



Bioinformatic analysis of DNA damage repair gene OGG1

GROUP 14

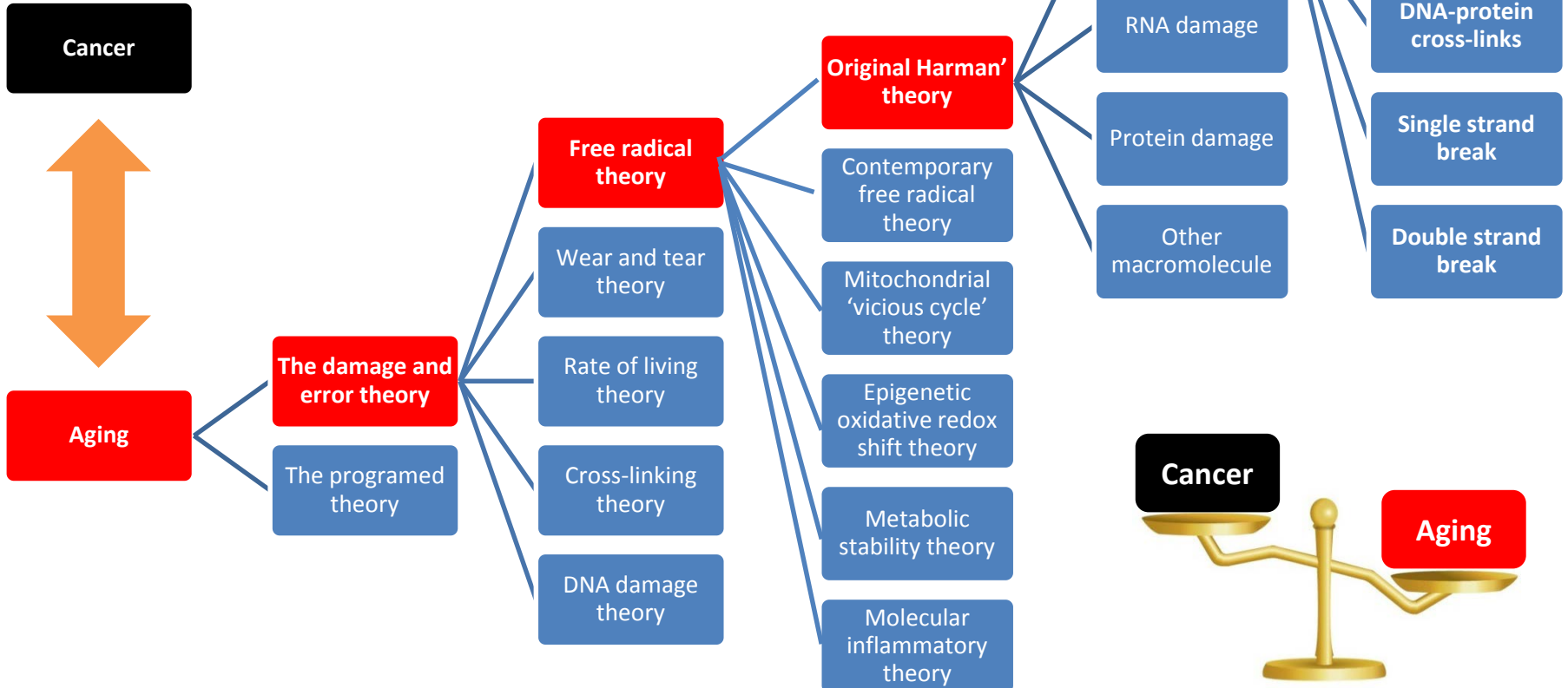
Han Yan Hanming Wang

Liqing Chi Lida Wu

2014.01.09

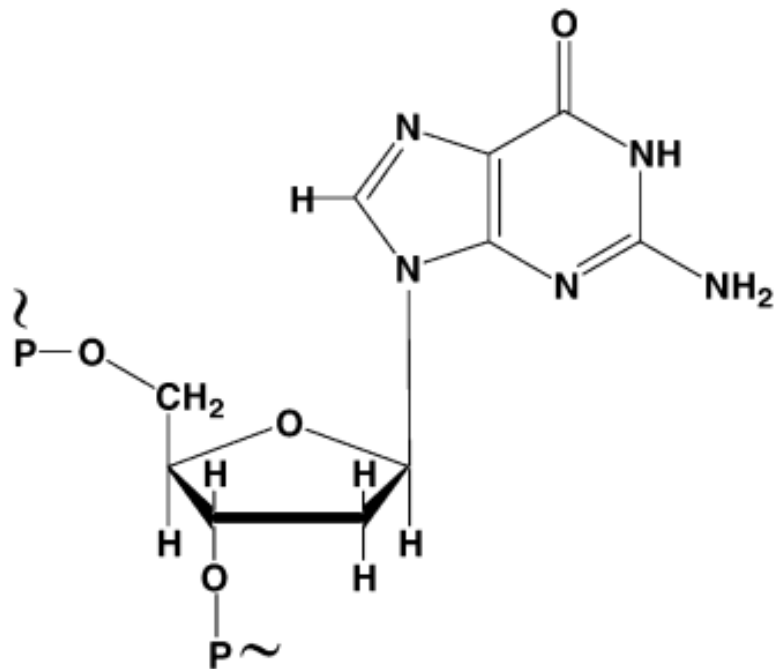


DNA Damage Accumulation Contributes to Cancer and Aging

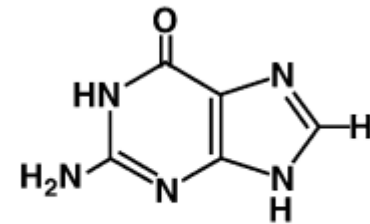


8-oxo-dGTP is the main DNA damage product in human

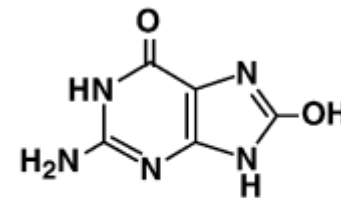
8-hydroxyguanine(8-OH-dG/8-oxo-dG)



Deoxyguanosine monophosphate



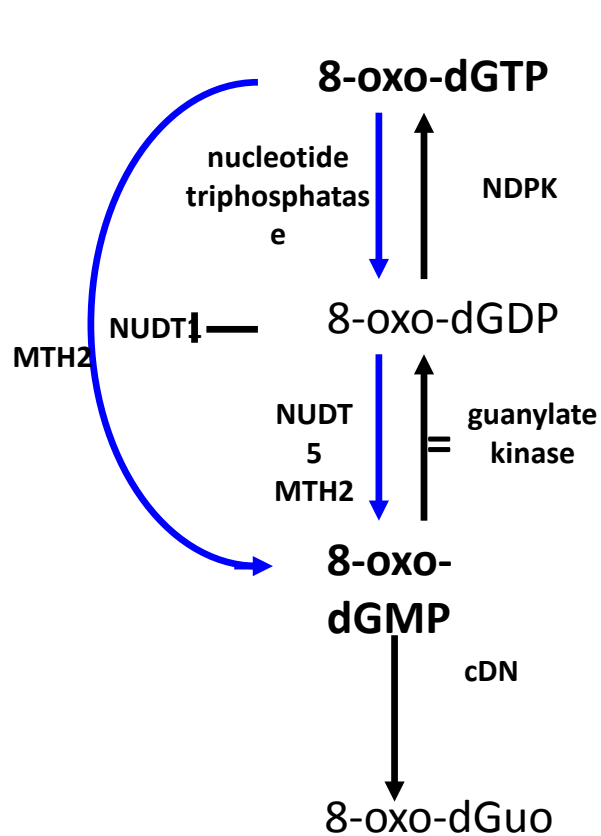
guanine



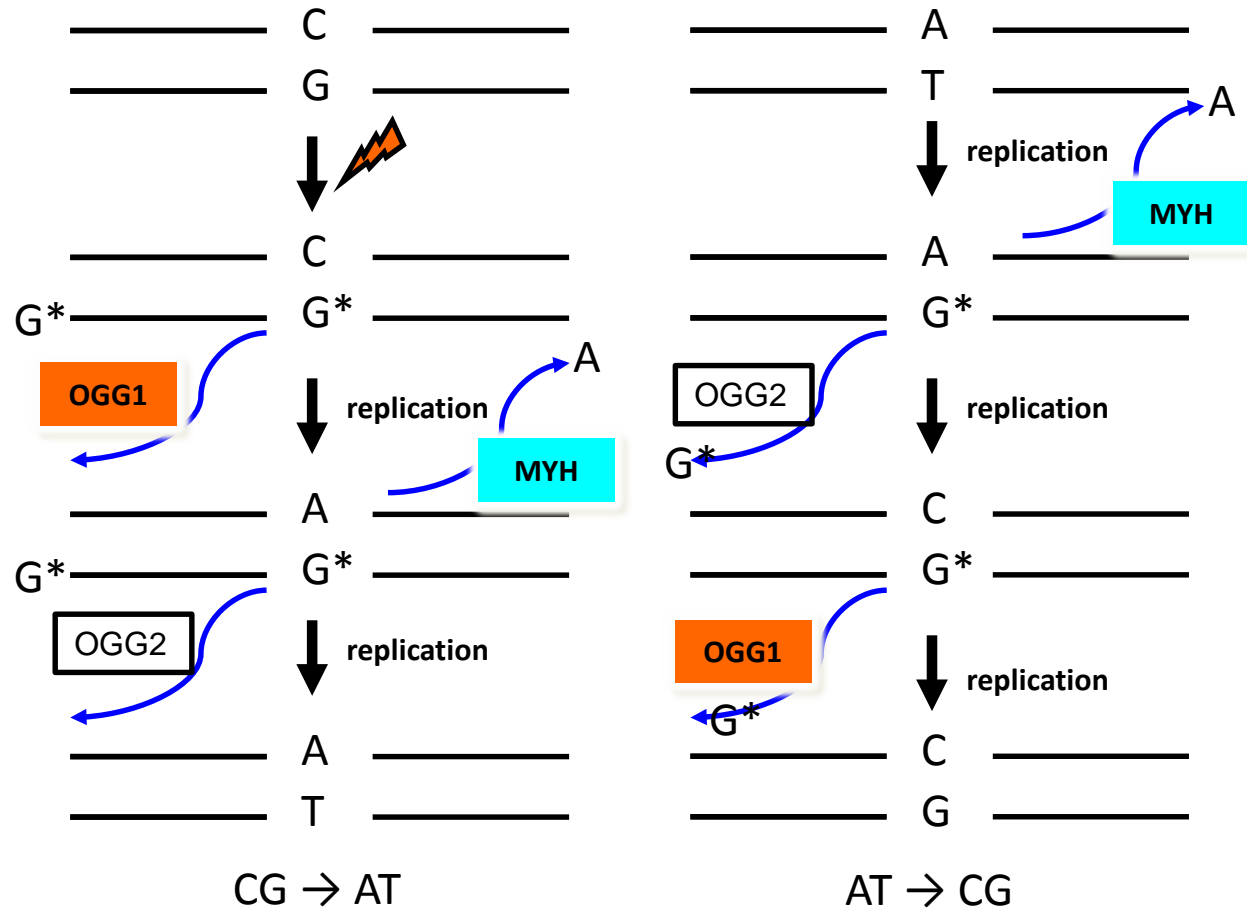
8-hydroxyguanine

Nucleotide pool sanitization and Base excision repair

Nucleotide pool sanitization



Base excision repair

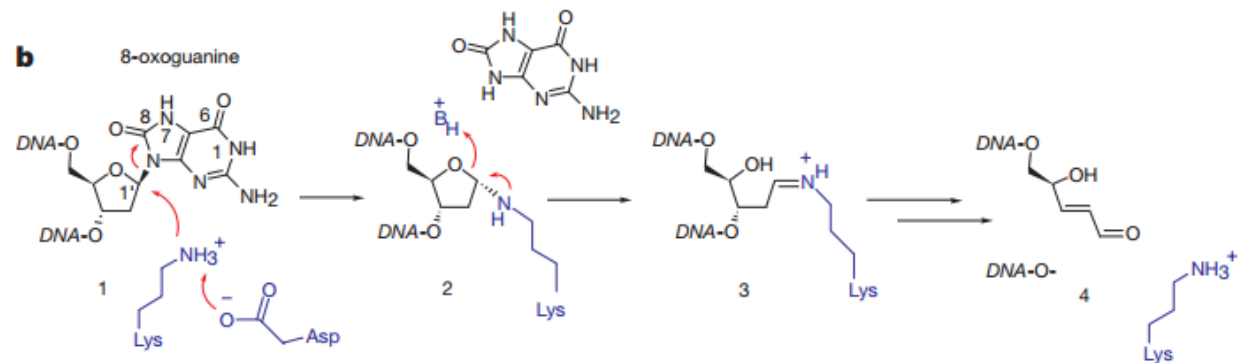
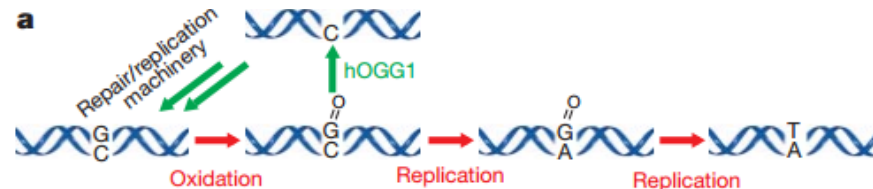


OGG1 cleans up 8-oxo-dGTP through BER (Base Excision Repair)

OGG1

General annotation (Comments)

Function	DNA repair enzyme that incises DNA at 8-oxoG residues. Excises 7,8-dihydro-8-oxoguanine and 2,6-diamino-4-hydroxy-5-N-methylformamidopyrimidine (FAPY) from damaged DNA. Has a beta-lyase activity that nicks DNA 3' to the lesion.
Catalytic activity	The C-O-P bond 3' to the apurinic or apyrimidinic site in DNA is broken by a beta-elimination reaction, leaving a 3'-terminal unsaturated sugar and a product with a terminal 5'-phosphate.
Subcellular location	Nucleus > nucleoplasm. Nucleus speckle . Nucleus matrix . Note: Together with APEX1 is recruited to nuclear speckles in UVA-irradiated cells. (Ref. 17) Isoform 1A : Nucleus (Ref. 17) . Isoform 2A : Mitochondrion (Ref. 17) .
Tissue specificity	Ubiquitous.

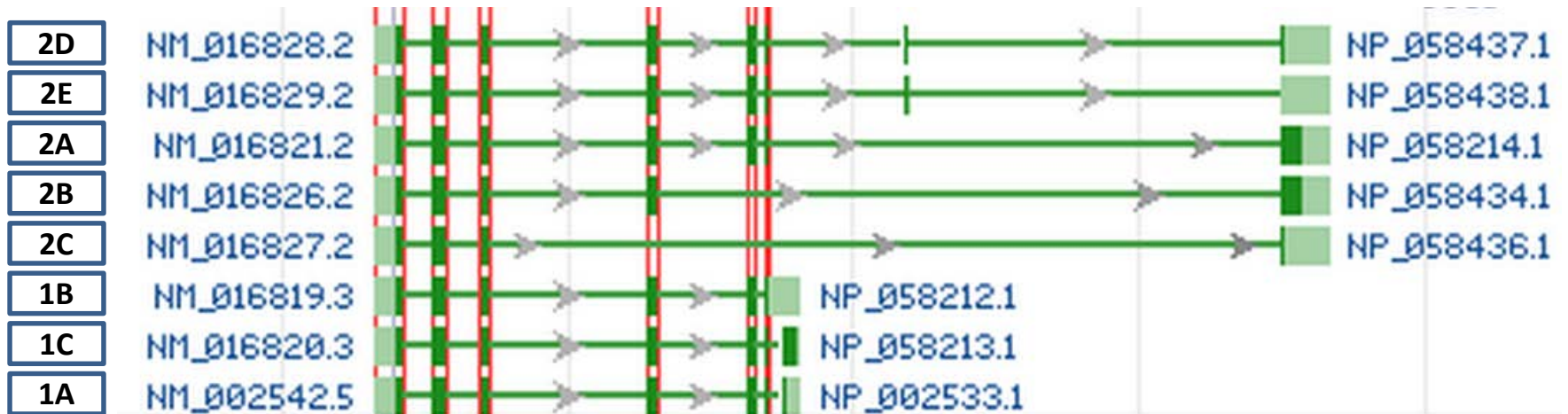


Glycosidase

Endonuclease

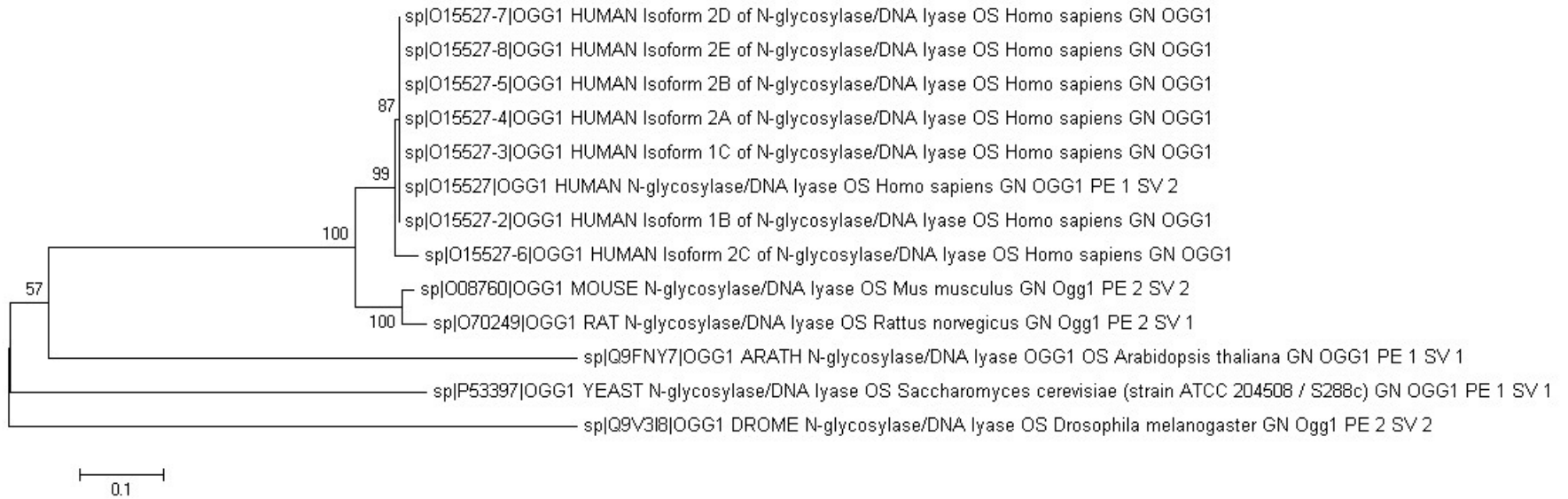
Different OGG1 transcripts (isoform1 and isoform2)

OGG1 transcripts



Different OGG1 transcripts (isoform1 and isoform2)

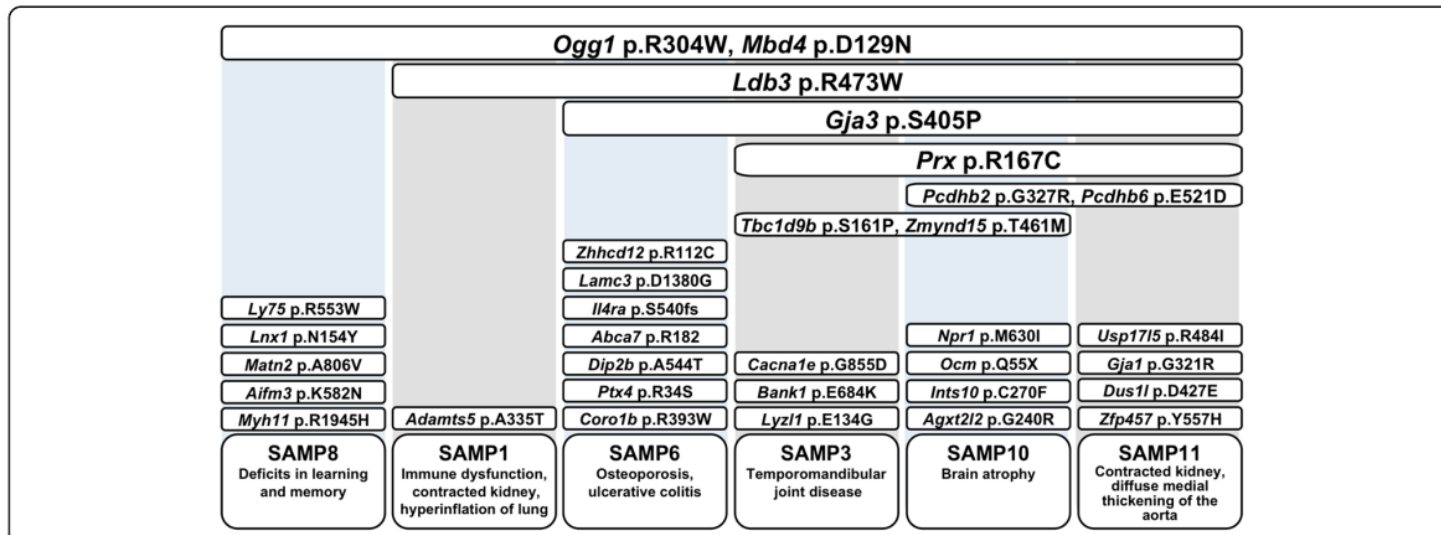
Phylogeny Tree



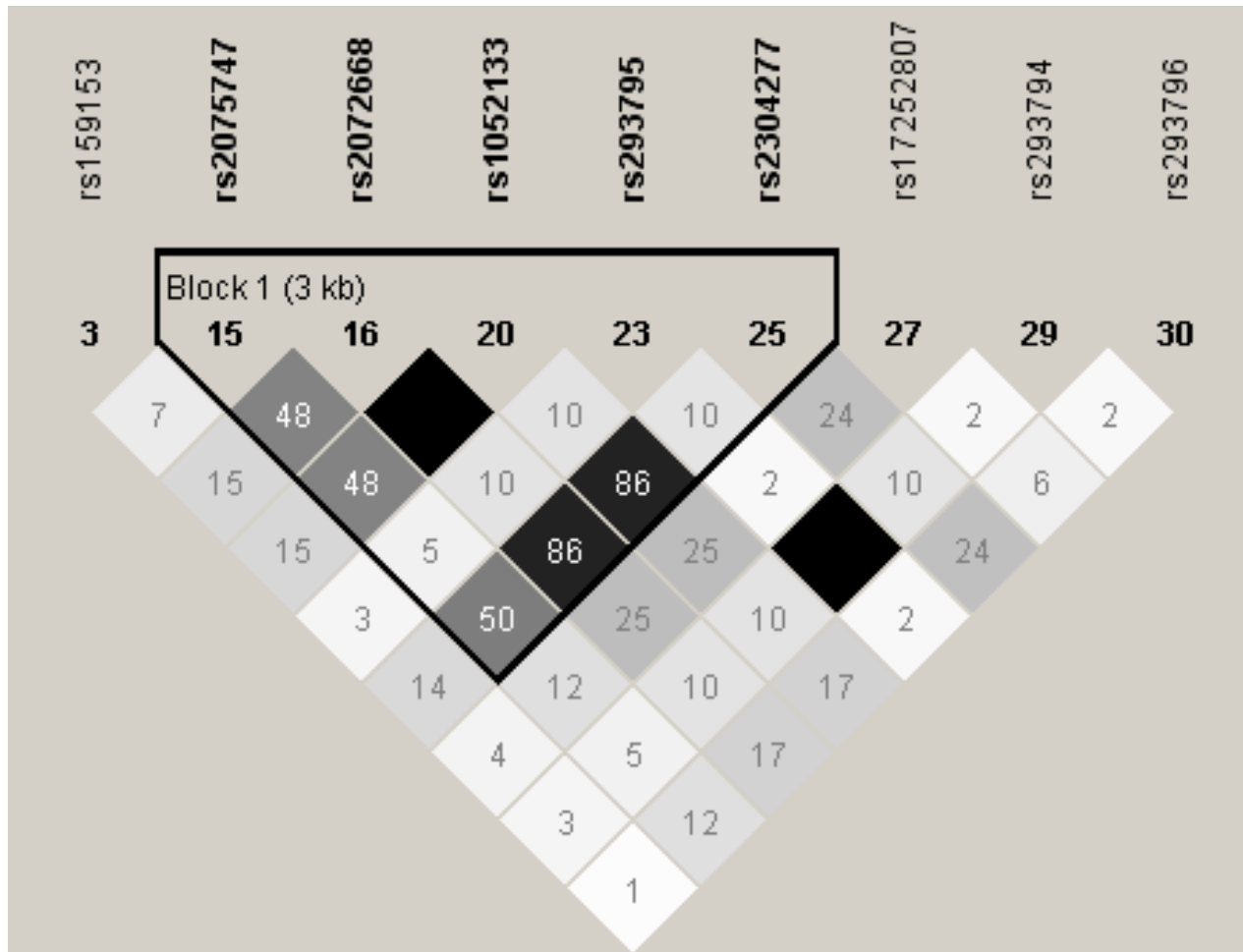
With Mega 6.0

OGG1 is associated with cancers and aging

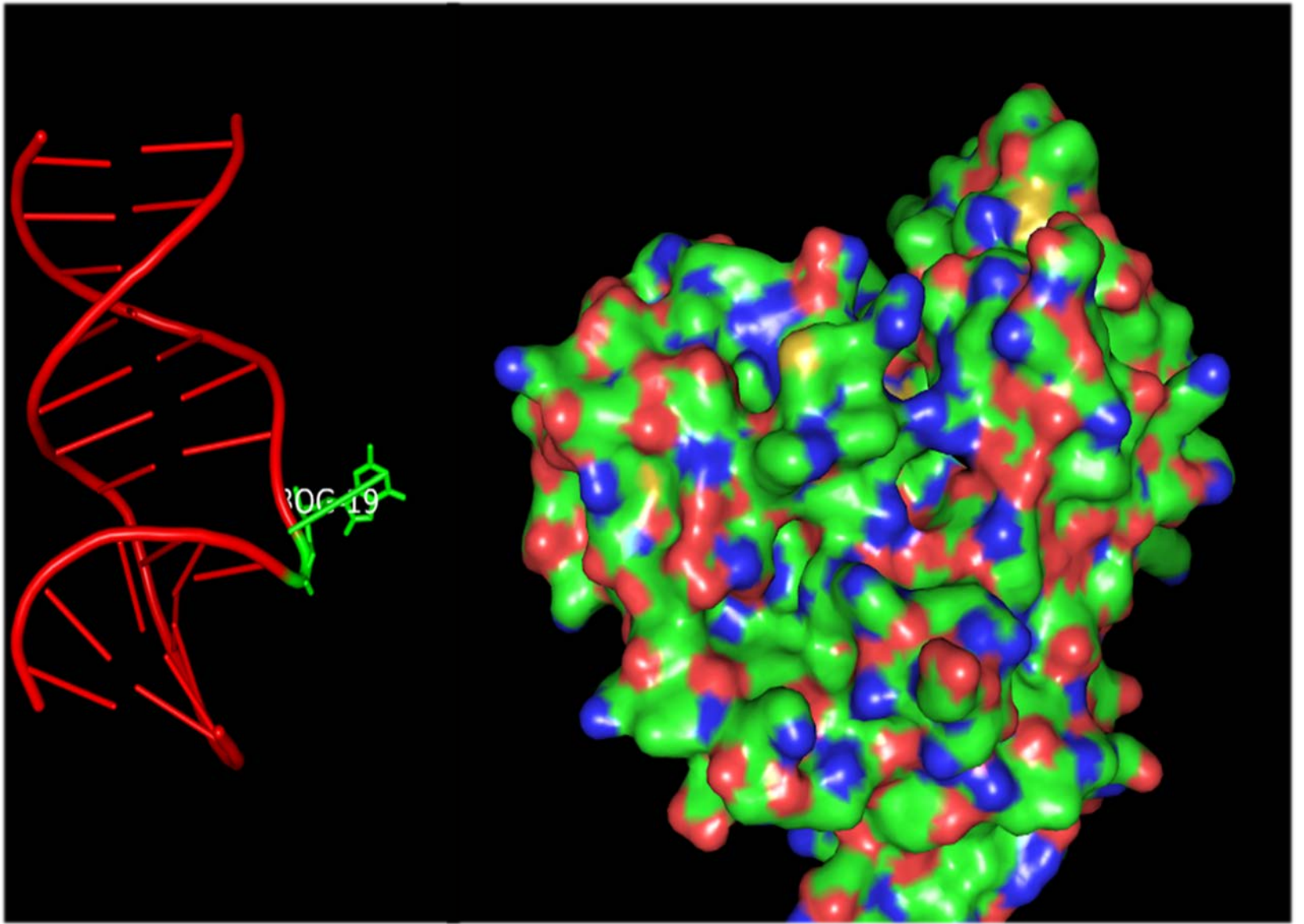
Site	Natural variants	Disease
12	G→E	Kidney cancer
46	R→Q	Renal cell carcinoma
85	A→S	Lung cancer
131	R→Q	Lung cancer
154	R→H	Gastric cancer
232	D→N	Kidney cancer
326	S→C	Various Cancers

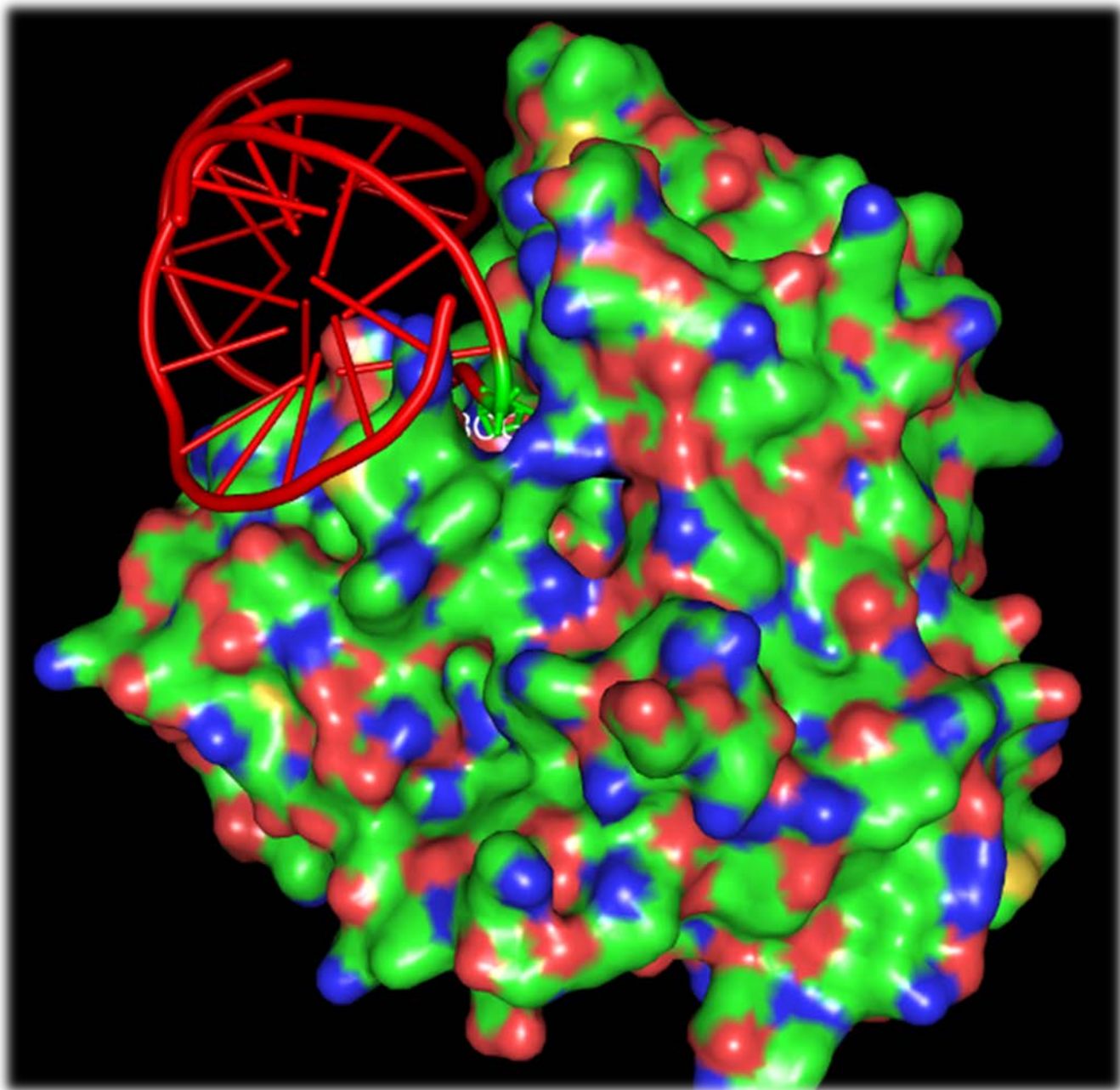


Genetic association study of OGG1 with centenarian trait



Selecting Tagging SNPs with Haploview

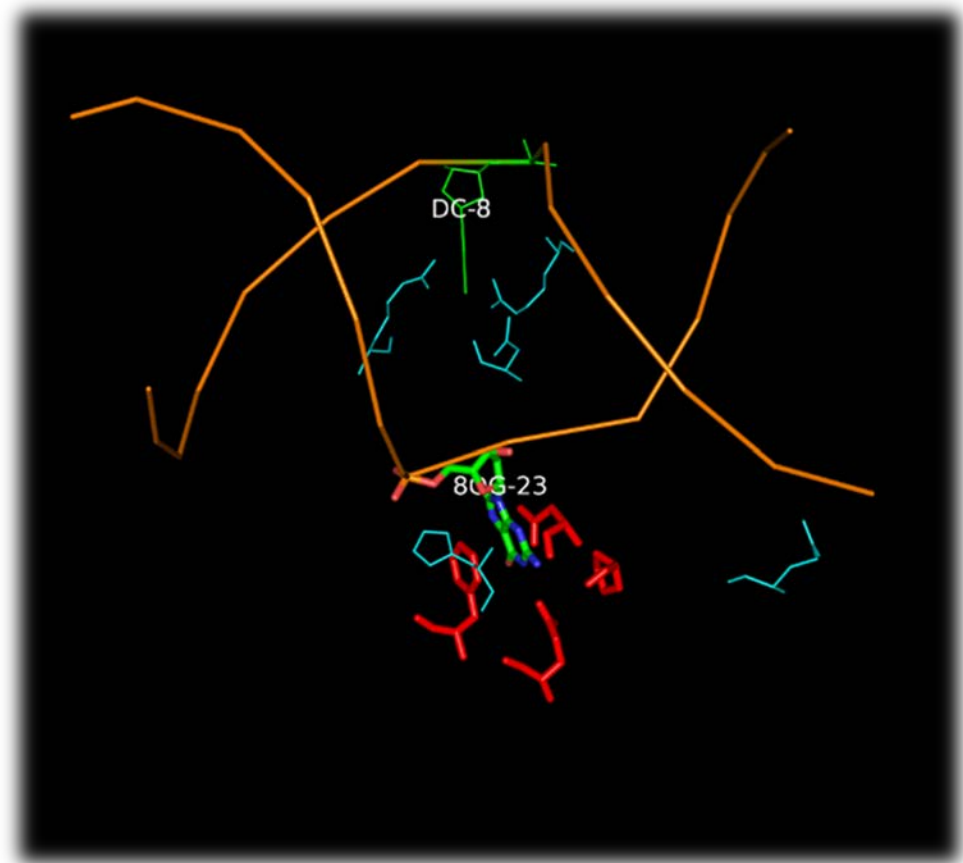




DNA Binding Sites

DNA binding sites

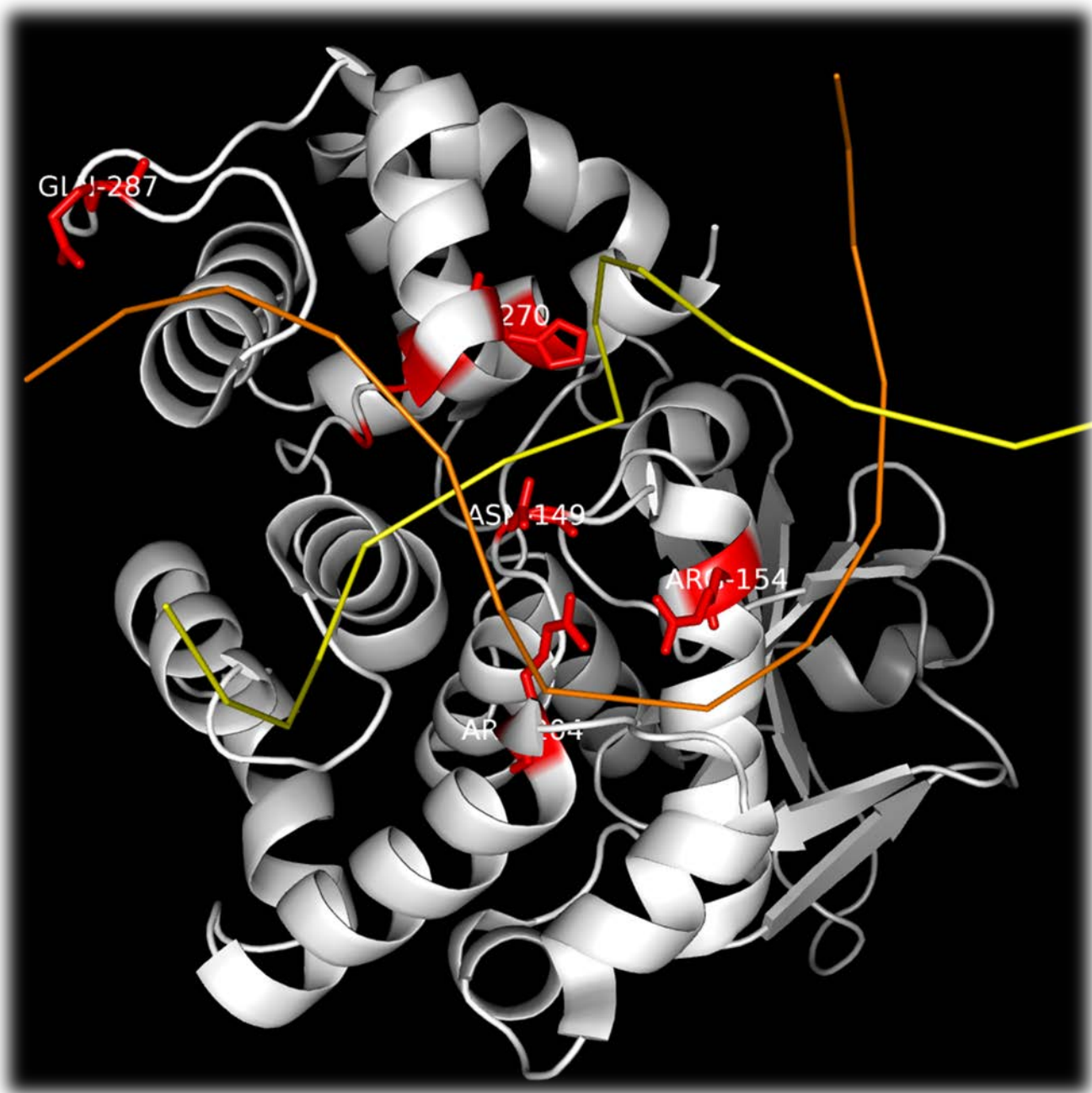
Position	Amino Acid	Binding
149	N	DNA
154	R	DNA
204	R	DNA
266	P	8-oxo-dGTP
268	D	8-oxo-dGTP
270	H	DNA
287	Q	DNA
315	Q	8-oxo-dGTP
319	F	8-oxo-dGTP

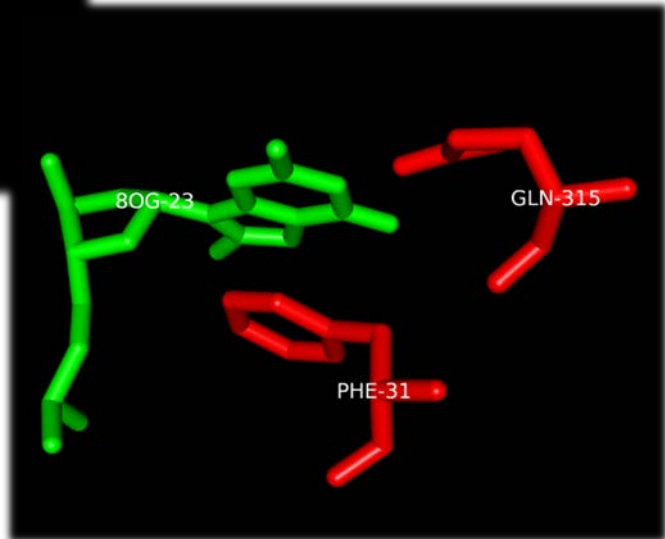
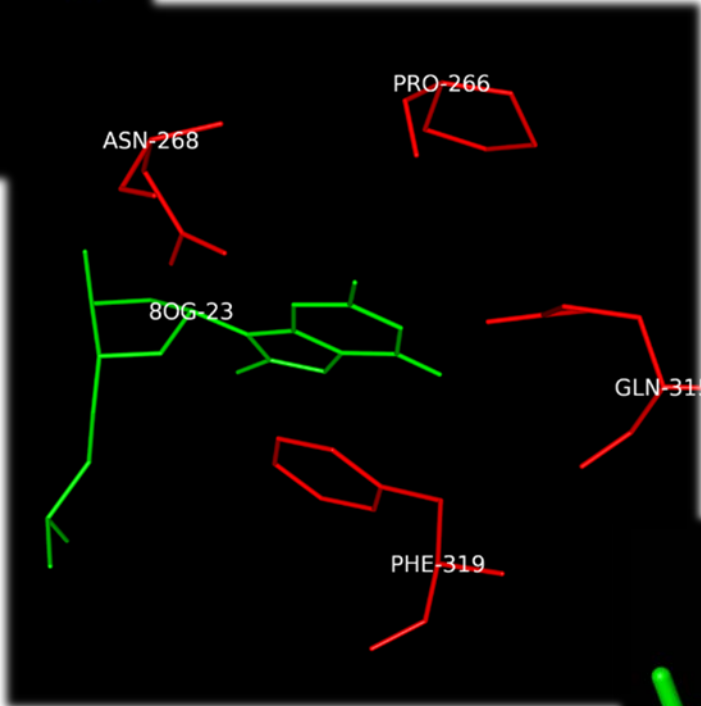
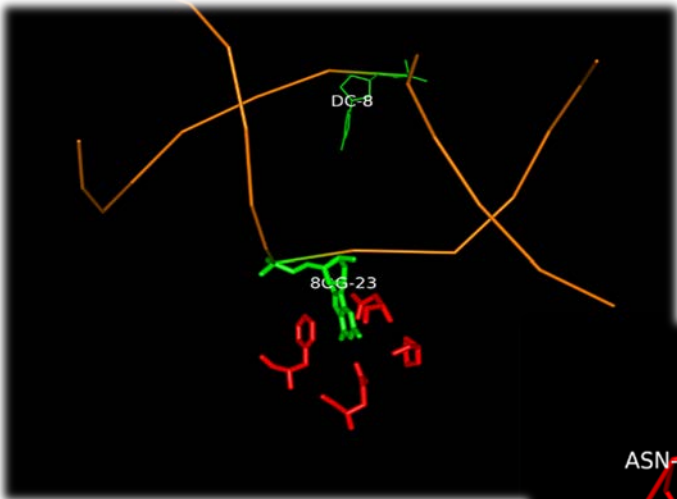


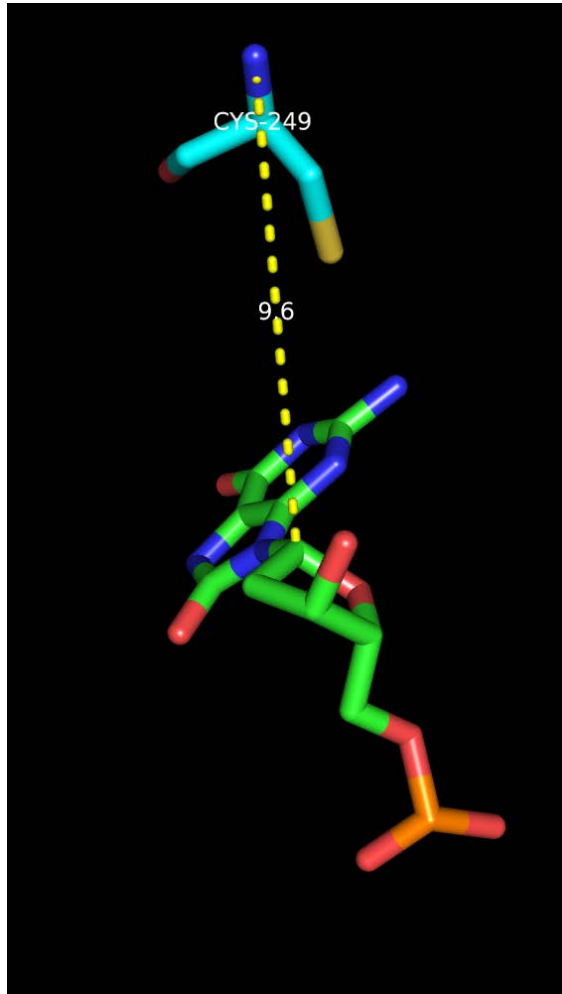
DNA Binding Sites

DNA binding sites are enriched in conserved sites

147	SSNNNIARITGMVERLCQAFGPRLIQLDDVTYHGFPNHLALA---GPEVEAHLRKLGLGY	203	O15527	OGG1_HUMAN
147	SSNNNIARITGMVERLCQAFGPRLIQLDDVTYHGFPNHLALA---GPEAETHLRKLGLGY	203	O08760	OGG1_MOUSE
147	SSNNNIARITGMVERLCQAFGPRLVQLDDVTYHGFPNHLALA---GPEVETHLRKLGLGY	203	O70249	OGG1_RAT
146	SCNNNIARISSMIEWFCATFGTKIGHFNGADAYTFPTINRFHDIPCEDLNAAQLRAAKFGY	205	Q9V3I8	OGG1_DROME
132	SSNNNISRITRMCNSLCSNFGNLITIDGVAYHSFPTSEELT---SRATEAKLRELGFY	188	P53397	OGG1_YEAST
160	SSNNNIARITKMVDF-VSSLGLHLGDIDGFQFPPSLDRLS----RVSEEEFRKAGFGY	214	Q9FNY7	OGG1_ARATH
	*.***** **: * : : * : : : ** . . : : : *			
204	RARYVSASARAILEEQGGL-----AWLQQ-LRESSYEEAHKALCILPGVGTQVADCICL	256	O15527	OGG1_HUMAN
204	RARYVRASAKAILEEQGFP-----AWLQQ-LRVAPYEEAHKALCTLPGVGAKVADCICL	256	O08760	OGG1_MOUSE
204	RARYVCASAKAILEEQGFP-----AWLQQ-LRVASYEEAHKALCTLPGVGTQVADCICL	256	O70249	OGG1_RAT
206	RAKFIAQTLQEIQKKG-GQ-----NWFIS-LKSMPFEKAREELTLLPGIGYKVADCICL	257	Q9V3I8	OGG1_DROME
189	RAKYIIEETARKLVNDKAEANITSDTTYLQSICKDAQYEDVREHLMSYNGVGPVKVADCVCL	248	P53397	OGG1_YEAST
215	RAKYITGTVNALQAKPGGGN-----EWLLS-LRKVELQEAVAALCTLPGVGPKVAACIAL	268	Q9FNY7	OGG1_ARATH
	**::: : . : . : : : : : : : * * : * * * * : . *			
257	MALDKPQAVPVDHVMWHIAQRDYSWHPPTSQAKGPS-----PQT-----NKELGNFF	303	O15527	OGG1_HUMAN
257	MALDKPQAVPVDHVMWQIAHRDYGWHPKTSQAKGPS-----PLA-----NKELGNFF	303	O08760	OGG1_MOUSE
257	MALDKPQAVPVDHVMWQIAHRDYGWQPKTSQTKGPS-----PLA-----NKELGNFF	303	O70249	OGG1_RAT
258	MSMGHLESVPVDHVIYRIAQNYLPHLTGQKQVTKK-----IYEEVSKHF	302	Q9V3I8	OGG1_DROME
249	MGLHMDGIVPVDHVSRIAKRDYQISANKNHLKELRTKYNALPISRKKINLELDHIRLML	308	P53397	OGG1_YEAST
269	FSLDQHSALPVDHVMWQIATNYLLPDLAGAKL-T-----PK-----LHGRVAEAF	312	Q9FNY7	OGG1_ARATH
	:: : : * * * * * : * * . : : :			
304	RSLWGFPYAGWQAVLFSADLRQSRHAQEP---AKRRKGS---KGPEG-----	345	O15527	OGG1_HUMAN
304	RNLWGFPYAGWQAVLFSADLRQPSLSREPP---AKRKKGS---KRPEG-----	345	O08760	OGG1_MOUSE
304	RNLWGFPYAGWQAVLFSADLRQQNLSREPP---AKRKKGS---KKTEG-----	345	O70249	OGG1_RAT
303	QKLHGKYAGWQAVLFSADLSQFQNTSTVACK---KKSNNK---KPKK-----	343	Q9V3I8	OGG1_DROME
309	FKKWGSYAGWQAVLFSKEIGGTSGSTTTG---TIKKRKWDMIKETEAIIVTKQMKLKVEL	365	P53397	OGG1_YEAST
313	VSKYGEYAGWQAVLFIAPQKTLTLLQSFSPINKLDESAEVNETSC---DTLKP----	365	Q9FNY7	OGG1_ARATH
	. * * * * * : * * : : :			







**CYS249 is associated with the
 β -elimination**

**How To Show the
Intermolecular Interactions?**

DNA Binding Sites

DNA binding prediction

Result: View:

NO HITCOUNT

```
#####  
# Program: helixturnhelix  
# Runcdate: Thu 9 Jan 2014 19:14:11  
# Commandline: helixturnhelix  
# -auto  
# -minsd 2.5  
# -sd 293.61  
# -mean 238.71  
# -sequence 695826  
# -outfile 696661  
# Report_format: motif  
# Report_file: 696661  
#####  
  
#-----  
#  
# Sequence: OGG1_HUMAN from: 1 to: 345  
# HitCount: 0  
#  
# Hits above +2.50 SD (972.73)  
#  
#-----  
  
Maximum_score_at at "*"   
  
#-----  
#-----  
  
#-----  
# Total_sequences: 1  
# Total_hitcount: 0  
#-----
```

With helixturnhelix

Natural variations are associated with cancers

Site	Natural variants	Disease
12	G→E	Kidney cancer
46	R→Q	Renal cell carcinoma
85	A→S	Lung cancer
131	R→Q	Lung cancer
154	R→H	Gastric cancer
232	D→N	Kidney cancer
326	S→C	Various Cancers

Natural variations are associated with cancers

Three conserved sites across species

46R

Species/Abbrv	Group Name									*	*		**	**	**	
1. sp 015527 OGG1 HUMAN N-glycosylase/		M	G	H	T	L	A	S	T	A	L	W	A	S	I	C
2. sp 015527-4 OGG1 HUMAN Isoform 2A c		M	G	H	T	L	A	S	T	A	L	W	A	S	I	C
3. sp 008760 OGG1 MOUSE N-glycosylase/		M	H	T	L	S	S	A	L	W	A	S	I	C	P	S
4. sp 070249 OGG1 RAT N-glycosylase/DM		M	H	T	L	I	S	A	L	W	A	S	I	C	P	S
5. sp Q9V3I8 OGG1 DROME N-glycosylase/		E	L	F	S	N	M	V	A	V	L	D	E	G	V	I
6. sp P53397 OGG1 YEAST N-glycosylase/		-	-	-	-	-	-	M	S	V	F	G	L	A	I	N
7. sp Q9FNY7 OGG1 ARATH N-glycosylase/		L	N	L	H	V	T	G	T	R	M	F	L	L	H	I

131R

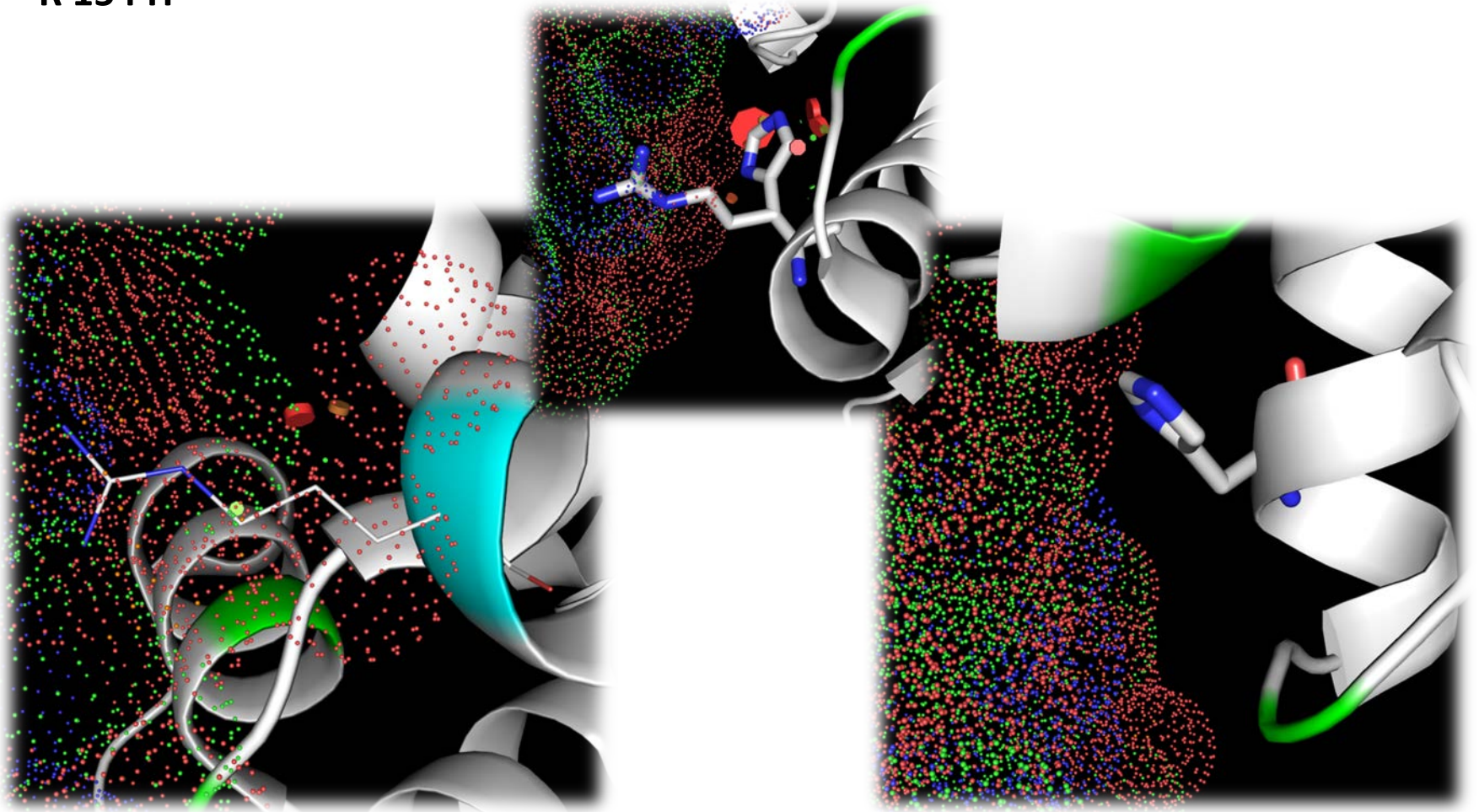
154R

Species/Abbrv	Group Name				*	*	*	*	*	*	*	*	*	*	*
1. sp 015527 OGG1 HUMAN N-glycosylase/		Q	F	Q	G	V	R	L	L	Q	D	P	I	E	C
2. sp 015527-4 OGG1 HUMAN Isoform 2A c		Q	F	Q	G	V	R	L	L	Q	D	P	I	E	C
3. sp 008760 OGG1 MOUSE N-glycosylase/		Q	F	Q	G	V	R	L	L	Q	D	P	I	E	C
4. sp 070249 OGG1 RAT N-glycosylase/DM		Q	F	Q	G	V	R	L	L	Q	D	P	I	E	C
5. sp Q9V3I8 OGG1 DROME N-glycosylase/		S	-	-	K	V	R	L	L	S	E	F	F	E	M
6. sp P53397 OGG1 YEAST N-glycosylase/		S	-	-	Q	G	I	R	I	L	A	E	P	M	E
7. sp Q9FNY7 OGG1 ARATH N-glycosylase/		S	H	L	G	A	R	V	L	Q	D	P	I	E	C

With MEGA 6.0

Natural variations are associated with cancers

R 154 H



Natural variations are associated with cancers

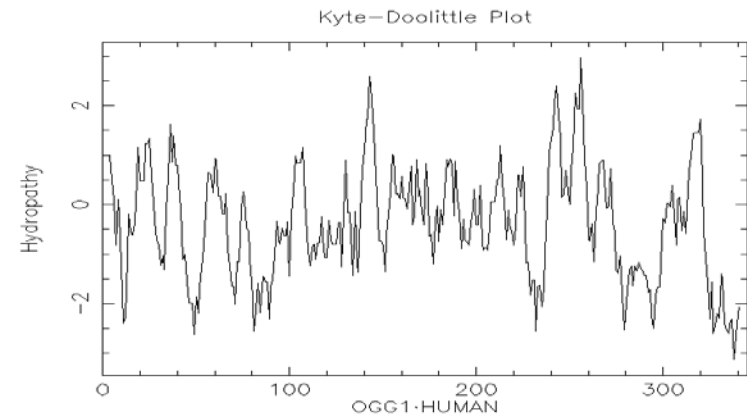
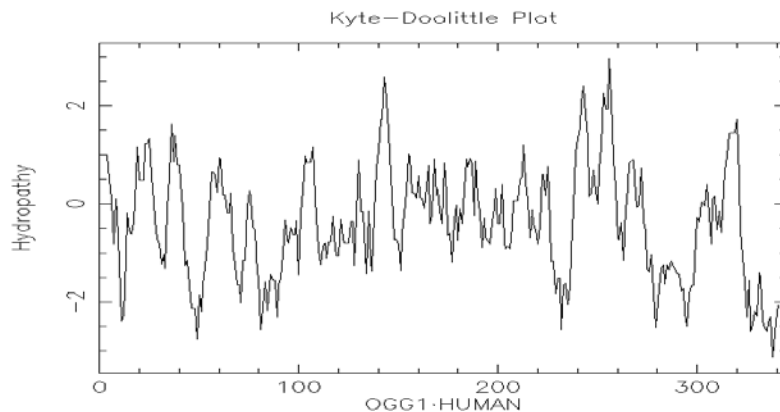
R 46 Q

44	S	0.400	44	S	0.200
45	F	0.200	45	F	0.000
46	R	0.200	46	Q	0.000
47	W	0.000	47	W	0.000
48	R	0.000	48	R	0.000
49	E	-0.200	49	E	-0.200
50	Q	0.100	50	Q	0.100

R 131 Q

125	Q	0.200	125	Q	0.200
126	K	0.200	126	K	0.200
127	F	0.200	127	F	0.000
128	Q	0.200	128	Q	0.000
129	G	0.200	129	G	0.000
130	V	0.400	130	V	0.200
131	R	0.400	131	Q	0.200
132	L	0.000	132	L	0.000
133	L	0.000	133	L	0.000
134	R	0.000	134	R	0.000
135	Q	-0.400	135	Q	-0.400

R 46 Q



With Kyte-Doolittle Plot

Subcellular location

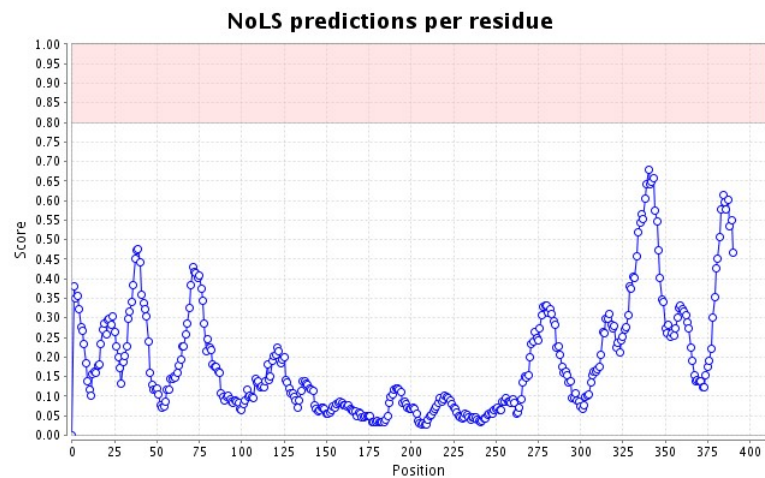
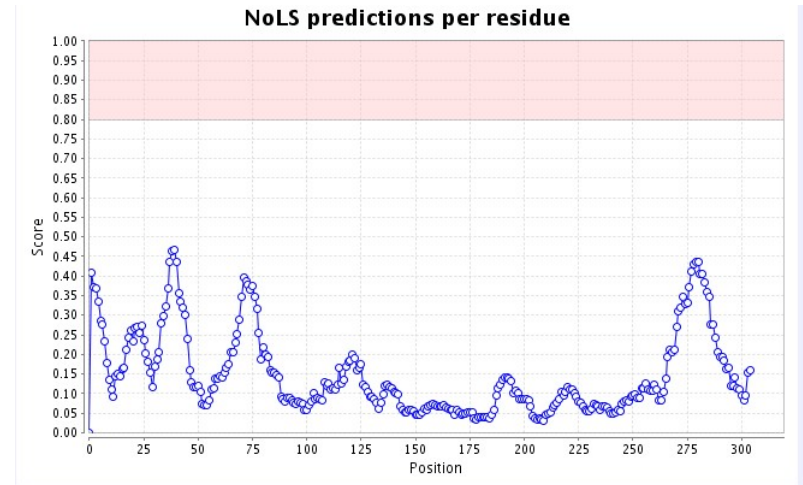
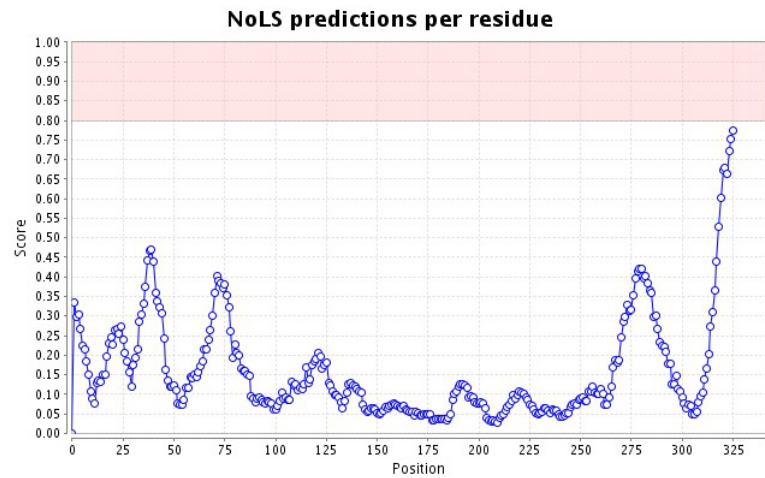
There is a NLS only in the C-terminal end of OGG1–1A, not 1B.



Kenichi Nishioka et al., *Molecular Biology of the Cell*, 1999

Subcellular location

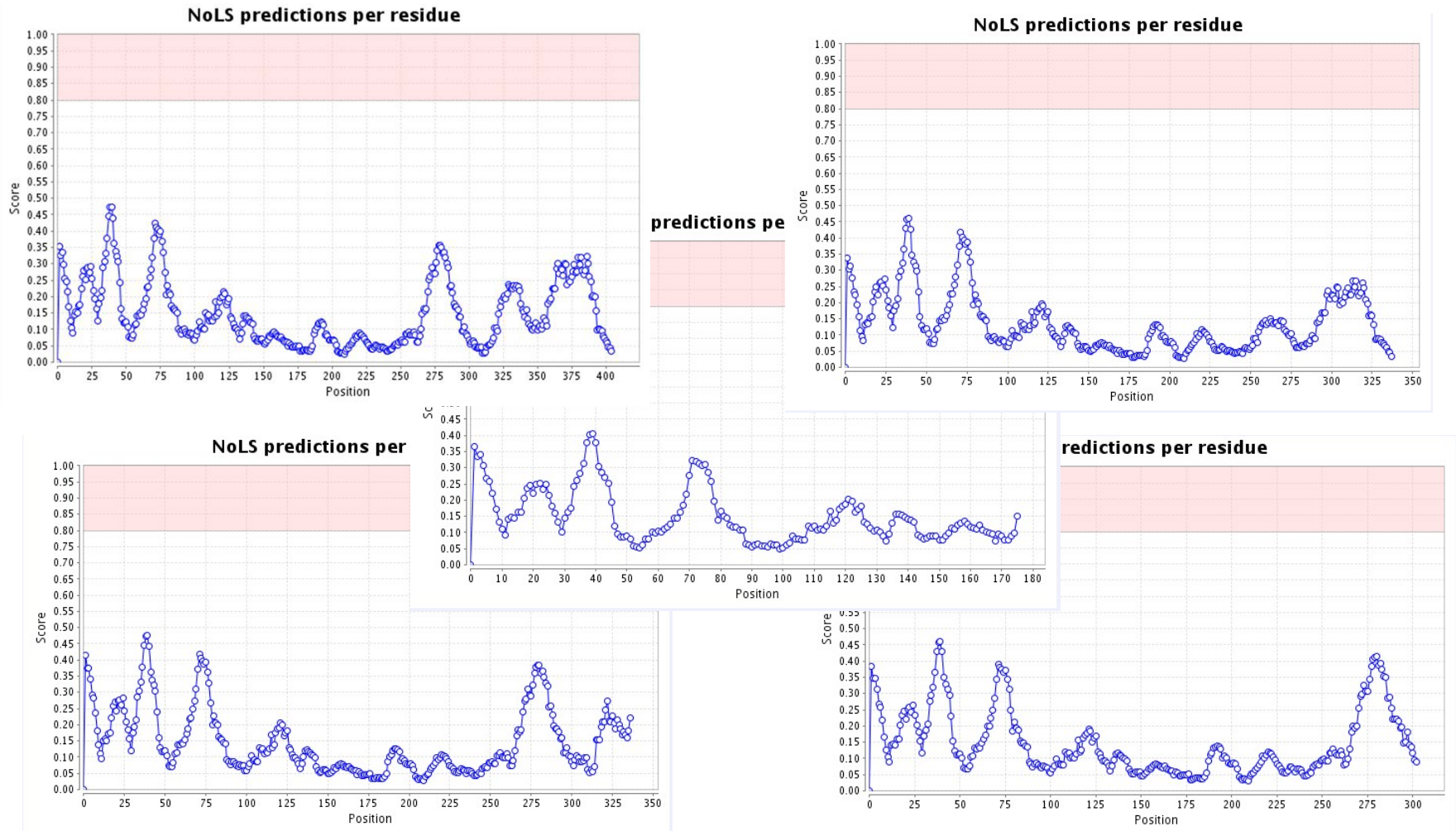
Different cellular location of OGG1 isoforms



332-338: KRRKGSK

Subcellular location

Different cellular location of OGG1 isoforms



THANK YOU