Schemat działania kompleksu Cas9-sgRNA i zależność wyniku naprawy od nukleotydu w pozycji -4.....	32
Rysunek 4.2 Schemat przedstawiający ludzki gen <i>HTT</i> oraz miejsca rozpoznawane przez poszczególne gRNA wykorzystane w pracy.....	47
Rysunek 4.6 Schemat procedury użycia inhibitorów chemicznych lub siRNA oraz transfekcji komórek HEK293T_41CAG plazmidem kodującym Cas9 i sgRNA.....	50
Rysunek 4.9 Schemat selekcji enzymów restrykcyjnych na podstawie ich efektywności trawienia DNA.....	52
Rysunek 4.10 Schemat przedstawiający sposób obliczania redukcji skróceń uzyskanych po edycji z udziałem sgRNA w próbach po zastosowaniu siRNA względem kontroli.....	55
Rysunek 5.1. Efekty naprawy DSB w regionie z powtórzeniami CAG w genie <i>HTT</i> w komórkach HEK293T_41CAG.....	64
Rysunek 5.1.1.1 Częstość występowania wariantów z określoną liczbą powtórzeń CAG w próbce edytowanej z udziałem gRNA1.....	65
Rysunek 5.1.1.2 Przykłady TINS zidentyfikowanych w wynikach sekwencjonowania regionu z powtórzeniami CAG w genie <i>HTT</i> po edycji z udziałem gRNA1.....	66
Rysunek 5.1.1.3 Przykłady sekwencji mikrohomologicznych zidentyfikowanych w wynikach sekwencjonowania regionu z powtórzeniami CAG w genie <i>HTT</i> po edycji z udziałem gRNA1, wykorzystane do łączenia końców DNA w czasie naprawy DSB.....	66
Rysunek 5.1.2 Częstość występowania wariantów z określoną liczbą powtórzeń CAG w próbce edytowanej z udziałem gRNA2.....	67
Rysunek 5.1.3 Częstość występowania wariantów z określoną liczbą powtórzeń CAG w próbce edytowanej z udziałem gRNA4.....	68

Rysunek 5.1.4 Częstość występowania wariantów z określoną liczbą powtórzeń CAG w próbie edytowanej jednocześnie z udziałem gRNA1 i gRNA4.....	69
Rysunek 5.2.1 Rozdział produktów PCR uzyskanych po amplifikacji regionu z powtórzeniami CAG w genie <i>HTT</i> na matrycy DNA izolowanego w opisanych punktach czasowych z komórek HEK293T_41CAG po transfekcji RNP z gRNA1 lub gRNA2.....	70
Rysunek 5.2.2 Jednonukleotydowa insercja obserwowana po edycji z udziałem gRNA1 po 24 godzinach od transfekcji komórek.....	71
Rysunek 5.3.1 Schematyczne przedstawienie lokalizacji starterów względem edytowanych <i>loci</i>	72
Rysunek 5.3.2 Przedstawienie miejsc flankujących region z powtórzeniami CAG w genie <i>HTT</i> oraz miejsce off-target dla gRNA1 wraz z użytymi enzymami restrykcyjnymi i odległościami miejsca cięcia enzymu od odpowiedniego miejsca edycji.....	75
Rysunek 5.4.1 Poziom i zasięg resekcji końców DNA w czasie po edycji z udziałem gRNA1 w regionie z powtórzeniami CAG w <i>HTT</i> oraz w miejscu off-target dla gRNA1 w komórkach HEK293T_41CAG.	77
Rysunek 5.4.2 Poziom i zasięg resekcji końców DNA w czasie po edycji z udziałem gRNA2 w regionie z powtórzeniami CAG w <i>HTT</i> w komórkach HEK293T_41CAG.....	81
Rysunek 5.4.3 Poziom i zasięg resekcji końców DNA w czasie po edycji z udziałem gRNA2 w regionie z powtórzeniami CAG w <i>HTT</i> w fibroblastach HD oraz produkty PCR powstałe po amplifikacji <i>locus HTT</i> na matrycy DNA z fibroblastów nieedytowanych i edytowanych z udziałem gRNA2.....	83
Rysunek 5.5 Elektroforegram przedstawiający rozdział produktów PCR po amplifikacji <i>locus</i> z powtórzeniami CAG w genie <i>HTT</i> w komórkach HEK293T_41CAG nieedytowanego oraz po edycji z udziałem sgRNA1 i sgRNA2.....	87
Rysunek 5.5.1.1 Efekt wyciszenia genów kodujących poszczególne białka naprawy DNA przy użyciu siRNA określony na poziomie mRNA lub białka.....	88
Rysunek 5.5.1.2 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w <i>HTT</i> w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA1, traktowanych inhibitorem CTIP oraz bez udziału inhibitora.....	89
Rysunek 5.5.1.3 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w <i>HTT</i> w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA1, traktowanych poszczególnymi inhibitorami oraz bez udziału inhibitorów.....	90
Rysunek 5.5.1.4 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w <i>HTT</i> w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA1, traktowanych inhibitorem POLθ oraz bez udziału inhibitora.....	91

Rysunek 5.5.1.5 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w <i>HTT</i> w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA1, traktowanych siRNA oraz bez udziału siRNA.....	92
Rysunek 5.5.1.6 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w <i>HTT</i> w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA1, traktowanych siRNA oraz bez udziału siRNA.....	93
Rysunek 5.5.1.7 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w <i>HTT</i> w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA1, traktowanych siRNA oraz bez udziału siRNA.....	94
Rysunek 5.5.1.8 Udział produktów amplifikacji regionu z powtórzeniami CAG w <i>HTT</i> zawierających skrócenia ciągu oraz produktów wyjściowych w komórkach HEK293T_41CAG po edycji z udziałem sgRNA1 oraz obliczona redukcja skróceń ciągu CAG po zastosowaniu inhibitora lub siRNA po odniesieniu do kontroli.....	95
Rysunek 5.5.2.1 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w <i>HTT</i> w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA2, traktowanych poszczególnymi inhibitorami oraz bez udziału inhibitorów.....	97
Rysunek 5.5.2.2 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w <i>HTT</i> w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA2, traktowanych inhibitorem białek KU70 i KU80 oraz bez udziału inhibitora.....	98
Rysunek 5.5.2.3 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w <i>HTT</i> w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA2, traktowanych inhibitorem POLθ oraz bez udziału inhibitora.....	99
Rysunek 5.5.2.4 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w <i>HTT</i> w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA2, traktowanych siRNA oraz bez udziału siRNA.....	100
Rysunek 5.5.2.5 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w <i>HTT</i> w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA2, traktowanych siRNA oraz bez udziału siRNA.....	101
Rysunek 5.5.2.6 Elektroforegramy przedstawiające rozdział produktów PCR po amplifikacji regionu z powtórzeniami CAG w <i>HTT</i> w komórkach HEK293T_41CAG edytowanych z udziałem sgRNA2, traktowanych siRNA oraz bez udziału siRNA.....	102
Rysunek 5.5.2.7 Udział produktów amplifikacji <i>locus</i> z powtórzeniami CAG w <i>HTT</i> zawierających skrócenia ciągu oraz produktów wyjściowych w komórkach HEK293T_41CAG po edycji z udziałem sgRNA2 oraz obliczona redukcja skróceń ciągu CAG po zastosowaniu inhibitora lub siRNA po odniesieniu do kontroli.....	103

Rysunek 5.5.2.8 Porównanie rozkładu pików na elektroforegramach z rozdziałem produktów PCR po amplifikacji regionu z powtórzeniami CAG w <i>HTT</i> dla prób edytowanych sgRNA1: kontrolnej i traktowanej inhibitorem MRE11.....	104
---	-----

Rysunek 5.6.1 Rozkład częstości występowania wariantów z określoną liczbą powtórzeń CAG w produktach PCR po amplifikacji regionu z powtórzeniami CAG w genie <i>HTT</i> , w komórkach HEK293T_41CAG edytowanych z udziałem gRNA1 z wyciszeniem ekspresji <i>POLQ</i> oraz bez wyciszenia (kontrola). Względna ekspresja genu <i>POLQ</i> po zastosowaniu siRNA.....	106
--	-----

Rysunek 5.6.2 Procentowy udział poszczególnych typów insercji w wynikach sekwencjonowania regionu z powtórzeniami CAG w <i>HTT</i> , w komórkach HEK293T_41CAG po edycji z udziałem gRNA1 z wyciszeniem ekspresji <i>POLQ</i> oraz bez wyciszenia.....	107
---	-----

Rysunek 6.1 Model naprawy DSB indukowanego na początku ciągu powtórzeń CAG w komórkach ludzkich.....	114
---	-----

Rysunek 6.2 Model naprawy DSB indukowanego wewnątrz ciągu powtórzeń CAG w komórkach ludzkich.....	116
--	-----

8.2 Tabele

Tabela 1.2 Zestawienie przykładowych chorób związanych z ekspansją powtórzeń mikrosatelitarnych.....	14
---	----

Tabela 3.2. Odczynniki używane w pracy do hodowli i transfekcji komórek.....	37
---	----

Tabela 3.3 Zestawienie oligonukleotydów wykorzystywanych w pracy.....	38
--	----

Tabela 3.4. Przeciwciała pierwszo i drugorzędowe wykorzystane w pracy.....	40
---	----

Tabela 3.5.1. Inhibitory białek naprawy DNA wykorzystane w pracy.....	40
--	----

Tabela 3.5.2. siRNA wykorzystane w pracy.....	40
--	----

Tabela 3.6. Enzymy restrykcyjne wykorzystane w pracy.....	41
--	----

Tabela 3.7 Składniki buforów i żeli do elektroforezy wykorzystywanych w pracy.....	45
---	----

Tabela 3.8. Zestawy odczynników wykorzystywane w pracy.....	46
--	----

Tabela 4.3. Szczegółowe warunki elektroporacji dla różnych typów i liczby komórek.....	48
---	----

Tabela 4.9 Enzymy restrykcyjne testowane w poszczególnych regionach.....	53
---	----

Tabela 5.3. Zestawienie wartości ΔCt i poziomu resekcji uzyskanych podczas testowania poszczególnych enzymów w określonych regionach w komórkach HEK293T_41CAG i w fibroblastach.....	73
--	----

Tabela 5.4.1.1 Wyniki poziomu resekcji w różnych punktach czasowych od transfekcji RNP z gRNA1 oraz kontroli w regionach flankujących region z powtórzeniami CAG w genie <i>HTT</i> w komórkach HEK293T_41CAG.....	78
Tabela 5.4.1.2 Wyniki poziomu resekcji w różnych punktach czasowych od transfekcji RNP z gRNA1 oraz kontroli w regionach flankujących <i>locus</i> off-target dla gRNA1 w komórkach HEK293T_41CAG.....	79
Tabela 5.4.2 Wyniki poziomu resekcji w różnych punktach czasowych od transfekcji RNP z gRNA2 oraz kontroli w regionach flankujących region z powtórzeniami CAG w genie <i>HTT</i> w komórkach HEK293T_41CAG.....	81
Tabela 5.4.3 Wyniki poziomu resekcji w różnych punktach czasowych od transfekcji RNP z gRNA2 oraz kontroli w regionach flankujących region z powtórzeniami CAG w genie <i>HTT</i> w fibroblastach.	84
Tabela 5.5 Białka naprawy DNA, dla których wykonano badanie ich udziału w skracaniu ciągu powtórzeń CAG w genie <i>HTT</i>	86

9. Literatura

1. Wright, S. E. & Todd, P. K. Native functions of short tandem repeats. *eLife* **12**, e84043 (2023).
2. Malik, I., Kelley, C. P., Wang, E. T. & Todd, P. K. Molecular mechanisms underlying nucleotide repeat expansion disorders. *Nat Rev Mol Cell Biol* **22**, 589–607 (2021).
3. Quilez, J. *et al.* Polymorphic tandem repeats within gene promoters act as modifiers of gene expression and DNA methylation in humans. *Nucleic Acids Research* **44**, 3750–3762 (2016).
4. Fotsing, S. F. *et al.* The impact of short tandem repeat variation on gene expression. *Nat Genet* **51**, 1652–1659 (2019).
5. Willems, T., Gymrek, M., Poznik, G. D., Tyler-Smith, C. & Erlich, Y. Population-Scale Sequencing Data Enable Precise Estimates of Y-STR Mutation Rates. *The American Journal of Human Genetics* **98**, 919–933 (2016).
6. Verbiest, M. *et al.* Mutation and selection processes regulating short tandem repeats give rise to genetic and phenotypic diversity across species. *Journal of Evolutionary Biology* **36**, 321–336 (2023).
7. Ihara, N. *et al.* A comprehensive genetic map of the cattle genome based on 3802 microsatellites. *Genome Res* **14**, 1987–1998 (2004).
8. Kayser, M. Forensic use of Y-chromosome DNA: a general overview. *Hum Genet* **136**, 621–635 (2017).
9. Khristich, A. N. & Mirkin, S. M. On the wrong DNA track: Molecular mechanisms of repeat-mediated genome instability. *J. Biol. Chem.* **295**, 4134–4170 (2020).
10. Chintalaphani, S. R., Pineda, S. S., Deveson, I. W. & Kumar, K. R. An update on the neurological short tandem repeat expansion disorders and the emergence of long-read sequencing diagnostics. *Acta Neuropathologica Communications* **9**, 98 (2021).
11. Swinnen, B., Robberecht, W. & Van Den Bosch, L. RNA toxicity in non-coding repeat expansion disorders. *The EMBO Journal* **39**, e101112 (2020).

12. Colak, D. *et al.* Promoter-Bound Trinucleotide Repeat mRNA Drives Epigenetic Silencing in Fragile X Syndrome. *Science* **343**, 1002–1005 (2014).
13. Nobile, V., Pucci, C., Chiurazzi, P., Neri, G. & Tabolacci, E. DNA Methylation, Mechanisms of FMR1 Inactivation and Therapeutic Perspectives for Fragile X Syndrome. *Biomolecules* **11**, 296 (2021).
14. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **72**, 245–256 (2011).
15. Jiang, J. & Ravits, J. Pathogenic Mechanisms and Therapy Development for C9orf72 Amyotrophic Lateral Sclerosis/Frontotemporal Dementia. *Neurotherapeutics* **16**, 1115–1132 (2019).
16. Glasmacher, S. A., Wong, C., Pearson, I. E. & Pal, S. Survival and Prognostic Factors in C9orf72 Repeat Expansion Carriers: A Systematic Review and Meta-analysis. *JAMA Neurology* **77**, 367–376 (2020).
17. Giesselmann, P. *et al.* Analysis of short tandem repeat expansions and their methylation state with nanopore sequencing. *Nat Biotechnol* **37**, 1478–1481 (2019).
18. Musova, Z. *et al.* Highly unstable sequence interruptions of the CTG repeat in the myotonic dystrophy gene. *Am J Med Genet A* **149A**, 1365–1374 (2009).
19. Pascual-Gilabert, M., Artero, R. & López-Castel, A. The myotonic dystrophy type 1 drug development pipeline: 2022 edition. *Drug Discovery Today* **28**, 103489 (2023).
20. Bauer, P. & Nukina, N. The pathogenic mechanisms of polyglutamine diseases and current therapeutic strategies. *Journal of neurochemistry* **110**, 1737–65 (2009).
21. Trushina, E. *et al.* Mutant Huntingtin Impairs Axonal Trafficking in Mammalian Neurons In Vivo and In Vitro. *Molecular and Cellular Biology* **24**, 8195–8209 (2004).
22. Temgire, P., Arthur, R. & Kumar, P. Neuroinflammation and the role of epigenetic-based therapies for Huntington’s disease management: the new paradigm. *Inflammopharmacol* **32**, 1791–1804 (2024).

23. Neueder, A. *et al.* The pathogenic exon 1 HTT protein is produced by incomplete splicing in Huntington's disease patients. *Sci Rep* **7**, 1307 (2017).
24. Sathasivam, K. *et al.* Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease. *PNAS* **110**, 2366–2370 (2013).
25. Neueder, A., Dumas, A. A., Benjamin, A. C. & Bates, G. P. Regulatory mechanisms of incomplete huntingtin mRNA splicing. *Nat Commun* **9**, 3955 (2018).
26. Schilling, J. *et al.* Deregulated Splicing Is a Major Mechanism of RNA-Induced Toxicity in Huntington's Disease. *Journal of Molecular Biology* **431**, 1869–1877 (2019).
27. Nalavade, R., Griesche, N., Ryan, D. P., Hildebrand, S. & Krauß, S. Mechanisms of RNA-induced toxicity in CAG repeat disorders. *Cell Death Dis* **4**, e752–e752 (2013).
28. Graham, R. K. *et al.* Cleavage at the Caspase-6 Site Is Required for Neuronal Dysfunction and Degeneration Due to Mutant Huntingtin. *Cell* **125**, 1179–1191 (2006).
29. Tebbenkamp, A. T. N. *et al.* Analysis of Proteolytic Processes and Enzymatic Activities in the Generation of Huntingtin N-Terminal Fragments in an HEK293 Cell Model. *PLOS ONE* **7**, e50750 (2012).
30. Bañez-Coronel, M. *et al.* RAN Translation in Huntington Disease. *Neuron* **88**, 667–677 (2015).
31. Malla, B., Guo, X., Senger, G., Chasapopoulou, Z. & Yildirim, F. A Systematic Review of Transcriptional Dysregulation in Huntington's Disease Studied by RNA Sequencing. *Front Genet* **12**, 751033 (2021).
32. Jiang, A., Handley, R. R., Lehnert, K. & Snell, R. G. From Pathogenesis to Therapeutics: A Review of 150 Years of Huntington's Disease Research. *International Journal of Molecular Sciences* **24**, 13021 (2023).
33. Latoszek, E. & Czeredys, M. Molecular Components of Store-Operated Calcium Channels in the Regulation of Neural Stem Cell Physiology, Neurogenesis, and the Pathology of Huntington's Disease. *Front. Cell Dev. Biol.* **9**, 657337 (2021).

34. Arrasate, M., Mitra, S., Schweitzer, E. S., Segal, M. R. & Finkbeiner, S. Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* **431**, 805–810 (2004).
35. Slow, E. J. *et al.* Absence of behavioral abnormalities and neurodegeneration in vivo despite widespread neuronal huntingtin inclusions. *Proceedings of the National Academy of Sciences* **102**, 11402–11407 (2005).
36. Hodgson, J. G. *et al.* A YAC Mouse Model for Huntington's Disease with Full-Length Mutant Huntingtin, Cytoplasmic Toxicity, and Selective Striatal Neurodegeneration. *Neuron* **23**, 181–192 (1999).
37. McMurray, C. T. Mechanisms of trinucleotide repeat instability during human development. *Nat Rev Genet* **11**, 786–799 (2010).
38. Wang, Y. *et al.* Concurrent Preimplantation Genetic Testing and Competence Assessment of Human Embryos by Transcriptome Sequencing. *Advanced Science* **11**, 2309817 (2024).
39. Kennedy, L. *et al.* Dramatic tissue-specific mutation length increases are an early molecular event in Huntington disease pathogenesis. *Human Molecular Genetics* **12**, 3359–3367 (2003).
40. Shelbourne, P. F. *et al.* Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain. *Human Molecular Genetics* **16**, 1133–1142 (2007).
41. Lahue, R. S. New developments in Huntington's disease and other triplet repeat diseases: DNA repair turns to the dark side. *Neuronal Signal* **4**, NS20200010 (2020).
42. Jones, L., Houlden, H. & Tabrizi, S. J. DNA repair in the trinucleotide repeat disorders. *The Lancet Neurology* **16**, 88–96 (2017).
43. Guo, J., Chen, L. & Li, G.-M. DNA mismatch repair in trinucleotide repeat instability. *Sci. China Life Sci.* **60**, 1087–1092 (2017).
44. Iyer, R. R. & Pluciennik, A. DNA Mismatch Repair and its Role in Huntington's Disease. *JHD* **10**, 75–94 (2021).

45. Pengo, M. & Squitieri, F. Beyond CAG Repeats: The Multifaceted Role of Genetics in Huntington Disease. *Genes* **15**, 807 (2024).
46. Ferguson, R., Goold, R., Coupland, L., Flower, M. & Tabrizi, S. J. Therapeutic validation of MMR-associated genetic modifiers in a human ex vivo model of Huntington's disease. 2023.12.05.570095 Preprint at <https://doi.org/10.1101/2023.12.05.570095> (2023).
47. Liu, D., Keijzers, G. & Rasmussen, L. J. DNA mismatch repair and its many roles in eukaryotic cells. *Mutation Research/Reviews in Mutation Research* **773**, 174–187 (2017).
48. Axford, M. M. *et al.* Detection of Slipped-DNAs at the Trinucleotide Repeats of the Myotonic Dystrophy Type I Disease Locus in Patient Tissues. *PLOS Genetics* **9**, e1003866 (2013).
49. Yamamoto, H. *et al.* Microsatellite instability: A 2024 update. *Cancer Science* **115**, 1738–1748 (2024).
50. Tomé, S. *et al.* MSH3 Polymorphisms and Protein Levels Affect CAG Repeat Instability in Huntington's Disease Mice. *PLOS Genetics* **9**, e1003280 (2013).
51. Keogh, N., Chan, K. Y., Li, G.-M. & Lahue, R. S. MutS β abundance and Msh3 ATP hydrolysis activity are important drivers of CTG•CAG repeat expansions. *Nucleic Acids Research* **45**, 10068–10078 (2017).
52. Hong, E. P. *et al.* Huntington's Disease Pathogenesis: Two Sequential Components. *Journal of Huntington's Disease* **10**, 35–51 (2021).
53. Dragileva, E. *et al.* Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. *Neurobiology of Disease* **33**, 37–47 (2009).
54. Wang, N. *et al.* Msh3 and Pms1 Set Neuronal CAG-repeat Migration Rate to Drive Selective Striatal and Cortical Pathogenesis in HD Mice. 2024.07.09.602815 Preprint at <https://doi.org/10.1101/2024.07.09.602815> (2024).
55. Aldous, S. G. *et al.* A CAG repeat threshold for therapeutics targeting somatic instability in Huntington's disease. *Brain* **147**, 1784–1798 (2024).

56. Mätlik, K. *et al.* Cell-type-specific CAG repeat expansions and toxicity of mutant Huntingtin in human striatum and cerebellum. *Nat Genet* 1–12 (2024) doi:10.1038/s41588-024-01653-6.
57. Handsaker, R. E. *et al.* Long somatic DNA-repeat expansion drives neurodegeneration in Huntington disease. 2024.05.17.592722 Preprint at <https://doi.org/10.1101/2024.05.17.592722> (2024).
58. Pinto, R. M. *et al.* Identification of genetic modifiers of Huntington’s disease somatic CAG repeat instability by in vivo CRISPR-Cas9 genome editing. 2024.06.08.597823 Preprint at <https://doi.org/10.1101/2024.06.08.597823> (2024).
59. Driscoll, R. *et al.* Dose-dependent reduction of somatic expansions but not Htt aggregates by di-valent siRNA-mediated silencing of MSH3 in HdhQ111 mice. *Sci Rep* **14**, 2061 (2024).
60. O’Reilly, D. *et al.* Di-valent siRNA-mediated silencing of MSH3 blocks somatic repeat expansion in mouse models of Huntington’s disease. *Molecular Therapy* **31**, 1661–1674 (2023).
61. Bunting, E. L. *et al.* Antisense oligonucleotide-mediated MSH3 suppression reduces somatic CAG repeat expansion in Huntington’s disease iPSC-derived striatal neurons. *Science Translational Medicine* **17**, eadn4600 (2025).
62. Phadte, A. S. *et al.* FAN1 removes triplet repeat extrusions via a PCNA- and RFC-dependent mechanism. *Proceedings of the National Academy of Sciences* **120**, e2302103120 (2023).
63. Loupe, J. M. *et al.* Promotion of somatic CAG repeat expansion by Fan1 knock-out in Huntington’s disease knock-in mice is blocked by Mlh1 knock-out. *Human Molecular Genetics* **29**, 3044–3053 (2020).
64. Goold, R. *et al.* FAN1 controls mismatch repair complex assembly via MLH1 retention to stabilize CAG repeat expansion in Huntington’s disease. *Cell Reports* **36**, 109649 (2021).
65. Porro, A. *et al.* FAN1-MLH1 interaction affects repair of DNA interstrand cross-links and slipped-CAG/CTG repeats. *Science Advances* (2021) doi:10.1126/sciadv.abf7906.

66. Kim, K.-H. *et al.* Genetic and Functional Analyses Point to FAN1 as the Source of Multiple Huntington Disease Modifier Effects. *The American Journal of Human Genetics* **107**, 96–110 (2020).
67. Goold, R. *et al.* FAN1 modifies Huntington's disease progression by stabilizing the expanded HTT CAG repeat. *Human Molecular Genetics* **28**, 650–661 (2019).
68. Cheng, Y.-C. *et al.* Increased FAN1 expression by mRNA-LNP attenuates CAG repeat expansion in Huntington patients' iPSC-derived astrocytes. Preprint at <https://doi.org/10.1101/2023.11.24.568451> (2023).
69. Bustamante, M. B. *et al.* C006 Insights into the DNA mismatch repair system exploiting a frameshift mutator cell-based assay to identify new targets and potential treatments for Huntington's disease. *J Neurol Neurosurg Psychiatry* **95**, A28–A28 (2024).
70. Ceccaldi, R., Rondinelli, B. & D'Andrea, A. D. Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends in Cell Biology* **26**, 52–64 (2016).
71. Bouwman, B. A. M. & Crosetto, N. Endogenous DNA Double-Strand Breaks during DNA Transactions: Emerging Insights and Methods for Genome-Wide Profiling. *Genes* **9**, 632 (2018).
72. De Bragança, S., Dillingham, M. S. & Moreno-Herrero, F. Recent insights into eukaryotic double-strand DNA break repair unveiled by single-molecule methods. *Trends in Genetics* **39**, 924–940 (2023).
73. Chang, H., Pannunzio, N., Adachi, N. & Lieber, M. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature reviews. Molecular cell biology* **18**, (2017).
74. Scully, R., Panday, A., Elango, R. & Willis, N. A. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nature Reviews Molecular Cell Biology* **20**, 698–714 (2019).
75. Xue, C. & Greene, E. C. DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing. *Trends in Genetics* (2021) doi:10.1016/j.tig.2021.02.008.

76. Zhou, Y. & Caron, P. Quantitation of DNA double-strand break resection intermediates in human cells. *Nucleic Acids Research* **42**, 11 (2014).
77. Jazayeri, A. *et al.* ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. *Nat Cell Biol* **8**, 37–45 (2006).
78. Keshavarzi, M. *et al.* MicroRNAs-Based Imaging Techniques in Cancer Diagnosis and Therapy. *Journal of Cellular Biochemistry* **118**, 4121–4128 (2017).
79. Peterson, S. E. *et al.* Cdk1 uncouples CtIP-dependent resection and Rad51 filament formation during M-phase double-strand break repair. *Journal of Cell Biology* **194**, 705–720 (2011).
80. Yun, M. H. & Hiom, K. CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle. *Nature* **459**, 460–463 (2009).
81. Ma, Y., Schwarz, K. & Lieber, M. R. The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps. *DNA Repair* **4**, 845–851 (2005).
82. Hustedt, N. & Durocher, D. The control of DNA repair by the cell cycle. *Nat Cell Biol* **19**, 1–9 (2017).
83. Gerodimos, C. A., Chang, H. H. Y., Watanabe, G. & Lieber, M. R. Effects of DNA end configuration on XRCC4-DNA ligase IV and its stimulation of Artemis activity. *J. Biol. Chem.* **292**, 13914–13924 (2017).
84. Price, B. D. & D’Andrea, A. D. Chromatin Remodeling at DNA Double-Strand Breaks. *Cell* **152**, 1344–1354 (2013).
85. Moyal, L. *et al.* Requirement of ATM-Dependent Monoubiquitylation of Histone H2B for Timely Repair of DNA Double-Strand Breaks. *Molecular Cell* **42**, 137 (2011).
86. Xie, A. *et al.* Distinct Roles of Chromatin-Associated Proteins MDC1 and 53BP1 in Mammalian Double-Strand Break Repair. *Molecular Cell* **28**, 1045–1057 (2007).
87. Aleksandrov, R., Hristova, R., Stoyanov, S. & Gospodinov, A. The Chromatin Response to Double-Strand DNA Breaks and Their Repair. *Cells* **9**, 1853 (2020).

88. Chang, H. H. Y., Watanabe, G. & Lieber, M. R. Unifying the DNA End-processing Roles of the Artemis Nuclease KU-DEPENDENT ARTEMIS RESECTION AT BLUNT DNA ENDS. *J. Biol. Chem.* **290**, 24036–24050 (2015).
89. Spagnolo, L., Rivera-Calzada, A., Pearl, L. H. & Llorca, O. Three-Dimensional Structure of the Human DNA-PKcs/Ku70/Ku80 Complex Assembled on DNA and Its Implications for DNA DSB Repair. *Molecular Cell* **22**, 511–519 (2006).
90. Chang, H. *et al.* Different DNA End Configurations Dictate Which NHEJ Components Are Most Important for Joining Efficiency. *The Journal of Biological Chemistry* **291**, 24377–24389 (2016).
91. Ma, Y., Pannicke, U., Schwarz, K. & Lieber, M. R. Hairpin Opening and Overhang Processing by an Artemis/DNA-Dependent Protein Kinase Complex in Nonhomologous End Joining and V(D)J Recombination. *Cell* **108**, 781–794 (2002).
92. Gu, J. *et al.* XRCC4:DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps. *The EMBO Journal* **26**, 1010–1023 (2007).
93. Ma, Y. *et al.* A Biochemically Defined System for Mammalian Nonhomologous DNA End Joining. *Molecular Cell* **16**, 701–713 (2004).
94. Belousova, E. A. & Lavrik, O. I. DNA polymerases β and λ and their roles in cell. *DNA Repair* **29**, 112–126 (2015).
95. Lavin, M. F. ATM and the Mre11 complex combine to recognize and signal DNA double-strand breaks. *Oncogene* **26**, 7749–7758 (2007).
96. Sartori, A. A. *et al.* Human CtIP promotes DNA end resection. *Nature* **450**, 509–514 (2007).
97. Mimitou, E. P. & Symington, L. S. Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature* **455**, 770–774 (2008).
98. Nimmonkar, A. V. *et al.* BLM–DNA2–RPA–MRN and EXO1–BLM–RPA–MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev.* **25**, 350–362 (2011).
99. Ceppi, I. *et al.* CtIP promotes the motor activity of DNA2 to accelerate long-range DNA end resection. *PNAS* **117**, 8859–8869 (2020).

100. Daley, J. M. *et al.* Enhancement of BLM-DNA2-Mediated Long-Range DNA End Resection by CtIP. *Cell Reports* **21**, 324–332 (2017).
101. Gall-Duncan, T. *et al.* Antagonistic roles of canonical and Alternative-RPA in disease-associated tandem CAG repeat instability. *Cell* **186**, 4898–4919.e25 (2023).
102. Jensen, R. B., Carreira, A. & Kowalczykowski, S. C. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* **467**, 678–683 (2010).
103. So, A. *et al.* RAD51 protects against nonconservative DNA double-strand break repair through a nonenzymatic function. *Nucleic Acids Research* **50**, 2651–2666 (2022).
104. Ahrabi, S. *et al.* A role for human homologous recombination factors in suppressing microhomology-mediated end joining. *Nucleic Acids Res* **44**, 5743–5757 (2016).
105. Moynahan, M. E., Pierce, A. J. & Jasin, M. BRCA2 Is Required for Homology-Directed Repair of Chromosomal Breaks. *Molecular Cell* **7**, 263–272 (2001).
106. Sy, S. M. H., Huen, M. S. Y. & Chen, J. PALB2 is an integral component of the BRCA complex required for homologous recombination repair. *Proceedings of the National Academy of Sciences* **106**, 7155–7160 (2009).
107. Zhao, W. *et al.* BRCA1–BARD1 promotes RAD51-mediated homologous DNA pairing. *Nature* **550**, 360–365 (2017).
108. McVey, M., Khodaverdian, V. Y., Meyer, D., Cerqueira, P. G. & Heyer, W.-D. Eukaryotic DNA Polymerases in Homologous Recombination. *Annu. Rev. Genet.* **50**, 393–421 (2016).
109. Kramara, J., Osia, B. & Malkova, A. Break-Induced Replication: The Where, The Why, and The How. *Trends in Genetics* **34**, 518–531 (2018).
110. Truong, L. N. *et al.* Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *PNAS* **110**, 7720–7725 (2013).

111. Seol, J.-H., Shim, E. Y. & Lee, S. E. Microhomology-mediated end joining: Good, bad and ugly. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **809**, 81–87 (2018).
112. Sallmyr, A. & Tomkinson, A. E. Repair of DNA double-strand breaks by mammalian alternative end-joining pathways. *Journal of Biological Chemistry* **293**, 10536–10546 (2018).
113. Wang, M. *et al.* PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Res* **34**, 6170–6182 (2006).
114. Luedeman, M. E. *et al.* Poly(ADP) ribose polymerase promotes DNA polymerase theta-mediated end joining by activation of end resection. *Nat Commun* **13**, 4547 (2022).
115. Caron, M.-C. *et al.* Poly(ADP-ribose) polymerase-1 antagonizes DNA resection at double-strand breaks. *Nat Commun* **10**, 2954 (2019).
116. Krüger, A., Bürkle, A., Hauser, K. & Mangerich, A. Real-time monitoring of PARP1-dependent PARylation by ATR-FTIR spectroscopy. *Nature Communications* **11**, 2174 (2020).
117. Pascal, J. M. The comings and goings of PARP-1 in response to DNA damage. *DNA Repair* **71**, 177–182 (2018).
118. Pleschke, J. M., Kleczkowska, H. E., Strohm, M. & Althaus, F. R. Poly(ADP-ribose) Binds to Specific Domains in DNA Damage Checkpoint Proteins *. *Journal of Biological Chemistry* **275**, 40974–40980 (2000).
119. Wyatt, D. W. *et al.* Essential Roles for Polymerase θ -Mediated End Joining in the Repair of Chromosome Breaks. *Molecular Cell* **63**, 662–673 (2016).
120. Black, S. J. *et al.* Molecular basis of microhomology-mediated end-joining by purified full-length Pol θ . *Nature Communications* **10**, 4423 (2019).
121. Carvajal-Garcia, J. *et al.* Mechanistic basis for microhomology identification and genome scarring by polymerase theta. *Proc Natl Acad Sci U S A* **117**, 8476–8485 (2020).

122. Schimmel, J., van Schendel, R., den Dunnen, J. T. & Tijsterman, M. Templated Insertions: A Smoking Gun for Polymerase Theta-Mediated End Joining. *Trends in Genetics* **35**, 632–644 (2019).
123. Zahn, K. E., Jensen, R. B., Wood, R. D. & Doublié, S. Human DNA polymerase θ harbors DNA end-trimming activity critical for DNA repair. *Molecular Cell* **81**, 1534-1547.e4 (2021).
124. Brambati, A., Barry, R. M. & Sfeir, A. DNA polymerase theta (Pol θ) – an error-prone polymerase necessary for genome stability. *Current Opinion in Genetics & Development* **60**, 119–126 (2020).
125. Ahmad, A. *et al.* ERCC1-XPF endonuclease facilitates DNA double-strand break repair. *Molecular and Cellular Biology* **28**, 5082–5092 (2008).
126. Mengwasser, K. E. *et al.* Genetic Screens Reveal *FEN1* and *APEX2* as *BRCA2* Synthetic Lethal Targets. *Molecular Cell* **73**, 885-899.e6 (2019).
127. Dehé, P.-M. & Gaillard, P.-H. L. Control of structure-specific endonucleases to maintain genome stability. *Nat Rev Mol Cell Biol* **18**, 315–330 (2017).
128. Liang, L. *et al.* Human DNA ligases I and III, but not ligase IV, are required for microhomology-mediated end joining of DNA double-strand breaks. *Nucleic Acids Research* **36**, 3297–3310 (2008).
129. Hanscom, T. *et al.* Characterization of sequence contexts that favor alternative end joining at Cas9-induced double-strand breaks. *Nucleic Acids Research* **50**, 7465–7478 (2022).
130. Kent, T., Mateos-Gomez, P. A., Sfeir, A. & Pomerantz, R. T. Polymerase θ is a robust terminal transferase that oscillates between three different mechanisms during end-joining. *eLife* **5**, e13740 (2016).
131. Frit, P., Barboule, N., Yuan, Y., Gomez, D. & Calsou, P. Alternative end-joining pathway(s): Bricolage at DNA breaks. *DNA Repair* **17**, 81–97 (2014).
132. Simsek, D. *et al.* DNA Ligase III Promotes Alternative Nonhomologous End-Joining during Chromosomal Translocation Formation. *PLOS Genetics* **7**, e1002080 (2011).

133. Bennardo, N., Cheng, A., Huang, N. & Stark, J. M. Alternative-NHEJ Is a Mechanistically Distinct Pathway of Mammalian Chromosome Break Repair. *PLOS Genetics* **4**, e1000110 (2008).
134. Mendez-Dorantes, C., Bhargava, R. & Stark, J. M. Repeat-mediated deletions can be induced by a chromosomal break far from a repeat, but multiple pathways suppress such rearrangements. *Genes Dev.* **32**, 524–536 (2018).
135. Bhargava, R., Onyango, D. O. & Stark, J. M. Regulation of Single-Strand Annealing and its Role in Genome Maintenance. *Trends in Genetics* **32**, 566–575 (2016).
136. Al-minawi, A. Z., Saleh-gohari, N. & Helleday, T. The ERCC1/XPF endonuclease is required for efficient single-strand annealing and gene conversion in mammalian cells. *Nucleic Acids Research* **36**, 1–9 (2008).
137. Ramakrishnan, S., Kockler, Z., Evans, R., Downing, B. D. & Malkova, A. Single-strand annealing between inverted DNA repeats: Pathway choice, participating proteins, and genome destabilizing consequences. *PLOS Genetics* **14**, e1007543 (2018).
138. Richardson, C. & Jasin, M. Frequent chromosomal translocations induced by DNA double-strand breaks. *Nature* **405**, 697–700 (2000).
139. Frangoul, H. *et al.* CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β -Thalassemia. *New England Journal of Medicine* **384**, 252–260 (2021).
140. Richard, G.-F. Shortening trinucleotide repeats using highly specific endonucleases: a possible approach to gene therapy? *Trends in Genetics* **31**, 177–186 (2015).
141. Carroll, D. Genome Engineering With Zinc-Finger Nucleases. *Genetics* **188**, 773–782 (2011).
142. Becker, S. & Boch, J. TALE and TALEN genome editing technologies. *Gene and Genome Editing* **2**, 100007 (2021).
143. Leenay, R. T. & Beisel, C. L. Deciphering, Communicating, and Engineering the CRISPR PAM. *Journal of Molecular Biology* **429**, 177–191 (2017).
144. Jiang, F. & Doudna, J. A. CRISPR–Cas9 Structures and Mechanisms. *Annu. Rev. Biophys.* **46**, 505–529 (2017).

145. Jinek, M. *et al.* A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816–821 (2012).
146. Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513**, 569–573 (2014).
147. Nishimasu, H. *et al.* Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* **156**, 935–949 (2014).
148. Karvelis, T. *et al.* Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. *Genome Biol* **16**, 253 (2015).
149. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67 (2014).
150. Sternberg, S. H., LaFrance, B., Kaplan, M. & Doudna, J. A. Conformational control of DNA target cleavage by CRISPR–Cas9. *Nature* **527**, 110–113 (2015).
151. Jinek, M. *et al.* Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science* **343**, 1247997–1247997 (2014).
152. Wu, X. *et al.* Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol* **32**, 670–676 (2014).
153. East, K. W. *et al.* Allosteric Motions of the CRISPR–Cas9 HNH Nuclease Probed by NMR and Molecular Dynamics. *J. Am. Chem. Soc.* **142**, 1348–1358 (2020).
154. Babu, K. *et al.* Coordinated Actions of Cas9 HNH and RuvC Nuclease Domains Are Regulated by the Bridge Helix and the Target DNA Sequence. *Biochemistry* **60**, 3783–3800 (2021).
155. Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* **31**, 839–843 (2013).
156. Jiang, F. *et al.* Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science* **351**, 867–871 (2016).
157. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819–823 (2013).

158. Duan, J. *et al.* Genome-wide identification of CRISPR/Cas9 off-targets in human genome. *Cell Research* **24**, 1009–1012 (2014).
159. Josephs, E. *et al.* Structure and specificity of the RNA-guided endonuclease Cas9 during DNA interrogation, target binding and cleavage. *Nucleic Acids Research* **43**, (2015).
160. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**, 822–826 (2013).
161. Chakrabarti, A. M. *et al.* Target-Specific Precision of CRISPR-Mediated Genome Editing. *Molecular Cell* **73**, 699–713.e6 (2019).
162. Lemos, B. R. *et al.* CRISPR/Cas9 cleavages in budding yeast reveal templated insertions and strand-specific insertion/deletion profiles. *PNAS* **115**, E2040–E2047 (2018).
163. Shou, J., Li, J., Liu, Y. & Wu, Q. Precise and Predictable CRISPR Chromosomal Rearrangements Reveal Principles of Cas9-Mediated Nucleotide Insertion. *Molecular Cell* **71**, 498–509.e4 (2018).
164. Zuo, Z. & Liu, J. Allosteric regulation of CRISPR-Cas9 for DNA-targeting and cleavage. *Current Opinion in Structural Biology* **62**, 166–174 (2020).
165. Taheri-Ghahfarokhi, A. *et al.* Decoding non-random mutational signatures at Cas9 targeted sites. *Nucleic Acids Res* **46**, 8417–8434 (2018).
166. van Overbeek, M. *et al.* DNA Repair Profiling Reveals Nonrandom Outcomes at Cas9-Mediated Breaks. *Molecular Cell* **63**, 633–646 (2016).
167. Allen, F. *et al.* Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat Biotechnol* **37**, 64–72 (2019).
168. Leenay, R. T. *et al.* Large dataset enables prediction of repair after CRISPR–Cas9 editing in primary T cells. *Nat Biotechnol* **37**, 1034–1037 (2019).
169. Chen, W. *et al.* Massively parallel profiling and predictive modeling of the outcomes of CRISPR/Cas9-mediated double-strand break repair. *Nucleic Acids Research* **47**, 7989–8003 (2019).

170. Sledzinski, P. Paving the way towards precise and safe CRISPR genome editing. *Biotechnology Advances* **20** (2021).
171. Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol* **32**, 279–284 (2014).
172. Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nature Biotechnology* **33**, 187–197 (2015).
173. Kocak, D. D. *et al.* Increasing the specificity of CRISPR systems with engineered RNA secondary structures. *Nat Biotechnol* **37**, 657–666 (2019).
174. Cromwell, C. R. *et al.* Incorporation of bridged nucleic acids into CRISPR RNAs improves Cas9 endonuclease specificity. *Nat Commun* **9**, 1448 (2018).
175. Ryan, D. E. *et al.* Improving CRISPR–Cas specificity with chemical modifications in single-guide RNAs. *Nucleic Acids Res* **46**, 792–803 (2018).
176. Rueda, F. O. *et al.* Mapping the sugar dependency for rational generation of a DNA-RNA hybrid-guided Cas9 endonuclease. *Nat Commun* **8**, 1–11 (2017).
177. Ran, F. A. *et al.* Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell* **154**, 1380–1389 (2013).
178. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152**, 1173–1183 (2013).
179. Slaymaker, I. M. *et al.* Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2016).
180. Kleinstiver, B. P. *et al.* Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485 (2015).
181. Lee, J. K. *et al.* Directed evolution of CRISPR-Cas9 to increase its specificity. *Nature Communications* **9**, 3048 (2018).

182. Sledzinski, P., Nowaczyk, M., Smielowska, M. I. & Olejniczak, M. CRISPR/Cas9-induced double-strand breaks in the huntingtin locus lead to CAG repeat contraction through DNA end resection and homology-mediated repair. *BMC Biology* **22**, 282 (2024).
183. Freudenreich, C. H., Kantrow, S. M. & Zakian, V. A. Expansion and length-dependent fragility of CTG repeats in yeast. *Science* **279**, 853–856 (1998).
184. Richard, G.-F., Dujon, B. & Haber, J. E. Double-strand break repair can lead to high frequencies of deletions within short CAG/CTG trinucleotide repeats. *Mol Gen Genet* **261**, 871–882 (1999).
185. Richard, G., Goellner, G. M., McMurray, C. T. & Haber, J. E. Recombination-induced CAG trinucleotide repeat expansions in yeast involve the MRE11–RAD50–XRS2 complex. *The EMBO Journal* **19**, 2381–2390 (2000).
186. Richard, G.-F. *et al.* Highly specific contractions of a single CAG/CTG trinucleotide repeat by TALEN in yeast. *PLoS ONE* **9**, e95611 (2014).
187. Mosbach, V., Poggi, L., Viterbo, D., Charpentier, M. & Richard, G.-F. TALEN-Induced Double-Strand Break Repair of CTG Trinucleotide Repeats. *Cell Reports* **22**, 2146–2159 (2018).
188. Mosbach, V. *et al.* Resection and repair of a Cas9 double-strand break at CTG trinucleotide repeats induces local and extensive chromosomal rearrangements. *bioRxiv* 782268 (2019) doi:10.1101/782268.
189. Nicolette, M. L. *et al.* Mre11–Rad50–Xrs2 and Sae2 promote 5' strand resection of DNA double-strand breaks. *Nat Struct Mol Biol* **17**, 1478–1485 (2010).
190. Poggi, L., Emmenegger, L., Descorps-Declère, S., Dumas, B. & Richard, G.-F. Differential efficacies of Cas nucleases on microsatellites involved in human disorders and associated off-target mutations. *Nucleic Acids Research* (2021) doi:10.1093/nar/gkab569.
191. Mittelman, D. *et al.* Zinc-finger directed double-strand breaks within CAG repeat tracts promote repeat instability in human cells. *PNAS* **106**, 9607–9612 (2009).
192. Cinesi, C., Aeschbach, L., Yang, B. & Dion, V. Contracting CAG/CTG repeats using the CRISPR-Cas9 nickase. *Nature Communications* **7**, 13272 (2016).

193. van Agtmaal, E. L. *et al.* CRISPR/Cas9-Induced (CTG·CAG)_n Repeat Instability in the Myotonic Dystrophy Type 1 Locus: Implications for Therapeutic Genome Editing. *Molecular Therapy* **25**, 24–43 (2017).
194. Dabrowska, M., Juzwa, W., Krzyzosiak, W. J. & Olejniczak, M. Precise Excision of the CAG Tract from the Huntingtin Gene by Cas9 Nickases. *Front. Neurosci.* **12**, 75 (2018).
195. Larin, M. *et al.* Cas9 nickase-mediated contractions of CAG/CTG repeats are transcription-dependent and replication-independent. *NAR Molecular Medicine* **1**, ugae013 (2024).
196. Murillo, A. *et al.* Cas9 nickase-mediated contraction of CAG/CTG repeats at multiple disease loci. 2024.02.19.580669 Preprint at <https://doi.org/10.1101/2024.02.19.580669> (2024).
197. Oura, S. *et al.* Precise CAG repeat contraction in a Huntington's Disease mouse model is enabled by gene editing with SpCas9-NG. *Commun Biol* **4**, 1–13 (2021).
198. Savouret, C. *et al.* CTG repeat instability and size variation timing in DNA repair-deficient mice. *The EMBO Journal* **22**, 2264–2273 (2003).
199. Sundararajan, R., Gellon, L., Zunder, R. M. & Freudenreich, C. H. Double-Strand Break Repair Pathways Protect against CAG/CTG Repeat Expansions, Contractions and Repeat-Mediated Chromosomal Fragility in *Saccharomyces cerevisiae*. *Genetics* **184**, 65–77 (2010).
200. Chatterjee, N., Lin, Y., Yotnda, P. & Wilson, J. H. Environmental Stress Induces Trinucleotide Repeat Mutagenesis in Human Cells by Alt-Nonhomologous End Joining Repair. *J Mol Biol* **428**, 2978–2980 (2016).
201. Kim, J. C., Harris, S. T., Dinter, T., Shah, K. A. & Mirkin, S. M. The role of break-induced replication in large-scale expansions of (CAG)_n/(CTG)_n repeats. *Nat Struct Mol Biol* **24**, 55–60 (2017).
202. Concannon, C. & Lahue, R. S. Nucleotide excision repair and the 26S proteasome function together to promote trinucleotide repeat expansions. *DNA Repair* **13**, 42–49 (2014).

203. Hubert, L., Jr, Lin, Y., Dion, V. & Wilson, J. H. Xpa deficiency reduces CAG trinucleotide repeat instability in neuronal tissues in a mouse model of SCA1. *Human Molecular Genetics* **20**, 4822–4830 (2011).
204. Lin, Y. & Wilson, J. H. Transcription-Induced CAG Repeat Contraction in Human Cells Is Mediated in Part by Transcription-Coupled Nucleotide Excision Repair. *Molecular and Cellular Biology* **27**, 6209–6217 (2007).
205. Szwarcrocka, S. T., Stączek, P. & Parniewski, P. Chromosomal model for analysis of a long CTG/CAG tract stability in wild-type Escherichia coli and its nucleotide excision repair mutants. *Can. J. Microbiol.* **53**, 860–868 (2007).
206. Jankowski, C., Nasar, F. & Nag, D. K. Meiotic instability of CAG repeat tracts occurs by double-strand break repair in yeast. *Proceedings of the National Academy of Sciences* **97**, 2134–2139 (2000).
207. Lin, Y., Dion, V. & Wilson, J. H. Transcription promotes contraction of CAG repeat tracts in human cells. *Nat Struct Mol Biol* **13**, 179–180 (2006).
208. Nakatani, R., Nakamori, M., Fujimura, H., Mochizuki, H. & Takahashi, M. P. Large expansion of CTG•CAG repeats is exacerbated by MutS β in human cells. *Sci Rep* **5**, 11020 (2015).
209. Crespan, E., Czabany, T., Maga, G. & Hübscher, U. Microhomology-mediated DNA strand annealing and elongation by human DNA polymerases λ and β on normal and repetitive DNA sequences. *Nucleic Acids Research* **40**, 5577–5590 (2012).
210. Crespan, E., Hübscher, U. & Maga, G. Expansion of CAG triplet repeats by human DNA polymerases λ and β in vitro, is regulated by flap endonuclease 1 and DNA ligase 1. *DNA Repair* **29**, 101–111 (2015).
211. Goula, A.-V. *et al.* The Nucleotide Sequence, DNA Damage Location, and Protein Stoichiometry Influence the Base Excision Repair Outcome at CAG/CTG Repeats. *Biochemistry* **51**, 3919–3932 (2012).

212. Dabrowska, M., Ciolak, A., Kozłowska, E., Fiszer, A. & Olejniczak, M. Generation of New Isogenic Models of Huntington's Disease Using CRISPR-Cas9 Technology. *International Journal of Molecular Sciences* **21**, 1854 (2020).
213. Dabrowska, M. *et al.* qEva-CRISPR: a method for quantitative evaluation of CRISPR/Cas-mediated genome editing in target and off-target sites. *Nucleic Acids Res* **46**, e101–e101 (2018).
214. Mosbach, V. *et al.* Resection and repair of a Cas9 double-strand break at CTG trinucleotide repeats induces local and extensive chromosomal deletions. *PLoS Genet* **16**, e1008924 (2020).
215. Zierhut, C. & Diffley, J. F. X. Break dosage, cell cycle stage and DNA replication influence DNA double strand break response. *The EMBO Journal* **27**, 1875–1885 (2008).
216. Bennett, C. F., Krainer, A. R. & Cleveland, D. W. Antisense Oligonucleotide Therapies for Neurodegenerative Diseases. *Annu Rev Neurosci* **42**, 385–406 (2019).
217. Henriques, C. Antisense oligonucleotide therapeutics in neurodegenerative diseases: the case of polyglutamine disorders. *Brain* (2019) doi:10.1093/BRAIN/AWZ328.
218. Schimmel, J. *et al.* Modulating mutational outcomes and improving precise gene editing at CRISPR-Cas9-induced breaks by chemical inhibition of end-joining pathways. *Cell Reports* **42**, 112019 (2023).
219. Wimberger, S. *et al.* Simultaneous inhibition of DNA-PK and Pol Θ improves integration efficiency and precision of genome editing. *Nat Commun* **14**, 4761 (2023).
220. Ramadoss, G. N. *et al.* Neuronal DNA repair reveals strategies to influence CRISPR editing outcomes. 2024.06.25.600517 Preprint at <https://doi.org/10.1101/2024.06.25.600517> (2025).
221. Sebesta, M. *et al.* Role of PCNA and TLS polymerases in D-loop extension during homologous recombination in humans. *DNA Repair (Amst)* **12**, 691–698 (2013).
222. Musio, A. The multiple facets of the SMC1A gene. *Gene* **743**, 144612 (2020).