Transcription is a potential threat to genome integrity, and transcription-associated DNA damage must be repaired for proper messenger RNA (mRNA) synthesis and for cells to transmit their genome intact into progeny. For a wide range of structurally diverse DNA lesions, cells employ the highly conserved nucleotide excision repair (NER) pathway to restore their genome back to its native form. Recent evidence suggests that NER factors function, in addition to the canonical DNA repair mechanism, in processes that facilitate mRNA synthesis or shape the 3D chromatin architecture. Here, these findings are critically discussed and a working model that explains the puzzling clinical heterogeneity of NER syndromes highlighting the relevance of physiological, transcription-associated DNA damage to mammalian development and disease is proposed.

1. Introduction: The Yin-Yang of DNA Damage

In Chinese philosophy, the shady (yin) and sunny (yang) sides of the Taijitu symbol represent two seemingly opposing halves that are complementary and interrelated in the natural world. Likewise, DNA damage events may occur accidentally, injuring indiscriminately the DNA helix, or in a scheduled and spatially restricted manner that may often be beneficial to the survival of the cell. If left unrepaired, random DNA lesions gradually build up in the mammalian genome inadvertently interfering with vital DNA-templated transactions, such as transcription, DNA replication, or homologous recombination (HR). Irreparable DNA insults may occasionally be fixed into mutations during DNA replication, driving cancer,[1] or obstruct ongoing transcription,[2] leading to cellular malfunction and the premature onset of age-related diseases.[3,4] Unlike stochastic DNA insults, physiological DNA lesions are scheduled (in terms of time) or targeted (in terms of genomic location) DNA damage events that are thought to facilitate DNA-dependent reactions during mammalian development or upon exposure to various stimuli. For instance, DNA double-strand breaks (DSBs) are purposely formed during meiotic HR to facilitate genetic diversity in gamete cells.[5] In developing B- and T-lymphocytes, DNA breaks are also required during V(D)J recombination to secure antigen-binding diversity of the immunoglobulin and T-cell receptor proteins.[6] Moreover, the error-prone repair of DNA lesions may enhance antibody diversity by triggering somatic hypermutation of immunoglobulin variable genes.[7] Targeted DNA damage events are an inevitable consequence of extreme shifts in transcription demands[8] and, as we will discuss later, they are required for fine-tuning gene expression, facilitating access of transcription factors and the RNAPII at promoter and enhancer regions or for enabling rapid intra- and inter-chromosomal interactions during DNA-templated transactions. Thus, whereas spontaneous DNA insults pose a threat to cell viability, scheduled DNA damage events are unavoidable during mammalian development and an essential prerequisite for cellular and organismal survival.

2. The Enigmatic Heterogeneity of NER Syndromes

For helix-distorting damage, mammalian cells employ the nucleotide excision repair (NER) pathway, a highly conserved mechanism that recognizes and removes helical distortions throughout the genome (global genome NER [GG-NER]) or selectively from the transcribed strand of active genes (transcription-coupled NER [TC-NER]).[9] Several recent reviews have already described in great detail the mechanism of NER, and it is also summarized briefly herein (Figure 1A).[9,25,26] Below, we will attempt to portray the enormous phenotypic heterogeneity seen in NER-deficient patients and highlight...
Figure 1. The NER mechanism and the puzzling heterogeneity of NER mutant mice. A) In NER, lesion recognition occurs genome-wide (as in GG-NER) or in a transcription-coupled mode (as in TC-NER). In GG-NER, the UV−DDB complex composed of DDB1 and DDB2 acts as an auxiliary damage recognition factor that binds directly to UV-induced helix-distorting DNA lesions, such as bulky adducts and pyrimidine dimers (shown in red)\(^\text{[10–12]}\) (1). In TC-NER, the principal DNA damage sensor is the elongating RNAPII that is blocked at damaged DNA sites on the actively transcribed strand of genes and recruits the DNA-dependent ATPase CSB protein (2). Upon damage, the higher affinity of CSB for RNAPII and its binding to CSA likely helps to backtrack RNAPII allowing the remaining NER factors to access the DNA lesion.\(^\text{[13–17]}\) Following damage recognition by the GG- or TC-NER sub-pathways, TFIIH is recruited to the DNA damage site\(^\text{[18]}\) (3). XPA binds the DNA on the 5′ side of the DNA insult. Together with RPA that binds the ssDNA opposite the DNA lesion, XPA stabilizes the damaged DNA for incision. RPA activates ERCC1–XPF and XPG that cleave the 5′ and 3′ side of the 24–32-nucleotide fragment, respectively containing the damaged DNA fragment\(^\text{[19, 20]}\) (4). Soon after the removal of the damaged DNA fragment, the polymerase activity of the DNA replication machinery fills in the single-stranded DNA gap and the new DNA fragment is then sealed by DNA ligase I or IIIα-XRCC1 (not shown). B) Mouse models with inborn NER defects associate with one or more syndromes (shown in red)\(^\text{[2] of 12}\) www.advancedsciencenews.com www.bioessays-journal.com © 2019 The Authors. BioEssays Published by WILEY Periodicals, Inc.
some of the genetic–phenotypic paradoxes seen in accompanied animal models.

Genetic pleiotropy refers to a single gene mutation that influences distinct and seemingly disparate phenotypic disorders. Instead, genetic heterogeneity occurs when a single phenotypic trait is linked to mutations on multiple alleles from the same or different loci. In humans, inborn mutations in NER genes are associated with syndromes that are both heterogeneous and pleiotropic. Defects in GG-NER give rise to the skin cancer-prone disorder xeroderma pigmentosum ([XP] affected proteins: XPA through XPG). Defects in TC-NER lead to a heterogeneous group of progeroid (accelerated aging) disorders that are phenotypically distinct from XP, including the Cockayne syndrome ([CS] affected proteins: CSA, CSB, UVSSA, XBP, XPD, XPF, and TTDA and certain mutations in the gene-encoding XPG) and trichothiodystrophy ([TTD] affected proteins: XBP, XPD, and P8). Although XP, CS, and TTD are derived from mutations in the genes involved in the same DNA repair mechanism or even in the same NER gene,[31] their clinical outcome is remarkably diverse. For example, two siblings, one an XP and the other a CS patient, carrying the same mutation in the XPC gene, or mice having a causative point mutation in the XPC gene,[38] XPV patients are progeria-like mice that also carry defects in other DNA repair proteins: CSA, CSB, XPA, XPD, XPE, and TTDA and certain mutations in the gene-encoding XPG) and trichothiodystrophy ([TTD] affected proteins: XBP, XPD, and P8).[37–40] Although XP, CS, and TTD are derived from mutations in the genes involved in the same DNA repair mechanism or even in the same NER gene,[31] their clinical outcome is remarkably diverse.

For example, two siblings, one an XP and the other a CS patient, carrying the same mutation in the XPC gene were previously shown to be affected differently in terms of the severity of the disease, indicating that a straightforward association between the genetic defect and the observed pathological features is not guaranteed.[32] In XP, patients carrying mutations in the XPC or XPE genes show extreme sensitivity to sunlight and present with freckling in sun-exposed skin areas. If left unprotected, XP patients have a high frequency of nonmelanoma skin cancer and melanomas that manifest at an early age (<9 years and ≈22 years, respectively) in UV-exposed areas, i.e., skin, eye, and tongue and occasionally also in non-UV-exposed organs of the central nervous system.[33] However, only 60% of XP patients are photosensitive, whereas the remaining 40% show no sunburn reactions but still present with increased freckle-like pigmentation.[134] The onset time of neurological abnormalities also varies dramatically from the age of 2 years to middle age.[33,36] A smaller percentage, i.e., <30% of the patients with defects in XPA, XBP, XPD, or XPG manifest neurological deficits of varying severity, including isolated hyperreflexia, progressive mental retardation, sensorineural deafness, spasticity, or seizures.[4] In XPC patients, the inactivating mutations are spread throughout the gene.[37] A rather distinct group of XP patients involves mutations in the gene-encoding XVP (also called low-fidelity Pol η).[138] XVP patients are proficient in NER but their symptoms are similar to those seen in XP, including UV-induced mutagenesis and skin cancer.[39,40] CS patients suffer from progressive neurodevelopmental defects along with an early (at birth) or late (late childhood or even adulthood) onset of growth defects and cachexia, microcephaly, retinal degeneration, and a reported life expectancy of ≈12 years. The CS pathological symptoms are more severe in the CS variant cerebro-oculo-facio-skeletal syndrome ([C0FS] life expectancy <2 years) or considerably milder as in the UV-sensitive syndrome ([UVsS]) affected proteins: CSB, CSA, and UVSSA where the stability of CSB is compromised.[41,42] Intriguingly, although CS patients manifest with cutaneous photosensitivity, they have no predisposition to skin cancer. The lower incidence of tumors in CS results from the lack of elevated UV-induced mutagenesis rather than from enhanced lethality.[43] Similar to CS, the clinical phenotypes of TTD patients are remarkably diverse in terms of severity and extent. TTD patients are photosensitive and present with the characteristic sulfur-deficient, dry and sparse hair and nails of TTD, congenital ichthyosis, physical and mental retardation and signs of premature aging, but no cancer.[44] In certain instances, the pathological features in TTD are uncoupled from the DNA repair defect. For example, TTD patients carrying mutations in the TTDA gene manifest with clinical symptoms associated with the photosensitive TTD disorder but are DNA repair-proficient and show no signs of photosensitivity or cancer proneness.[45,46] Recently, it was discovered that a TTD causative TFIIIEβ mutation did not affect the GG-NER or TC-NER sub-pathways of NER, uncoupling any defects in transcription from DNA repair in TTD.[47]

### 3. NER Defects: From Man to Mice

Mice with engineered mutations in NER genes mimic most, but not all, of the pleiotropic and heterogeneous pathological symptoms seen in NER syndromes (Table 1).[62] The CG-NER-deficient Xpc−/− mice or the complete NER-defective Xpa−/− or Xpc−/− mice[49] double mutant animals present a higher frequency of UV-induced skin cancers but show none of the progeroid CS features or the progressive neurodegenerative complications associated with XP (Figure 1B; as indicated).[48,51,63] The TC-NER-defective Csbm/m animals associate with a mild phenotype[40] (similar to the UVsS) and need to be crossed with either Xpa−/− or Xpc−/− mice (i.e., Csbm/m/Xpa−/− or Csbm/m/Xpc−/− mice) to reliably recapitulate the severe CS variant COFS (Figure 1B; as indicated).[48] Both XpdTTD and Xpd−/− mice carrying a causative point mutation in Xpd gene fail to show the complete clinical spectrum of CS features associated with the phenotypic resemblance and severity of the disease. Mice carrying defects in GG-NER or TC-NER show heterogeneous phenotypes that manifest with severe (red-colored circles), moderate (pink-colored circles), or no pathological (white-colored circles) symptoms during mammalian development or at later stages in murine life. The completely NER-defective Xpc−/− animals show none of the clinical features seen in CS or XP and differ dramatically from those seen in Xpg−/−, Xpg−/−, or Ercc1−/− mice that also carry defects in other DNA repair systems. The phenotype of Ercc1−/− mice having a 7-aminoacid carboxy-terminal deletion is milder than that of Ercc1−/− or Xpg−/− animals[21] but more severe than that of XpgΔex15 mice having a deletion in exon 15 of the Xpg gene. Instead, the double mutant XpgΔex15/Xpa−/− animals manifest with severe growth retardation,[22] Csbm/m mice present with only mild clinicopathological symptoms but their phenotype aggravates substantially when crossed with the NER-defective Xpa−/− animals, i.e., Csbm/m/Xpa−/− mice. Similar phenotypes are also seen in Csbm/m/Xpc−/− animals but not when Xpc−/− mice are crossed with Xpa−/− animals. Xab2 or Rpo1 gene disruption (Xab2−/− and Rpo1−/− animals, respectively) causes embryonic lethality in mice (not shown).[23,24] COFS, cerebro-oculo-facio-skeletal syndrome; ICL, interstrand crosslink repair; FA, Fanconi anemia; XFE, XPF–ERCC1 progeria.
with TTD and XP/CS, respectively. With the exception of two distinct mutations in Xpd, the structural scaffold that maintains TFIIH integrity during transcription, XP−/−, Xpd−/−, or Ercc1−/− mice are perhaps the only single-NER mutant animals (also defective in other DNA repair systems) to reiterate the severe developmental abnormalities and premature aging features of patients with mutations in the corresponding genes. Most other single-mutant animals, e.g., Xpc−/−, Csb−/−, or Xpa−/− mice result in phenotypically healthy mice with minor to moderate pathological symptoms.

Current models are based on the differential roles of NER factors in TC-NER, GG-NER, ICL, and DSB repair, or in the transcription-coupled-associated base excision repair (BER) to explain the complex and heterogeneous constellation of clinical symptoms in NER. Moreover, the diversity of DNA lesions encountered, the position of the DNA damage itself in the genome along with the proliferative status, and type of cells are expected to further contribute to the “genetic–phenotypic” paradox seen in NER. From a genetic point of view, however, it may be difficult to reconcile why the completely NER-defective Xpa−/− animals show none of the clinical features (with the exception of UV-induced skin cancer) seen in CS or XP (Figure 1B; as indicated) or why Xpa−/− animals differ substantially from the Csb−/−/Xpa−/− mice that are also defective in both sub-pathways of NER (Figure 1B). It is also puzzling that the causative TC-NER defect in Csb−/− mice aggravates immensely the phenotype of the NER-deficient Xpa−/− animals as in Csb−/−/Xpa−/− mice. The accumulation of spontaneous DNA lesions across the mammalian genome may well explain the high incidence of UV-induced skin tumors in Xpa−/− animals or the progressive onset of neurological symptoms seen in XPA patients over time. However, the undeviating onset and severity of developmental abnormalities seen in all isogenic animals carrying the same inborn mutation in NER contrasts with the random spatiotemporal occurrence of DNA damage events within the genome. The reported frequency of DNA insults (=70 000 DNA lesions per day per cell in humans) and presumed random position in the mammalian genome make it unlikely that all genes poised for transcription are damaged during mammalian development or that the same genes are, at all times, selectively damaged. The contrast of the disease onset and the random nature of DNA damage could be better rationalized when one considers that certain NER factors are involved in the repair of transcription-associated DNA lesions that are scheduled in time as in, e.g., developmental gene expression programs and topologically restricted as in, e.g., promoters.

## 4. Linking NER to Transcription Demands

### 4.1. NER Factors on Gene Promoters

During the last three decades, there has been ample evidence that NER and transcription are closely intertwined processes. The first hint came when the basal transcription factor TFIIH was found to facilitate initiation, promoter escape, and early elongation steps of RNAPII transcription upon DNA damage. The Xpb and Xpd genes encode two helicase subunits of TFIIH that unwind DNA in the vicinity of DNA-damaged sites for repair or on promoters to facilitate the synthesis of the primary transcript. Xpb may also inhibit transcription initiation by interfering with promoter melting, thus revealing an additional level of transcriptional regulation. It was several years after the functional implication of TFIIH in NER that the remaining NER factors, i.e., XPC, CSB, XPA with RPA, XPG, and the ERCC1–XPF complex were shown to be sequentially recruited to the promoters of nuclear receptor genes upon transcription stimulation in vitro.
4.2. XPC and Stem Cell Pluripotency

Using a series of chromatographic and purification steps, the XPC/RAD23B/CETN2 was also shown to be part of a multi-subunit stem cell coactivator (SCC) complex required for coactivating Oct4/Sox2-dependent transcription of Nanog for stem cell pluripotency (Figure 2B). The effect of XPC or RAD23B knockdown was more profound in the reprogramming of mouse embryonic fibroblasts than in the maintenance of embryonic stem cells. Similar to HeLa cells, the complex in transcription was distinct from that involved in DNA repair. The XPC–SCC could co-activate Nanog transcription even when XPC was mutated abolishing GG-NER and occurred independently of DNA binding, indicating the possible secondary interactions with sequence-specific transcription factors. In further support, Zhang et al. proposed the regions of contact between the XPC complex and OCT4, SOX2, XPA, and TFIIH. XPC may also impinge on gene expression by facilitating TDG-mediated DNA demethylation during somatic cell reprogramming (Figure 2C). Arguably, the functional role of this complex in promoters reflects their transient and, in this context unproductive, interaction with TFIIH that is involved in both NER and transcription. In this respect, recruitment of NER proteins would correspond to the amount of TFIIH bound on promoters and, by analogy, to the level of transcription activity (in personal communication with Prof. Jan Hoeijmakers). Although this possibility cannot be excluded, it cannot fully explain why P15 XpdTTD livers are—unlike the Csbm/m/Xpa−/− or Ercc1−/− mice—not growth-defective and manifest distinct gene expression changes from those seen in P15 Csbm/m/Xpa−/− or Ercc1−/− mice.

Figure 2. The functional role of NER factors outside the canonical DNA repair mechanism. NER factors may function in transcription outside the canonical NER pathway. A) NER factors recruit on promoters to facilitate transcription initiation in vitro or in vivo. B) The XPC/RAD23B/CETN2 is known to be part of a SCC complex required for coactivating Oct4/Sox2-dependent transcription of Nanog for stem cell pluripotency. C) XPC also facilitates TDG-mediated DNA demethylation during somatic cell reprogramming independently of its function in DNA repair. D) XPG and ERCC1–XPF may also promote active DNA demethylation of nuclear receptor (NR) genes (as in C) or facilitate gene looping for optimal gene transcription. E) ERCC1–XPF also interacts with CTCF to facilitate the postnatal silencing of imprinted genes. Black lollipops, methylated cytosine; white lollipops, nonmethylated cytosine.
role of XPC in stem cell pluripotency contrasts with the lack of any apparent developmental defects seen in Xpc−/− mice.\(^5\) It is also puzzling that ablation of the C-terminal region of Xpc gene abrogating the interaction sites of XPC with RAD23 and CETN2 has minimal impact on gene expression.\(^89\) More importantly, the truncated Xpc embryonic stem cells retain proper gene expression patterns and pluripotency to contribute chimeric embryos.\(^89\) Thus, the true biological relevance of XPC in developmental gene expression programs remains elusive. XPG–TFIIH has also been shown to interact with factors involved in transcription elongation, suggesting that defects in this association could account for certain CS features in XPG/CS patients.\(^90\)

### 4.3. NER Factors and Indirect Roles in Transcription

The XPG–TFIIH complex is a promoter and coding regions of both epidermal growth factor (EGF)-induced (FOS) and housekeeping (EEF1A1) genes, whereas XPG knockdown dampens FOS transcription whose messenger RNA (mRNA) levels are also decreased in XPG/CS cells carrying a truncated XPG. XPG may further influence gene expression by being targeted to distinct demethylation sites replacing methylated cytosine with an unmethylated nucleotide to relieve epigenetic silencing.\(^91\) Lastly, inborn defects in NER may selectively hamper the expression of genes with long primary transcript lengths. Long primary transcripts of actively transcribed genes are thought to increase their likelihood to get hit by DNA lesions and by analogy the risk of RNAPII blockage at damaged DNA sites. In support, the great majority of genes with significantly decreased mRNA levels in Ercc1Δ/− mouse livers have substantially long primary transcript lengths.\(^92\) Since it is mostly introns rather than exons that contribute to the total gene size, it is thought-provoking that, in humans, introns are 14 times shorter in highly expressed genes than in genes that are poorly expressed.\(^93\) In this scenario, a selective evolutionary pressure would favor genes with shorter introns to minimize the cost of transcription but also the risk of transcription-blocking DNA lesions interfering with RNAPII during ongoing mRNA synthesis.

### 5. DNA Damage: Linking NER to mRNA Synthesis

During zygotic genome activation, the process of mRNA synthesis may be abrupt, leading to genome instability.\(^94\) Indeed, DNA damage events are known to be randomly or purposely generated during transcription,\(^90\) supporting the notion that mRNA synthesis is, in fact, a risky process (Figure 3).

#### 5.1. NER and Transcription-Associated Mutagenesis

Almost 50 years ago, transcription was found to be mutagenic.\(^96\) Active transcription increases the rate of C to T transitions in the non-transcribed strand of genes in Escherichia coli and in humans\(^97,98\) and correlates positively with the occurrence of spontaneous mutations in yeast.\(^99\) In human cells, RNAPII may bypass a cyclobA or CPD lesion, which are relevant substrates for NER, by inserting a uridine opposite the lesion or by generating deletions of a few nucleotides (7–21 bp long) opposite the inserted uridine leading to transcription mutagenesis.\(^100\) In mouse embryonic stem cells, the stalling of the transcription complex on a UV-induced photolesion increases the spontaneous deamination of cytosine to uracil, leading to single or tandem CC to TT transitions in the transcribed strand (Figure 3A) or, if a stalled transcription complex collides with the DNA replication machinery, to intragenic deletions via error-prone end joining.\(^101\) TC-NER provides protection against both classes of transcription-associated mutagenesis. In addition to BER, a role for CSA, CSB, and XPG in the transcription-coupled repair of oxidative DNA lesions has been suggested,\(^102,103\) and certain CS features are thought to originate from a defect in the repair of oxidative DNA damage interfering with transcription.\(^11,25\) Whether 8-oxoG, the most abundant oxidative DNA lesion, is a substrate for TC-NER is, however, a subject of active research and awaits further investigation. However, as CSB cells cannot forward-translocate RNAPII, this is thought to result in persistent stalling of RNAPII at oxidative DNA damage sites, thereby preventing access for BER proteins to remove the lesion.\(^104\)

#### 5.2. Distinct NER Factors and Transcription-Associated DSBs

Unlike spontaneously generated UV-induced or oxidative DNA lesions, DNA breaks appear to be less random and play a physiologic role during the process of mRNA synthesis (Figure 3B). As RNAPII advances along the DNA template, it induces a positive supercoiling on the 3′-end and negative supercoiling on the 5′-end. Negative and/or positive DNA supercoiling also builds up ahead of and/or behind RNAPII when the enzyme encounters “topological locks,” such as those generated by DNA–histone interactions.\(^105\) DNA topoisomerase (TOP)2 resolves some of these topological constraints by binding preferentially to nucleosome-free DNA to generate transient, site-specific DSBs on promoters.\(^106–109\) The latter creates a more permissive chromatin setting that is thought to facilitate access of transcription factors and the RNAPII at promoter and enhancer regions. In agreement, neuronal activity requires the presence of DSBs on promoters to trigger the expression of Fos and Npas4 genes in the absence of any exogenously derived genotoxic insult or stimulus.\(^110\) Likewise, the autoimmune regulator protein (AIRE) interacts with TOP2a to promote DSBs on promoters, activating DNA–PK and other partners that attract chromatin-remodeling complexes to energize the transcription of low-expressed genes.\(^111,112\) Interestingly, Helicobacter pylori induces transcription-dependent XPF/XPG-mediated DSBs that promote gene expression and cell survival.\(^113\) Furthermore, the liganded androgen receptor triggers site-specific breakpoints at intronic translocation regions to facilitate rapid intra- and inter-chromosomal interactions.\(^114\) DNA breaks are also required for chromatin de-compaction and to initiate global gene expression during zygotic genome activation in worms.\(^115\) At present, the relevance of transcription-associated DNA breaks to NER remains obscure. However, it is intriguing that DNA breaks
accumulate in UV-irradiated cells of XPD/CS patients and, surprisingly so, at sites distant from the DNA damage itself.\[^{116}\]

Inhibition of RNAPII transcription with transcription inhibitors dramatically reduces the number of UV-induced breaks. Moreover, the breaks required active XPA, pointing to the possible involvement of NER in the induction of these lesions (likely introduced by XPG and XPF–ERCC1) at sites of ongoing transcription.\[^{[117]}\]

ERCC1–XPF itself could also participate in the repair of transcription-associated DSBs through its direct involvement in certain HR and nonhomologous end-joining repair mechanisms outside NER (summarized in Manandhar et al.\[^{118}\]). Using next-generation sequencing approaches and currently available methods that precisely measure the location and frequency of DSBs along the mammalian genome, it would be useful to test whether, upon transcription induction, ERCC1–XPF or XPG (and other NER factors) recruit on promoters genome-wide and whether recruitment of these factors coincide with the induction of DSBs on defined genomic regions.\[^{[77, 84]}\]

5.3. Distinct NER Factors and “R-loops”

Finally, longer tracts of RNA–DNA hybrids, known as “R-loops,” liaise between spontaneous DNA damage events and transcription (Figure 3C).\[^{[120]}\] Naturally occurring R-loops generated on promoters or termination regions are frequently formed when a nascent RNA molecule is hybridized with the
DNA template before the two strands of the DNA duplex reanneal, leaving the non-template DNA single-stranded. R-loops expose long stretches of ssDNA leading to the spontaneous formation of DSBs or to transcription-associated mutagenesis. Transient DNA–RNA hybrids are found in organisms ranging from bacteria to humans and play physiological roles in, e.g., transcription or class-switching recombination but are likely also generated when an RNAPII is stalled at transcription-blocking DNA lesions. ERCC1–XPF and XPG actively process R-loops. Intriguingly, R-loop processing requires CSB (involved in TC-NER) but not XPC (involved in GG-NER). XPF (the nuclease active site of ERCC1–XPF) and XPG are thought to generate nicks or gaps on both the transcribed and non-transcribed strand and at both ends of the R-loop. More recently, an alternative mechanism was proposed where Rad52 recruits on DSBs during active transcription in a DNA/RNA hybrid-dependent manner promoting XPG- (but not XPF-) mediated R-loop processing and transcription-associated HR. Thus, it is uncertain whether the cleavage of R-loops by XPG is directly related to NER. Likewise, cleavage of R-loops by ERCC1–XPF may require functional NER (where targeting would be mediated by, e.g., XPA or TFIIH), or else the ERCC1–XPF is targeted by MSH2–MSH3 or SLX4 outside the canonical NER mechanism. In this respect, one could test whether, R-loops are successfully cleaved in, e.g., XPA-defective cells.

6. NER Factors and Chromatin Looping: The CTCF Link

A great majority of genes involved in developmental and tissue-specific gene expression programs are found in gene-dense or repressive chromatin regions. Their cis-acting regulatory sequences, however, are often located several kilobases apart from one or several unrelated genes in between. The timely expression of a gene relies on the formation of an “active chromatin hub” to allow the transcription unit to physically communicate with its cis-regulatory elements. Such chromatin hubs depend on protein complexes to loop out intervening sequences or inactive genes and bring promoters and cis- or trans-acting sequences into physical proximity. The CCCTC-binding factor CTCF is a highly conserved 11-zinc finger protein that is involved in transcription activation and repression, chromatin architecture, genome imprinting, X chromosome inactivation, and insulation. Distinct NER factors associated with chromatin organizers and remodelers to shape the chromatin architecture for optimal gene expression, developmental gene silencing, or for promoting the accessibility of DNA repair proteins to sites of DNA damage. XPG and XPF are essential for establishing CTCF-dependent chromatin looping between the promoter and terminator of the activated RARb2 gene (Figure 2D). Abrogation of XPG or XPF due to gene silencing or a mutation in their catalytic sites substantially hampers the recruitment of CTCF in chromatin, the formation of appropriate chromatin loops, and the fine tuning of RARb2 mRNA synthesis. In further support, a recent in vivo biotinylation tagging approach revealed that the heterodimer ERCC1–XPF complex interacts with CTCF, the cohesin subunits SMC1A and SMC3, and with MBD2 to facilitate the postnatal silencing of imprinted genes during murine development (Figure 2E). Abrogation of ERCC1 in Ercc1−/− mice or exposure of primary embryonic fibroblasts to the DNA interstrand cross-linker mitomycin triggers the localization of CTCF to heterochromatin and the dissociation of the CTCF–cohesin complex and ATRX from promoters and imprinted control regions. These findings are also in agreement with the newly identified role for CTCF in facilitating the repair of DNA DSBs through HR. CTCF also interacts with CSB. CTCF and CSB regulate each other’s chromatin association in response to oxidative stress. Whereas oxidative stress enhances the CSB–CTCF interaction in 293 T cells, CSB itself facilitates CTCF–DNA interactions in vitro and regulates CTCF–chromatin interactions in cells treated with menadione, an inducer of oxidative stress. Lastly, the crosstalk between NER and chromatin organizers could also explain how, in spite of the low mobility of TFIIH in cortex neurons, the protein is still capable of responding rapidly to new transcription demands or local DNA damage during differentiation or lineage-specific gene expression programs.

7. Concluding Remarks and Future Aspects

Delineating the functional role of protein complexes involved in NER, transcription, and the chromatin architecture is fascinating but also challenging owing to the fact that these processes are intrinsically intertwined. A causal link that could keep these otherwise functionally distinct mechanisms together is DNA damage itself. Indeed, whereas stochastic DNA lesions represent a threatening menace during the lifetime of a cell, physiological DNA-damaging events, such as those occurring on promoters during abrupt switches in transcriptional load, are often unavoidable and indispensable for the proper execution of developmental gene expression programs. The latter would require that NER factors are in close cooperation with protein complexes involved in transcription or the chromatin architecture. Failure to do so (as in the case of the NER-deficient patients and mice), could hamper unequivocally the expression of genes during developmental transitions. The severity and time to onset of developmental defects would reflect the individual roles of certain NER proteins in transcription or in the repair of tightly scheduled, transcription-associated, DNA lesions as well as the time- and cell-type-specific requirements of the transcriptional programs involved. This scenario offers a likely explanation for the recurring onset of heterogeneous, but importantly nonrandom, developmental abnormalities associated with NER syndromes.

In spite of the recent progress, the precise biochemical functions of NER proteins in transcription-associated DNA damage and repair, in fine-tuning gene expression or in shaping the chromatin architecture remain obscure. With the advent of more sophisticated, functional animal models, e.g., biotin-tagged or cell-specific knockout mice coupled to high-throughput mass spectrometry, and next-generation sequencing approaches, however, we may soon be able to gain more insights into the role of NER-associated protein complexes and gene targets in distinct cell types at any stage.
during mammalian development. Moreover, the recent advances in genome-wide chromosome conformation capture approaches for characterizing the 3D architectures of genomes\textsuperscript{[135]} will allow us to gain insights into the functional relevance of distinct NER factors in genome-wide looping formation. As it is now possible to map genome-wide and at nucleotide resolution, the formation and repair of DNA lesions\textsuperscript{[136–140]} it may also soon be possible to test whether the recruitment of NER-associated protein complexes on promoters marks the simultaneous presence of, e.g., DNA breaks on these genomic regions. Further work will also be necessary to explore how and to what extent, e.g., DNA breaks or R-loops are functionally linked to transcription demands during mammalian development or with disease onset. It may only be then that we might fully appreciate the impact and relevance of unavoidable, transcription-associated DNA damage events to the developmental abnormalities associated with NER- and other DNA repair-deficient syndromes, also shedding some light on the multiple pathological complications associated with old age.

Acknowledgements

The “Chromatin3D” (GA622934), the Horizon 2020 ERC Consolidator grant “DefiNER” (GA64663), the Hellenic Foundation for Research and Innovation (HfRI), and the General Secretariat for Research and Technology (GSRT) under grant agreement No. 1059 and the Fondation Sante Grant supported this work. G.C. was supported by the IKY postdoctoral research fellowship program (MIS: 5001552), co-financed by the European Social Fund-ESF and the Greek government.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

developmental defects, DNA damage, DNA repair, human diseases, NER, nucleotide excision repair, mouse models, transcription

Received: October 1, 2018
Revised: December 10, 2018
Published online: March 28, 2019

[1] V. Pages, R. P. Fuchs, Oncogene 2002, 21, 8957.
[2] G. A. Garinis, B. Schumacher, Cell Cycle 2009, 8, 2134.
[3] G. A. Garinis, L. M. Uittenboogaard, T. Stacheck, M. Froustier, W. van Ickxen, T. M. Breit, H. van Steeg, L. H. Mullenders, G. T. van der Horst, J. C. Brüning, Nat. Cell Biol. 2009, 11, 604.
[4] G. A. Garinis, G. T. van der Horst, J. Vlij, J. H. Hoeijmakers, Nat. Cell Biol. 2008, 10, 1241.
[5] D. Zickler, N. Kleckner, Cold Spring Harb. Perspect. Biol. 2015, 7.
[6] C. Boboila, F. W. Alt, B. Schwer, Adv. Immunol. 2012, 116, 1.
[7] J. M. Di Noia, M. S. Neuberger, Annu. Rev. Biochem. 2007, 76, 1.
[8] H. Gaillard, A. Aguillera, Annu. Rev. Biochem. 2016, 85, 291.
[9] J. A. Marteijn, H. Lans, W. Vermeulen, J. H. Hoeijmakers, Nat. Rev. Mol. Cell Biol. 2014, 15, 465.
[10] K. Sugasawa, J. M. Ng, C. Masutani, S. Iwai, P. J. van der Spek, A. P. Eker, F. Hanaoka, D. Bootsma, J. H. Hoeijmakers, Mol. Cell 1998, 2, 223.
[11] M. Araki, C. Masutani, M. Takemura, A. Uchida, K. Sugasawa, J. Kondoh, Y. Okhuma, F. Hanaoka, J. Biol. Chem. 2001, 276, 18665.
[12] T. Riedl, F. Hanaoka, J. M. Egly, EMBO J. 2003, 22, 5293.
[13] P. C. Hanawalt, G. Spivac, Nat. Rev. Mol. Cell Biol. 2009, 10, 958.
[14] E. Citerio, S. Rademakers, G. T. van der Horst, A. J. van Gool, J. H. Hoeijmakers, W. Vermeulen, J. Biol. Chem. 1998, 273, 11844.
[15] J. A. Eisen, K. S. Sweder, P. C. Hanawalt, Nucleic Acids Res. 1995, 23, 2715.
[16] C. Troelstra, A. van Gool, J. de Wit, W. Vermeulen, D. Bootsma, J. H. Hoeijmakers, Cell 1992, 71, 939.
[17] N. Beerens, J. H. Hoeijmakers, R. Kanaar, W. Vermeulen, C. Wyman, J. Biol. Chem. 2005, 280, 4272.
[18] C. L. Li, F. M. Golebiowski, Y. Onishi, N. L. Samara, K. Sugasawa, W. Yang, Mol. Cell 2015, 59, 1025.
[19] R. M. Overmeer, J. Moser, M. Volker, H. Kool, A. E. Tomkinson, A. A. van Zeeland, L. H. Mullenders, M. Froustier, J. Cell Biol. 2011, 192, 401.
[20] A. F. Fagbemi, B. Orielli, O. D. Scharer, DNA Repair 2011, 10, 722.
[21] M. E. Dolle, R. V. Kuiper, M. Roodebergen, J. Robinson, S. de Vlugt, S. W. Wijnhoven, R. B. Beems, L. de la Fonteine, B. de Wither, I. van der Pluijm, L. J. Niedernhofer, P. Hasty, J. Vijg, J. H. Hoeijmakers, H. van Steeg, Pathobiol. Aging Age Relat. Dis. 2011, 1.
[22] N. Shioi, M. Mori, S. Kito, Y. N. Harada, K. Tanaka, T. Shiomi, DNA Repair 2005, 4, 351.
[23] Y. Wang, C. D. Putnam, M. F. Kane, W. Zhang, L. Edelmann, R. Russell, D. V. Carrion, L. Chin, R. Kucherlapati, R. D. Kolodner, W. Edelmann, Nat. Genet. 2005, 37, 750.
[24] R. Yonematsu, M. Minami, Y. Nakatsu, M. Takeuchi, I. Kuraoka, Y. Matsuda, Y. Higashi, H. Kondoh, K. Tanaka, DNA Repair 2005, 4, 479.
[25] J. H. Hoeijmakers, N. Engl. J. Med. 2009, 361, 1475.
[131] F. Lang, X. Li, W. Zheng, Z. Li, D. Lu, G. Chen, D. Gong, L. Yang, J. Fu, P. Shi, J. Zhou, Proc. Natl. Acad. Sci. U. S. A. 2017, 114, 10912.

[132] R. J. Lake, E. L. Boetefuer, K. J. Won, H. Y. Fan, Nucleic Acids Res. 2016, 44, 2125.

[133] G. Giglia-Mari, A. F. Theil, P. O. Mari, S. Mourgues, J. Nonnekens, L. O. Andrieux, J. de Wit, C. Miquel, N. Wijgers, A. Maas, M. Fousteri, J. H. Hoeijmakers, W. Vermeulen, PLoS Biol. 2009, 7, e1000220.

[134] I. Karakasilioti, I. Kamilier, G. Chatzinikolaou, T. Kosteas, E. Vergadi, A. R. Robinson, I. Tsarnardinos, T. A. Rozgaja, S. Siakoul, C. Tsatsanis, L. J. Niedernhofer, G. A. Garinis, Cell Metab. 2013, 18, 403.

[135] Z. Duan, M. Andronescu, K. Schultz, C. Lee, J. Shendure, S. Fields, W. S. Noble, C. Anthony Blau, Methods 2012, 58, 277.

[136] P. E. Garcia-Nieto, E. K. Schwartz, D. A. King, J. Paulsen, P. Collas, R. E. Herrera, A. J. Morrison, EMBO J. 2017, 36, 2829.

[137] J. Hu, S. Adar, C. P. Selby, J. D. Lieb, A. Sancar, Genes Dev. 2015, 29, 948.

[138] A. G. Zavala, R. T. Morris, J. J. Wyrick, M. J. Smerdon, Nucleic Acids Res. 2014, 42, 893.

[139] D. S. Bryan, M. Ransom, B. Adane, K. York, J. R. Hesselberth, Genome Res. 2014, 24, 1534.

[140] N. Crosetto, A. Mitra, M. J. Silva, M. Bienko, N. Dojer, Q. Wang, E. Karaca, R. Chiarle, M. Skrzypczak, K. Ginals, P. Pasero, M. Rowicka, I. Dikic, Nat. Methods 2013, 10, 361.