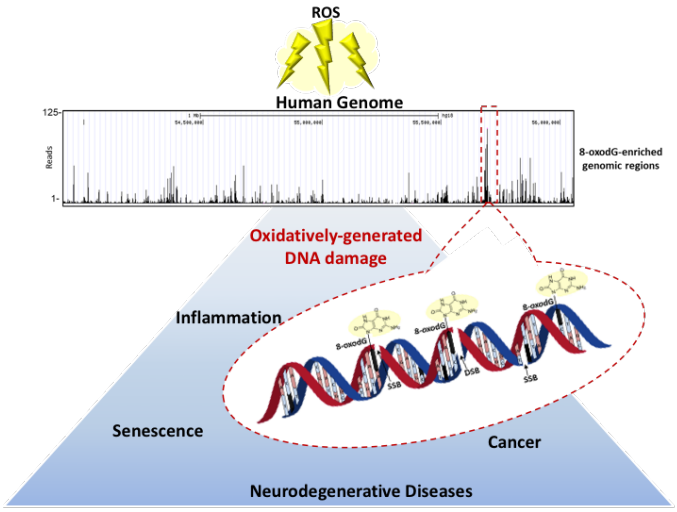
UNIVERSITY OF NAPOLI "FEDERICO II"

Doctorate in Molecular Medicine and Medical Biotechnology

XXXIII CYCLE

Francesca Gorini

Oxidative DNA Damage associates with Transcription and/or Replication processes



Year 2021

UNIVERSITY OF NAPOLI "FEDERICO II"

Doctorate in Molecular Medicine and Medical Biotechnology

XXXIII CYCLE

Oxidative DNA Damage associates with Transcription and/or Replication processes

Tutor
Prof. Stefano Amente

Candidate
Francesca Gorini

Year 2021

Table of contents

List of Abbreviations

ROS Reactive Oxygen Species

$^{\bullet}$OH Hydroxyl radical

H_2O_2 Hydrogen peroxide

O_2 Oxigen

$O_2^{\bullet-}$ Superoxide anion

NO Nitric oxide

NOX NADPH oxidases

$ONOO^-$ Peroxynitrite

SOD superoxide dismutase

MnSOD manganese- dependent superoxide dismutase

LSD1 Lysine-Specific histone Demethylases 1

LSD2 Lysine-Specific histone Demethylases 2

FAD flavin adenine dinucleotide

FADH2 Di-hydroflavine-adenine dinucleotide

NADPH nicotinamide adenine dinucleotide phosphate

OS oxidative stress

dG deoxyguanosine

8-oxodG 8-oxo-7,8-dihydro-2′-deoxyguanosine

BER Base Excision Repair

AP apurinic/apyrimidinic

OGG1 8-oxodG DNA glycosylase 1

3'dRP 3′- phospho-α,β-unsaturated aldehyde terminus

APE1 apurinic/apyrimidinic endonuclease 1

SSB single-strand break

XRCC1 X-ray repair cross-complementing 1

FEN1 Flap Endonuclease 1

PCNA Proliferating Cell Nuclear Antigen

RF-C Replication factor C

MYH MutY glycosylase Homologue

NEIL Nei-like

RP-A PCNA replication protein A

MMR Mis-Match Repair

TC-NER Transcription-Coupled Nucleotide Excision Repair

hMSH2 MutS homologues 2

hMSH6 MutS homologues 6

CSB Cockayne syndrome B

CSA Cockayne syndrome A

XPC Xeroderma pigmentosum complementation group C

8-oxodGTP 8-Oxo-2'-deoxyguanosine-5'-Triphosphate

PQSs G-quadruplex forming sequences

TFs transcription factors

VEGF vascular endothelial growth factor

NTHL Nth Like DNA Glycosylase

HIF1-α Hypoxia-inducible factor 1-alpha

TSS transcription start site

H3K9me2 histone 3 di-methylated lysine 9

ERE estrogen responsive elements

SNPs single nucleotide polymorphisms

LADs lamina-associated domain

G4 G-quadruplex

ARP aldehyde reactive probe

HIPS Hydrazino-iso-Pictet-Spengler probe

DMEM Dulbecco's Modified Eagle Medium

PBS Phosphate-buffered saline

DIP DNA Immuno-Precipitation

ChIP Chromatin Immuno-Precipitation

FACS Fluorescence-activated cell sorting

GRO Genomic run-on

MACS Model-based Analysis for ChIP-Seq

FW Forward

REV Reverse

ATAC Assay for Transposase-Accessible Chromatin

GO Gene Ontology

GSEA Gene Set Enrichment Analysis

MSigDB Molecular Signatures Database

FDR False Discovery Rate

TSS Transcriptional Start Site

TES Transcriptional End Site

MAZ Myc-associated zinc finger

PRDM1 PR domain zinc finger protein 1

BLIMP-1 B lymphocyte-induced maturation protein-1

β-IFN beta-interferon

ORC1 Origin Recognition Complex

ORIs DNA replication Origins

Abstract

8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxodG) is the most common marker of oxidative stress and its accumulation within the genome has been associated with major human health issues such as cancer, aging, cardiovascular and neurodegenerative diseases. 8-oxodG is considered the main source of dC:dG to dA:dT transversion. The majority of 8-oxodG is preferentially repaired by OGG1 glycosylase/AP (apurinic/apyrimidinic) lyase-initiated BER (Base Excision Repair) pathway.

It has been widely demonstrated that the 8-oxodG accumulation in the promoter regions of specific genes can stimulate transcription *via* the BER pathway. Based on this knowledge, we wondered whether the oxidative DNA damage could globally be correlated with the transcription process.

Here, we report the genome-wide distribution of 8-oxodG in human non-tumorigenic epithelial breast cells (MCF10A) using OxiDIP-Seq which combines the immunoprecipitation anti–8-oxodG antibodies with high-throughput sequencing. We show a specific 8-oxodG bimodal distribution within promoter regions and identify a subset of human promoters that accumulate 8-oxodG under steady-state condition. Furthermore, the comparison between OxiDIP-Seq and ChIP-Seq of Ser5- and Ser2-phosphorylated isoforms of RNA Polymerase II or GRO-Seq strongly suggest the association between 8-oxodG accumulation and transcription process. Besides, by performing OxiDIP-seq in quiescent (G0) cells, we classified the oxidized promoters in two subsets. The growing-specific promoters accumulate 8-oxodG through DNA replication-dependent events while the persistently oxidized promoters accumulate 8-oxodG through replication-independent and transcription- associated events.

I. Introduction

1.1 Reactive oxygen species (ROS), generated by endogenous and exogenous agents, cause Oxidative DNA damage

1.1.1 ROS production by endogenous and exogenous agents

DNA is a dynamic molecule that changes over time as a result of the alterations that involve DNA backbone and the normal nucleobases (Lindahl 1993).

It has been estimated that approximately 20.000 DNA nucleobase lesions accumulate in the human genome per cell every day (Jackson 2009).

Physiological processes cause some of these alterations which, if are not repaired or are repaired incorrectly, lead to DNA damage. This damage is correlated with pathological processes such as neurodegeneration (Kim 2015), ageing (Beckman 1998), and cancer (Klaunig 2004).

Among the DNA damage agents, the reactive oxygen species (ROS), generated by various exogenous sources or endogenous processes and/or environmental agents compromise the structure and the function of DNA (Figure 1) (Di Meo 2016).

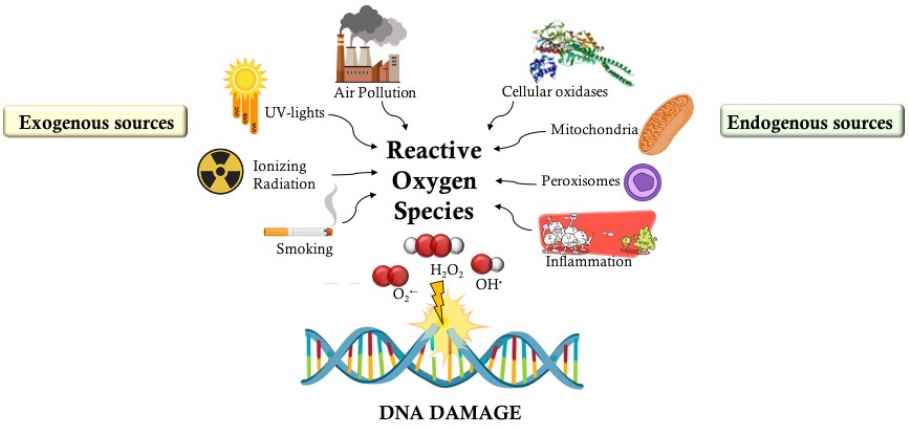


Figure 1: Exogenous and Endogenous ROS sources (Created with bio-render.com).

Specifically, the exposition to ionizing (e.g., γ-irradiation) (Azzam 2012) and non-ionizing (e.g. UV-C, UV-B, UV-A) irradiation together with air pollutants (e.g. cigarette smoke, industrial contaminants) constitute the major exogenous source of ROS (Fry 2005; Mena 2009; van Loon 2010). The exposure of the cell to gamma-irradiation produces directly a variety of radical and non-radical species from ionization of intracellular water (e.g. aqueous electron, $^{\bullet}OH$, H_2O_2) (Kohen 2002). Furthermore, the non-ionizing irradiation produces, even if indirectly, a whole range of ROS including O_2, H_2O_2, and $O_2^{\bullet-}$. Then, the UV radiations determine the hemolytic cleavage of H_2O_2 that yields $^{\bullet}OH$ radicals. Finally, also the air pollutants contain many types of nitric oxide (NO) derivates representing an important source of ROS that attack and damage the DNA (Azzam 2012; Kohen 2002).

Although the extremely high exposure of the organism to exogenous ROS determines an important DNA damage, the production of endogenous sources, being a continuous process during the life span of the cells, causes a more extensive damage (Davalli 2018). Specifically, endogenous ROS are mainly generated from cellular metabolism in the

mitochondria and peroxisomes (Figure 2) (Matsuzawa 2017; Rouhier 2008; Sabharwal 2013).

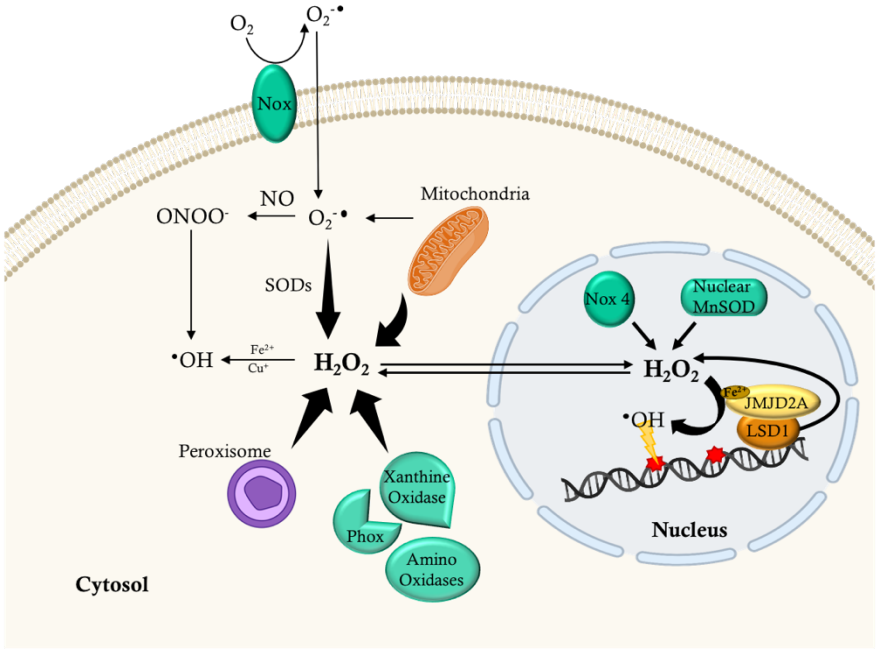


Figure 2: The reactive oxygen species (ROS) are produced in different cell compartments. Adapted from (Gorini 2021).

The formation of reactive oxygen species in the cell can be from the transmembrane NADPH oxidases (NOX) or mitochondria that produce superoxide anion ($O_2^{\bullet -}$) (Hansford 1997). It diffuses in the cytoplasm where can be converted either to peroxynitrite ($ONOO^-$) in the presence of NO or can be converted to H_2O_2 spontaneously or through catalysis by superoxide dismutase (SOD) (Fenton 1894; Fukai 2011). The H_2O_2 may also be generated from different oxidases (Amino oxidases, Phox and Xanthine oxidase) or from peroxisome or mitochondria. Then, the H_2O_2 can be reduced to H_2O, or partially reduced to $^{\bullet}OH$, through the metal-catalyzed Haber-Weiss reaction in the presence of reduced transition metals [e.g. Fe^{2+} (Fenton reaction) and Cu^{+}] (Fenton 1894). The $^{\bullet}OH$ molecules are also formed from the decay of the reactive

nitrogen species such as $ONOO^-$. ˙OH is the main strong cellular oxidant, but it remains unclear how cytosol-produced ˙OH can travel into the nucleus. Thus, it has been proposed that H_2O_2, being a molecule much more able to diffuse can move from the site of production into the nucleus (Aust 1999; Giorgio 2007). Alternatively, the H_2O_2 can be also produced directly in the nucleus where, through the Fenton, or other metal-dependent reactions, can give rise to local ˙OH, and hence DNA damage. In the nucleus, the nuclear oxidases (e.g. NOX4 and MnSOD, Lysine-Specific histone Demethylases 1 and 2, LSD1 and LSD2) produce H_2O_2 (Fang 2010; Kuroda 2005; Shi 2004; Tsang 2014). In particular, LSD1 and LSD2 are known to demethylate the mono and di-methylated Lys4 and Lys9 of histone H3 through a flavin adenine dinucleotide (FAD)-dependent oxidative reaction. During the reaction, the FAD cofactor is first reduced to $FADH_2$ and then newly oxidized to FAD by oxygen generating formaldehyde and H_2O_2 (Fang 2010; Shi 2004).

Then, the Fenton reaction converts the nuclear H_2O_2 to •OH, probably involving iron-containing complexes associated with the DNA structure. Supporting this hypothesis, it has been demonstrated that LSD1 can form complexes with members of the Jumonji-type demethylases, a superfamily of oxygenases containing a Fe^{2+} ion in their catalytic domain. The H_2O_2, produced during LSD1 reaction, is reduced in the hydroxide ion by the oxidation of the Jumonji-contained Fe^{2+} ion. In close proximity of DNA molecules, this event may be responsible for the oxidation of nucleobases (Pezone 2020).

1.1.2 ROS induce Oxidative DNA Damage

ROS play a physiological role in cell function contributing to cell homeostasis by modulating the activities of key regulatory molecules,

such as protein kinases and phosphatases, G proteins, and transcription factors (Davalli 2016; López-Otín 2013; Waypa 2016; Winterbourn 2008). To counter the excessive production of ROS, cells possess antioxidant systems which control ROS levels alleviating their effects by decreasing primary DNA damage that reduces the risk of mutation and tumor initiation.

The oxidative signal is reversed by enzymatic proteins that act in concert with non-enzymatic proteins and together with low molecular weight antioxidants (Matsuzawa 2017; Nicolussi 2017; Rouhier 2008; Sabharwal 2013; Vašák 2011). The antioxidant enzymes are peroxisomal catalase, SODs, glutathione peroxidase, and ascorbate peroxidase and generally remove O_2^-, H_2O_2 (Loren 2017). As non-enzymatic proteins are essentially known two potent systems Thioredoxins/ Thioredoxins reductase and Glutaredoxins/ Glutaredoxins reductase, which reduce disulfides to free thiol groups at the expense of nicotinamide adenine dinucleotide phosphate (NADPH) depletion (Matsuzawa 2017; Nicolussi 2017; Rouhier 2008; Sabharwal 2013; Vašák 2011). Finally, low molecular weight antioxidants are glutathione, ascorbate, carotenoids and melatonin.

To maintain a healthy cell status, oxidant generation and removal by antioxidants should be in a state of equilibrium (Figure 3) (Davalli 2018; Schieber 2014).

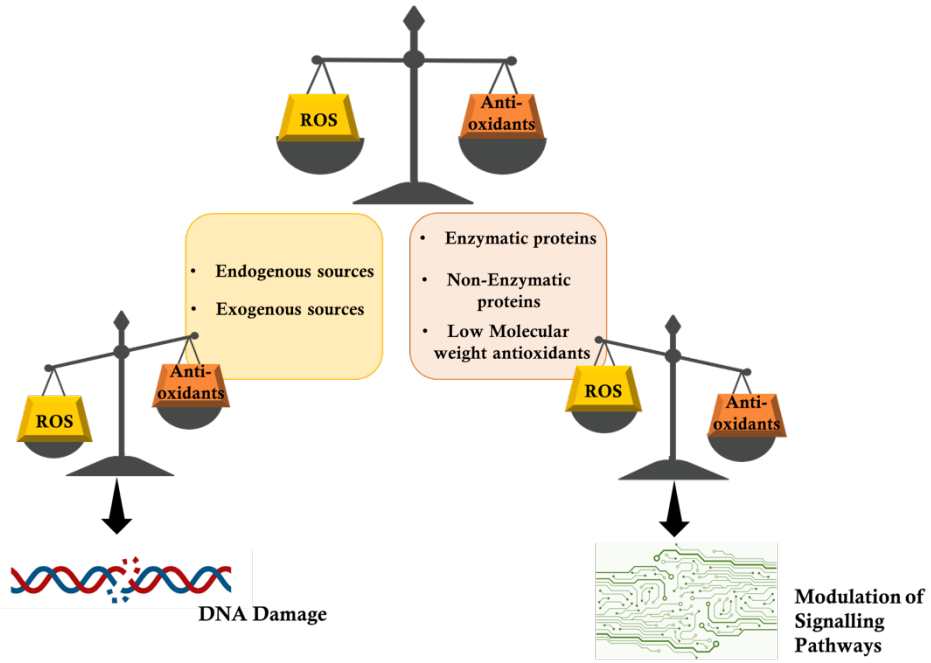


Figure 3: The balance between ROS production and antioxidant defense maintains redox homeostasis in the cells. Adapted from (Gorini 2021).

Imbalance of this equilibrium in favor of the oxidants disrupts the redox homeostasis and causes oxidative stress (OS) that, with an increase of ROS levels, induces improper signaling and increased oxidation of cell macromolecules (Davalli 2018; Schieber 2014; Sies 2017). In particular, ROS can generate various types of damage ranging from nucleobase and sugar modifications to modification and breakage of the phosphate backbone, having a dramatic impact on DNA (Cooke 2003; Cooke 2007; Evans 2004; Fleming 2017).

1.1.3 8-oxodG is a major product of DNA oxidation

Among DNA damage, the nucleobase and sugar damages, DNA strand breaks, DNA-DNA and DNA-protein crosslinks are the major caused by ROS.

All four genomic bases are susceptible to damage by ROS, but the low redox potential of deoxyguanosine (dG) makes this base especially

vulnerable to oxidation (Baik 2001; Cooke 2003; Evans 2004; van Loon 2010).

In particular, 8-oxo-7,8-dihydro-2′-deoxyguanosine (8-oxodG) is the primary product of dG oxidation and it is one of the most abundant and best characterized lesions created by ROS (Figure 4A) (Lindahl 2000; Montelone 2006).

This chemically altered base is the result of the C8 oxidation in the imidazole ring of deoxyguanosine.

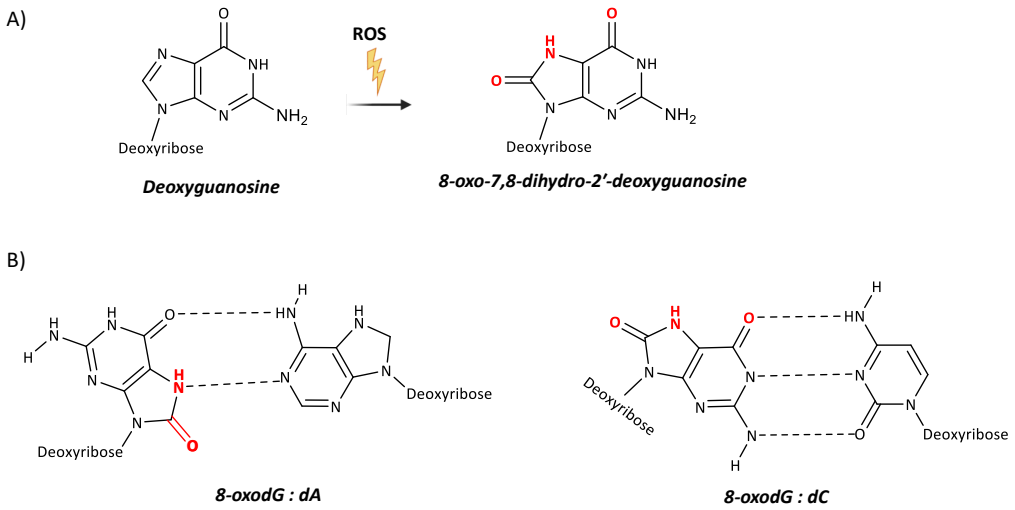


Figure 4: Schematic diagram of 8-oxodG chemical structures (A) and its altered base-pairs (B). Adapted from (Gorini 2021)

8-oxodG is considered a premutagenic DNA lesion because when present in the DNA during the replication process as it causes dC:dG to dA:dT transversion (Batra 2010; Boiteux 2017; Koga 2013; Maga 2007; Shibutani 1991).

Whereas the unmodified deoxyguanine assumes the *anti* conformation of glycosidic torsion angle, 8-oxodG favors a *syn* conformation due to steric repulsion between O8 and deoxyribose. The 8-oxodG base-pairing partner depends on the *syn* and *anti* conformation. Indeed, in the syn conformation, 8-oxodG mimics dT and determines an 8-oxodG

(syn):dA (anti) Hoogsteen base mispairing during DNA replication (Batra 2010; Boiteux 2017; Koga 2013; Maga 2007; Shibutani 1991). This mispair structurally mimics the dT:dA base pair and, because does not determine distortion of the DNA helix structure, escapes the DNA repair (Hsu 2004). Conversely, in *anti* conformation 8-oxodG pairs with a complementary dC and this induces a modest distortion of the DNA helix structure which is recognized by DNA repair that removes the 8-oxodG (Figure 4B).

1.2 8-oxodG recognition and repair through BER pathway

Cells have evolved the Base Excision Repair (BER) as the main mechanism of pre-replicative excision of 8-oxodG (Figure 5) (Lindahl 1990; Lindahl 2000). BER can proceed via two different subsets of proteins involved in two repair sub-pathways which operate independently: the classic short-patch mechanism, characterized by the replacement of the single damaged nucleotide, and the long patch, which processes 2–12 nucleotides (Fortini 1999,2007; Frosina 1996; Klungland 1997). The DNA glycosylase can be monofunctional, bifunctional with β-lyase activity or bifunctional with β/δ lyase activity (Fortini 1999; Markkanen 2017; van Loon 2010). The difference between monofunctional and bifunctional glycosylase is that the first one excised only the nucleobase releasing the modified nucleobase and creating an apurinic/apyrimidinic (AP) site, while the bifunctional perform the excision of the modified nucleobase and subsequently hydrolyze, via the associated AP lyase activity, the DNA backbone. This produces a 3' α,β-unsaturated aldehyde adjacent to a 5' phosphate following a β-elimination reaction. Some bifunctional DNA glycosylases can also convert the 3' aldehyde to a 3' phosphate, performing

a δ-elimination step (Fortini 1999; Fortini 2007; Frosina 1996; Markkanen, 2017; van Loon 2010).

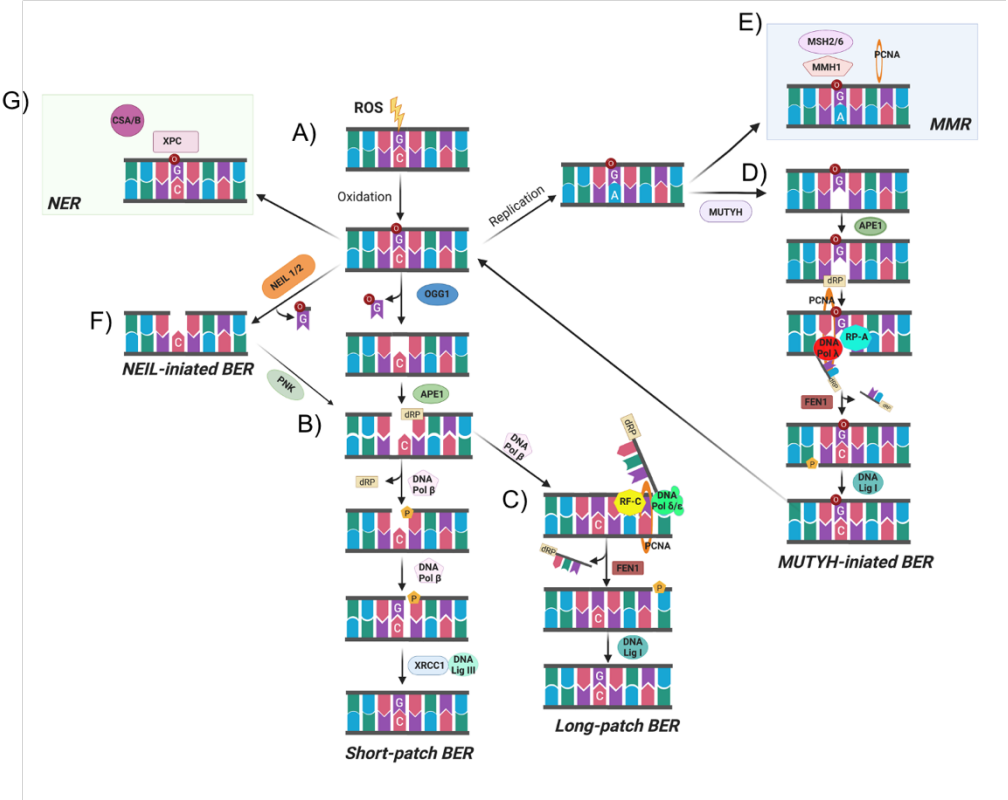


Figure 5: Summary of 8-oxodG repair pathways. (Gorini 2021)

The majority of 8-oxodG removal is due to short patch BER initiated by bifunctional glycosylase 8-oxodG DNA glycosylase 1 (OGG1), which specifically recognizes and then excises the 8-oxodG from the sugar-phosphate backbone, creating an abasic (apurinic/apyrimidinic, AP) site. OGG1, via its AP lyase activity, cleaves the DNA forming a 3′- phospho-α,β-unsaturated aldehyde terminus (3'dRP) and 5'-phosphate AP site. Then, the apurinic/apyrimidinic endonuclease 1 (APE1) cleaves the DNA phosphate backbone generating a DNA single-strand break (SSB). At this point, the short patch BER carries on with repair of SSBs; specifically, the DNA polymerase β utilizes its dRP-lyase activity to remove the downstream 5′sugar phosphate and repairs a single-nucleotide gap through the insertion

of one nucleotide. Next, the DNA ligase III/ X-ray repair cross-complementing protein 1 (XRCC1) complex fixed the 3′OH group of the newly inserted nucleotide with the downstream 5′P finishing the short-patch BER (Fortini 1999; Frosina 1996; van Loon 2010). When the 5'-termini cannot be fixed by SSB end-processing enzymes, the repair proceeds through the long-patch pathway by replacing a stretch of 2–12 nucleotides.

Two DNA polymerases participate in long-patch BER: the DNA polymerase β incorporates the first nucleotide into the nick and then the DNA polymerase δ performs strand displacement synthesis in the elongation step. In this scenario, the Flap endonuclease 1 (FEN1) in association with the DNA polymerases the Proliferating Cell Nuclear Antigen (PCNA), loaded by the replication factor C (RF-C), recognizes and excises the displaced oligonucleotide. Finally, DNA ligase I fills the remaining nick in the DNA backbone to coordinate the final step of long-patch BER (Fortini 1999; Klungland 1997; van Loon 2010).

Besides OGG1, other glycosylases, such as MutY glycosylase Homologue (MYH) and Nei-like (NEIL) -1 and -2, participate to 8-oxodG repair via the BER pathway (Boiteux 2017; Das 2006; Dou 2003; Markkanen 2013).

During DNA replication the DNA polymerase δ/e could mis-incorporate dAMP opposite 8-oxodG, the resulting mismatch is repaired by MYH-initiated long-patch BER (Boiteux 2017; Markkanen 2013). Before another round of replication, the monofunctional MYH excises, unconventionally, the undamaged adenine, and the resulting DNA ends are subsequently processed by APE1, as in canonical BER, resulting in a nick with 3′OH and 5′dRP moieties. Then, during the elongation step, the complex DNA polymerase λ-PCNA replication protein A (RP-A) incorporates the correct dC opposite 8-oxodG and additional one nucleotide. After the lesion bypass, there is a disassociation between RP-A and DNA pol λ, and FEN1, interacting with PCNA, cleaves the 5′ flap. Subsequently, the DNA ligase I,

through the interaction with PCNA, binds the created nicked intermediate and ligates the 5'P ends (Markkanen 2017; Svilar 2011).

In addition to bifunctional OGG1 and monofunctional MYH, the two bifunctional NEIL 1 and 2 repair 8-oxodG via BER (Boiteux 2017; Das 2006; Dou 2003). NEIL1 and NEIL2 excise 8-oxodG by catalyzing the β,δ-elimination of the abasic site and leave a 3′-phosphate at the resulting strand break. The 3′-phosphate in the gap is removed by the phosphatase polynucleotide kinase (PNK), generating a substrate for DNA polymerase β. The repair is completed with the ligation step mediated by the complex XRCC1/LigIII (Boiteux 2017; Das 2006; Dou 2003).

It has also been reported that exists an intricate network of other repair pathways for oxidatively DNA damage. Besides BER, Mis-Match Repair (MMR) (Yang 2001) and Transcription-Coupled Nucleotide Excision Repair (TC-NER) (Boiteux 2017; Fousteri 2008), plays an important role in 8-oxodG repair. In particular, MutS homologues 2 (hMSH2) and 6 (hMSH6) as factors of MMR machinery plays a role in the post-replicative removal of 8-oxodG opposite dA interacting with PCNA (Colussi 2002; Li 2008; Parker 2001).

Additionally, the Cockayne syndrome B (CSB) and A (CSA) proteins, together with Xeroderma pigmentosum complementation group C (XPC) protein, as components of the TC-NER pathway, have a role in the repair of 8-oxodG (D'Errico 2007, 2006; Wong 2007).

1.3 Unrepaired 8-oxodG represents a main source of genome instability

When 8-oxodG is incomplete repaired or not repaired, it may contribute to genome instability via several mechanisms, as well as the accumulation of SSBs unrepaired intermediates.

A single 8-oxodG was able to generate a spectrum of mutations, but predominantly dC:dG to dA:dT transversions and single-base deletions (Boiteux 2017; Bruner 2000; Dulak 2013).

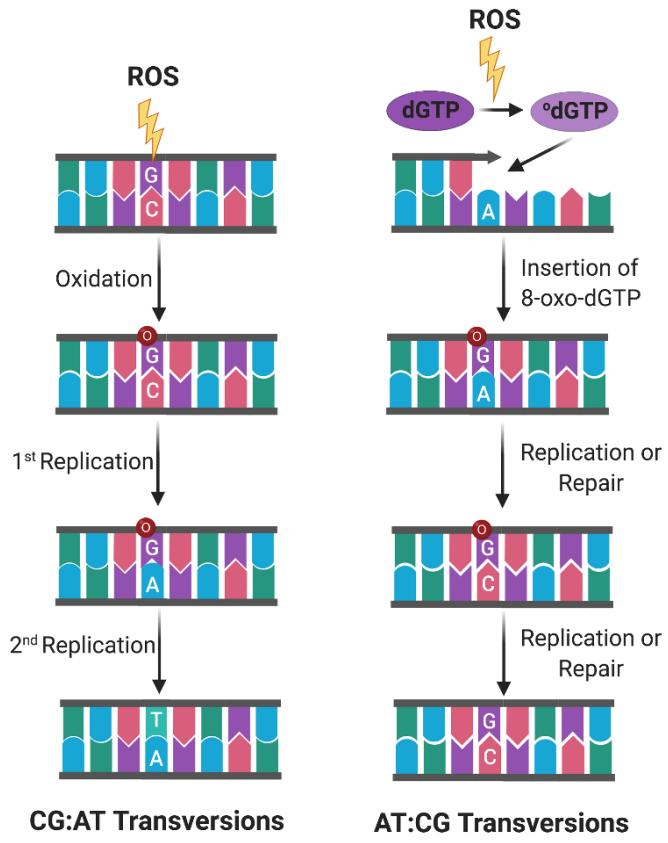


Figure 6: Two mechanisms of 8-oxodG induced transversions: (A) dC:dG to dA:dT transversions, (B) dA:dT to dC:dG transversions (Gorini 2021).

In the syn conformation, 8-oxodG has the ability to mimic dT and if it is erroneously bypassed, during DNA replication a dA:8-oxodG base pair is formed. If the dA:8-oxodG mispair is left unrepaired during a second round of replication it will produce a dA:dT base-pair leading a dC:dG to dA:dT transversion mutation (Figure 6A). In a study that examined the coding regions of 518 cancer-related genes in 210 different human cancers, the

dC:dG to dA:dT transversion has been identified to be the predominant somatic mutation (Greenman 2007). Independent studies also demonstrated that the dC:dG to dA:dT transversion represents one-third of the 22,910 somatic substitutions identified in a lung cancer cell line and the second most predominant mutational signature in melanoma cells (Pleasance 2010b, 2010a).

When in the deoxyribonucleotide pool, 8-oxodGTP is erroneously incorporated during DNA replication gives rise to the 8-oxodG:dA mismatch. If replication or inappropriate MYH excision of 2′-deoxyadenosine occurs, the 8-oxodG pairs with cytidine, in the next round of replication. Subsequently, 8-oxodG:dC mispair is a substrate for OGG1-initiated repair which, in this context, produces the dA:dT to dC:dG mutation(Figure 6B) (Inoue 1998; Maki 1992; Satou 2009, 2007; Suzuki 2017). Such dA:dT to dC:dG mutations are found in esophageal adenocarcinoma and inflammatory Barrett's esophagus (Cs 2015; Murugaesu 2015).

Moreover, 8- oxodG accumulation in the genome has been associated with cancer initiation and progression and has been proposed as a prognostic factor in breast cancer. Genetic knock-out mouse models of the BER pathway proteins have been generated to determine their role in genome maintenance. It has been demonstrated that in OGG1+/- and -/- mice there is an increase in nuclear and mitochondrial 8-oxodG levels with an elevated dC:dG to dA:dT transversion rates at 18 months after birth (Kakehashi 2017; Klungland 1997; Minowa 2000; Sakumi 2003; Xie 2004). In addition, when exposed to the genotoxic agents OGG1+/- and -/- mice show a multiorgan enhanced susceptibility for cancer development. Similarly, MYH+/- mice have been generated and are viable and fertile and also show only a slight predisposition to develop intestinal cancer (Sakamoto 2007; Xie 2004). Moreover, either OGG1-/- and MYH-/- mice exhibit a strong predisposition for lung and ovarian cancers and lymphomas 118. Finally,

lacking other proteins involved in the repair of 8-oxodG, such as APE1, Pol β, XRCC1, DNA ligase I and III display embryonic lethality. In conclusion, these data suggest that the impairment of the function of one or more than one intermediates of the BER pathway has a crucial role in the maintenance of genomic stability and in the prevention of certain cancers (Bentley 1996; Ludwig 1998; Maga 2007; Puebla-Osorio 2006; Sobol 1996; Tebbs 1999; Xanthoudakis 1996).

1.4 The role of 8-oxodG as an epigenetic marker

The mechanisms of 8-oxodG formation and its repair could play a role in the regulation of gene expression and may have possibly epigenetic-like features in cells during the response to oxidative stress (Fleming 2017; Giorgio 2020; Poetsch 2020; Wang 2018).

While the 8-oxodG pro-mutagenic aspects have been widely reviewed, the flip side of oxidative DNA damage relating its involvement in possibly beneficial gene regulatory roles must be discussed.

Several evidences showed that the interaction of 8-oxodG and AP sites with G-quadruplex structures is important for the epigenetic gene regulation. In details, the oxidative modification of G in G-quadruplex forming sequences (PQSs) present in the coding strand of gene promoters leads to an increase of the transcription level (Fleming 2017; Fleming and Burrows 2017). When the OGG1 recognizes and excises 8-oxodG from the duplex it creates an AP site which unmasks the PQS, leading to the formation of a G-quadruplex (Fleming 2017; Fleming and Burrows 2017). Following the G-quadruplex structure extends the APE1 binding on the AP site and favors the recruitment of other transcription factors (TFs) for gene activation. This mechanism has been proposed for the expression of different genes such as VEGF, PCNA, NTHL, HIF1-α and, more recently, to the DNA repair gene NEIL3 (Fleming 2019, 2017; Pastukh 2007; Redstone 2019; Zhu 2018).

Furthermore, increased levels of 8-oxodG have been found in binding regions of other TFs that interact with specific enzymes that produce ROS as intermediates, as shown for the Estrogen Receptor α and c-Myc. Both these transcriptional factors recruit LSD1 during their regulation (Amente 2010, 2013a, 2013b; Majello 2019; Perillo 2008; Pezone 2020; Sengupta 2020; Zuchegna 2014).

The LSD1 demethylation reaction determines local oxidation of DNA around the transcription start site (TSS) region of Myc target genes (Figure 7) (Amente 2010;Amente 2010).

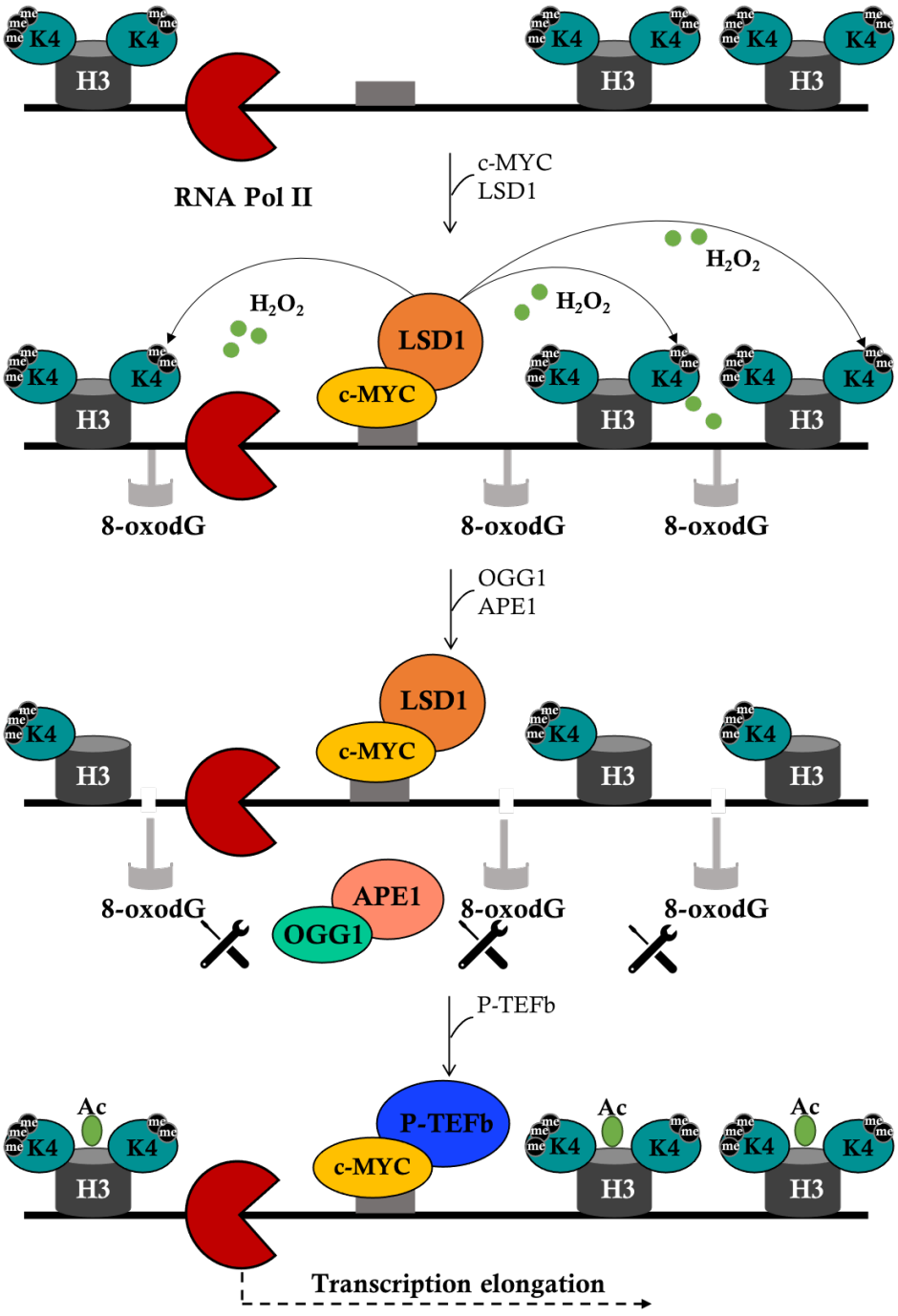


Figure 7: Model of "scheduled" DNA damage induced upon c- Myc mediated transcription. Adapted from (Amente 2010)

This event triggers the recruitment of BER enzymes such as OGG1 and APE1. The last one seems to play an important role as a linker between gene

transcription and repair, being able to interact with some RNA Polymerase II co-activators, as p300 (Amente 2010; Amente 2010).

A quite similar mechanism has been reported for Estrogen receptor-mediated transcription (Figure 8) (Perillo 2008).

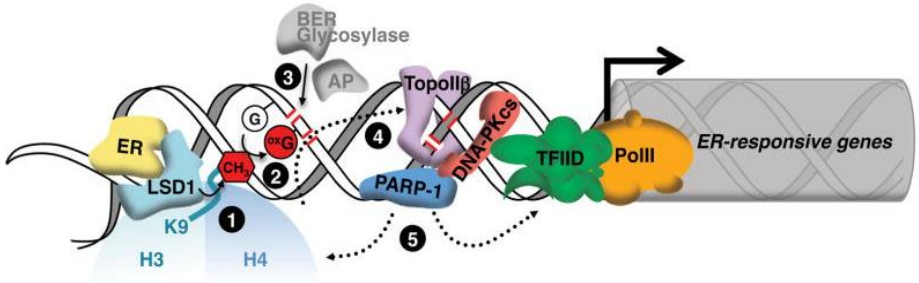


Figure 8: "Scheduled" DNA damage induced by Estrogen receptor-mediated transcription (Perillo 2008).

In this case, the estradiol-estrogen receptor complex mediates the recruitment of LSD1 which remove a methyl group from the di-methylated lysine 9 on the histone 3 (H3K9me2). The peroxide produced by LSD1 oxidizes guanines in ERE (estrogen responsive elements) regions and as consequence OGG1 is recruited in site. The DNA nicks generated upon the removal of 8-oxodG by the BER machinery act as entry points for the Topoisomerase II β, and induce chromatin relaxation and the recruitment of initiation complexes (Perillo 2008).

Finally, 8-oxodG could act also in concert with the DNA methylation performing its epigenetic function. Indeed, even if the mechanisms are not yet well defined, the repair of 8-oxodG is linked to DNA methylation. Taken together, these findings reveal that 8-oxodG clearly has potential roles in gene regulation.

1.5 Methods to genome-wide mapping 8-oxodG

It is seemingly well-known that the dynamic equilibrium between local rate of 8-oxodG production and its local repair efficiency results in 8-oxodG unevenly distribution along the genome.

Nakabeppu et al. provided the first strategies for the genomic view of the 8-oxodG distribution by using a monoclonal antibody for the immunofluorescence detection of 8-oxodG (Ohno 2006). They delineated the 8-oxodG distribution pattern on human metaphase chromosomes from human peripheral lymphocytes. Specifically, a high density of 8-oxodG is mapped, with a megabase resolution, within the boundary regions of R and/or G chromosomal bands as well as within regions with a high frequency of recombination and single nucleotide polymorphisms (SNPs). This suggested that 8-oxodG could largely contribute to the genomic diversity in human beings (Ohno 2006).

Subsequently, *Toyokuni et al.* combined the immunoprecipitation of 8-oxodG-containing DNA with microarray hybridization in order to map 8-oxodG in the genome of normal rat kidney cells. They observed that 8-oxodG is preferentially enriched within the gene deserts and not associated with the transcription activity of genes (Yoshihara 2014). Moreover, they showed that there was a strong correlation between the 8-oxodG distribution and lamina-associated domain (LADs) suggesting that the susceptibility to oxidative modification is determined by the spatial location of genomic DNA in the nucleus (Yoshihara 2014). Recently, novel genome-wide strategies, that combine oxidation-enriched library preparation protocols with next-generation sequencing approaches, have arisen to map steady-state levels of 8-oxodG in the yeast, mouse and human genomes (Figure 9) (Ding 2017; Poetsch 2018; Wu 2018).

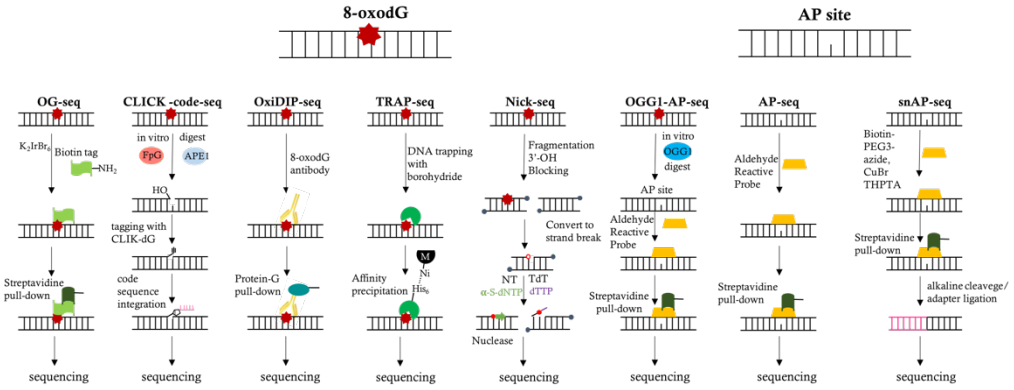


Figure 9: Scheme of methods to map 8-oxodG genome-wide. Adapted from (Poetsch 2020).

Burrow's laboratory developed a method to sequence the 8-oxodG sites in mouse genome, called **OG-Seq** which take the advantage of the sensitivity of 8-oxodG to hyperoxidation leading to covalent biotinylation of 8-oxodG (Ding 2017). This approach delineated an 8-oxodG map, with a 150 bp genomic resolution, in both wild-type and OGG1-/- MEFs. The study showed an 8-oxodG enrichment at specific gene loci (including promoters, 5′-UTRs, 3′-UTRs, exons and introns) compared with the intergenic regions. Moreover, the authors established that more G-quadruplex (G4) and 5′-GG-3′ reactive sequences than they expected were inside 8-oxodG-peaks (Ding 2017).

In order to mark each 8-oxodG position in the Saccharomyces cerevisiae genome, *Wu et al.* used an innovative technique **Click-code-Seq** that joins three specific elements: the specificity of DNA repair enzymes, the efficiency of a click DNA ligation reaction and high-throughput sequencing (Wu 2018). Thanks to Click-code-Seq, it was demonstrated that 8-oxodG accumulates at sites of high nucleosome occupancy and that the first dG in a 5′-GG-3′ dinucleotide is most easily oxidized (Wu 2018).

Also *Poetsch et al.*, to better understand the correlation between DNA repair and mutation distributions developed **AP-seq** as new method to measure AP

sites as the first repair intermediate approximately 250-bp resolution on genome-wide scale (Poetsch 2018). In this study, it is used an aldehyde reactive probe (ARP) which reacts with the aldehyde group of the AP site and consequently introduces a covalent biotin tag into the DNA at the damage site (Poetsch 2018).

Then the biotin-tagged DNA fragments are enriched through a streptavidine pull-down and sequenced using high-throughput technology. Similarly to AP-seq, for the **OGG1-AP-seq** it is used ARP. But in this case, the AP sites are masked with methoxyamine before in vitro glycosylation with OGG1 leading the specific detection of 8-oxoG through conversion into a secondary AP site (Poetsch 2018).

Recently *Amente et al.*, developed a highly sensitive methodology that enhanced a fragmented mammalian genome with an 8-oxodG-specific antibody following next-generation sequencing (**OxiDIP-seq)** in human and mouse cells with a resolution of about 200-300 bp (Amente 2019). This is a pull-down-based approach to enrich 8-oxodG-containing DNA fragments. Indeed, the study revealed that 42% of the identified 8-oxodG peaks mapped within gene loci, specifically in the promoter and in the gene body regions (Amente 2019). Moreover, the authors demonstrated that there is a G4-enrichment at 8-oxodG-containing regions and there is a complex association between 8-oxodG and GC content. Promoter regions with high (> 47%) GC content display low levels of 8-oxodG (Gorini 2020). This suggested that other mechanisms, such as the epigenetics involved in the regulation of transcription and replication regulation, may be involved in the accumulation of 8-oxodG (Amente 2019; Gorini 2020).

Balasubramaniam's group mapped AP sites, as intermediate in the repair of 8-oxodG, in the human genome at single-nucleotide resolution through the "**snAP-seq**" procedure (Liu 2019). This approach attacks the biotin to AP sites through Click chemistry using the Hydrazino-iso-Pictet-Spengler (HIPS) probe. Appling the snAP-seq directly to HeLa cells, it was not

possible to detect a consensus position between the mapped AP sites suggesting that AP sites do not accumulate site-specifically at single-nucleotide level. In this case, the authors identified DNA stretches where AP sites accumulate in control and APE1 knocked-down cell using a peak calling bioinformatic tool (Liu 2019). According to previous reports (Ding 2017; Liu 2019; Poetsch 2018), the genomic locations with open chromatin are enriched for AP sites, suggesting that these regions are more prone to the formation of AP sites, compared to other forms of DNA damage (Liu 2019).

Moreover, *Fang and Zou* in their paper published the "**enTRAP-seq**" protocol to identify 8-oxodG in mouse embryonic fibroblasts with a resolution of 250bp (Fang 2020). The 8-oxodG-containing DNA fragments are trapped in OGG1 mutant (K249Q) and enriched via affinity purification. Notably, enTRAP-seq determined an enrichment of 8-oxodG in regulatory elements such as promoters, 5'UTR, CpG islands and G4 (Fang 2020).

In conclusion, a highly sensitive and quantitative approach (named "**Nick-seq**") to map, at single-nucleotide resolution, oxidation-induced AP sites in DNA from E. coli treated with a sub-lethal dose of hydrogen peroxide, is recently reported by *Dedon et al* (Cao 2020).

The authors demonstrated that the oxidation-induced AP sites are specifically associated with DNA regions undergoing replication or transcription during H_2O_2 stress. They also suggested that transcriptionally active and single-strand DNA regions are vulnerable to oxidatively-induced DNA damage (Cao 2020).

	OG-seq	Click-code-seq	OxiDIP-seq	OGG1-AP-seq	AP-seq	snAP-seq
Damage recognition	✗	✓	✗	✗	✗	✓
Single nucleotide resolution	✓	✓	✗	✓	✓	✓
Potential single nucleotide	✗/✓	✗	✗	✓	✓	✗
Damage recognition before sonication	✓	✓	✓	✓	✓	✓
Ease of use	✓	✓	✓	✓	✓	✓
Reagents commercially avaiable	✓	✗	✓	✓	✓	✗
Potential problems with specificity	Induced oxidation and probe side reactions	Specificity of FpG for 8-oxodG and incomplete masking of AP sites/strand breaks	General antibody specificity	Specificity of OGG1 for 8-oxodG and incomplete masking of AP sites	Side reactions with other aldehydes, e.g 5-fU,if present	Side reactions with other aldehydes. Incolplete depletion of 5-fU

Figure 10: Table of properties of the different methods to map 8-oxodG genome wide. Adapted from (Poetsch 2020).

All the above-described experimental genome-wide methodology used to map oxidatively generated DNA damage possess benefits and weaknesses. Differences in methods, with their advantages and disadvantages, have been summarized in Figure 10.

II. Aims of the study

8-oxodG is widely considered as the most common marker of oxidative stress. Its accumulation within the genome has been associated with major human health issues but new discoveries revealed that there is more to it than that. Indeed, there is growing evidence about 8-oxodG potential role as an epigenetic marker. In particular, 8-oxodG has a role in the epigenetic regulation of transcription, but this role could be intrinsically associated with its nature as DNA damage.

Indeed, it is well known that the transcription activation can require "scheduled" oxidative DNA damage, thus providing to DNA lesions not only a harmful role, but also effects that are indispensable for an important cell function that is transcription.

Thus, the aim of this work is firstly focused to provide an accurate map of 8-oxodG accumulation to determine if was randomly distributed along human genome. Then, we wondered if the pre-mutagenic 8-oxodG accumulation was correlated with structural/functional landscape of the genome.

III. Materials and Methods

Cell culture and treatments

MCF10A cells were cultured in a 1:1 mixture DMEM-F12 supplemented with 5% horse serum, 10 μg/ml insulin, 5 μg/ml hydrocortisone, 20 ng/ml epidermal growth factor, 100 ng/ml cholera enterotoxin and incubated at 37°C in a humidified atmosphere with 5% CO2 (Ambrosio 2016). MCF10A cells were arrested in G0 by growth for 2 days in a minimal medium containing 1:1 mixture DMEM-F12 supplemented with 5% horse serum .

Flow cytometry analysis and 8-oxodG genomic quantification

DNA profiles were analyzed as follows: cells were fixed in 70% ethanol at –20◦C, then stained in hypotonic solution of PBS, 50 μg/ml propidium iodide, 50 μg/ml and 0.00125% Nonident-P40 for 30 min at room temperature. For Ki67 quantification cells were, after fixing, permeabilized with 0.1% Triton X-100/PBS and blocked in 5% bovine serum albumin/PBS. Cells were incubated with the primary antibody anti-Ki67 and with the secondary antibody Alexa647 donkey anti-goat (Invitrogen) before propidium iodide staining. For 8-oxodG quantification, cells were fixed and permeabilized as indicated for Ki67, treated with 50 μg/ml RNAse incubated with anti-8-oxodG (Millipore, 1:200 diluted 5% bovine serum albumin/PBS) and the secondary antibodies Alexa488 anti-goat before propidium iodide staining. Cytofluorimetric acquisition and analysis were performed on a Becton Dickinson FACScalibur flow cytometer using FACSDiva and Cyflogic for analysis.

OxiDIP-sequencing and quantitative 8-oxodG immunoprecipitation assays

Genomic DNA from growing or G0 arrested MCF10A cells was extracted by using Dneasy Blood&Tissue kit (Cat. no. 69504, QIAGEN). 10 μg of

genomic DNA per immuno-precipitation were sonicated in 100 µl TE buffer (100 mM Tris–HCl pH 8.0, 0.5 M EDTA pH 8.0) to generate random fragments ranging in size between 200 and 800 bp using Bioruptor Plus UCD-300. 4 µg of fragmented DNA in 500 µl TE Buffer were denatured for 5 min at 95◦ C and immuno-precipitated overnight at 4◦ C with 4 µl of polyclonal antibodies against 8-Hydroxydeoxyguanosine (AB5830 Millipore) in a final volume of 500 µl IP buffer (110 mM NaH_2PO_4, 110 mM Na_2HPO_4 ph 7.4, 0.15 M NaCl, 0.05% Triton X-100, 100 mM Tris–HCl pH 8.0, 0.5 M EDTA pH 8.0) under constant rotation. The immuno-precipitated complex was incubated with 50 µl Dynabeads Protein G (Cat. No. 10003D, ThermoFisher Scientific, previously saturated with 0.5% bovine serum albumin diluted in PBS) for 3h at 4 ◦C, under constant rotation, and washed three times with 1 ml washing buffer (110 mM NaH_2PO_4, 110mM Na_2HPO_4 pH 7.4, 0.15 M NaCl, 0.05% Triton X- 100). The beads–antibody–DNA complexes were then disrupted by incubation in 200 µl Lysis buffer (50 mM Tris– HCl pH 8, 10 mM EDTA pH 8, 1% SDS, 0.5 mg/ml Proteinase K) for 4 h at 37◦C, and 1 h at 52◦C following addition 100 µl Lysis buffer. The immuno-precipitated DNA was purified by using MinElute PCR Purification kit (Cat. No. 28004, QIAGEN) in a final volume of 72 µl EB buffer (provided in the kit). All the steps of the OxiDIP-Seq protocol, including the washes of the immunocomplexes, were carried out in low-light conditions. Furthermore, 50 µM N- tert-butyl-α-phenylnitrone (stock solution: 28 mM in H2O; B7263, Sigma) was added to each Dneasy Blood&Tissue buffer, IP and washing buffers, to preserve the oxidized DNA (Lu 2004). Conversion of ssDNA to dsDNA was obtained by Random Primers DNA Labeling System (Cat. No. 18187-013, ThermoFisher Scientific).

For qPCR analysis, 3µl of 8oxodG immunoprecipitated DNA (antibody AB5830, Millipore) was analyzed in duplicate by quantitative PCR, using

SYBR Green 2X PCR Master Mix (Applied Biosystems). The following primer sets were used:
Positive region (genomic position chr2: 233294905 - 233294981)
FW 5'-CCAACATCTTAAATTTGTCAACTCTC;
REV 5'-TGCTGGCAGAAGTGTGATTT.
Negative region (genomic position chr2: 232053796-232053862)
FW 5'- AAGCTGGAGGCAGAGTGG;
REV 5'- TCTGACAACCCTGTTCACTACC.

Preparation of OxiDIP sequencing libraries

Library preparation was performed as described in (Amente 2019) using 2 ng of DIP or Input DNA. Prior to sequencing, libraries were quantified using Qubit dsDNA HS Assay Kit (Invitrogen) and quality-controlled using Agilent Bioanalyzer. 50 bp single-end sequencing was performed using Illumina HiSeq 2000 platform according to standard operating procedures.

Read processing and identification of 8-oxodG-enriched regions

Reads were quality checked and filtered with NGS-QC Toolkit (Patel 2012). Alignments were performed with BWA (Li 2010) to hg18 reference genome using default parameters. SAMtools (Li 2009) and Bedtools (Quinlan 2010) were used for filtering and format conversion steps. The identification of peaks from uniquely mapped reads, after removal of PCR duplicates, was performed using MACS (Zhang 2008) ($P < 1e{-}5$ and fold enrichment > 7) and DNA Input was used as control. UCSC genome browser was employed for data visualization. Uniquely mapped reads of the 8-oxodG signal were normalized over genomic input ($\log_2$ 8-oxodG/Input ratio) using the bamCompare tool from Deeptools suite (Ramírez 2014), with SES method (Diaz 2012) as scaling factor. This should account for GC content sequencing bias, which would affect the pull-down samples and the inputs alike, as well as for the bias linked to different amounts of DNA.

OxiDIP-Seq peaks were annotated using PAVIS (Huang 2013). The hg18 genomic coordinates of peaks identified in MCF10A cells were converted to hg38 coordinates before annotation by using the UCSC tool liftover. Relative peak enrichment was determined with the Fisher test of the Bedtools suite.

Metagene analysis and heatmaps were generated using the computeMatrix and plotHeatmap tools from Deeptools suite with default parameters. Signal profile plots were derived using R starting from the matrices generated by the computeMatrix tool.

Pol II-S2P ChIP-sequencing

Chromatin extracts of growing or G0 arrested MCF10A cells were performed as described (Ambrosio 2015).

Rabbit polyclonal against Phospho RNA Polymerase II (S2) (A300-654A) was used.

ChIP-seq libraries were prepared from 2 ng of ChIP (or Input) DNA with TruSeq ChIP Sample Prep Kit (Illumina) according to the manufacturer's instructions. 50bp single-end sequencing was performed using Illumina HiSeq 2000 platform. For qPCR analysis, 3µl of Pol II-S2P immunoprecipitated DNA (antibody A300-654A, Bethil) was analyzed in duplicate by quantitative PCR, using SYBR Green 2X PCR Master Mix (Applied Biosystems).

The following primer sets were used:

Positive region (genomic position chr5: 172126206 - 172126502)

FW 5'- CTGGTACATTTCCACGGAGAG;

REV 5'- AAGTGCTATGGGAAGCAGAAA.

Negative region (genomic position chr3: 73471477 - 73471656)

FW 5'- TGCTGTCATCTTGCTTACCC;

REV 5'- ACTGGACTAGCTGCTTCATTC.

ChIP-seq and ATAC-seq analysis

OGG1 ChIP-seq data was downloaded from GEO (GSE89017); ChIP-seq of PARP1, Pol II-S5P, were downloaded from GEO (GSE93040); ChIP-seq of H3K4me3, H3K4me1, H3K9ac, H3K27ac were from GEO (GSE85158). Pol II-S2P ChIP-seq data was performed in Amente's laboratory according to the above-described protocol.

ChIP-seq reads were quality checked and filtered with NGS-QC-Toolkit. Alignments to hg18 reference genome were performed with BWA using default parameters. SAMtools and Bedtools were used to perform filtering steps and format conversions. Uniquely mapped reads of each signal were normalized over genomic input (log2 ChIP/Input ratio) using the bamCompare tool from Deeptools suite (Ramírez 2014), using SES method (Diaz 2012) as scaling factor as described above.

Metagene analysis and heatmaps were generated using computeMatrix and plotHeatmap tools from Deeptools suite with default parameters, signal profile plots were derived using custom script in R starting from the matrices generated by the computeMatrix tool. Preprocessed and normalized data for ATAC-seq was publicly available from GEO (GSE89013). Genomic coordinates were converted to hg18 reference coordinates using UCSC liftover tool. Metagene analysis was performed using computeMatrix tool from the Deeptools suite with default parameters, while plots were generated using custom script in R starting from the data matrix produced by computeMatrix tool.

GRO-seq analysis

GRO-seq data was obtained from Array-Express (E- MTAB-742). FASTQ files were aligned to hg18 reference genome using Bowtie (Langmead 2009) with default parameters and allowing a maximum of two mismatches for the identification of uniquely mapping regions. GRO-seq read quantifications were performed using HTSeq (Anders 2015) and transcription levels of the

human genes were converted into transcripts per million (TPM) of mapped reads. Bidirectional transcription at TSSs was determined analyzing the GRO-seq signal with the computeMatrix tool from the Deeptools suite with default parameters and R was used for plots generation.

Reference, damaged and control gene/promoter datasets

The 21074 human genes used in this study were obtained from the hg18 Refseq genes catalogue removing all of the transcripts having the same genomic coordinates ('chromosome', 'start' and 'end') and keeping those showing alternative TSSs. Genes <2.5 kb in length were also removed. The promoter regions were then defined as the 1 kb regions flanking the TSSs of the above selected (21 074) human genes. Bedtools suite was used to intersect the identified 8- oxodG high-confidence peaks (Amente 2019) with the above-described promoters. The genes containing at least one 8-oxodG high-confidence peak in their promoter region were defined as 8- oxodG-positive (oxidized) promoters in growing cells. Control genes were defined as genes with negligible levels of 8-oxodG at promoter regions in growing cells. In particular, the 8-oxodG signal was binned over the considered human promoters using a 100 bp bin size. Then, the promoter level of 8-oxodG signal was defined as the maximum 8-oxodG value computed overall the bins of each promoter by using the computeMatrix tool from Deeptools suite with default parameters and custom R scripts. Finally, the subset of control genes was defined as the genes where the value of 8-oxodG promoter signal over the Input DNA was lower than 1-fold, corresponding to the bottom 15% of the over- all 8-oxodG promoter signal distribution.

Comparison of oxidized promoters identified in G0 and growing cells

Starting from the set of peaks identified from cells in the G0 condition we defined a set of (n = 811) G0 promoters by following the same procedure as for the growing cells (see previous paragraph).

We defined the class of growing-specific promoters as the promoters marked as oxidized in growing cell condition only and the set of G0-specific promoters as the promoters marked as oxidized in G0 condition only. Common oxidized promoters were defined as the intersection between G0 oxidized promoters and growing oxidized promoters. We analyzed the presence of DE genes between growing and G0 condition in these two classes by counting how many genes from the class of common oxidized promoters and the class of growing-specific promoters were also marked as differentially expressed between growing and G0 cells (~ 11% and ~ 9% respectively).

GC content enrichment at the promoter regions

For the quantification of GC content, the hg18.gc5Base track was retrieved from UCSC and the average GC content was assessed at oxidized and control promoters using the computeMatrix tool from the Deeptools suite with default parameters.

To investigate the relationship between GC content and 8-oxodG levels, oxidized promoters were analyzed at two resolution levels by first dividing the corresponding regions in bins of 50 bp (Figure 14C) or 10 bp (Figure 14D) and then calculating the average GC content per bin using the computeMatrix tool from the Deeptools suite with default parameters.

Statistical analysis

Linear correlations between Pol II-S5P/Pol II-S2P/OGG1 signal, transcription levels, and 8-oxodG signals, as well as between biological replicates of Input and OxiDIP-seq experiments, were tested by means of Pearson's correlation test on the gene loci of the oxidized promoters using multi-BamSummary and plotCorrelation tools from the Deep- tools suite with default parameters. Heatmaps were generated with R. Statistical significance of the observed differences in transcription levels and gene

lengths between the gene clusters were evaluated by means of two tailed t-test with heteroskedasticity assumption. Fisher's exact test was used to test the statistical significance of the distribution of each promoter class among the identified gene clusters. Mean value and standard deviation of each genomic signal at the promoter/gene loci were calculated with compute-Matrix tool from Deeptools suite, with default parameters, while the standard error was calculated in R. This study was conducted using 0.05 as the significance threshold and all statistical analyses were performed with R version 3.5.

Pathway enrichment and HOMER motif analysis

Gene Ontology (GO) functional annotation of pathways (Puissant 2013), or groups of functionally related genes was performed using GSEA. GSEA assigns to each gene set an enrichment score (ES) (Puissant 2013). The significance of the ES is assessed by means of permutation based on p-value and adjusted for multiple testing through FDR correction.

GSEA was run on the collections of 50 hallmark gene sets (h) from version 7.4 of the Molecular Signatures Database Gene sets with an FDR ≤ 0.25 and a nominal $P \leq 0.05$ were considered significant (Puissant 2013).

The search for motif enrichment in specific oxidized promoters was performed by using the HOMER tool. We used the *findMotifs.pl* that performs a number of operations to provide a basic analysis of motif and functional enrichment. To run findMotifs.pl we used the above defined human promoter set. HOMER tool creates HTML output pages for its de novo analysis, containing a sorted list of non-redundant motifs ranked by their enrichment p-values (Heinz 2010).

IV. Results

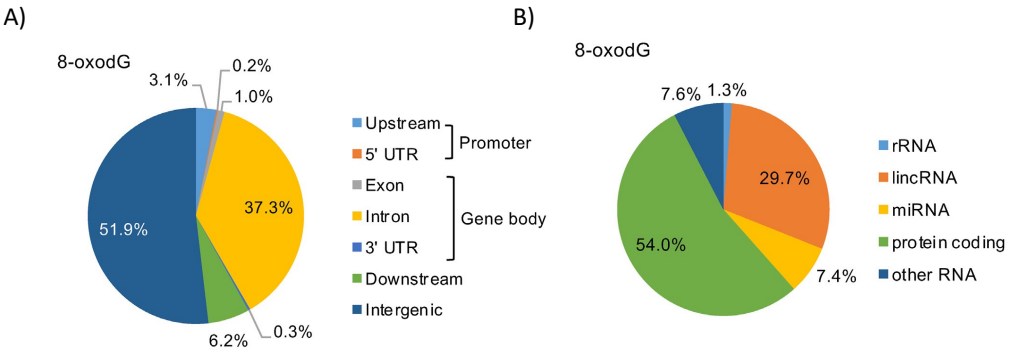
4.1 8-oxodG maps within gene loci and co-localizes with OGG1 and PARP1 at the promoter regions of human genes

In our previous study (Amente 2019), we carried out OxiDIP-Seq in *vitro* immortalized non-tumorigenic human epithelial breast cells (MFC10A) in order to profile the genome-wide distribution of 8-oxodG. The OxiDIP-Seq data were performed in growing cells in two biological replicates. The genomic DNA was first extracted from asynchronous MCF10A cells, then was fragmented by sonication and denatured. After that, the DNA fragments were immunoprecipitated using efficient anti–8-oxodG antibodies and the immunoprecipitated samples and input DNA were sequenced through next generation sequencing. The method combines the high-throughput sequencing to DNA immunoprecipitation for genome-wide profiling of 8-oxodG

The obtained sequence tags were aligned to the human genome and a peak calling analysis was performed by using MACS (Model-based Analysis for ChIP-Seq) algorithm. MACS identifies regions in the genome that have been enriched with aligned reads in the OxiDIP sample relative to the genome background. In this way, it detects areas enriched in 8-oxodG residues deriving from sequenced and aligned immunoprecipitated fragments. Specifically, this analysis identified *52 298* genomic regions enriched in 8-oxodG.

Next, in order to determine whether the genomic distribution of 8-oxodG peaks was correlated to specific regions of the human genome, we annotated peaks by using Peak Annotation and Visualization (PAVIS) tool. First, PAVIS analysis displays the number of annotated peaks and relative enrichment level in each genomic feature category including transcription start site, intron, exon or 5′/3′-untranslated region. Then, PAVIS analysis

provides a summary plot of the relative proportion of peaks in each category reporting the relative enrichment of peaks in these functionally distinct categories.



C)

Genetic element	8-oxodG enrichment
Gene loci	p<2.2e-16
Promoter	p<2.2e-16
Gene body	p<2.2e-16
Protein coding gene	p<2.2e-16
lncRNA	p<2.2e-16
miRNA	p=5.8e-3
intergenic	p=1

Figure 11: Pie charts show the annotated 8-oxodG peaks based on their distribution in the context of genomic features (A) and in the context of gene types (B). The genomic features are divided in genic regions with Upstream (-5kb), 5'UTR, Exons; Introns; 3'UTR; Downstream (+5kb) and intergenic regions (A). The gene types are classified in ribosomal RNA (rRNA), long intergenic non-coding RNA (lincRNA), microRNA (miRNA), mRNA (protein coding), other RNA (B). (C) Table shows p-value of enrichment analysis (Bedtools Fisher test) of the 8-oxodG peaks within the regions indicated in the pie charts (A,B).

Thus, we analyzed the genomic distribution of 8-oxodG peaks in the context of genomic features and we found that 42 % of 8-oxodG peaks localize within the gene loci (i.e., promoter and gene body). Furthermore, relative peak enrichment was determined with Fisher test of Bedtools suite, revealing

that 8-oxodG peaks were enriched within both gene body and promoter regions P < 2.2e–16 (Figure 11 A,C).

In conclusion, PAVIS analysis allows us also to annotate 8-oxodG distribution among different types of genes. Interestingly, we found that 54% and 30% of the 8-oxodG peaks were enriched within protein-coding and long non-coding genes, respectively (*P* < 2.2e–16; Figure 11B,C).

These data reveal that 8-oxodG is preferentially mapped in gene-related regions and particularly, that the protein-coding genes were the most oxidized. It has known that 8-oxodG has a role in the epigenetic regulation of transcription. In order to further study the association between 8-oxodG and gene transcription, in this work we will analyze the 8-oxodG signals within promoter regions.

Indeed, the 8-oxodG enrichment within promoter regions prompt us to identify 8-oxodG-positive promoters. To this purpose, using Bedtools suite the 8-oxodG high-confidence peaks (n = 52 298) were intersected with promoters of RefSeq human genes (n = 21 074). We defined the promoter regions as the 1 kb regions flanking the TSSs of the above selected human genes. We identified *1456* genes containing at least one 8-oxodG high-confidence peak in their promoter regions. These promoters were hereafter called "oxidized promoters".

Then, we performed a metagene analysis of 8-oxodG signals along the genes associated with oxidized promoters in growing cells.

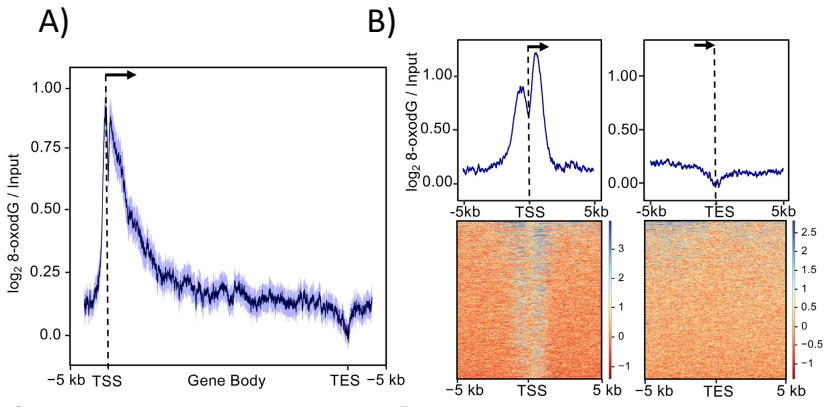


Figure 12: (A) 8-oxodG distribution over the corresponding gene loci (from –5 kb upstream the transcription start site, TSS, to +5 kb from transcription end site, TES) in growing MCF10A cells. The 95% confidence interval (2 standard error) of the mean is indicated by the light blue shaded area. (B) Mean-density profile (top) and heatmap (bottom) of the normalized 8-oxodG signal ±5 kb from TSS (left) or ±5 kb from TES (right) of genes with oxidized promoters. The arrows indicate the direction of transcription.

The analysis of 8-oxodG signals, over the corresponding gene loci (n = 1456), was performed from 5 kb upstream of the Transcription Start Site (TSS) to 5 kb downstream of the Transcription End Site (TES). It confirmed the enrichment of 8-oxodG within the promoter region and revealed that 8-oxodG is enriched also within the gene body (Figure 12A).

Focusing on the ±5 kb region flanking the TSS and ±5 kb region flanking the TES of the same genes (n = 1456), the analysis showed a promoter-specific bimodal distribution of 8-oxodG peaking at about +600 and – 600 bp from the TSS (Figure 12 B).

We then investigated the presence of OGG1, which specifically recognizes and excises 8-oxodG. We analyzed the OGG1 signal over the genes marked by oxidized promoters (n = 1456), using a publicly available ChIP-Seq dataset.

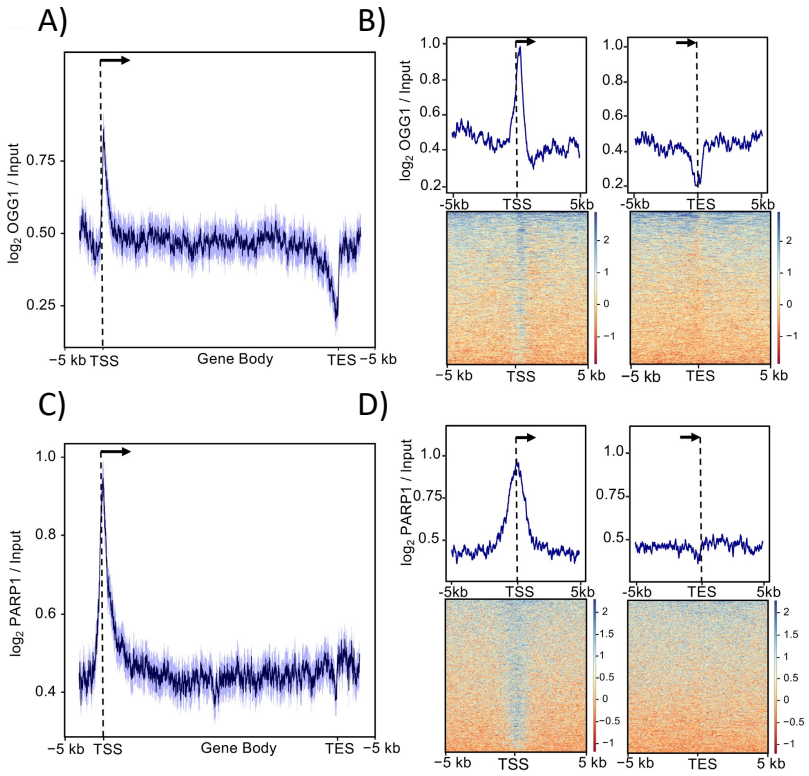


Figure 13: OGG1 (A) and PARP1 distribution (C) over the corresponding gene loci (from –5 kb upstream the transcription start site, TSS, to +5 kb from transcription end site, TES) in growing MCF10A cells. The 95% confidence interval (2 standard error) of the mean is indicated by the light blue shaded area. Mean-density profile (top) and heatmap (bottom) of the normalized 8-oxodG signal ±5 kb from TSS (left) or ±5 kb from TES (right) of genes with oxidized promoters. The arrows indicate the direction of transcription. (B,D).

First, we compared the OGG1 ChIP signal with the 8-oxodG high-confidence peaks (n = 52 298) and we founded a high correlation using Pearson correlation test (r = 0.80, P = 2.2e–16).

Consistently, the metagene analysis of OGG1 over the oxidized promoters (n = 1456) and their respective flanking regions showed a peak at the promoter that is slightly shifted downstream of the TSS (Figure 13 A,B).

In conclusion, because PARP1 is considered a modulator of BER pathway and a molecular nick-sensor in the repair induced by oxidatively-generated

DNA damage (Reynolds 2015), we performed a metagene analysis of the distribution of PARP1 signal over the oxidized promoters (n = 1456) and their flanking regions. The distribution of PARP1 signal over the genes with oxidized promoters (n = 1456) showed a clear peak at their TSS (Figure 13 C,D).

All together, these data show that 8-oxodG accumulation at gene promoters associates with OGG1 and PARP1 recruitment.

4.2 The 8-oxodG distribution at promoter regions correlates with GC content

In order to determine the relationship between the GC content and 8-oxodG accumulation at promoter regions, we quantified the percentage of GC content at oxidized and control promoters. The subsets of oxidized and control promoters were properly chosen. In particular, as previously reported, the oxidized promoters are the 1456 genes that contains at least one 8-oxodG high-confidence peak in their promoter regions. Instead, the control promoters are defined as genes with negligible levels of 8-oxodG at promoter regions, and they were specifically selected as described in Materials and methods.

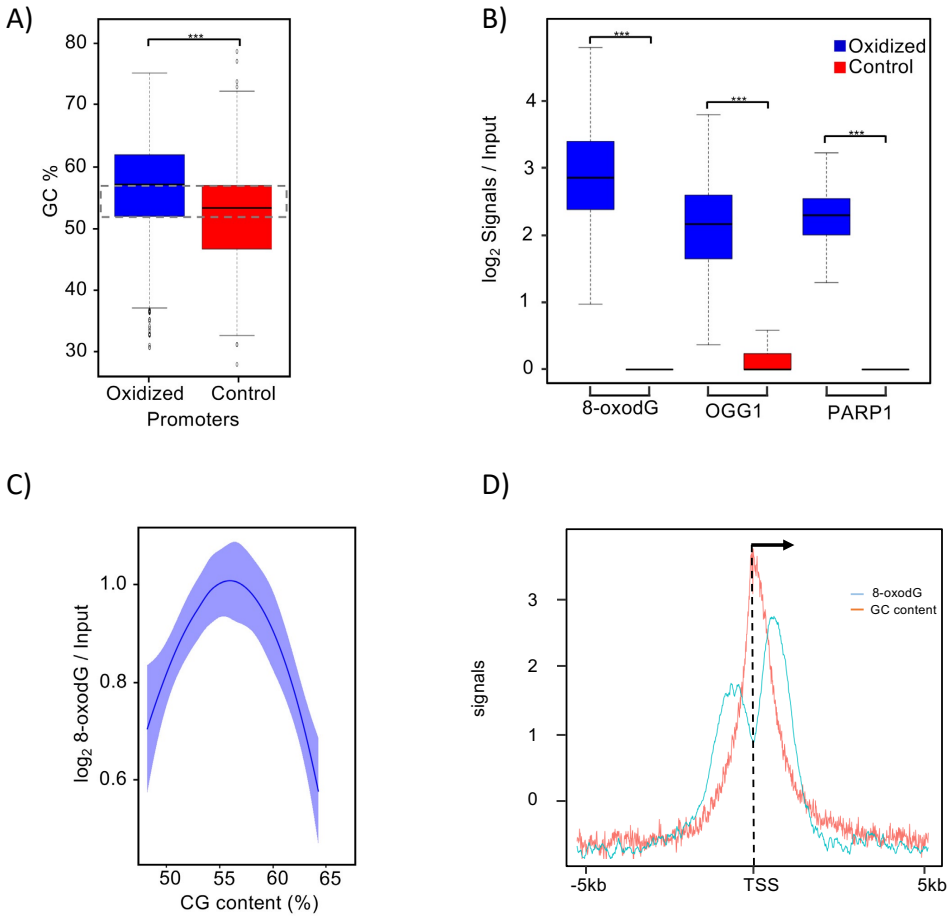


Figure 14: A) The GC content (%) distribution in box plot was measured at oxidized (blue) and control (red) promoters (Bonferroni adjusted pairwise t-test; ***P = 2.2e–16). The regions of the box plots where the oxidized and control promoters have comparable GC content (52–57%) are indicated with the dashed rectangle. (B) The distribution of the normalized signal of 8-oxodG, OGG1 and PARP1 measured at oxidized (blue) and control (red) promoters with comparable GC content (52–57%) was represented in box plot (Bonferroni adjusted pairwise t-test; ***P < 2.2e–16). The dependencies between 8-oxodG and GC content at oxidized promoters was visualized at 50 bp resolution(C) and at 10bp resolution (D). The normalized 8-oxodG signal and average GC content (color code, as indicated) ±5 kb from TSS of genes with oxidized promoters was showed in the mean-density profile (D).

In the box plot represented in Figure 14A, we can observe the GC content distribution in control and oxidized promoters. We demonstrated that the GC content at control promoters is marginal reduced comparing with oxidized ones. This result shows a high statistical significance (P = 2.2e−16). Then, we selected the subsets of oxidized and control promoters that have been shown the comparable GC content (52–57%) and we indicated them with a dashed rectangle as reported in Figure 14A. In these subsets of promoters, we analyzed the 8-oxodG, OGG1 and PARP1 signals. Consistently, the box plots in Figure 14 B reveal that only the oxidized promoters show a specific enrichment of 8-oxodG, OGG1 and PARP1 signals normalized on their input.

We analyzed the GC content and 8-oxodG signal at the oxidized promoters at two resolution levels. First, we divided the corresponding regions in bins of 50 bp (Figure 14C) and then calculated the average GC content per bin. The 8-oxodG levels increase as the GC content arises from 50 to 56% and then sharply drop down when GC content arises to 65% (Figure 14 C).

Furthermore, GC content and 8-oxodG signal at the oxidized promoters were analyzed with a higher resolution (10 bp windows size). The 8-oxodG and GC content signals showed different profiles at TSS (Figure 14 D). The GC content (in red in Figure 14D) showed a peak at TSS; the 8-oxodG enrichment (in blue in Figure 14D) dropped down at TSS and generated two peaks. In particular, the 8-oxodG peak downstream of the TSS showed the strongest signal.

Overall, these findings demonstrate that the 8-oxodG distribution at gene promoters does not depend only by GC content supporting the hypothesis that the 8-oxodG accumulation in the TSS region could be linked also to the epigenetic mechanism.

4.3 The association between 8-oxodG positive promoters and transcription process

In order to assess whether 8-oxodG accumulation is associated with gene transcription, we decided to compare 8-oxodG signals to either ChIP-Seq signals of Ser5- and Ser2-phosphorylated isoforms of RNA Polymerase II (Pol II-S5P and Pol II-S2P), or gene transcription measured by global run-on sequencing GRO-Seq. We chose to analyze both Pol II-S5P and Pol II-S2P because they mark respectively initiating and elongating RNA Polymerase II.

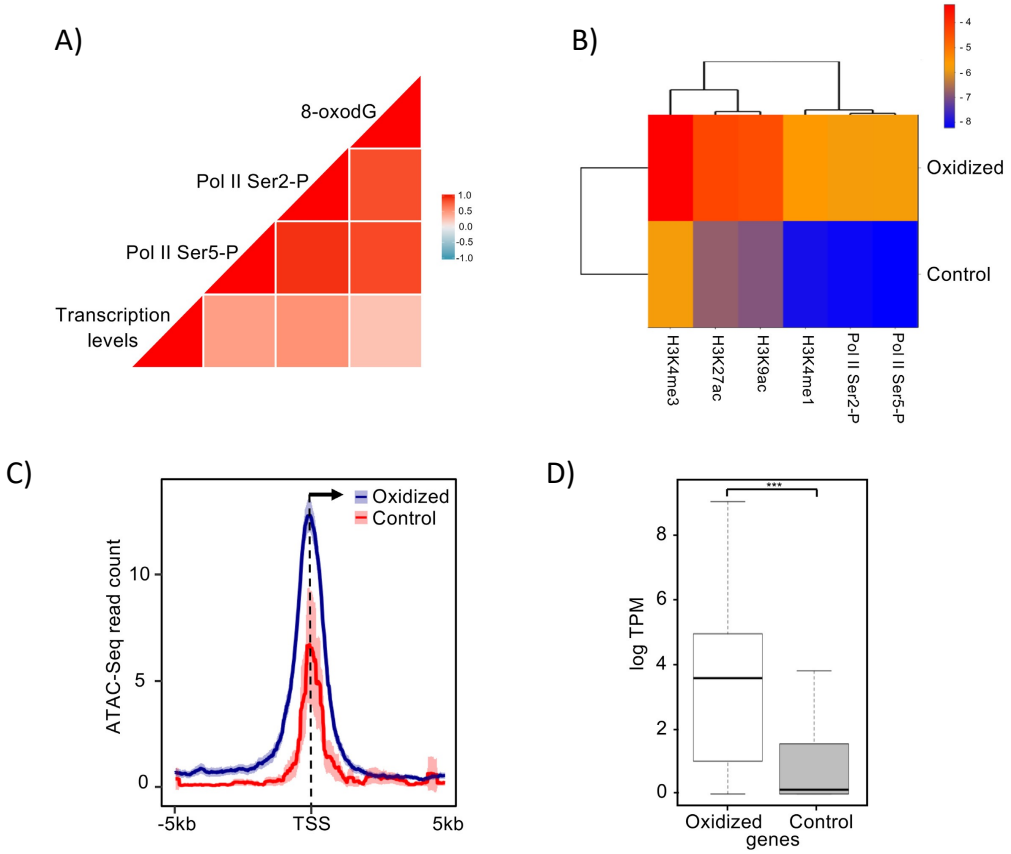


Figure 15: (A) The Pearson correlation coefficients between transcription (GRO-Seq data), Pol II-S5P and Pol II-S2P (ChIP-Seq signals) and 8-oxodG (Oxi-DIP-Seq signal), at genes with oxidized promoters were distributed in an heatmap. (B) The H3K4me3, H3K27ac, H3K9ac, H3K4me1, Pol II-S5P and Pol II-S2P (log2 signal/Input DNA) levels at the oxidized and control promoters were reported in

the heatmap. (C) The normalized ATAC-Seq signal ±5 kb from the TSS of oxidized (blue) and control (red) promoters were visualized in a mean-density profile.The 95% confidence interval (2 standard error) of the mean is indicated by the light blue and light red shaded areas. The arrow indicates the direction of transcription. (D) The Box plot showed the distribution of genes with oxidized (white) and control (grey) promoters associated with transcription levels measured by GRO-Seq (logTPM) (Bonferroni adjusted pairwise t-test; ***P < 2.2e–16).

We performed a Pearson correlation analysis to evaluate the relationship between 8-oxodG accumulation and the transcriptional activity. The correlation coefficients were distributed into the heatmap (Figure 15A) that showed a high-to-moderate correlation between 8-oxodG levels and Pol II-S5P and S2P (r = 0.9) and also transcription levels (r = 0.3).

Moreover, it has been suggested that 8-oxodG correlates with open chromatin, thus we compared both the control and oxidized promoters with open chromatin markers (H3K4me3, H3K27ac and H3K4ac), in a new heatmap (Figure 15B). Coherently with previous studies (Poetsch 2018), we found that the oxidized promoters are associated with the open chromatin markers (Figure 15B).

In order to further determine the chromatin accessibility at the oxidized promoters I analyzed the ATAC-Seq signals (Liu 2017). Consistently, the oxidized promoters showed a higher chromatin accessibility (Figure 15C) and higher levels of Pol II-S5P and S2P (Figure 15B) compared to the control promoters.

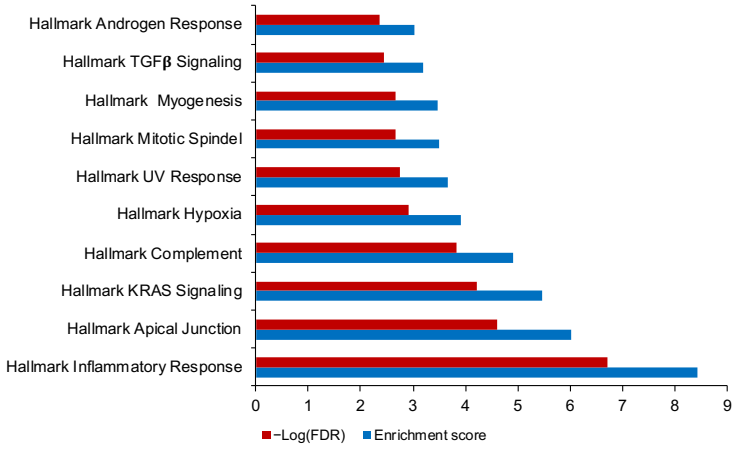
Moreover, the GRO-Seq analysis revealed that the genes associated with oxidized promoters show higher transcription levels than control ones measured as log of Transcript per Million (median transcription levels of 3.6 and 0.1 logTPM, respectively, P < 2.2e– 16; Figure 15D).

Collectively, these data suggest that the 8-oxodG accumulation in the genes associated with oxidized promoters correlates with transcription process at genome-wide level.

4.4 The oxidized promoters are involved in inflammatory response

In order to investigate the pathways involving the genes associated with oxidized promoters we performed a gene set enrichment analysis (GSEA). First, we evaluated the overlaps between the gene associated with oxidized promoters and the gene sets in Hallmarks collection available in Molecular Signatures Database (MSigDB). The enrichment score value of genes associated with the corresponding pathway was expressed as -log10 of *p*-value (determined by Fisher's exact test). The false discovery rate (FDR) was used to measure the statistical significance of the analysis.

A)



B)

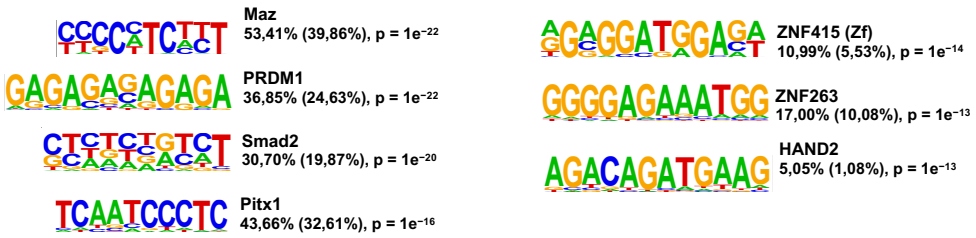


Figure 16: A) The Top ten pathway associated with the oxidized promoter with the highest enrichment score. The p-value of the corresponding pathway is determined by Fisher's exact test and it is used to calculate the enrichment score value

(expressed as -log10(pvalue). The statistical significance of the association between the oxidized promoter and the corresponding pathways is expressed as -log of the False Discovery Rate (FDR). B) HOMER analysis of putative motif within the oxidized promoters.

The Gene Ontology analysis revealed that the gene sets which were enriched in the oxidized promoters included the gene related to the inflammatory response. In addition, the Hallmark of Apical Junction and KRAS signaling pathway were highly represented.

In order to confirm these results, we performed a promoter-based motif finding analysis using HOMER Motif Analysis. HOMER finds the motifs that are enriched in the genes associated with oxidized promoters relative to other human promoters. The oxidized promoters showed significant enrichment for the motifs of MAZ and PRDM1 ($P = 1e^{-22}$). In particular, Myc-associated zinc finger (MAZ) is a transcription factor highly upregulated in chronic inflammatory disease (Triner 2018). Instead, PR domain zinc finger protein 1 (PRDM1) also known as B lymphocyte-induced maturation protein-1 (BLIMP-1) plays a significant role in B cell development and antibody production and acts as a repressor of beta-interferon (β-IFN) gene expression (Di Zazzo 2013).

Overall, these findings suggest that the genes associated with oxidized promoters are related with the inflammatory response pathway.

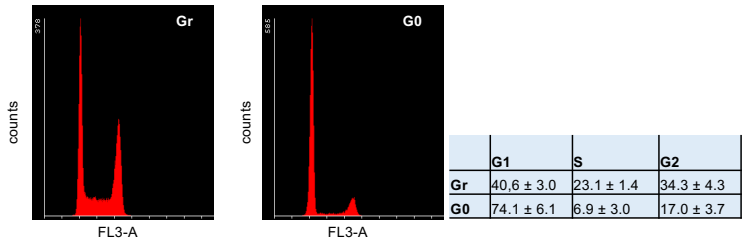
4.5 The accumulation of 8-oxodG at promoters is associated with DNA replication and/or transcription

It has been demonstrated that the TSS of transcribed genes is associated with the recruitment of the Origin Recognition Complex (ORC1) and with the firing of DNA replication Origins (ORIs) (Chen 2019; Dellino 2013).

These findings prompted us to wonder whether the 8-oxodG accumulation at the promoter regions is due to the DNA replication process, alone or in combination with transcription.

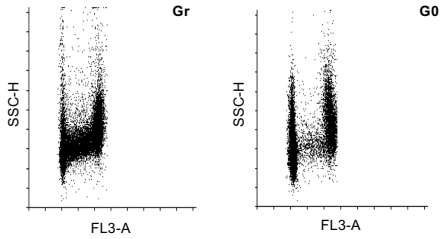
In order to study the association between 8-oxodG accumulation and DNA replication process, the MCF10A cells were first deprived by growth factors. We performed a FACS analysis (Figure 17A) and used a Ki67 as proliferation marker (Figure 17B).

A)



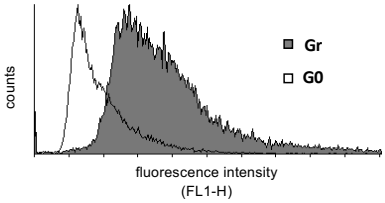
	G1	S	G2
Gr	40,6 ± 3.0	23.1 ± 1.4	34.3 ± 4.3
G0	74.1 ± 6.1	6.9 ± 3.0	17.0 ± 3.7

B)



	G1	S	G2
Gr	41.7 ± 1.3	19.3 ± 1.8	40.0 ± 1.2
G0	67.0 ± 2.6	5.0 ± 1.6	28.1 ± 1.3

C)



Gr	387.4 ± 39.1
G0	194.4 ± 19.1

Figure 17: Flow cytofluorimetry analysis of DNA content (propidium iodide stained, panel A) and of Ki-67 (immuno-stained, panel B) in both asynchronously growing and G1- arrested (G0) MCF10A cells. (C) Quantification of 8-oxodG at genomic level of both growing or G0 MCF10A cells through flow cytofluorimetry analysis of immuno-stained 8-oxodG cells. Plots represented the results of an experiment whereas data in the tables reported the results of independent experiments with SD (n = 2).

Thus, we confirmed that MCF10A were arrested in G1 phase (G0 =74,1 ± 6,1 panel A; G0 = 67,0 ± 2,6 panel B).

Then, the 8-oxodG global levels were measured in cycling and quiescent (G0) MCF10A cells and it was demonstrated a 2-fold reduction of 8-oxodG levels in quiescent than growing cells (Figure 17C).

Once we confirmed that MCF10A cells were arrested in G1, we performed OxiDIP-Seq in the quiescent cells in two biological replicates (Pearson correlation test between two biological replicates , r=0.98). We identified *23 641* genomic regions enriched in 8-oxodG. After that, using Bedtools suite the 8-oxodG high-confidence peaks (n = 23 641) were intersected with promoter of RefSeq human genes (n = 21 074) in order to identify the oxidized promoters in quiescent cells (G0). We obtained *811* genes that contain 832 8-oxodG high-confidence peaks in their promoter regions.

As previously described for growing MCF10A cells, we performed a metagene analysis of 8-oxodG signals on the oxidized promoters in quiescent cells.

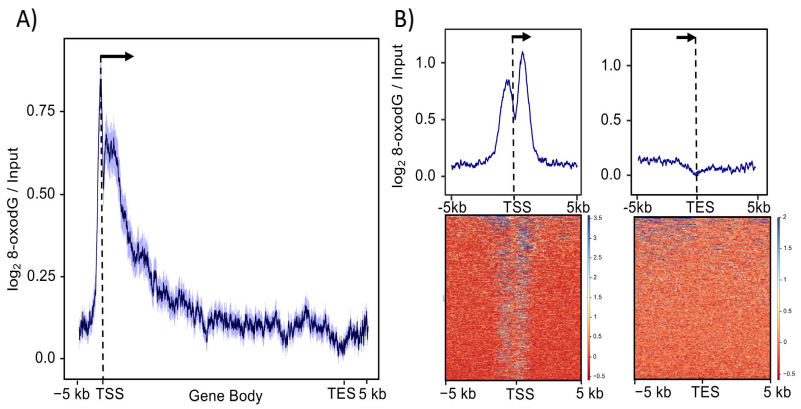


Figure 18: (A) 8-oxodG distribution over the corresponding gene loci (from □ 5 kb upstream the transcription start site, TSS, to +5 kb from transcription end site, TES) in quiescent (G0) MCF10A cells. The 95% confidence interval (2 standard error) of the mean is indicated by the light blue shaded area. (B) Mean-density profile (top) and heatmap (bottom) of the normalized 8-oxodG signal ±5 kb from TSS (left) or ±5 kb from TES (right) of genes with oxidized promoters. The arrows indicate the direction of transcription.

The analysis of 8-oxodG signals, over the corresponding gene loci (n = 811) confirmed the presence of DNA oxidation specifically within the promoter regions (Figure 18A). Furthermore, 8-oxodG signal at promoters showed the same bimodal distribution observed in cycling MCF10A cells (Figure 18B), with only a slight reduction of 8-oxodG signal.

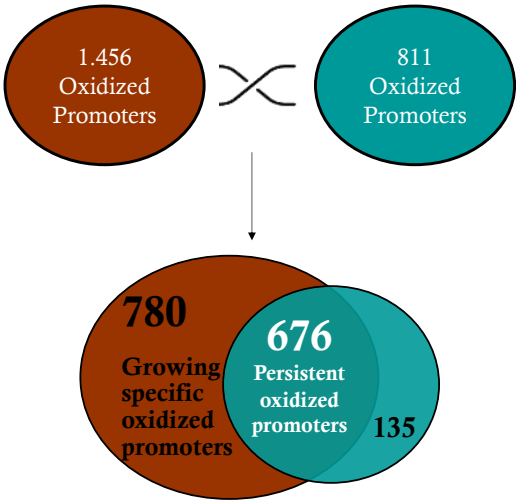


Figure 19: Subsets of oxidized promoters detected through intersection analyses.

In order to determine a subset of promoters that showed persistent 8-oxodG signal upon growth factors withdrawal, we intersected the *1456* oxidized promoters, identified in growing condition, with the *811* oxidized promoters, identified in quiescent cells. Interestingly, we identified *676* promoters that showed persistent 8-oxodG signal upon growth factors withdrawal (oxidized promoters 'common' to growing and quiescent cells; Figure 19). These promoters were hereafter called "persistently oxidized promoters".
Instead, the 54% of the oxidized promoters identified at steady-state in the growing cells were lost in the G0 cells (n = 780/1456; hereafter called 'growing-specific' oxidized promoters).

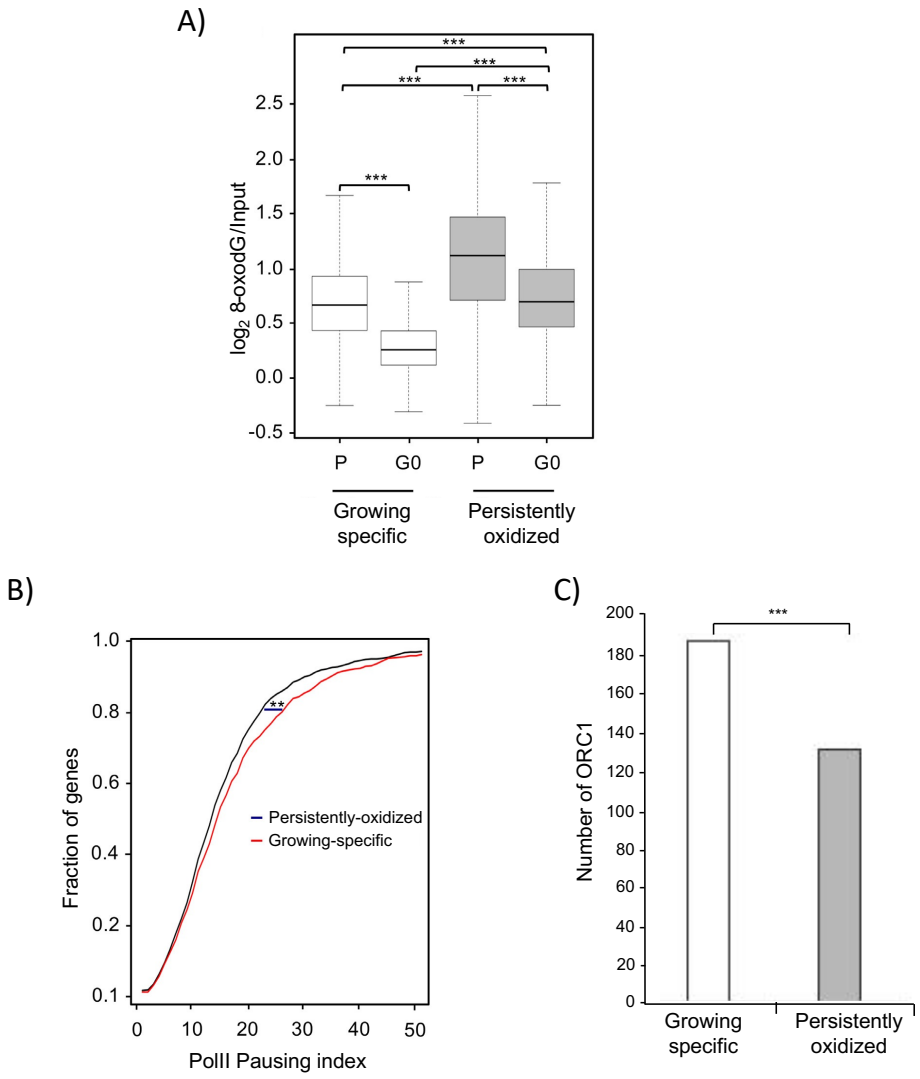


Figure 20: (A) The normalized 8-oxodG signal measured at promoter regions of the growing-specific (white) and persistently (grey) oxidized promoters in proliferating (P) and quiescent (G0) MCF10A cells are visualized in a box plot (Bonferroni adjusted pairwise t-test; ***P < 2.2e-16) (B) Pol II pausing index (calculated using the Pol II-S2P ChIP-Seq data) was determined based on the fraction of genes with growing-specific (red) or persistently (black) oxidized promoters in the proliferating MCF10A cells (Kruskal–Wallis test; **P = 2.5e-2). (C) Number of ORC1 were plotted in both growing-specific (white) and persistently (grey) oxidized promoters (***P < 3.5e-2).

Then, we compared the 8-oxodG signal levels in growing and quiescent condition in both subsets of promoters (growing-specific and persistently oxidized). Interestingly, the 8-oxodG levels in growing condition at the persistently promoters were higher than those observed at the growing-specific ones (Figure 19A; compare P of persistently versus P of growing-specific). Furthermore, the persistently oxidized promoters showed smaller 8-oxodG signal drop upon growth arrest compared with the growing specific oxidized promoters (Figure 19A). Altogether, these data suggest that 8-oxodG accumulation at promoters showing persistent 8-oxodG signal in G0 cells is mainly associated with replication-independent transcription-associated events.

Furthermore, the RNA:DNA hybrids can form and/or be stabilized as a consequence of physiological Pol II pausing (Skourti-Stathaki 2011), or following collisions between transcription-replication machineries (Helmrich 2011), at promoter regions of human genes.

Thus, we calculated the Pol II pausing index of the genes associated with oxidized promoters as the ratio of promoter to gene body signals of Pol II-S2P. Interestingly, the Pol II pausing index measured at the genes associated with the growing-specific oxidized promoters was higher than the genes with persistent oxidatively-damaged promoters (Figure 19B; P = 2.5e– 2). These findings suggest that paused Pol II contributes to the 8-oxodG levels observed in these loci. In conclusion we measured the occurrence of DNA replication origins at both growing-specific and persistently oxidized promoters. As shown in Figure 19C the ORC1 were enriched at growing specific promoters compared to persistently (186 versus 130, P = 3.5e– 2). Thus, frequent transcription-replication clashes might contribute to the oxidation levels observed at the growing-specific oxidized promoters.

Collectively, these data led us to identify two promoter classes: the "growing-specific" and "persistently-oxidized" promoters. The 8-oxodG accumulation in these two subsets of promoters is due to either DNA

replication-dependent or replication-independent and transcription-associated events.

4.6 Spontaneous accumulation of 8-oxodG at persistently oxidized promoters is associated with transcription process

In order to demonstrate that the transcription *per se* has a role in the accumulation of 8-oxodG in the persistently oxidized promoters, we performed a Pol II-S2P ChIP-Seq under both growing and quiescent growth conditions. First, we correlated the 8-oxodG signals with Pol II-S2P signals obtained respectively from growing and G1-arrested MCF10A cells.

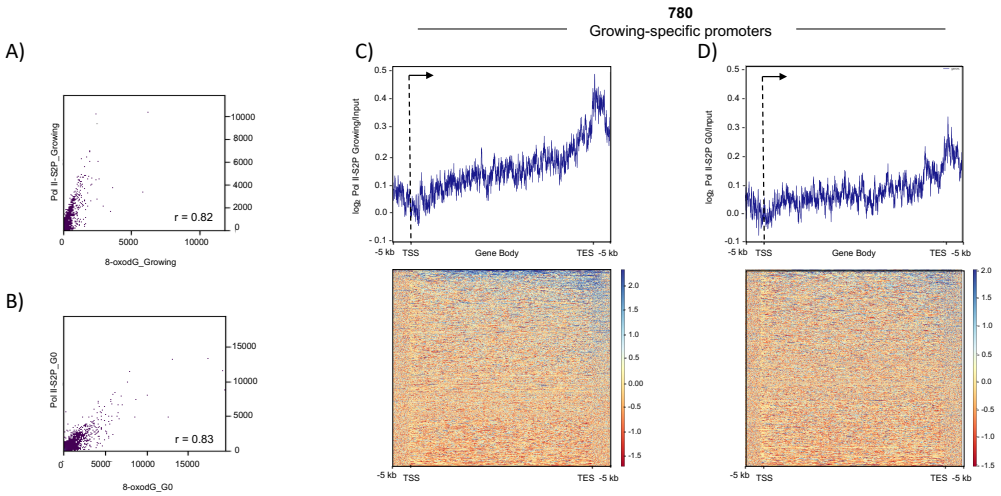


Figure 21: Scatter plot showing the correlation between Pol II-S2P ChIP signals and 8-oxodG signals in growing (panel A) and quiescent (G0; panel B) MCF10A cells, reported as Pearson's correlation coefficient (r = 0,82 panel A; r = 0,83 panel B). (C,D) Mean-density profile (top) and heatmap (bottom) of Pol II S2P occupancy from -5 kb upstream TSS to -5 kb downstream TES in growing-specific oxidized promoters (n =780) in growing (panel C) and quiescent (G0; panel D) condition. The arrows indicate the direction of transcription.

The results showed a high correlation among oxidation levels and Pol II-S2P occupancy, expressed as Pearson coefficient, in both growing (r = 0,82; Figure 21A) and quiescent (r = 0,83; Figure 21B) samples. To further investigate the role of transcription in the specific genes associated with oxidized promoters we performed a metagene analysis of Pol II-S2P signals on both subsets of genes associated with oxidized promoters (780 growing-specific; 676 persistently oxidized). As observed in Figure 21C, in growing condition the Pol II-S2P signals showed high levels along the gene body of the 780 genes associated with oxidized promoters. Interestingly, in quiescent condition the Pol II-S2P occupancy is still considerable, even if there is a reduction of the signal (Figure 21D). These results are consistent with the idea that growing specific promoters accumulate 8-oxodG through DNA replication-dependent event.

Instead, when we performed the same metagene analysis of Pol II-S2P signals obtained under growing and quiescent growth conditions on genes associated with persistently oxidized promoters (676 genes) the results were different.

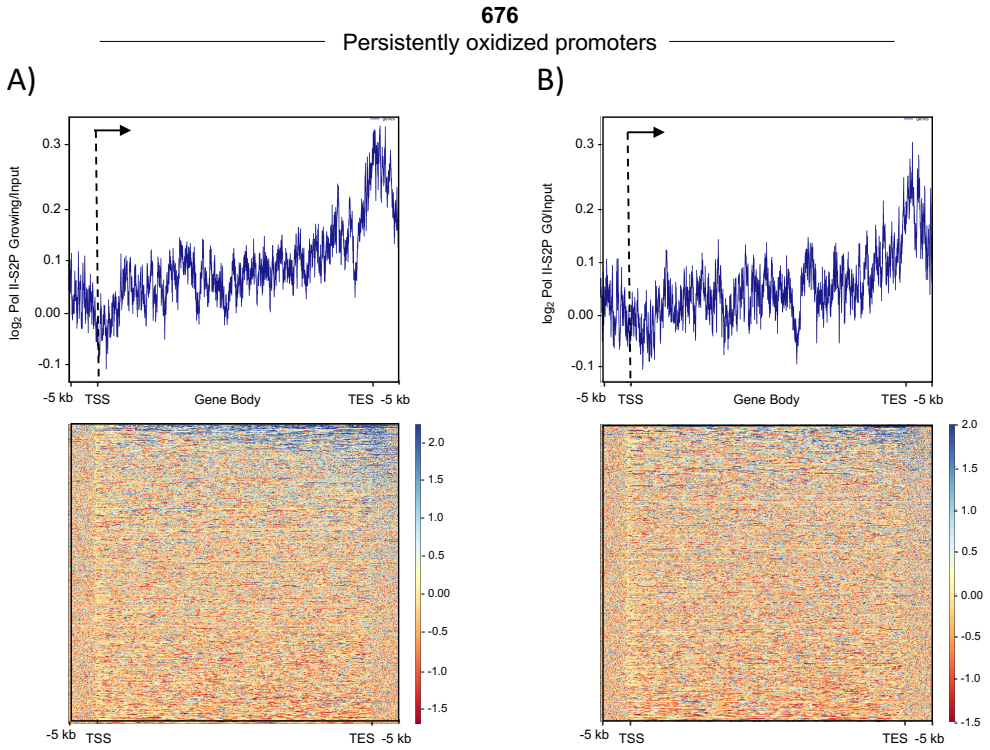


Figure 22: Mean-density profile (top) and heatmap (bottom) of Pol II S2P occupancy from -5 kb upstream TSS to -5 kb downstream TES in persistently oxidized promoters (n = 676) under growing (panel A) and quiescent (G0; panel B) cell conditions. The arrows indicate the direction of transcription.

Indeed, as we observed in Figure 22, the Pol II-S2P signals along the gene body of the genes associated with persistently oxidized promoters (n = 676) showed the same profile in both growing and G0 MCF10A cells conditions. These results are consistent with the idea that the persistently oxidized promoters accumulate 8-oxodG through replication-independent transcription-associated events.

Overall, these data strongly confirm a crucial role of the transcription process for the accumulation of the oxidative DNA damage, even if these data alone cannot establish which is causality between the two mechanisms.

V. Discussion

In the present study, we applied an innovative methodology (OxiDIP-Seq), which combines the high-throughput next-generation sequencing to 8-oxodG immunoprecipitation, for the genome-wide mapping of DNA oxidation. In diploid human mammary-epithelial cells (MCF10A) in unperturbed conditions, we found ~52 000 genomic regions enriched in 8-oxodG. The 42% of these regions showed an accumulation of oxidation in gene-related regions (i.e., promoters and gene body) and among these, in protein-coding genes (Amente 2019). The 8-oxodG enrichment within gene-related regions prompted us to identify human promoters containing 8-oxodG peaks.

The oxidized promoters, identified in cycling MCF10A cells, showed two 8-oxodG sharp peaks, immediately upstream and downstream of the TSS, defining a promoter-specific bimodal distribution of 8-oxodG which seems to suggest the presence of bidirectional transcription.

Because the TSS is a GC-rich region we wondered if the 8-oxodG accumulation was correlated with a higher GC content (Ginno 2013; Kellner 2015). Indeed, we investigated the relationship between 8-oxodG accumulation and GC content at gene promoters. It has been found that the percentage of GC content measured at oxidized and control promoters was very similar. This analysis suggested that the GC content could have only a marginal influence in the accumulation of 8-oxodG at promoter regions. However, analysis at a higher resolution level determined a rather intricate relationship. Indeed, the promoter regions with high GC content showed a decrease of 8-oxodG levels while promoter regions with similar GC content showed different 8-oxodG levels. Overall, these findings demonstrated that the accumulation of oxidatively-modified nucleobases within promoter

regions does not only depend on GC content but could be linked also to the epigenetic mechanism.

Moreover, we observed that OGG1, one of the major DNA glycosylases/AP lyases involved in the BER pathway (Wang 2018), was recruited to the oxidatively-damaged promoters. Also PARP1, known as a negative modulator of the BER pathway and molecular nick-sensor in the repair induced by oxidatively-generated DNA damage (Hegde 2012; Morales 2014; Reynolds 2015), was recruited to the oxidized promoters. The OGG1 and PARP1 recruitment, as BER intermediates, at oxidized promoters suggested that the cells can sense and repair these 8-oxodGs. Moreover, there is various evidence on the coupling of BER and transcriptional activation which leads to the hypothesis that 8-oxodG accumulation may be an epigenetic modification that regulates gene expression.

Previous works showed that the accumulation of 8-oxodG in the promoters of specific genes can stimulate transcription through the BER pathway. In particular, *Perillo et al.* reported how DNA breaks, generated upon processing of 8oxodG residues, were necessary for transcription activation of estrogen-responsive genes (Perillo 2008); similarly, *Amente et al.* demonstrated that the repair of the 8-oxodG drives the transcription of Myc target genes (Amente 2010).

Contrariwise, we use a genome-wide approach to study the association between oxidative DNA damage and transcription globally, no longer restricted to a specific subset of genes.

Comparing 8-oxodG signals to either ChIP-seq signals of Pol II-S5P and Pol II-S2P or GRO-seq signals, we found a high-to-moderate correlation between 8-oxodG accumulation and gene transcription. The high chromatin accessibility at the oxidized promoters further confirmed the strong association between the oxidatively DNA damage and the transcription process.

Among the genes associated with the oxidized promoters, we found that they were essentially enriched for MAZ and PRDM1 motif. The both of aforementioned factors are involved in the inflammatory process; in particular, the Myc-associated zinc finger (MAZ) is a transcription factor highly upregulated in chronic inflammatory disease (Triner 2018). Instead, PR domain zinc finger protein 1 (PRDM1) also known as B lymphocyte-induced maturation protein-1 (BLIMP-1) plays a significant role in B cell development and antibody production and acts as a repressor of beta-interferon (β-IFN) gene expression (Di Zazzo 2013). The Gene Ontology analysis further confirms that the gene sets which were principally enriched in the oxidized promoters included the gene related to the inflammatory response. Different studies support an interdependent relationship between inflammation and oxidative stress involving the activation of specific transcription factors such as NF- κB (Biswas 2016; Castellani 2014; Mittal 2014). Our work demonstrated at a genome-wide level the link between inflammation and 8-oxodG accumulation.

Finally, mapping of 8-oxodG-enriched regions in G0-arrested cells determined the identification of two subsets of oxidatively-damaged promoters classified in "*growing-specific"* and "*persistently"* oxidized promoters. In particular, we demonstrated that growing-specific oxidized promoters showed a strong 8-oxodG signal reduction, when the DNA replication is absent. Furthermore, the RNAPII Ser2-P occupancy at growing-specific oxidized promoters showed a reduced signal intensity in the quiescent condition compared with growing condition. These results are consistent with the idea that growing-specific promoters accumulate 8-oxodG through DNA replication-dependent events. Coherently, we also found high levels of physiological Pol II pausing and ORC1 binding in a significant fraction of the growing-specific oxidized promoters, thus suggesting an increased probability of DNA replication-transcription clashes.

By contrast, persistently oxidized promoters showed the highest 8-oxodG signals at the steady-state, underwent milder signal reduction. Interestingly the RNAPII Ser2-P occupancy at persistently oxidized promoters showed the same signal intensity in the quiescent condition compared with the growing condition.

Moreover, the expression levels of the vast majority of genes associated with both promoter classes did not show significant changes upon growth arrest and this further supports an epigenetic role of the 8-oxodGs in gene expression processes (Ba 2014; Fleming 2017; Fleming and Burrows 2017; Li 2013; Olinski 2018; Zarakowska 2014). Thus, while transcription seems to be the main contributing factor to the 8-oxodG accumulation observed at the persistently oxidized promoters, DNA replication-associated events (alone, or in combination with transcription) are responsible for the 8-oxodG accumulation at the growing-specific oxidized promoters.

In conclusion, our findings add further evidence in support of the idea that the oxidatively-generated DNA damage at gene promoters could act in combination with the gene transcription through an epigenetic mechanism.

VI. Conclusions

Here, we report a study of the correlation between oxidative DNA damage and replication and/or transcription processes.

We provided OxiDIP-Seq methodology to precisely determine the 8-oxodG distribution in human genomes. By using a genome-wide approach, we found that 8-oxodG is preferentially mapped in gene-related regions and in particular the protein-coding genes were the most oxidized. The 8-oxodG enrichment within promoter regions suggested us to further investigate the correlation between DNA damage and transcription.

In this study, we identified 8-oxodG-positive promoters in proliferative human non-tumorigenic epithelial breast cells (MCF10A). The OxiDIP-Seq, and ChIP-seq of Ser5- and Ser2-phosphorylated isoforms of RNA Polymerase II (Pol II-S5P and Pol II-S2P) and GRO-Seq supported the evidence that 8-oxodG accumulation correlates with transcription processes. Moreover, the gene ontology analysis showed that the genes associated with oxidized promoters are involved in the inflammatory response.

The OxiDIP-Seq in quiescent (G0) cells identified the 8-oxodG-positive promoters in absence of DNA replication.

Thus, we classified the oxidized promoters into two subsets: the growing-specific and persistently oxidized promoters.

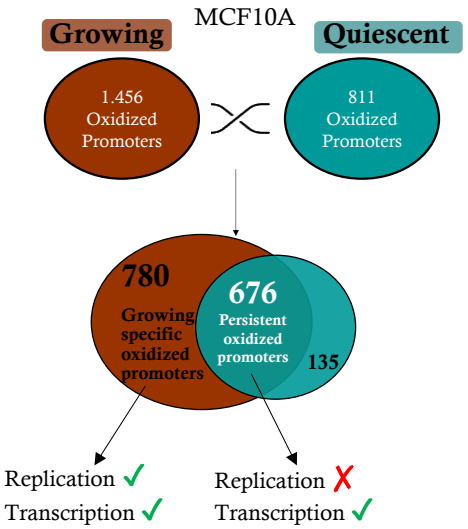


Figure 23: Schematic representation of oxidized promoters classified in the two subsets.

In conclusion the ChIP-seq of Pol II-S2P in both proliferative and quiescent (G0) cells allowed us to conclude that, while DNA replication (alone or in combination with transcription) influenced the 8-oxodG accumulation at growing-specific oxidized promoters, instead the transcription is responsible *per se* for the 8-oxodG accumulation at the persistently oxidized promoters. Finally, new studies are required to add new insights to a putative molecular mechanism through which the transcriptional apparatus could govern 8-oxodG accumulation at promoter regions.

VII. Acknowledgement

I would like to thank Dr Giovanni Scala for the bioinformatics and statistical analyses support.

VIII. List of Publications

1. Majello B, Gorini F, Saccà CD, Amente S. **Expanding the Role of the Histone Lysine-Specific Demethylase LSD1 in Cancer.** Cancers (Basel). 2019 Mar 7;11(3). pii: E324. doi: 10.3390/cancers11030324

2. Saccà CD, Gorini F, Ambrosio S, Amente, Majello B. **Targeting histone lysine-specific demethylase KDM1A/LSD1 to control epithelial-mesenchymal transition program in breast cancers.** J Cancer Metastasis Treat 2019; 5:15. 10.20517/2394-4722.2018.95.11 Mar 2019

3. Saccà CD, Gorini F, Ambrosio S, Amente S, Faicchia D, Matarese G, Lania L, Majello B. **Inhibition of lysine-specific demethylase LSD1 induces senescence in Glioblastoma cells through a HIF-1α-dependent pathway.** Biochim Biophys Acta Gene Regul Mech. 2019 May;1862(5):535-546. doi: 10.1016/j.bbagrm.2019.03.004. Epub 2019 Apr 2.

4. Amente S, Di Palo G, Scala G, Castrignanò T, Gorini F, Cocozza S, Moresano A, Pucci P, Ma B, Stepanov I, Lania L, Pelicci PG, Dellino GI, Majello B. **Genome-wide mapping of 8-oxo-7,8-dihydro-2'-deoxyguanosine reveals accumulation of oxidatively-generated damage at DNA replication origins within transcribed long genes of mammalian cells.** Nucleic Acids Res. 2019 Jan 10;47(1):221-236. doi: 10.1093/nar/gky1152.

5. Gorini F, Scala G, Di Palo G, Dellino GI, Cocozza S, Pelicci PG, Lania L, Majello B, Amente S. **The genomic landscape of 8-oxodG reveals enrichment at specific inherently fragile promoters.** Nucleic Acids Res.2020 March 2; https://doi.org/10.1093/nar/gkaa175

6. Gorini F, Scala G, Cooke M.S, Majello B, Amente S .**Towards a comprehensive view of 8-oxo-7,8- dihydro-2'-deoxyguanosine: highlighting the intertwined roles of DNA damage and epigenetics in genomic instability.** DNA repair 2020. November 26; https://doi.org/10.1016/j.dnarep.2020.103027

IX. References

Ambrosio S, Amente S, Napolitano G, Di Palo G, Lania L, Majello B. 2015. MYC impairs resolution of site-specific DNA double-strand breaks repair. *Mutat Res* **774**:6–13. doi:10.1016/j.mrfmmm.2015.02.005

Ambrosio S, Di Palo G, Napolitano G, Amente S, Dellino GI, Faretta M, Pelicci PG, Lania L, Majello B. 2016. Cell cycle-dependent resolution of DNA double-strand breaks. *Oncotarget* **7**:4949–4960. doi:10.18632/oncotarget.6644

Amente S., Bertoni A, Morano A, Lania L, Avvedimento EV, Majello B. 2010. LSD1-mediated demethylation of histone H3 lysine 4 triggers Myc-induced transcription. *Oncogene* **29**:3691–3702. doi:10.1038/onc.2010.120

Amente S, Di Palo G, Scala G, Castrignanò T, Gorini F, Cocozza S, Moresano A, Pucci P, Ma B, Stepanov I, Lania L, Pelicci PG, Dellino GI, Majello B. 2019. Genome-wide mapping of 8-oxo-7,8-dihydro-2'-deoxyguanosine reveals accumulation of oxidatively-generated damage at DNA replication origins within transcribed long genes of mammalian cells. *Nucleic Acids Res* **47**:221–236. doi:10.1093/nar/gky1152

Amente Stefano, Lania L, Avvedimento EV, Majello B. 2010. DNA oxidation drives Myc mediated transcription. *Cell Cycle* **9**:3002–3004. doi:10.4161/cc.9.15.12499

Amente S, Lania L, Majello B. 2013a. The histone LSD1 demethylase in stemness and cancer transcription programs. *Biochim Biophys Acta* **1829**:981–986. doi:10.1016/j.bbagrm.2013.05.002

Amente S, Lania L, Majello B. 2013b. The histone LSD1 demethylase in stemness and cancer transcription programs. *Biochim Biophys Acta* **1829**:981–986. doi:10.1016/j.bbagrm.2013.05.002

Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**:166–169. doi:10.1093/bioinformatics/btu638

Aust AE, Eveleigh JF. 1999. Mechanisms of DNA oxidation. *Proc Soc Exp Biol Med* **222**:246–252. doi:10.1046/j.1525-1373.1999.d01-141.x

Azzam EI, Jay-Gerin J-P, Pain D. 2012. Ionizing radiation-induced metabolic oxidative stress and prolonged cell injury. *Cancer Lett* **327**:48–60. doi:10.1016/j.canlet.2011.12.012

Ba X, Bacsi A, Luo J, Aguilera-Aguirre L, Zeng X, Radak Z, Brasier AR, Boldogh I. 2014. 8-oxoguanine DNA glycosylase-1 augments

proinflammatory gene expression by facilitating the recruitment of site-specific transcription factors. *J Immunol* **192**:2384–2394. doi:10.4049/jimmunol.1302472

Baik M-H, Silverman JS, Yang IV, Ropp PA, Szalai VA, Yang W, Thorp HH. 2001. Using Density Functional Theory To Design DNA Base Analogues with Low Oxidation Potentials. *J Phys Chem B* **105**:6437–6444. doi:10.1021/jp010643g

Batra VK, Beard WA, Hou EW, Pedersen LC, Prasad R, Wilson SH. 2010. Mutagenic conformation of 8-oxo-7,8-dihydro-2'-dGTP in the confines of a DNA polymerase active site. *Nat Struct Mol Biol* **17**:889–890. doi:10.1038/nsmb.1852

Beckman KB, Ames BN. 1998. The Free Radical Theory of Aging Matures. *Physiological Reviews* **78**:547–581. doi:10.1152/physrev.1998.78.2.547

Bentley D, Selfridge J, Millar JK, Samuel K, Hole N, Ansell JD, Melton DW. 1996. DNA ligase I is required for fetal liver erythropoiesis but is not essential for mammalian cell viability. *Nat Genet* **13**:489–491. doi:10.1038/ng0896-489

Biswas SK. 2016. Does the Interdependence between Oxidative Stress and Inflammation Explain the Antioxidant Paradox? *Oxidative Medicine and Cellular Longevity* **2016**:e5698931. doi:10.1155/2016/5698931

Boiteux S, Coste F, Castaing B. 2017. Repair of 8-oxo-7,8-dihydroguanine in prokaryotic and eukaryotic cells: Properties and biological roles of the Fpg and OGG1 DNA N-glycosylases. *Free Radic Biol Med* **107**:179–201. doi:10.1016/j.freeradbiomed.2016.11.042

Bruner SD, Norman DP, Verdine GL. 2000. Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature* **403**:859–866. doi:10.1038/35002510

Cao B, Wu X, Zhou J, Wu H, Liu L, Zhang Q, DeMott MS, Gu C, Wang L, You D, Dedon PC. 2020. Nick-seq for single-nucleotide resolution genomic maps of DNA modifications and damage. *Nucleic Acids Res* **48**:6715–6725. doi:10.1093/nar/gkaa473

Castellani P, Balza E, Rubartelli A. 2014. Inflammation, DAMPs, tumor development, and progression: a vicious circle orchestrated by redox signaling. *Antioxid Redox Signal* **20**:1086–1097. doi:10.1089/ars.2012.5164

Chen Y-H, Keegan S, Kahli M, Tonzi P, Fenyö D, Huang TT, Smith DJ. 2019. Transcription shapes DNA replication initiation and termination in human cells. *Nat Struct Mol Biol* **26**:67–77. doi:10.1038/s41594-018-0171-0

Colussi C, Parlanti E, Degan P, Aquilina G, Barnes D, Macpherson P, Karran P, Crescenzi M, Dogliotti E, Bignami M. 2002. The mammalian mismatch repair pathway removes DNA 8-oxodGMP incorporated from the oxidized dNTP pool. *Curr Biol* **12**:912–918. doi:10.1016/s0960-9822(02)00863-1

Cooke MS, Evans MD. 2007. 8-Oxo-deoxyguanosine: reduce, reuse, recycle? *Proc Natl Acad Sci U S A* **104**:13535–13536. doi:10.1073/pnas.0706878104

Cooke MS, Evans MD, Dizdaroglu M, Lunec J. 2003. Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB J* **17**:1195–1214. doi:10.1096/fj.02-0752rev

Cs R-I, J B, A W, Rk C, H N, M O, S M, M di P, S I, M H, Jmj W, Ag L, Z K, M R, S H, D B, Rc F. 2015. Whole-genome sequencing provides new insights into the clonal architecture of Barrett's esophagus and esophageal adenocarcinoma. *Nat Genet* **47**:1038–1046. doi:10.1038/ng.3357

Das A, Wiederhold L, Leppard JB, Kedar P, Prasad R, Wang H, Boldogh I, Karimi-Busheri F, Weinfeld M, Tomkinson AE, Wilson SH, Mitra S, Hazra TK. 2006. NEIL2-initiated, APE-independent repair of oxidized bases in DNA: Evidence for a repair complex in human cells. *DNA Repair (Amst)* **5**:1439–1448. doi:10.1016/j.dnarep.2006.07.003

Davalli P, Marverti G, Lauriola A, D'Arca D. 2018. Targeting Oxidatively Induced DNA Damage Response in Cancer: Opportunities for Novel Cancer Therapies. *Oxid Med Cell Longev* **2018**. doi:10.1155/2018/2389523

Davalli P, Mitic T, Caporali A, Lauriola A, D'Arca D. 2016. ROS, Cell Senescence, and Novel Molecular Mechanisms in Aging and Age-Related Diseases. *Oxid Med Cell Longev* **2016**:3565127. doi:10.1155/2016/3565127

Dellino GI, Cittaro D, Piccioni R, Luzi L, Banfi S, Segalla S, Cesaroni M, Mendoza-Maldonado R, Giacca M, Pelicci PG. 2013. Genome-wide mapping of human DNA-replication origins: levels of transcription at ORC1 sites regulate origin selection and replication timing. *Genome Res* **23**:1–11. doi:10.1101/gr.142331.112

D'Errico M, Parlanti E, Teson M, de Jesus BMB, Degan P, Calcagnile A, Jaruga P, Bjørås M, Crescenzi M, Pedrini AM, Egly J-M, Zambruno G, Stefanini M, Dizdaroglu M, Dogliotti E. 2006. New functions of XPC in the protection of human skin cells from oxidative damage. *EMBO J* **25**:4305–4315. doi:10.1038/sj.emboj.7601277

D'Errico M, Parlanti E, Teson M, Degan P, Lemma T, Calcagnile A, Iavarone I, Jaruga P, Ropolo M, Pedrini AM, Orioli D, Frosina G, Zambruno G, Dizdaroglu M, Stefanini M, Dogliotti E. 2007. The role of CSA in the response to oxidative DNA damage in human cells. *Oncogene* **26**:4336–4343. doi:10.1038/sj.onc.1210232

Di Meo S, Reed TT, Venditti P, Victor VM. 2016. Role of ROS and RNS Sources in Physiological and Pathological Conditions. *Oxidative Medicine and Cellular Longevity* **2016**:e1245049. doi:10.1155/2016/1245049

Di Zazzo E, De Rosa C, Abbondanza C, Moncharmont B. 2013. PRDM Proteins: Molecular Mechanisms in Signal Transduction and Transcriptional Regulation. *Biology (Basel)* **2**:107–141. doi:10.3390/biology2010107

Diaz A, Park K, Lim DA, Song JS. 2012. Normalization, bias correction, and peak calling for ChIP-seq. *Stat Appl Genet Mol Biol* **11**:Article 9. doi:10.1515/1544-6115.1750

Ding Y, Fleming AM, Burrows CJ. 2017. Sequencing the Mouse Genome for the Oxidatively Modified Base 8-Oxo-7,8-dihydroguanine by OG-Seq. *J Am Chem Soc* **139**:2569–2572. doi:10.1021/jacs.6b12604

Dou H, Mitra S, Hazra TK. 2003. Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. *J Biol Chem* **278**:49679–49684. doi:10.1074/jbc.M308658200

Dulak AM, Stojanov P, Peng S, Lawrence MS, Fox C, Stewart C, Bandla S, Imamura Y, Schumacher SE, Shefler E, McKenna A, Carter SL, Cibulskis K, Sivachenko A, Saksena G, Voet D, Ramos AH, Auclair D, Thompson K, Sougnez C, Onofrio RC, Guiducci C, Beroukhim R, Zhou Z, Lin L, Lin J, Reddy R, Chang A, Landrenau R, Pennathur A, Ogino S, Luketich JD, Golub TR, Gabriel SB, Lander ES, Beer DG, Godfrey TE, Getz G, Bass AJ. 2013. Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational complexity. *Nat Genet* **45**:478–486. doi:10.1038/ng.2591

Evans MD, Cooke MS. 2004. Factors contributing to the outcome of oxidative damage to nucleic acids. *Bioessays* **26**:533–542. doi:10.1002/bies.20027

Evans MD, Dizdaroglu M, Cooke MS. 2004. Oxidative DNA damage and disease: induction, repair and significance. *Mutat Res* **567**:1–61. doi:10.1016/j.mrrev.2003.11.001

Fang R, Barbera AJ, Xu Y, Rutenberg M, Leonor T, Bi Q, Lan F, Mei P, Yuan G-C, Lian C, Peng J, Cheng D, Sui G, Kaiser UB, Shi Y, Shi YG. 2010. Human LSD2/KDM1b/AOF1 regulates gene transcription by modulating intragenic H3K4me2 methylation. *Mol Cell* **39**:222–233. doi:10.1016/j.molcel.2010.07.008

Fang Y, Zou P. 2020. Genome-Wide Mapping of Oxidative DNA Damage via Engineering of 8-Oxoguanine DNA Glycosylase. *Biochemistry* **59**:85–89. doi:10.1021/acs.biochem.9b00782

Fenton HJH. 1894. LXXIII.—Oxidation of tartaric acid in presence of iron. *J Chem Soc, Trans* **65**:899–910. doi:10.1039/CT8946500899

Fleming AM, Burrows CJ. 2017. 8-Oxo-7,8-dihydroguanine, friend and foe: Epigenetic-like regulator versus initiator of mutagenesis. *DNA Repair (Amst)* **56**:75–83. doi:10.1016/j.dnarep.2017.06.009

Fleming AM, Ding Y, Burrows CJ. 2017. Oxidative DNA damage is epigenetic by regulating gene transcription via base excision repair.

Proc Natl Acad Sci U S A **114**:2604–2609. doi:10.1073/pnas.1619809114

Fleming AM, Zhu J, Howpay Manage SA, Burrows CJ. 2019. Human NEIL3 Gene Expression Regulated by Epigenetic-Like Oxidative DNA Modification. *J Am Chem Soc* **141**:11036–11049. doi:10.1021/jacs.9b01847

Fortini P, Dogliotti E. 2007. Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways. *DNA Repair (Amst)* **6**:398–409. doi:10.1016/j.dnarep.2006.10.008

Fortini P, Parlanti E, Sidorkina OM, Laval J, Dogliotti E. 1999. The type of DNA glycosylase determines the base excision repair pathway in mammalian cells. *J Biol Chem* **274**:15230–15236. doi:10.1074/jbc.274.21.15230

Fousteri M, Mullenders LHF. 2008. Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects. *Cell Res* **18**:73–84. doi:10.1038/cr.2008.6

Frosina G, Fortini P, Rossi O, Carrozzino F, Raspaglio G, Cox LS, Lane DP, Abbondandolo A, Dogliotti E. 1996. Two pathways for base excision repair in mammalian cells. *J Biol Chem* **271**:9573–9578. doi:10.1074/jbc.271.16.9573

Fry RC, Begley TJ, Samson LD. 2005. Genome-wide responses to DNA-damaging agents. *Annu Rev Microbiol* **59**:357–377. doi:10.1146/annurev.micro.59.031805.133658

Fukai T, Ushio-Fukai M. 2011. Superoxide dismutases: role in redox signaling, vascular function, and diseases. *Antioxid Redox Signal* **15**:1583–1606. doi:10.1089/ars.2011.3999

Ginno PA, Lim YW, Lott PL, Korf I, Chédin F. 2013. GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination. *Genome Res* **23**:1590–1600. doi:10.1101/gr.158436.113

Giorgio M, Dellino GI, Gambino V, Roda N, Pelicci PG. 2020. On the epigenetic role of guanosine oxidation. *Redox Biol* **29**:101398. doi:10.1016/j.redox.2019.101398

Giorgio M, Trinei M, Migliaccio E, Pelicci PG. 2007. Hydrogen peroxide: a metabolic by-product or a common mediator of ageing signals? *Nat Rev Mol Cell Biol* **8**:722–728. doi:10.1038/nrm2240

Gorini F, Scala G, Cooke MS, Majello B, Amente S. 2021. Towards a comprehensive view of 8-oxo-7,8-dihydro-2'-deoxyguanosine: Highlighting the intertwined roles of DNA damage and epigenetics in genomic instability. *DNA Repair (Amst)* **97**:103027. doi:10.1016/j.dnarep.2020.103027

Gorini F, Scala G, Di Palo G, Dellino GI, Cocozza S, Pelicci PG, Lania L, Majello B, Amente S. 2020. The genomic landscape of 8-oxodG

reveals enrichment at specific inherently fragile promoters. *Nucleic Acids Res* **48**:4309–4324. doi:10.1093/nar/gkaa175

Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, Davies H, Teague J, Butler A, Stevens C, Edkins S, O'Meara S, Vastrik I, Schmidt EE, Avis T, Barthorpe S, Bhamra G, Buck G, Choudhury B, Clements J, Cole J, Dicks E, Forbes S, Gray K, Halliday K, Harrison R, Hills K, Hinton J, Jenkinson A, Jones D, Menzies A, Mironenko T, Perry J, Raine K, Richardson D, Shepherd R, Small A, Tofts C, Varian J, Webb T, West S, Widaa S, Yates A, Cahill DP, Louis DN, Goldstraw P, Nicholson AG, Brasseur F, Looijenga L, Weber BL, Chiew Y-E, DeFazio A, Greaves MF, Green AR, Campbell P, Birney E, Easton DF, Chenevix-Trench G, Tan M-H, Khoo SK, Teh BT, Yuen ST, Leung SY, Wooster R, Futreal PA, Stratton MR. 2007. Patterns of somatic mutation in human cancer genomes. *Nature* **446**:153–158. doi:10.1038/nature05610

Hegde ML, Izumi T, Mitra S. 2012. Oxidized base damage and single-strand break repair in mammalian genomes: role of disordered regions and posttranslational modifications in early enzymes. *Prog Mol Biol Transl Sci* **110**:123–153. doi:10.1016/B978-0-12-387665-2.00006-7

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Molecular Cell* **38**:576–589. doi:10.1016/j.molcel.2010.05.004

Helmrich A, Ballarino M, Tora L. 2011. Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Mol Cell* **44**:966–977. doi:10.1016/j.molcel.2011.10.013

Hsu GW, Ober M, Carell T, Beese LS. 2004. Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase. *Nature* **431**:217–221. doi:10.1038/nature02908

Huang W, Loganantharaj R, Schroeder B, Fargo D, Li L. 2013. PAVIS: a tool for Peak Annotation and Visualization. *Bioinformatics* **29**:3097–3099. doi:10.1093/bioinformatics/btt520

Inoue M, Kamiya H, Fujikawa K, Ootsuyama Y, Murata-Kamiya N, Osaki T, Yasumoto K, Kasai H. 1998. Induction of chromosomal gene mutations in Escherichia coli by direct incorporation of oxidatively damaged nucleotides. New evaluation method for mutagenesis by damaged DNA precursors in vivo. *J Biol Chem* **273**:11069–11074. doi:10.1074/jbc.273.18.11069

Jackson SP, Bartek J. 2009. The DNA-damage response in human biology and disease. *Nature* **461**:1071–1078. doi:10.1038/nature08467

Kakehashi A, Ishii N, Okuno T, Fujioka M, Gi M, Wanibuchi H. 2017. Enhanced Susceptibility of Ogg1 Mutant Mice to Multiorgan Carcinogenesis. *Int J Mol Sci* **18**. doi:10.3390/ijms18081801

Kellner WA, Bell JSK, Vertino PM. 2015. GC skew defines distinct RNA polymerase pause sites in CpG island promoters. *Genome Res* **25**:1600–1609. doi:10.1101/gr.189068.114

Kim GH, Kim JE, Rhie SJ, Yoon S. 2015. The Role of Oxidative Stress in Neurodegenerative Diseases. *Exp Neurobiol* **24**:325–340. doi:10.5607/en.2015.24.4.325

Klaunig JE, Kamendulis LM. 2004. The role of oxidative stress in carcinogenesis. *Annu Rev Pharmacol Toxicol* **44**:239–267. doi:10.1146/annurev.pharmtox.44.101802.121851

Klungland A, Lindahl T. 1997. Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO J* **16**:3341–3348. doi:10.1093/emboj/16.11.3341

Koga Y, Taniguchi Y, Sasaki S. 2013. Synthesis of the oligoribonucleotides incorporating 8-oxo-guanosine and evaluation of their base pairing properties. *Nucleosides Nucleotides Nucleic Acids* **32**:124–136. doi:10.1080/15257770.2013.767461

Kohen R, Nyska A. 2002. Oxidation of biological systems: oxidative stress phenomena, antioxidants, redox reactions, and methods for their quantification. *Toxicol Pathol* **30**:620–650. doi:10.1080/01926230290166724

Kuroda J, Nakagawa K, Yamasaki T, Nakamura K, Takeya R, Kuribayashi F, Imajoh-Ohmi S, Igarashi K, Shibata Y, Sueishi K, Sumimoto H. 2005. The superoxide-producing NAD(P)H oxidase Nox4 in the nucleus of human vascular endothelial cells. *Genes Cells* **10**:1139–1151. doi:10.1111/j.1365-2443.2005.00907.x

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**:R25. doi:10.1186/gb-2009-10-3-r25

Li G-M. 2008. Mechanisms and functions of DNA mismatch repair. *Cell Res* **18**:85–98. doi:10.1038/cr.2007.115

Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**:589–595. doi:10.1093/bioinformatics/btp698

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**:2078–2079. doi:10.1093/bioinformatics/btp352

Li J, Braganza A, Sobol RW. 2013. Base excision repair facilitates a functional relationship between Guanine oxidation and histone

demethylation. *Antioxid Redox Signal* **18**:2429–2443. doi:10.1089/ars.2012.5107

Lindahl T. 1993. Instability and decay of the primary structure of DNA. *Nature* **362**:709–715. doi:10.1038/362709a0

Lindahl T. 1990. Repair of intrinsic DNA lesions. *Mutat Res* **238**:305–311. doi:10.1016/0165-1110(90)90022-4

Lindahl T, Barnes DE. 2000. Repair of endogenous DNA damage. *Cold Spring Harb Symp Quant Biol* **65**:127–133. doi:10.1101/sqb.2000.65.127

Liu Y, Walavalkar NM, Dozmorov MG, Rich SS, Civelek M, Guertin MJ. 2017. Identification of breast cancer associated variants that modulate transcription factor binding. *PLoS Genet* **13**:e1006761. doi:10.1371/journal.pgen.1006761

Liu ZJ, Martínez Cuesta S, van Delft P, Balasubramanian S. 2019. Sequencing abasic sites in DNA at single-nucleotide resolution. *Nat Chem* **11**:629–637. doi:10.1038/s41557-019-0279-9

López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. 2013. The hallmarks of aging. *Cell* **153**:1194–1217. doi:10.1016/j.cell.2013.05.039

Loren P, Sánchez R, Arias M-E, Felmer R, Risopatrón J, Cheuquemán C. 2017. Melatonin Scavenger Properties against Oxidative and Nitrosative Stress: Impact on Gamete Handling and In Vitro Embryo Production in Humans and Other Mammals. *Int J Mol Sci* **18**. doi:10.3390/ijms18061119

Lu T, Pan Y, Kao S-Y, Li C, Kohane I, Chan J, Yankner BA. 2004. Gene regulation and DNA damage in the ageing human brain. *Nature* **429**:883–891. doi:10.1038/nature02661

Ludwig DL, MacInnes MA, Takiguchi Y, Purtymun PE, Henrie M, Flannery M, Meneses J, Pedersen RA, Chen DJ. 1998. A murine AP-endonuclease gene-targeted deficiency with post-implantation embryonic progression and ionizing radiation sensitivity. *Mutat Res* **409**:17–29. doi:10.1016/s0921-8777(98)00039-1

Maga G, Villani G, Crespan E, Wimmer U, Ferrari E, Bertocci B, Hübscher U. 2007. 8-oxo-guanine bypass by human DNA polymerases in the presence of auxiliary proteins. *Nature* **447**:606–608. doi:10.1038/nature05843

Majello B, Gorini F, Saccà CD, Amente S. 2019. Expanding the Role of the Histone Lysine-Specific Demethylase LSD1 in Cancer. *Cancers (Basel)* **11**. doi:10.3390/cancers11030324

Maki H, Sekiguchi M. 1992. MutT protein specifically hydrolyses a potent mutagenic substrate for DNA synthesis. *Nature* **355**:273–275. doi:10.1038/355273a0

Markkanen E. 2017. Not breathing is not an option: How to deal with oxidative DNA damage. *DNA Repair (Amst)* **59**:82–105. doi:10.1016/j.dnarep.2017.09.007

Markkanen E, Dorn J, Hübscher U. 2013. MUTYH DNA glycosylase: the rationale for removing undamaged bases from the DNA. *Front Genet* **4**:18. doi:10.3389/fgene.2013.00018

Matsuzawa A. 2017. Thioredoxin and redox signaling: Roles of the thioredoxin system in control of cell fate. *Arch Biochem Biophys* **617**:101–105. doi:10.1016/j.abb.2016.09.011

Mena S, Ortega A, Estrela JM. 2009. Oxidative stress in environmental-induced carcinogenesis. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Oxidative Stress and Mechanisms of Environmental Toxicity **674**:36–44. doi:10.1016/j.mrgentox.2008.09.017

Minowa O, Arai T, Hirano M, Monden Y, Nakai S, Fukuda M, Itoh M, Takano H, Hippou Y, Aburatani H, Masumura K, Nohmi T, Nishimura S, Noda T. 2000. Mmh/Ogg1 gene inactivation results in accumulation of 8-hydroxyguanine in mice. *Proc Natl Acad Sci U S A* **97**:4156–4161. doi:10.1073/pnas.050404497

Mittal M, Siddiqui MR, Tran K, Reddy SP, Malik AB. 2014. Reactive oxygen species in inflammation and tissue injury. *Antioxid Redox Signal* **20**:1126–1167. doi:10.1089/ars.2012.5149

Montelone BA. 2006. DNA Repair and Mutagenesis. Second Edition. By Errol C Friedberg, Graham C Walker, Wolfram Siede, Richard D Wood, Roger A Schultz, and Tom Ellenberger. *The Quarterly Review of Biology* **81**:273–273. doi:10.1086/509407

Morales J, Li L, Fattah FJ, Dong Y, Bey EA, Patel M, Gao J, Boothman DA. 2014. Review of poly (ADP-ribose) polymerase (PARP) mechanisms of action and rationale for targeting in cancer and other diseases. *Crit Rev Eukaryot Gene Expr* **24**:15–28. doi:10.1615/critreveukaryotgeneexpr.2013006875

Murugaesu N, Wilson GA, Birkbak NJ, Watkins T, McGranahan N, Kumar S, Abbassi-Ghadi N, Salm M, Mitter R, Horswell S, Rowan A, Phillimore B, Biggs J, Begum S, Matthews N, Hochhauser D, Hanna GB, Swanton C. 2015. Tracking the genomic evolution of esophageal adenocarcinoma through neoadjuvant chemotherapy. *Cancer Discov* **5**:821–831. doi:10.1158/2159-8290.CD-15-0412

Nicolussi A, D'Inzeo S, Capalbo C, Giannini G, Coppa A. 2017. The role of peroxiredoxins in cancer. *Mol Clin Oncol* **6**:139–153. doi:10.3892/mco.2017.1129

Ohno M, Miura T, Furuichi M, Tominaga Y, Tsuchimoto D, Sakumi K, Nakabeppu Y. 2006. A genome-wide distribution of 8-oxoguanine correlates with the preferred regions for recombination and single

nucleotide polymorphism in the human genome. *Genome Res* **16**:567–575. doi:10.1101/gr.4769606

Olinski R, Gackowski D, Cooke MS. 2018. Endogenously generated DNA nucleobase modifications source, and significance as possible biomarkers of malignant transformation risk, and role in anticancer therapy. *Biochim Biophys Acta Rev Cancer* **1869**:29–41. doi:10.1016/j.bbcan.2017.11.002

Parker A, Gu Y, Mahoney W, Lee SH, Singh KK, Lu AL. 2001. Human homolog of the MutY repair protein (hMYH) physically interacts with proteins involved in long patch DNA base excision repair. *J Biol Chem* **276**:5547–5555. doi:10.1074/jbc.M008463200

Pastukh V, Ruchko M, Gorodnya O, Wilson GL, Gillespie MN. 2007. Sequence-specific oxidative base modifications in hypoxia-inducible genes. *Free Radic Biol Med* **43**:1616–1626. doi:10.1016/j.freeradbiomed.2007.08.027

Patel RK, Jain M. 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* **7**:e30619. doi:10.1371/journal.pone.0030619

Perillo B, Ombra MN, Bertoni A, Cuozzo C, Sacchetti S, Sasso A, Chiariotti L, Malorni A, Abbondanza C, Avvedimento EV. 2008. DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression. *Science* **319**:202–206. doi:10.1126/science.1147674

Pezone A, Taddei ML, Tramontano A, Dolcini J, Boffo FL, De Rosa M, Parri M, Stinziani S, Comito G, Porcellini A, Raugei G, Gackowski D, Zarakowska E, Olinski R, Gabrielli A, Chiarugi P, Avvedimento EV. 2020. Targeted DNA oxidation by LSD1-SMAD2/3 primes TGF-β1/ EMT genes for activation or repression. *Nucleic Acids Res* **48**:8943–8958. doi:10.1093/nar/gkaa599

Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, Greenman CD, Varela I, Lin M-L, Ordóñez GR, Bignell GR, Ye K, Alipaz J, Bauer MJ, Beare D, Butler A, Carter RJ, Chen L, Cox AJ, Edkins S, Kokko-Gonzales PI, Gormley NA, Grocock RJ, Haudenschild CD, Hims MM, James T, Jia M, Kingsbury Z, Leroy C, Marshall J, Menzies A, Mudie LJ, Ning Z, Royce T, Schulz-Trieglaff OB, Spiridou A, Stebbings LA, Szajkowski L, Teague J, Williamson D, Chin L, Ross MT, Campbell PJ, Bentley DR, Futreal PA, Stratton MR. 2010a. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* **463**:191–196. doi:10.1038/nature08658

Pleasance ED, Stephens PJ, O'Meara S, McBride DJ, Meynert A, Jones D, Lin M-L, Beare D, Lau KW, Greenman C, Varela I, Nik-Zainal S, Davies HR, Ordoñez GR, Mudie LJ, Latimer C, Edkins S, Stebbings L, Chen L, Jia M, Leroy C, Marshall J, Menzies A, Butler A, Teague

JW, Mangion J, Sun YA, McLaughlin SF, Peckham HE, Tsung EF, Costa GL, Lee CC, Minna JD, Gazdar A, Birney E, Rhodes MD, McKernan KJ, Stratton MR, Futreal PA, Campbell PJ. 2010b. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* **463**:184–190. doi:10.1038/nature08629

Poetsch AR. 2020. The genomics of oxidative DNA damage, repair, and resulting mutagenesis. *Comput Struct Biotechnol J* **18**:207–219. doi:10.1016/j.csbj.2019.12.013

Poetsch AR, Boulton SJ, Luscombe NM. 2018. Genomic landscape of oxidative DNA damage and repair reveals regioselective protection from mutagenesis. *Genome Biol* **19**. doi:10.1186/s13059-018-1582-2

Puebla-Osorio N, Lacey DB, Alt FW, Zhu C. 2006. Early embryonic lethality due to targeted inactivation of DNA ligase III. *Mol Cell Biol* **26**:3935–3941. doi:10.1128/MCB.26.10.3935-3941.2006

Puissant A, Frumm SM, Alexe G, Bassil CF, Qi J, Chanthery YH, Nekritz EA, Zeid R, Gustafson WC, Greninger P, Garnett MJ, McDermott U, Benes CH, Kung AL, Weiss WA, Bradner JE, Stegmaier K. 2013. Targeting MYCN in neuroblastoma by BET bromodomain inhibition. *Cancer Discov* **3**:308–323. doi:10.1158/2159-8290.CD-12-0418

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**:841–842. doi:10.1093/bioinformatics/btq033

Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**:W187-191. doi:10.1093/nar/gku365

Redstone SCJ, Fleming AM, Burrows CJ. 2019. Oxidative Modification of the Potential G-Quadruplex Sequence in the PCNA Gene Promoter Can Turn on Transcription. *Chem Res Toxicol* **32**:437–446. doi:10.1021/acs.chemrestox.8b00332

Reynolds P, Cooper S, Lomax M, O'Neill P. 2015. Disruption of PARP1 function inhibits base excision repair of a sub-set of DNA lesions. *Nucleic Acids Res* **43**:4028–4038. doi:10.1093/nar/gkv250

Rouhier N, Lemaire SD, Jacquot J-P. 2008. The role of glutathione in photosynthetic organisms: emerging functions for glutaredoxins and glutathionylation. *Annu Rev Plant Biol* **59**:143–166. doi:10.1146/annurev.arplant.59.032607.092811

Sabharwal SS, Waypa GB, Marks JD, Schumacker PT. 2013. Peroxiredoxin-5 targeted to the mitochondrial intermembrane space attenuates hypoxia-induced reactive oxygen species signalling. *Biochem J* **456**:337–346. doi:10.1042/BJ20130740

Sakamoto K, Tominaga Y, Yamauchi K, Nakatsu Y, Sakumi K, Yoshiyama K, Egashira A, Kura S, Yao T, Tsuneyoshi M, Maki H, Nakabeppu Y, Tsuzuki T. 2007. MUTYH-null mice are susceptible to

spontaneous and oxidative stress induced intestinal tumorigenesis. *Cancer Res* **67**:6599–6604. doi:10.1158/0008-5472.CAN-06-4802

Sakumi K, Tominaga Y, Furuichi M, Xu P, Tsuzuki T, Sekiguchi M, Nakabeppu Y. 2003. Ogg1 knockout-associated lung tumorigenesis and its suppression by Mth1 gene disruption. *Cancer Res* **63**:902–905.

Satou K, Hori M, Kawai K, Kasai H, Harashima H, Kamiya H. 2009. Involvement of specialized DNA polymerases in mutagenesis by 8-hydroxy-dGTP in human cells. *DNA Repair (Amst)* **8**:637–642. doi:10.1016/j.dnarep.2008.12.009

Satou K, Kawai K, Kasai H, Harashima H, Kamiya H. 2007. Mutagenic effects of 8-hydroxy-dGTP in live mammalian cells. *Free Radic Biol Med* **42**:1552–1560. doi:10.1016/j.freeradbiomed.2007.02.024

Schieber M, Chandel NS. 2014. ROS function in redox signaling and oxidative stress. *Curr Biol* **24**:R453-462. doi:10.1016/j.cub.2014.03.034

Sengupta S, Wang H, Yang C, Szczesny B, Hegde ML, Mitra S. 2020. Ligand-induced gene activation is associated with oxidative genome damage whose repair is required for transcription. *Proc Natl Acad Sci U S A* **117**:22183–22192. doi:10.1073/pnas.1919445117

Shi Yujiang, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Yang. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**:941–953. doi:10.1016/j.cell.2004.12.012

Shibutani S, Takeshita M, Grollman AP. 1991. Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature* **349**:431–434. doi:10.1038/349431a0

Sies H, Berndt C, Jones DP. 2017. Oxidative Stress. *Annu Rev Biochem* **86**:715–748. doi:10.1146/annurev-biochem-061516-045037

Skourti-Stathaki K, Proudfoot NJ, Gromak N. 2011. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol Cell* **42**:794–805. doi:10.1016/j.molcel.2011.04.026

Sobol RW, Horton JK, Kühn R, Gu H, Singhal RK, Prasad R, Rajewsky K, Wilson SH. 1996. Requirement of mammalian DNA polymerase-beta in base-excision repair. *Nature* **379**:183–186. doi:10.1038/379183a0

Suzuki T, Kamiya H. 2017. Mutations induced by 8-hydroxyguanine (8-oxo-7,8-dihydroguanine), a representative oxidized base, in mammalian cells. *Genes Environ* **39**:2. doi:10.1186/s41021-016-0051-y

Svilar D, Goellner EM, Almeida KH, Sobol RW. 2011. Base excision repair and lesion-dependent subpathways for repair of oxidative DNA damage. *Antioxid Redox Signal* **14**:2491–2507. doi:10.1089/ars.2010.3466

Tebbs RS, Flannery ML, Meneses JJ, Hartmann A, Tucker JD, Thompson LH, Cleaver JE, Pedersen RA. 1999. Requirement for the Xrcc1 DNA base excision repair gene during early mouse development. *Dev Biol* **208**:513–529. doi:10.1006/dbio.1999.9232

Triner D, Castillo C, Hakim JB, Xue X, Greenson JK, Nuñez G, Chen GY, Colacino JA, Shah YM. 2018. Myc-Associated Zinc Finger Protein Regulates the Proinflammatory Response in Colitis and Colon Cancer via STAT3 Signaling. *Mol Cell Biol* **38**. doi:10.1128/MCB.00386-18

Tsang CK, Liu Y, Thomas J, Zhang Y, Zheng XFS. 2014. Superoxide dismutase 1 acts as a nuclear transcription factor to regulate oxidative stress resistance. *Nat Commun* **5**:3446. doi:10.1038/ncomms4446

van Loon B, Markkanen E, Hübscher U. 2010. Oxygen as a friend and enemy: How to combat the mutational potential of 8-oxo-guanine. *DNA Repair* **9**:604–616. doi:10.1016/j.dnarep.2010.03.004

Vašák M, Meloni G. 2011. Chemistry and biology of mammalian metallothioneins. *J Biol Inorg Chem* **16**:1067–1078. doi:10.1007/s00775-011-0799-2

Wang R, Hao W, Pan L, Boldogh I, Ba X. 2018. The roles of base excision repair enzyme OGG1 in gene expression. *Cell Mol Life Sci* **75**:3741–3750. doi:10.1007/s00018-018-2887-8

Waypa GB, Smith KA, Schumacker PT. 2016. O2 sensing, mitochondria and ROS signaling: The fog is lifting. *Mol Aspects Med* **47–48**:76–89. doi:10.1016/j.mam.2016.01.002

Winterbourn CC, Hampton MB. 2008. Thiol chemistry and specificity in redox signaling. *Free Radic Biol Med* **45**:549–561. doi:10.1016/j.freeradbiomed.2008.05.004

Wong H-K, Muftuoglu M, Beck G, Imam SZ, Bohr VA, Wilson DM. 2007. Cockayne syndrome B protein stimulates apurinic endonuclease 1 activity and protects against agents that introduce base excision repair intermediates. *Nucleic Acids Res* **35**:4103–4113. doi:10.1093/nar/gkm404

Wu J, McKeague M, Sturla SJ. 2018. Nucleotide-Resolution Genome-Wide Mapping of Oxidative DNA Damage by Click-Code-Seq. *J Am Chem Soc* **140**:9783–9787. doi:10.1021/jacs.8b03715

Xanthoudakis S, Smeyne RJ, Wallace JD, Curran T. 1996. The redox/DNA repair protein, Ref-1, is essential for early embryonic development in mice. *Proc Natl Acad Sci U S A* **93**:8919–8923. doi:10.1073/pnas.93.17.8919

Xie Y, Yang H, Cunanan C, Okamoto K, Shibata D, Pan J, Barnes DE, Lindahl T, McIlhatton M, Fishel R, Miller JH. 2004. Deficiencies in mouse Myh and Ogg1 result in tumor predisposition and G to T mutations in codon 12 of the K-ras oncogene in lung tumors. *Cancer Res* **64**:3096–3102. doi:10.1158/0008-5472.can-03-3834

Yang H, Clendenin WM, Wong D, Demple B, Slupska MM, Chiang JH, Miller JH. 2001. Enhanced activity of adenine-DNA glycosylase (Myh) by apurinic/apyrimidinic endonuclease (Ape1) in mammalian base excision repair of an A/GO mismatch. *Nucleic Acids Res* **29**:743–752. doi:10.1093/nar/29.3.743

Yoshihara M, Jiang L, Akatsuka S, Suyama M, Toyokuni S. 2014. Genome-wide profiling of 8-oxoguanine reveals its association with spatial positioning in nucleus. *DNA Res* **21**:603–612. doi:10.1093/dnares/dsu023

Zarakowska E, Gackowski D, Foksinski M, Olinski R. 2014. Are 8-oxoguanine (8-oxoGua) and 5-hydroxymethyluracil (5-hmUra) oxidatively damaged DNA bases or transcription (epigenetic) marks? *Mutat Res Genet Toxicol Environ Mutagen* **764–765**:58–63. doi:10.1016/j.mrgentox.2013.09.002

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**:R137. doi:10.1186/gb-2008-9-9-r137

Zhu J, Fleming AM, Burrows CJ. 2018. The RAD17 Promoter Sequence Contains a Potential Tail-Dependent G-Quadruplex That Downregulates Gene Expression upon Oxidative Modification. *ACS Chem Biol* **13**:2577–2584. doi:10.1021/acschembio.8b00522

Zuchegna C, Aceto F, Bertoni A, Romano A, Perillo B, Laccetti P, Gottesman ME, Avvedimento EV, Porcellini A. 2014. Mechanism of retinoic acid-induced transcription: histone code, DNA oxidation and formation of chromatin loops. *Nucleic Acids Res* **42**:11040–11055. doi:10.1093/nar/gku823