Article:
Del Mondo, Angelo, Iovinella, Manuela, Petriccione, Milena et al. (4 more authors) (2019) A Spotlight on Rad52 in Cyanidiophytina (Rhodophyta) : A Relic in Algal Heritage. Plants. 46.

https://doi.org/10.3390/plants8020046

A SPOTLIGHT ON *RAD52* IN *CYANIDIOPHYTINA (RHODOPHYTA)*: A RELIC IN ALGAL HERITAGE

Angelo Del Mondo[1], Manuela Iovinella[2], Milena Petriccione[3], Angelina Nunziata[3], Seth J. Davis[2], Diana Cioppa[1] and Claudia Ciniglia[4*]

[1]Department of Biology, University of Naples Federico II, Via Cinthia 21, 80126 Naples, Italy.

[2] Department of Biology, University of York, YO105DD York UK

[3] C.R.E.A.– Council for Agricultural Research and Economics – Research Centre for Olive, Citrus and Tree Fruit (OFA); Via Torrino 2; 81100 Caserta, Italy

[4]Department of Environmental, Biological and Pharmaceutical Science and Technology, University of Campania "L. Vanvitelli", Caserta, Italy

angelo.delmondo@unina.it

mi676@york.ac.uk

milena.petriccione@crea.gov.it

angelina.nunziata@crea.gov.it

seth.davis@york.ac.uk

claudia.ciniglia@unicampania.it

*Corresponding author: Claudia Ciniglia

Abstract

RADiation sensitive52 (RAD52) protein catalyzes the pairing between two homologous DNA sequences double-strand break repair and meiotic recombination, mediating RAD51 loading onto single-stranded DNA ends, and initiating homologous recombination and catalyzing DNA annealing. This article reports for the first time the presence of RAD52 homologs in the thermo-acidophilic Cyanidiophyceae whose genomes have undergone extensive sequencing. Database mining, phylogenetic inference, prediction of protein structure and evaluation of gene expression were performed in order to determine the functionality of RAD52 protein in Cyanidiophyceae. Our findings support that RAD52 gene and protein have an ancient origin, though it has been subsequently lost in all green algae and land plants. Its current function in Cyanidiophytina could be related to stress damage response for thriving in hot and acidic environments as well as to the genetic variability of these algae, in which – conversely to extant Rhodophyta - sexual mating was never observed.

Keywords RAD52, Homologous recombination, Cyanidiophytina*, Galdieria,* extremophiles

Introduction

Cyanidiophytina are unicellular red algae living in volcanic and post volcanic areas, where temperatures rise above 50°C, and high sulphuric acid concentrations, generated by the oxidation of sulphur gaseous emissions, greatly reduce the pH to values (pH 0.5-3.0) prohibitive for the majority of eukaryotic life forms [1–6]. The class includes three genera, the walled *Galdieria (G. sulphuraria, G. phlegrea, G.*

52 *maxima)* and *Cyanidium* (*C. caldarium, C. chilense*) and the naked *Cyanidioschyzon*
53 *(C. merolae).*
54 The long evolutionary history of Cyanidiophytina began around 1.5 BYA ([7–9],
55 before the formation of the supercontinent Rodinia (1.3-0.9 BYA), which resulted in
56 an increase in volcanic activity that would have favored the diversification and
57 dispersal of these thermoacidophilic algae [7–9].
58 According to Gross and Bhattacharya [10], the rising oxygenic atmosphere would
59 have exerted a selective pressure for efficient repair of ROS/UV-damaged DNA,
60 driving ultimately the evolution of sex, through cell-cell fusions, chromosome
61 movement, and emergence of the nuclear envelope, with the concurrent evolution of
62 meiosis and eukaryogenesis.
63 The occurrence of meiotic genes is not only related to genetic variation but it is also
64 involved in DNA repair [11]: one of the most threatening forms of DNA damage is
65 the break of the double helix (DSB), as both strands of the DNA duplex are impaired
66 simultaneously. The RAD52 epistasis group is implicated in various cellular
67 processes, such as recombinational repair and chromosome pairing in meisos, thus
68 guaranteeing the genome integrity; in particular, the RADiation sensitive52 (RAD52)
69 protein catalyzes the pairing between two homologous DNA sequences double-strand
70 break repair and meiotic recombination mediating the loading of RAD51 onto single-
71 stranded DNA ends, and thereby initiating homologous recombination and catalyzing
72 DNA annealing [12] RAD52 is recruited to the Replication Protein A (RPA)-single-
73 stranded DNA nucleoprotein complex, formed upon DSB induction and
74 exonucleolytic ends resection, and mediates its replacement by RAD51. RAD51 then
75 catalyzes strand invasion and D-loop formation. Eventually, RAD52 may assist in
76 capturing the second DNA end and promote its annealing to the D-loop, thus leading
77 to the formation of a Holliday junction [13].

78 RAD52 Epistasis Group also includes RAD50, RAD51, RAD54, RAD55, RAD57,
79 RAD59, RDH54, MRE11; they all cooperate in the process of homologous
80 recombination, playing an essential role in the mitotic and meiotic cell cycles, also
81 affecting the response to DNA damaging agents [12]. Homologues of the RAD52
82 group of genes have been identified in many eukaryotes, including animals and fungi
83 [14] and in some cases in prokaryotes [15] indicating high conservation of the
84 recombinational repair pathway. The lack of RAD52 in the vast majority of
85 photosynthetic protists, sexuated or not, is intriguing, considering its role in
86 homologous recombination process and its relatively high conservation across
87 eukaryotes. Even more unexpected is the presence of this key gene in the asexual red
88 algae *G. sulphuraria* and *C. merolae* genomes along with its absence in other
89 available genomes from sexuated Rhodophyta such as *Porphyra* and *Chondrus*.
90 The present paper displays the characterization of RAD52 homologs in *Galdieria*
91 *sulphuraria* genomes. The correspondence of the homologs to yeast and animal of the
92 RAD52 proteins was also provided. An in-depth sequence analysis of this protein
93 from 17 *Galdieria* strains was performed in order to delineate its evolutionary
94 relationship and phyletic horizon in available genomes. To exclude a relictic nature of
95 RAD52 sequences in *Galdieria*, selective pressures acting on the sequences were
96 detected by analysis of non-synonymous nucleotide substitutions over the number of
97 synonymous substitutions (Ka/Ks) [16–18]. The phylogenetic analyses were
98 combined with preliminary gene expression data on *Galdieria* in order to verify the
99 increasing of RAD52 mRNA expression during saline stress inducing DSBs.
100

101 **RESULTS AND DISCUSSION**

102 **RAD52 origin and distribution**

103 RAD52 gene homolog was identified in *G. sulphuraria* 074 genome (Gasu_26690,
104 Accession number M2XIH5). To support the identification of RAD52 homologs
105 within the genome of all analyzed taxa, a phylobayesian inference on protein
106 sequences was built (Fig. 1). Analyses showed that all the algal aminoacid sequences
107 were strongly supported as homologs of RAD52 excluding then being with RAD59
108 paralog; by the survey of the sequences, RAD52 appears to be sporadically distributed
109 both among bacteria and eukaryotes. RAD52 protein is commonly present in
110 Bacteria; among phototrophic bacteria, RAD52 was confirmed only for
111 *Synechococcus sp.* (Cyanophyta), and clusterized with significant posterior
112 probability (0.99) with *Spirochaete, Hyphomicrobium denitrificans* and
113 *Phaeomarinobacter ectocarpi*. Non-ambiguous blast hits included also Haptophyta
114 (*Emiliania huxleyi*), and Heterokontophyta (*Ectocarpus silicolosus, Phaeodactylum*
115 *tricornutum, Thalassiosira oceanica, Thalassiosira pseudonana*).
116 Within the phylogenetic tree, cyanidophycean RAD52 proteins formed a moderately
117 supported clade with the red algal group of Florideophyceae (*Gelidium*,
118 *Gracilariopsis* and *Calliarthron*), as sister clade of the RAD52 from Heterokonts
119 (*Phaeodactylum tricornutum, Thalassiosira oceanica, Thalassiosira pseudonana),*
120 with *Ectocarpus* positioned outside of this branch. Noteworthy, all these algal phyla
121 evolved through a secondary endosymbiosis in which a primary red algal cell would
122 have been acquired by a eukaryotic lineage [19]. Previous phylogenetic analyses
123 supported for a monophyletic origin of the plastids in cryptophytes, hapotophytes and
124 heterokonts. According to Oliveira and Bhattacharya [20], the

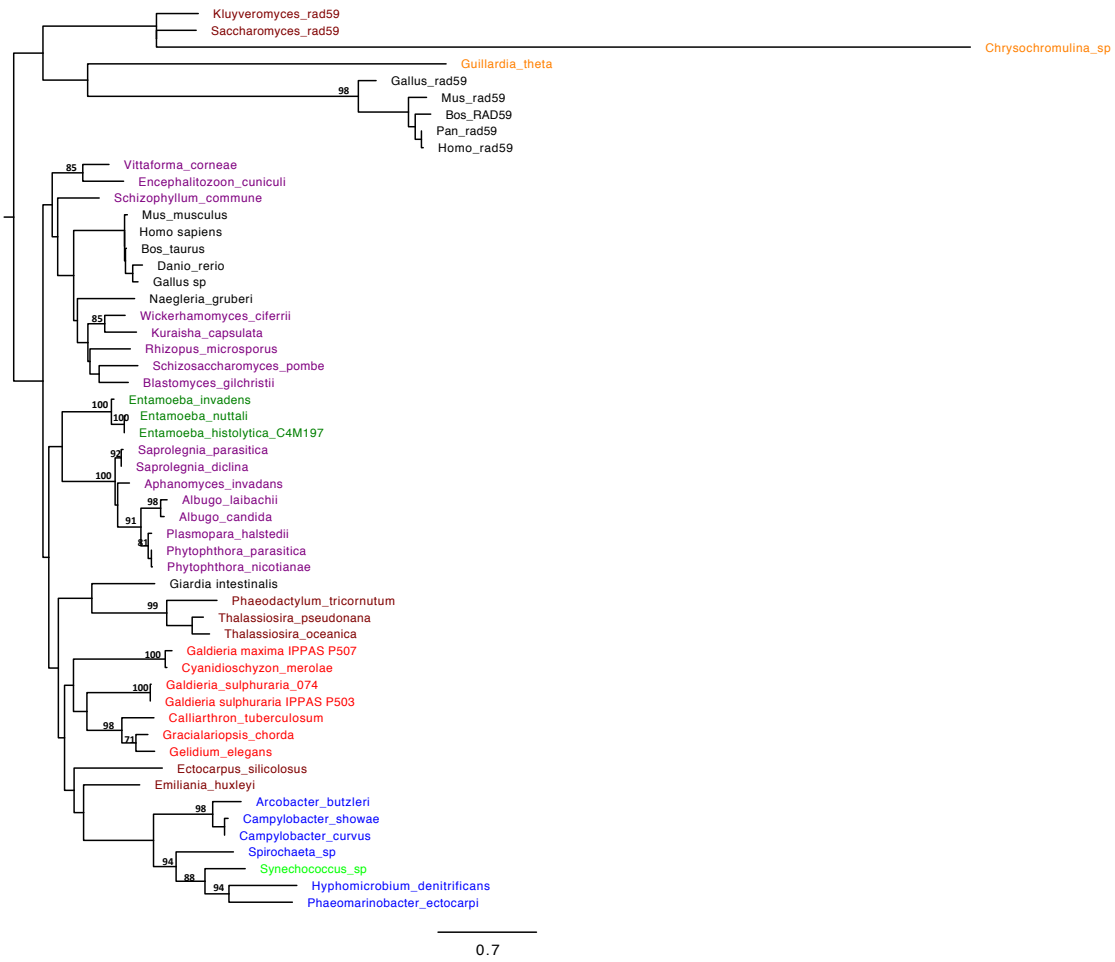


Fig. 1. RAD52 homologs, rooted with the RAD52 paralogs outgroup. 140 aligned amino acid sites from 54 taxa were analyzed; this consensus topology derived from >21.000 trees, $\alpha = 1.86$ ($1.45 < \alpha < 2.28$), pI = 7.269E-3 (7.4239E-8< pI < 0.0217) and lnL = –8952.79.

plastids of heterokonts would be most closely related to members of *Cyanidium-Galdieria* group, and not directly related to cryptophytes and haptophyte plastids, thus suggesting for these last an independent origin from different members of Bangiophycidae [20].

According to our investigations, the homology search for RAD52 in green algal genomes gave no results, as well as for Land Plants, Glaucophyta and Euglenophyta. However, the databases of protein, genomic, and transcribed (EST) sequences from the NCBI queried by Samach et al (2011) would have provided the evidence of RAD52-like proteins in several plants (monocotids and dicotids), as well as in some ferns and in filamentous (*Spyrogira pratensis*) and multicellular chlorophytes (*Chara vulgaris*). A gene duplication would have occurred according to Samach et al. [21] genome investigations: the green protists *S. pratensis* and *C. vulgaris* would possess only the paralog RAD52-1, whilst the gene would be lacking in Stramenopiles, Rhodophytes and unicellular Chlorophytes.

The level of similarity among RAD52 *G. sulphuraria* sequences ranged from 72 to 100%; the clustering reflects the phylogeny built on rbcL genes [5]: *G. sulphuraria* from Euroasiatic geothermal sites clusterized in an independent lineage (posterior probability= 0.89), but forming two well supported separate subclades: subclade I, including *G. sulphuraria* from Java and Russia (bp= 100%); subclade II, including both *G. sulphuraria* from Taiwan and *G. sulphuraria* from Iceland (bp= 100%). A second lineage included American accessions of *G. sulphuraria* clusterizing with Japanese and New Zealand strains, but into two well supported subclades (Fig. 2).

Fig. 2. Maximum likelihood tree for 24 newly sequenced *Galdieria* Rad52 gene. Only bootstrap values > 60% were reported.

189 **Support for functional homology of RAD52 protein in Cyanidiophytina**

190 The structure of RAD52 from Cyanidiophyceae was modeled on the base of the N-
191 terminal domain of human RAD52 [22]. In Figs. 3 and 4 results from Selecton
192 analysis are reported and related to information gained by I-Tasser. Results are shown
193 concerning M8 model. Ka/Ks ratio was never higher than 1, evidencing that no
194 divergent selection was detectable on analysed fragments. Values by MEC model
195 were not substantially different (data not shown). The longest conserved sequence
196 was made up of 36 residues that constitute 2 α-helix lining in the inner surface of the
197 DNA binding groove of the protein. Many other highly conserved residues were in
198 the first three β-sheets that constitute the outer surface of the DNA binding groove. In
199 β-sheets, conserved residues were flanked by non-conserved ones. All five AA (I4,
200 M9, Q59, K60 and V63) predicted as DNA binding by I-Tasser had highly conserved
201 pattern (evidenced by a yellow square in Fig.3 and a yellow halo in Fig. 4d, e). For
202 these residues, posterior probability evidenced a confidence interval for Ka/Ks
203 estimated between 2.60E-05 and 3.50E-01 for I4 and between 3.20E-04 and 2.40E-
204 01for all the others. Residues evidenced by a red square in Fig. 3 and a yellow in Fig.
205 4d,e are those predicted as DNA binding sites by Kagawa [22] (K129, R130 and
206 R133) and were highly conserved as well. The second part of the sequence, not
207 involved in the DNA binding groove formation, seemed not to be under puryfing
208 selection during *Galdieria* speciation. In Fig. 4c, the predicted model by I-Tasser was
209 shown, based on Singleton et al. [23] partial model for human RAD52 (Fig. 4a).
210 All these features supported the functional homology between RAD52 from
211 Cyanidiophyceae and the known RAD52 protein. To evaluate the functionality of
212 RAD52 and its role in repairing DNA damage by inducing homologous
213 recombination, the gene expression profile of RAD52 of *G. maxima* under salt-
214 stressed conditions was analyzed using real-time quantitative PCR (qPCR). RNAs

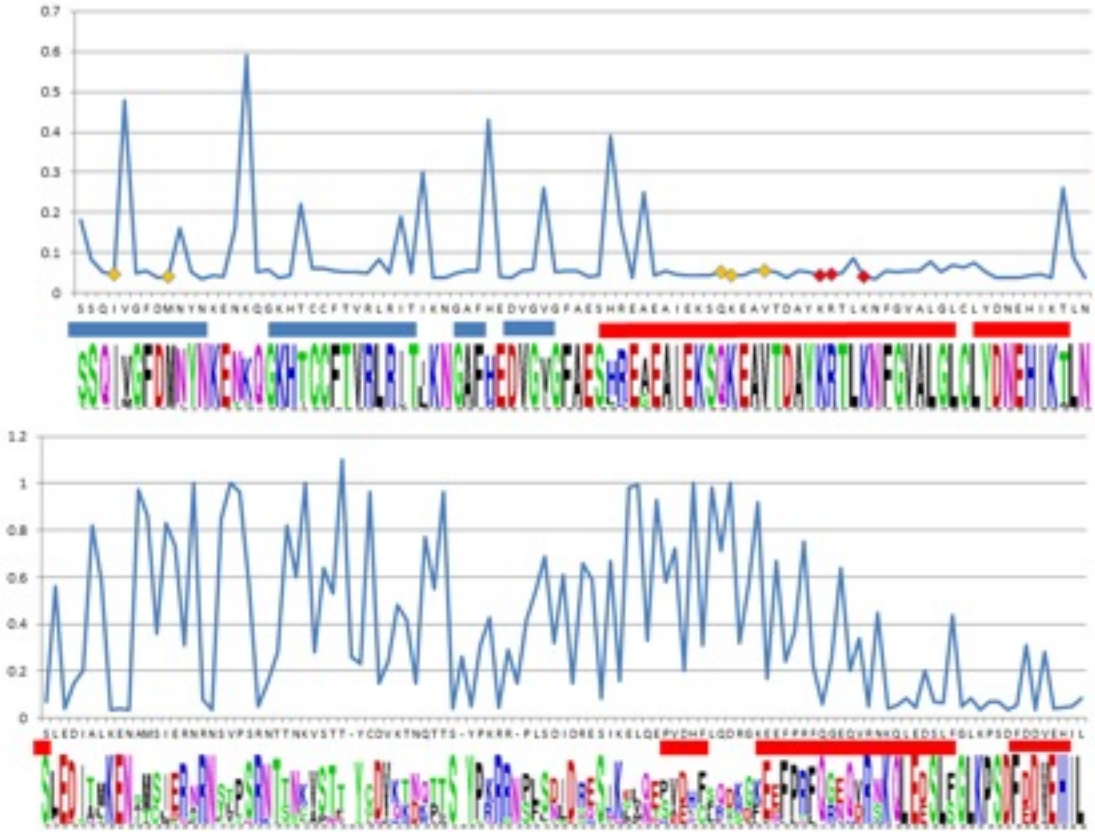
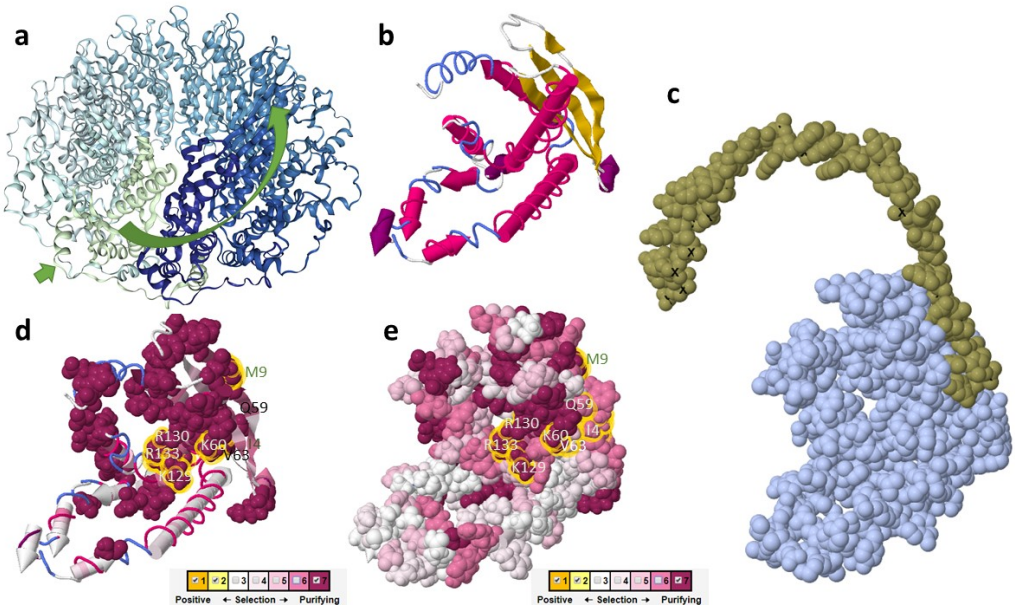


Fig. 3. Point value of Ka/Ks ratio along amino acidic sequence indicated by the Web-logo graphics. Values gained under M8 model. Amino acid participating in a β-sheet formation are underlined in blue, while α-helix are underlined in red. All the five AA (I4, M9, Q59, K60 and V63) predicted as DNA binding by I-Tasser are evidenced by a yellow square on the diagram. Residues evidenced by a red square on the diagram are those predicted as DNA binding sites by Kagawa [22] (; K129, R130 and R133)

229

230



231

232 Fig. 4. Three-dimensional representation of the structure predicted by I-TASSER
233 integrated with Selecton results; *a*, structure of human RAD 52 is reported with
234 the DNA binding groove evidenced and chains represented in different colours;
235 *b*, structure predicted by I-Tasser for the reference sequence used in the Selecton
236 analysis; *c,* DNA binding site as predicted by I-Tasser; *d*, Selecton results in M8
237 model reported on the predicted structure, 3D structures are represented as
238 cartoons with only strongly negatively selected sites highlighted. DNA binding
239 AA are highlited with yellow halos; *e*, Selecton results in M8 model reported on
240 the predicted structure, 3D structures are represented as spacefill. DNA binding
241 AA are highlited with yellow halos.

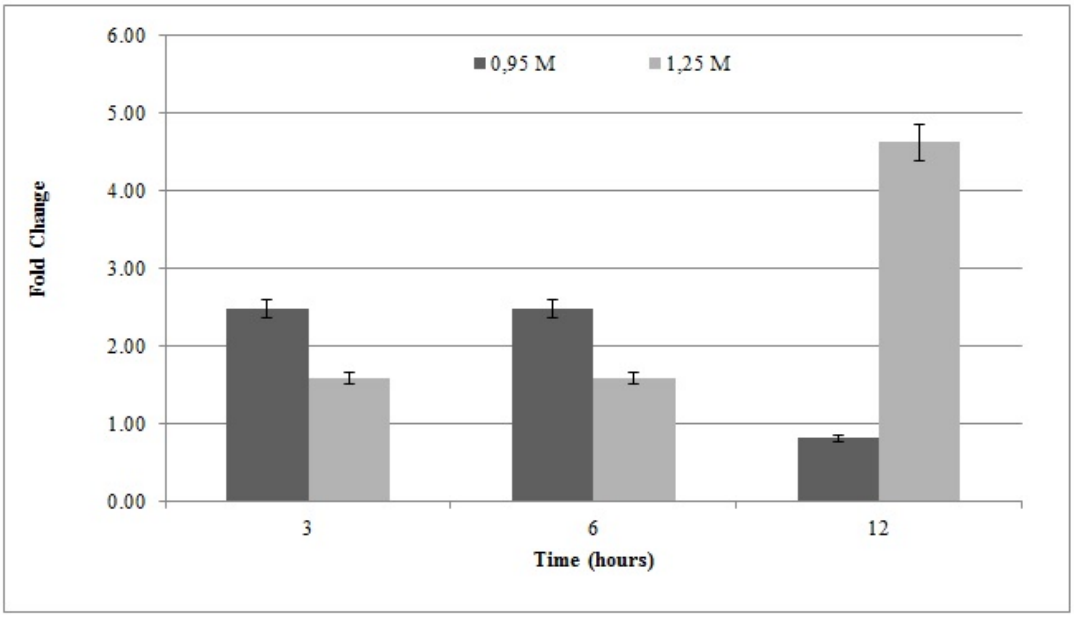
242

243

244

245

246 were extracted at multiple points (3,6 and 12 hours) from *G. maxima* cells under sub-
247 lethal and lethal NaCl (0,95M and 1,25M). RAD52 mRNA transcription levels
248 increased after salt-exposition at 1.25M NaCl with a significant up-regulation at 12
249 hours whereas at 0.95M NaCl the fold increase was higher compared to the control up
250 to 6 hours exposition but then a drastic decrease is observed after 12 hours (Fig. 5).
251 Accordingly with our expectations, RAD52 gene is present and plays an important
252 role in *Galdieria*. The observation of functional conserved residues in a RAD52
253 protein alignment showed that the catalytic activity of the protein may be conserved
254 not only in *Galdieria* but also in the other related algal organisms.
255
256



257
258 Fig. 5 .RAD52 gene expression in *G. sulphuraria ACUF 074* cells cultured under 0.95
259 M (dark grey bars) and 1.25M (light grey bars) NaCl. The mRNA levels were
260 normalized with respect to the level of mRNA for the reference genes (EF1α and
261 H2B). Bars show means ±SE from three independent experiments (n=3).
262
263

264 **Putative role of RAD52 protein in Cyanidiophytina**

265 The findings herewith reported show RAD52 homologs in the polyextremophilic red
266 algae Cyanidiophyceae; the conservation of predicted structures and of the amino acid
267 residues implicated in DNA binding strongly supports the hypothesis of a common
268 function between RAD52 from Cyanidiophyceae and the N-terminal domains of
269 RAD52 from previously described proteins. Cyanidiophyceae are likely to be the
270 oldest eukaryote with a RAD52 protein, in which it surely co-operates in DNA
271 damage response and maybe in other meiosis-like mechanism of genetic variability
272 (not shown); although RAD52 protein is lost for the most part in algae, it looks to be
273 conserved in algal lineages derived from an event of secondary endosymbiosis
274 involving a red alga, in which probably the ancestral RAD52 gene of the internalized
275 rhodophyte was re-arranged and conserved. Because of its key role in DNA repair
276 mechanism, RAD52 could have been retained as a relic heritage in some
277 photosynthetic eukaryotes still living in primordial-like environments, while lost in
278 others, even in closely related Rhodophyta with intricate life cycles. Being RAD52
279 gene crucial in meiotic machinery as well, its presence is probably also a hint for
280 looking at sexual behavior in putatively asexual Cyanidiophytina, inhabiting in
281 Archean environments where eukaryogenesis and meiosis co-evolved to reduce the
282 injuries in DNA of a rising oxygen atmosphere.

283 Interestingly, RAD52 sequences demonstrated to have undergone purifying selection
284 on all the part of the sequence involved in interaction with ssDNA and dsDNA. As
285 expected, mutations in such sites may reduce fitness and are therefore more likely to
286 be removed from the population (purified sites) [24]. In the remaining part of the
287 sequence, instead, several K, R and Y residues are conserved, interspersed in a
288 variable amino acidic context. As evidenced in human, these parts of the sequence are
289 responsible of the globular structure of each module or RAD52 and of the interactions

290 between modules. In such regions of the protein, a certain sequence variability is
291 compatible with the maintaining of the function.
292
293 **MATERIAL AND METHODS**
294 **Bioinformatics and phylogenetic analysis**
295 RAD52 nucleotide sequences of *G. sulphuraria* 074 (Java, Indonesia) and
296 *Cyanidioschyzon merolae* 10D (Japan) were retrieved from genome databases [25,26]
297 (http://www.ncbi.nlm.nih.gov/genbank) while 24 additional unannotated nucleotide
298 sequences of RAD52 from different *Galdieria* strains (10 *G. sulphuraria*, 14
299 *Galdieria* sp.) were obtained by MySeq Illumina data. RAD52 from *C. merolae* 10D
300 was retrieved from genome database and used as outgroup. For DNA extraction used
301 for Illumina, DNA was extracted by resuspending a stationary phase algal paste with
302 DNA extraction buffer [27]. DNA was incubated for 1 hr at 65℃, centrifuged and the
303 supernatant was precipitated by the addition of 1:1 isopropanol. The resultant pellet
304 was suspended in Qiagen buffer PB, then applied to a miniprep column and washed
305 according to manufacturers' details. DNA was eluted by adding pre-heated elution
306 buffer provided by Quiagen to the column in 4 sequential elution steps. The
307 sequencing was carried out as reported by Willing et al.[28]. After trimming, Illumina
308 MiSeq reads were assembled using Spades v3.1 [29].
309 RAD52 amino acid sequences were searched using the National Center for
310 Biotechnology Information (NCBI, http://blast.ncbi.nlm.nih.gov/Blast.cgi) by
311 querying protein, genomic and EST sequences on BLAST. A total of 45 RAD52
312 protein sequences from different organisms including algae, fungi, animals and
313 bacteria were recruited, and used to generate a multiple sequence alignment, together
314 with 9 RAD59 protein sequences as an outgroup. Among Cyanidiophytina, RAD52
315 protein sequences were retrieved from genome databases of *G. sulphuraria* 074 (Java,

316 Indonesia), *Cyanidioschyzon merolae* 10D (Japan) (Tables 1, 2)
317 (http://www.ncbi.nlm.nih.gov/genbank); [25,26] and *G. phlegrea* [30].
318 Phylogenetic inference of the evolutionary relationships of RAD52 from
319 Cyanidiophyceae and its homologs obtained from public databases was used to verify
320 the orthology of the protein; multiple alignment of amino acid sequences was
321 performed by ClustalW [31], trimmed and adjusted by eye. Only unambiguously
322 aligned amino acid sites were used for phylogenetic analyses. RAD52 phylogeny was
323 rooted by outgroup by using a RAD52 paralogue, RAD59. Bayesian analyses (BA)
324 were performed for combined and individual datasets with MrBayes v.3.1.1 [32]
325 using the Metropolis coupled Markov chain Monte Carlo (MC3) with the GTR + Γ +
326 I model. For each matrix, one million generations of two independent runs were
327 performed with sampling trees generated every 100 generations. The burnin period
328 was identified graphically by tracking the likelihoods at each generation to determine
329 whether they reached a plateau.
330 Maximum likelihood (ML) phylogenetic analysis was performed using the GTR + Γ
331 + I model implemented in RAxML software [33]. Statistical support for each branch
332 was obtained from 1000 bootstrap replications using the same substitution model and
333 RAxML program settings. The RAD52 evolutionary history of *Galdieria* strains was
334 inferred using Maximum likelihood (ML) method, based on Hasegawa-Kishino-Yano
335 model [34]. A discrete gamma distribution was used to model evolutionary rate
336 differences among sites. Bootstrap analyses were performed as previously described.
337
338
339
340
341

Taxa	GenBank ID
RAD52	
Albugo candida	635369772
Albugo laibachii	325180256
Aphanomyces invadans	673048395
Arcobacter butzleri	315478862
Blastomyces gilchristii	261192601
Bos taurus	528951193
Calliarthron tubercolosum	SRP005182
Campylobacter curvus	516863234
Campylobacter showae	489037738
Candidatus Phaeomarinobacter ectocarpii	918662481
Cyanidioschyzon merolae	544217672
Danio rerio	66269435
Ectocarpus silicolosus	298704860
Emiliania huxleyi	551599108
Encephalitozoon cuniculi	85014303
Entamoeba histolytica	67476176
Entamoeba invadens	471202697
Entamoeba nuttali	672809564
Galdieria sulphuraria IPPAS P507	
Galdieria sulphuraria IPPAS P503	MK21733250
Galdieria sp. ACUF074	MK217340
Gallus gallus	730466
Gracilaripsis chorda	NBIV01000177
Homo sapiens	863018
Hyphomicrobium denitrificans	505409238
Kuraisha capsulata	584391207
Mus musculus	261824011
Naegleria gruberi	290981385
Phaeodactylum tricornutum	219126773
Phytophthora nicotianae	970651832
Phytophthora parasitica	566015423
Plasmopara halstedii	953492183
Rhizopus microsporus	729702307
Saprolegnia diclina	669164116
Saprolegnia parasitica	813177361
Schizophyllum commune	302678737
Schizosaccharomyces pombe	19112088
Spirochaeta sp.	917473204
Synechococcus sp.	494162898

342

Taxa	GenBank ID
RAD52	
Thalassiosira oceanica	397635710
Thalassiosira pseudonana	220968365
Vittaforma corneae	667640414
Wickerhamomyces ciferrii	754409763
RAD59	
Bos taurus	61864423
Chrysochromulina sp.	922864786
Gallus gallus	45383087
Guillardia theta	551643257
Homo sapiens	21717826
Kluyveromyces lactis	49643317
Mus musculus	13385116
Pan troglodytes	55645233
Saccharomyces cerevisiae	6320144

343

344 Table 1. Accession numbers of RAD52 aminoacidic sequences used in this study

345

346

347

348

349

350

351

352

353

354

355

356

357

358

Strain	Strain code	Accession number
Galdieria sulphuraria	ACUF141G	MK217324
	ACUF141Y	MK217328
	ACUF141DG	MK217331
	ACUF142	MK217329
	ACUF388	MK217344
	ACUF402	MK217342
	ACUF427	MK217345
	ACUF455	MK217343
	SAG108.79	MK217327
	SAG21.92	MK217346
Galdieria sp.	IPPAS_P503	MK217325
	CCMEE5720	MK217326
	CCMEE5639	MK217330
	CCMEE5716	MK217332
	CCMEE5658	MK217333
	CCMEE5664	MK217334
	CCMEE5665	MK217335
	CCMEE5672	MK217336
	CCMEE5680	MK217337
	CCMEE5715	MK217338
	CCMEE5712	MK217339
	ACUF074	MK217340
	IPPAS_P502	MK217341
	THAL033	MK217347
Cyanidioschyzon merolae	10D	XM_005538923

Table 2. Accession number of RAD52 nucleotide sequences from Cyanidiophyceae used in this study

375 **2.2 *In silico* protein structure analysis**

376 The Selecton 2.4 Server (http://selecton.tau.ac.il/) was used to detect selection
377 affecting specific sites. The server program measures the Ka/Ks rate on each amino
378 acid residue [35–37]. Both M8 and MEC models were used. In M8 model, each
379 substitution that implies a different coded amino-acid is considered as non
380 synonymous, by contrast the mechanistic empirical combination model (MEC) takes
381 into account the differences between amino acid replacement probabilities, expanding
382 a 20 × 20 amino acid replacement rate matrix (such as the commonly used JTT
383 matrix) into a 61 × 61 sense-codon rate matrix. Confidence interval of Ka/Ks values
384 at each site were determined by posterior probability. The I-Tasser server
385 (http://zhanglab.ccmb.med. umich.edu/I-TASSER) was used to predict the 3D
386 structure of the domain and to map DNA binding sites especially conserved on the
387 examined sequences. A multi-alignment representation was draft by using WebLogo
388 application (http://weblogo.berkeley.edu/logo.cgi) and FirstGlance in JMolwas used
389 to visualize the 3D structure (http://bioinformatics.org/firstglance/fgij//index.htm).
390

391 **Rad52 gene expression under salt stress**

392 The functionality of RAD52 gene was also investigated by analyzing the gene
393 expression profile of the selected meiotic gene under osmotic stress conditions; *G.*
394 *sulphuraria* ACUF 074 was maintained in liquid culture in Allen medium [38], pH
395 1.5 at 37°C under a continuous irradiance of 60 µmol photons.$m^{-2}s^{-1}$. When in
396 exponential growth stage, the culture was supplemented with different NaCl
397 concentrations (0.16-2.5M). The growth rate was monitored until the stationary phase
398 and evaluated spectrophotometrically at 550nm. All test were prepared in triplicate.
399 Two NaCl stressed *G. sulphuraria* cultures with a sub-lethal (0,95M) and a lethal
400 (1,25M) salt concentration were then used to evaluate RAD52 mRNA levels after 3, 6

401 and 12 hours from the salt addiction. A qRT-PCR assay was performed on *G.*
402 *sulphuraria* ACUF 074. Total RNA was isolated by PureLink RNA Mini Kit (Thermo
403 Fisher Scientific, Waltham, MA USA), according to the manufacturer's instructions.
404 The RNA concentration was quantified by measuring the absorbance at 260 nm using
405 a Jasco V-530 UV/VIS spectrophotometer (Tokyo, Japan). The purity of all of the
406 RNA samples was assessed at an absorbance ratio of OD260/280 and OD260/230,
407 while its structural integrity was checked by agarose gel electrophoresis. Only high-
408 quality RNA with OD 260/280 and OD 260/230 >2 was used for subsequent steps.
409 Single-stranded cDNA was synthesized from 100 ng of total RNA using an
410 SuperScript® VILO™ cDNA Synthesis Kit (Thermo Fisher Scientific, Waltham, MA
411 USA), according to the manufacturer's instructions. EF1α and H2B were used as
412 housekeeping genes [39]. The amplification efficiency of each gene was determined
413 using a pool representing all of the cDNA samples. First, all of the primers were
414 examined by end-point PCR, all of the chosen target were expressed, and specific
415 amplification was confirmed by a single band of appropriate size in a 2% agarose gel
416 after electrophoresis. In a second step, the pool was used to generate a five-point
417 standard curve based on a ten-fold dilution series. The amplification efficiency (E)
418 and correlation coefficient (R^2) of the primers were calculated from the slope of the
419 standard curve according to the equation [40]:

420 $$E(\%) = (10^{(-1/slope)} - 1) \times 100$$

421 Quantitative Real-time-PCR was performed using a CFX Connect Real-time PCR
422 Detection System (Bio-Rad, Milan, Italy) to analyse the specific expression of each
423 reference/target gene. cDNA was amplified in 96-well plates using the
424 SsoAdvanced™ SYBR® Green Supermix (Bio-Rad, Milan, Italy), 15 ng of cDNA
425 and 300 nM specific sense and anti-sense primers in a final volume of 20 µl for each
426 well. Thermal cycling was performed, starting with an initial step at 95°C for 180 s,

427 followed by 40 cycles of denaturation at 95°C for 10 s and primer-dependent
428 annealing for 30 s. Each run was completed with a melting curve analysis to confirm
429 the specificity of amplification and lack of primer dimers.
430

431 **References**

432 1. Ciniglia, C.; Yoon, H.S.; Pollio, A.; Pinto, G.; Bhattacharya, D. Hidden
433 biodiversity of the extremophilic Cyanidiales red algae. *Mol. Ecol.* **2004**, *13*.
434 2. Pinto, G.; Ciniglia, C.; Cascone, C.; Pollio, A. Species Composition of
435 Cyanidiales Assemblages in Pisciarelli (Campi Flegrei, Italy) and Description
436 of Galdieria Phlegrea SP. NOV. In; Springer, Dordrecht, 2007; pp. 487–502.
437 3. Ciniglia, C.; Yang, E.C.; Pollio, A.; Pinto, G.; Iovinella, M.; Vitale, L.; Yoon,
438 H.S. Cyanidiophyceae in Iceland: plastid *rbc* L gene elucidates origin and
439 dispersal of extremophilic *Galdieria sulphuraria* and *G. maxima*
440 (Galdieriaceae, Rhodophyta). *Phycologia* **2014**, *53*, 542–551.
441 4. Cennamo, P.; Ciniglia, C. The algal diversity in the Phlegrean Fields
442 (Campania, Italy) archeological districts: an overview. *Upl. - J. Urban*
443 *Planning, Landsc. Environ. Des.* **2017**, *2*, 97–106.
444 5. Iovinella, M.; Eren, A.; Pinto, G.; Pollio, A.; Davis, S.J.; Cennamo, P.;
445 Ciniglia, C. Cryptic dispersal of Cyanidiophytina (Rhodophyta) in non-acidic
446 environments from Turkey. *Extremophiles* **2018**, *22*, 713–723.
447 6. Eren, A.; Iovinella, M.; Yoon, H.S.; Cennamo, P.; de Stefano, M.; de Castro,
448 O.; Ciniglia, C. Genetic structure of Galdieria populations from Iceland. *Polar*
449 *Biol.* **2018**, *41*.
450 7. Yang, E.C.; Boo, S.M.; Bhattacharya, D.; Saunders, G.W.; Knoll, A.H.;
451 Fredericq, S.; Graf, L.; Yoon, H.S. Divergence time estimates and the
452 evolution of major lineages in the florideophyte red algae. *Sci. Rep.* **2016**, *6*,

453 21361.
454 8. Müller, K.M.; Oliveira, M.C.; Sheath, R.G.; Bhattacharya, D. Ribosomal DNA
455 phylogeny of the Bangiophycidae (Rhodophyta) and the origin of secondary
456 plastids. *Am. J. Bot.* **2001**, *88*, 1390–1400.
457 9. Yoon, H.S.; Hackett, J.D.; Pinto, G.; Bhattacharya, D. The Single, Ancient
458 Origin of Chromist Plastids. *J. Phycol.* **2002**, *38*, 40–40.
459 10. Gross, J.; Bhattacharya, D. Uniting sex and eukaryote origins in an emerging
460 oxygenic world. *Biol. Direct* **2010**, *5*, 53.
461 11. Argueso, J.L.; Westmoreland, J.; Mieczkowski, P.A.; Gawel, M.; Petes, T.D.;
462 Resnick, M.A. Double-strand breaks associated with repetitive DNA can
463 reshape the genome. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 11845–50.
464 12. Symington, L.S. Role of RAD52 epistasis group genes in homologous
465 recombination and double-strand break repair. *Microbiol. Mol. Biol. Rev.* **2002**,
466 *66*, 630–70, table of contents.
467 13. Mortensen, U.H.; Bendixen, C.; Sunjevaric, I.; Rothstein, R. DNA strand
468 annealing is promoted by the yeast Rad52 protein. *Proc. Natl. Acad. Sci. U. S.*
469 *A.* **1996**, *93*, 10729–34.
470 14. Iyer, L.M.; Koonin, E. V; Aravind, L. Classification and evolutionary history
471 of the single-strand annealing proteins, RecT, Redβ, ERF and RAD52. *BMC*
472 *Genomics* **2002**, *3*, 8.
473 15. Aravind, L.; Walker, D.R.; Koonin, E. V. Conserved domains in DNA repair
474 proteins and evolution of repair systems. *Nucleic Acids Res.* **1999**, *27*, 1223–
475 1242.
476 16. Kreitman, M. M ETHODS TO D ETECT S ELECTION IN P OPULATIONS
477 WITH A PPLICATIONS TO THE H UMAN. *Annu. Rev. Genomics Hum.*
478 *Genet.* **2000**, *1*, 539–559.

479 17. MacCallum, C.; Hill, E. Being Positive about Selection. *PLoS Biol.* **2006**, *4*,
480 e87.
481 18. Roth, C.; Liberles, D.A. A systematic search for positive selection in higher
482 plants (Embryophytes). *BMC Plant Biol.* **2006**, *6*, 12.
483 19. Keeling, P.J. The endosymbiotic origin, diversification and fate of plastids.
484 *Philos. Trans. R. Soc. B Biol. Sci.* **2010**, *365*, 729–748.
485 20. Oliveira, M.C.; Bhattacharya, D. PHYLOGENY OF THE
486 BANGIOPHYCIDAE (RHODOPHYTA) AND THE SECONDARY
487 ENDOSYMBIOTIC ORIGIN OF ALGAL PLASTIDS. *J. Phycol.* **2000**, *36*,
488 52–52.
489 21. Samach, A.; Melamed-Bessudo, C.; Avivi-Ragolski, N.; Pietrokovski, S.;
490 Levy, A.A. Identification of plant RAD52 homologs and characterization of the
491 Arabidopsis thaliana RAD52-like genes. *Plant Cell* **2011**, *23*, 4266–79.
492 22. Kagawa, W.; Kurumizaka, H.; Ishitani, R.; Fukai, S.; Nureki, O.; Shibata, T.;
493 Yokoyama, S. Crystal Structure of the Homologous-Pairing Domain from the
494 Human Rad52 Recombinase in the Undecameric Form. *Mol. Cell* **2002**, *10*,
495 359–371.
496 23. Singleton, M.R.; Wentzell, L.M.; Liu, Y.; West, S.C.; Wigley, D.B. Structure
497 of the single-strand annealing domain of human RAD52 protein. *Proc. Natl.*
498 *Acad. Sci. U. S. A.* **2002**, *99*, 13492–7.
499 24. D., G. *Fundamental of Molecular Evolution*; Sinauer Press, Ed.; 2000;
500 25. Schönknecht, G.; Chen, W.-H.; Ternes, C.M.; Barbier, G.G.; Shrestha, R.P.;
501 Stanke, M.; Bräutigam, A.; Baker, B.J.; Banfield, J.F.; Garavito, R.M.; et al.
502 Gene transfer from bacteria and archaea facilitated evolution of an
503 extremophilic eukaryote. *Science* **2013**, *339*, 1207–10.
504 26. Matsuzaki, M.; Misumi, O.; Shin-i, T.; Maruyama, S.; Takahara, M.;

505 Miyagishima, S.; Mori, T.; Nishida, K.; Yagisawa, F.; Nishida, K.; et al.
506 Genome sequence of the ultrasmall unicellular red alga Cyanidioschyzon
507 merolae 10D. *Nature* **2004**, *428*, 653–657.
508 27. Davis, A.M.; Iovinella, M.; James, S.; Robshaw, T.; Dodson, J.H.; Herrero-
509 Davila, L.; Clark, J.H.; Agapiou, M.; McQueen-Mason, S.; Pinto, G.; et al.
510 Using MinION nanopore sequencing to generate a de novo eukaryotic draft
511 genome: preliminary physiological and genomic description of the
512 extremophilic red alga Galdieria sulphuraria strain SAG 107.79 - White Rose
513 Research Online Available online: http://eprints.whiterose.ac.uk/105094/
514 (accessed on Dec 18, 2018).
515 28. Willing, E.-M.; Rawat, V.; Mandáková, T.; Maumus, F.; James, G.V.;
516 Nordström, K.J.V.; Becker, C.; Warthmann, N.; Chica, C.; Szarzynska, B.; et
517 al. Genome expansion of Arabis alpina linked with retrotransposition and
518 reduced symmetric DNA methylation. *Nat. Plants* **2015**, *1*, 14023.
519 29. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov,
520 A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D.; et al. SPAdes: A
521 New Genome Assembly Algorithm and Its Applications to Single-Cell
522 Sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477.
523 30. Qiu, H.; Price, D.C.; Weber, A.P.M.; Reeb, V.; Yang, E.C.; Lee, J.M.; Kim,
524 S.Y.; Yoon, H.S.; Bhattacharya, D. Adaptation through horizontal gene transfer
525 in the cryptoendolithic red alga Galdieria phlegrea. *Curr. Biol.* **2013**, *23*, R865-
526 6.
527 31. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.;
528 McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal
529 W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948.
530 32. Ronquist, F.; Huelsenbeck, J.P. MrBayes 3: Bayesian phylogenetic inference

531 under mixed models. *Bioinformatics* **2003**, *19*, 1572–1574.
532 33. Stamatakis, A.; Hoover, P.; Rougemont, J. A Rapid Bootstrap Algorithm for
533 the RAxML Web Servers. *Syst. Biol.* **2008**, *57*, 758–771.
534 34. HASEGAWA, M.; YANO, T.; KISHINO, H. A new molecular clock of
535 mitochondrial DNA and the evolution of hominoids. *Proc. Japan Acad. Ser. B*
536 *Phys. Biol. Sci.* **1984**, *60*, 95–98.
537 35. Nielsen, R.; Yang, Z. Likelihood models for detecting positively selected
538 amino acid sites and applications to the HIV-1 envelope gene. *Genetics* **1998**,
539 *148*, 929–36.
540 36. Yang, Z.; Bielawski, J.P. Statistical methods for detecting molecular
541 adaptation. *Trends Ecol. Evol.* **2000**, *15*, 496–503.
542 37. Stern, A.; Doron-Faigenboim, A.; Erez, E.; Martz, E.; Bacharach, E.; Pupko, T.
543 Selecton 2007: advanced models for detecting positive and purifying selection
544 using a Bayesian inference approach. *Nucleic Acids Res.* **2007**, *35*, W506–
545 W511.
546 38. Allen, M.M. & Stainer, R.Y. Selective isolation of blue-green algae from water
547 and soil. *J Gen Microbiol* **1968**, *51*, 203–209.
548 39. Carfagna, S.; Bottone, C.; Cataletto, P.R.; Petriccione, M.; Pinto, G.; Salbitani,
549 G.; Vona, V.; Pollio, A.; Ciniglia, C. Impact of sulfur starvation in autotrophic
550 and heterotrophic cultures of the Extremophilic Microalga Galdieria Phlegrea
551 (Cyanidiophyceae). *Plant Cell Physiol.* **2016**, *57*.
552 40. Radonić, A.; Thulke, S.; Mackay, I.M.; Landt, O.; Siegert, W.; Nitsche, A.
553 Guideline to reference gene selection for quantitative real-time PCR. *Biochem.*
554 *Biophys. Res. Commun.* **2004**, *313*, 856–862.
555
556

557

558