



HAL
open science

Data on microsatellite markers in *Colletotrichum gloeosporioides* s.l., polymorphism levels and diversity range

Laurent Penet, Sophie Briand, Dalila Pétro, François Bussière, Sébastien Guyader

► To cite this version:

Laurent Penet, Sophie Briand, Dalila Pétro, François Bussière, Sébastien Guyader. Data on microsatellite markers in *Colletotrichum gloeosporioides* s.l., polymorphism levels and diversity range. *Data in Brief*, 2017, 12, pp.644-648. 10.1016/j.dib.2017.05.012 . hal-01608357

HAL Id: hal-01608357

<https://hal.science/hal-01608357v1>

Submitted on 25 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Data on microsatellite markers in *Colletotrichum gloeosporioides* s.l., polymorphism levels and diversity range



Laurent Penet^{*}, Sophie Briand, Dalila Petro, François Bussi re, S bastien Guyader

INRA, UR1321, ASTRO Agrosyst mes tropicaux, F-97170, Petit-Bourg (Guadeloupe) France

ARTICLE INFO

Article history:

Received 8 March 2017

Received in revised form

12 April 2017

Accepted 4 May 2017

Available online 11 May 2017

Keywords:

Colletotrichum gloeosporioides

Anthracnose disease

Microsatellites

Molecular markers

ABSTRACT

Colletotrichum gloeosporioides is a species complex of fungi belonging to the *Glomerellaceae* family (Ascomycota). It has a global worldwide occurrence and while sometimes described as a plant endophytic commensal, it also often demonstrates pathogenicity on crops and is responsible for anthracnose disease in many cultivated species. Thirty-nine polymorphic microsatellites were isolated and their polymorphism levels were determined in 95 strains from Guadeloupe (Lesser Antilles), mostly isolated from Water Yam (*Dioscorea alata*). The average allele number per polymorphic locus was 12.3 (decreasing to 4.3 at 5% frequency threshold, indicative of dramatic amounts of rare polymorphisms), with a range of 2–29 alleles. The microsatellite markers data will facilitate genetic diversity analyses and population genetics studies for the species complex.

  2017 Published by Elsevier Inc. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

Specifications Table

Subject area	Biology
More specific subject area	Microsatellite markers data (primers and expected diversity levels)

^{*} Corresponding author.

E-mail address: laurent.penet@inra.fr (L. Penet).

<http://dx.doi.org/10.1016/j.dib.2017.05.012>

2352-3409/  2017 Published by Elsevier Inc. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

Type of data	Table
How data was acquired	ABI PRISM 3730XL automated sequencer (MACROGEN)
Data format	Raw (primers information) and partially analyzed (diversity indices)
Experimental factors	Genomic DNA
Experimental features	Isolation of microsatellite markers and amplification test
Data source location	Guadeloupe 4°44.0694' N 53°46.881' W
Data accessibility	This manuscript (Table 1), primers are also available from probe data bank @ NCBI (www.ncbi.nlm.nih.gov/probe/)

Value of the data

- Large set of potentially polymorphic microsatellite markers in *Colletotrichum gloeosporioides*.
 - Diversity and genetic structure analyses at both fine and broad geographic scales.
 - Pathogenic strains genetic profiling.
 - Further *Colletotrichum gloeosporioides* species delineation (complimentary to sequencing data).
 - Origin of crop inocula and host origin analyses.
-

1. Data

This dataset is a list of 39 microsatellite markers from the worldwide pathogenic species complex *Colletotrichum gloeosporioides*, including primers and basic information relative to diversity levels expected at each locus. *Colletotrichum* fungi are diversified [1], with species ranging from genuine endophytic commensals to biotrophic parasites or even saprophytic pathogens [2]. Species of this genus are thus often associated with crop diseases, and especially anthracnose in plants [3–5]. Taxonomic studies are currently investigating sequence based delineation of species (DNA barcoding, e.g. [6–8]), but reaching consensus is still undergoing [9]. Defining co-dominant and highly polymorphic molecular markers such as microsatellites available for diversity studies and cross geographical or ecological comparisons would be a valuable tool for the study of this species complex and would allow introducing genetic data complementary to the current genomic approaches [9]. Also, these markers might allow differentiating genetic pools that could reflect host adaptation or even possibly identify new species within strain pools (structuration via reduced gene flow, e.g. [6]). We successfully developed 39 microsatellite markers for this wide geographical and ecological range pathogen (Table 1).

2. Experimental design, materials and methods

Genomic DNA was extracted from seven strains of *Colletotrichum gloeosporioides*. Six microsatellite-enriched genomic libraries were produced following [10]. DNA was digested with RsaI and fragments of 500 bp were ligated into a pCR 4-TOPO vector. These were then used to transform One Shot TOP10 chemically competent *Escherichia coli*, producing a total of 1158 positives clones and 128 were sequenced on an ABI PRISM 3730XL automated sequencer, using T3 and T7 primers. Consensus sequences were obtained using ChromasPro 1.34 software [11]. Of these sequences, 21 were of poor quality, 24 did not show microsatellite region, 24 were sister clones, and 59 showed microsatellites (motifs of three repetitions or more). Forty-nine primers pairs were thus designed using Primer-3 [12] and PrimerSelect of DNASTar [13].

Table 1

Characteristics for the 39 study microsatellite loci. Probe accession reference can be retrieved at www.ncbi.nlm.nih.gov/probe/. Size range in bp, includes rare alleles of high size. #A is the number of alleles in our 95 strains. F > x% is the number of alleles at a frequency higher than x% in the study sample. Ae is the efficient allele number (1/(1-Nei index)). In bold, 15 loci with amplification levels greater than 80%.

Locus name	Probe accession reference (NCBI)	Repeat motif	Forward Primer 5' → 3'	Reverse Primer 5' → 3'	Amplif. Success	Size range	#A	f > 5%	f > 1%	Ae	Nei index
Cg12	Pr032825007	tgg	GCAATGGAGCATGCAACTAA	TGGGTACTCTCACATACACG	76%	150–267	10	4	6	4.44	0.77
Cg14	Pr032825015	tcg	TCATTGTGCGCCATTTCTACG	GTCTCTGGCGGCTATGTTTC	37%	156–159	2	2	2	1.44	0.31
Cg16	Pr032825021	cac	ACAAGCAGTTTCTGGCTGCT	TGATGATGTCCGGTAGGTCCA	85%	102–276	9	3	7	3.41	0.71
Cg19	Pr032825026	gcc	GCGTTGTGAAATTGGACCTC	AGTCTTCAGCGTCAATCCGT	47%	103–224	10	3	6	3.24	0.69
Cg37	Pr032825027	gac	TTGCTGAAGCATACCGTGAG	AAGGTTTGAATTGTGTCCGC	45%	90–103	4	3	3	2.75	0.64
Cg53	Pr032825028	ttg	ACACCAGGAGAAACTCACCG	GGACCAGACAAGGACCAAA	69%	231–324	14	5	10	7.07	0.86
Cg57	Pr032825029	tcg	CCGTCTATTGGGGTAGCTGA	TGGTCATGGTGCATTTGAAG	100%	197–239	15	6	12	8.20	0.88
Cg68	Pr032825030	tcc	TGGTCTGCTTCTCGACATG	AGCCAAAGAGACCAAGCAAGA	87%	109–325	15	5	11	7.25	0.86
Cg71	Pr032825031	aac	TGATGGTTGTCATGGGATTC	GATCATGTCTCCATCCGCTC	47%	91–250	18	3	9	7.36	0.86
Cg83	Pr032825032	gt	GGATTTGTGCTGTGGGCTAT	GGACAAGAGAATGGAAGGACA	45%	122–218	16	9	16	6.71	0.85
Cg90	Pr032825033	gt	TAGCGTGATCGGAATGCGT	AGTGAATCGAATTGAAGGC	74%	176–294	12	4	8	5.78	0.83
Cg91	Pr032825034	ga	GGTTGCGACCAATGATCC	GACTCCGGTGAATAAGCCA	56%	94–136	6	3	6	2.36	0.58
Cg92	Pr032825035	tc	CATTTCCACAGCCACAC	GCAGCAGGTGTGAGAAGAGA	82%	92–250	29	7	13	17.3	0.94
Cg93	Pr032825036	tgg	TCTGTGTTGTGATGGTGACG	GCCCCAACCTTCTCTACTT	45%	86–234	11	6	7	5.45	0.82
Cg95	Pr032825037	ca	GGAGGTGGTTCGATAGTCGT	GGTTCGTTTGACACCACAA	80%	134–192	10	4	7	4.39	0.77
Cg96	Pr032825038	ga	ACGCGGAGGCATTACAGAG	GGAGTCCAATGTTGTGCGTA	92%	102–258	13	4	7	3.29	0.70
Cg97	Pr032825039	at	TTGTTGTGAAAGGAAAGGTTGA	AATCCCACGGGAGAATACAT	41%	112–152	4	2	3	2.11	0.53
Cg98	Pr032825040	tg	CGAGGCAAGCTGTAGCAGTA	TTCGTATTGTCTCCGTTCCC	56%	134–396	19	3	7	9.33	0.89
Cg100	Pr032825002	ag	GATGCATCTCGGGAGACC	CAATTCCCCACGAACATCTC	77%	78–128	9	5	7	4.46	0.78
Cg109	Pr032825003	gt	TCAAAAGACACGACCACGAC	CCATGGATGTGAGCATCATT	74%	130–190	7	3	6	3.94	0.75
Cg110	Pr032825004	ac	TGATACTGCGATGACAACCG	GGCATGGAGAGTCCGAACCTA	91%	94–152	8	3	5	3.07	0.67
Cg115	Pr032825005	cg	CACTGCTAGATGAGGCTTTGA	ACAAGTCCGAGCGGGAAGAC	89%	92–182	11	3	5	3.06	0.67
Cg116	Pr032825006	ca	CAATCTTATCCCAGGCTTTC	GCGGGGTTCAAGTACAGATA	62%	96–196	15	5	9	8.99	0.89
Cg120	Pr032825008	ac	AGTTTTCTGCTGAATCTCGCG	GCGAGCCATAGCCAAAGTAG	80%	86–176	26	6	16	13.2	0.92
Cg122	Pr032825009	ag	CTTTCGGTCAAGGTGTTG09	CTGGTCCCTTCAAATCTCC	73%	78–285	17	5	7	7.58	0.87
Cg127	Pr032825010	ac	GCTTGGTGGTTTAGCCAGTG	TTTGCCATATCCATGCTCTG	98%	208–268	10	4	6	3.57	0.72
Cg129	Pr032825011	ac	GACTCCAGCCACCACAAGAT	GATGCCGTTGTACCAGATCC	94%	72–240	18	6	10	5.58	0.82
Cg131	Pr032825012	ca	GAATCGATTGGGACGACG	AGCCGCTGCAGTTCCTTAA	83%	92–160	11	4	5	4.02	0.75
Cg136	Pr032825013	gt	AAGTCTCGAGTGGTGAAGGTG	TGCACTGACCTGACTGCTTT	66%	86–194	15	5	9	6.77	0.85
Cg137	Pr032825014	ga	GACCAGGTTGCAAGTCCGATA	GGATGAGGAGTAGAGGCGT	52%	162–262	21	3	9	11.6	0.91
Cg144	Pr032825016	ct	GCCTCCACCATCTATGGACT	CGGACTAGCGTAAGGCAAGA	34%	94–112	8	3	7	5.38	0.81
Cg149	Pr032825017	ga	ACCAGGGAATAAGGGGATG	CTGCCACTTCGGGTACCTT	78%	86–204	24	5	13	9.38	0.89
Cg150	Pr032825018	gt	TACCAGGGGTGGCAGCTC	GGTCCAGGACTCAAGCTC	75%	90–232	17	5	11	7.11	0.86
Cg156	Pr032825019	gtg	ACAGGGACATCCTGTTCACG	CCAGAATCGTCCGTTCAAGT	85%	87–285	15	5	7	4.65	0.79
Cg159	Pr032825020	cct	GTCCATCTTACCGCTGTTT	CGGTATCAACAACAACAATCA	83%	79–91	3	3	3	2.21	0.55
Cg161	Pr032825022	tacc	GGAGACAAGGAGGGGGC	GGATGCGGCTAGTAGGTAAGG	88%	87–267	15	4	10	3.18	0.69
Cg162	Pr032825023	aggt	GCCTTTGTGCTGCATGTAAG	TGCGCAAGATTCACTACTG	46%	134–150	4	3	4	3.71	0.73
Cg163	Pr032825024	accgc	CACAATACACAATCCACCACG	AGGATGCTATGCAGTCCAGG	62%	142–165	6	5	5	4.35	0.77
Cg164	Pr032825025	ctaca	ACACCGAACAGCGATCCTA	GACCTGAAGCGTGAATAG	49%	283–298	4	3	4	2.91	0.66

The primers were optimized for amplification, testing annealing temperature (44.5–64.2 °C), MgCl₂ concentration (1–3.5 mM), and polymerase chain reaction cycles (25–35). PCR conditions consisted of a denaturation stage at 95 °C for 5 min followed by 40 cycles at 95 °C for 30 s, 59 °C for 30 s, 72 °C for 30 s. Thirty-nine loci successfully amplified, all within expected sizes. In a further sample of 95 strains, polymorphism was assessed. High variability in alleles and Nei index were observed (Table 1). We report amplification success in single PCR runs, to help researchers chose loci more specifically. Indeed, *Colletotrichum gloeosporioides* demonstrate high phenotypic plasticity, possibly involving flexible DNA methylation, and amplification might vary depending on methylation state. We thus recommend choosing among these loci with a subsample study first.

In this polymorphism assessment, our strains were sampled from *Dioscorea alata* in Guadeloupe, where anthracnose is the main threat [14] and impacted agro-diversity [15]. Comparisons at wider geographical scales might enlighten important population processes: local dispersal [16], up to migration at greater scales [17], as well as genetic differentiation levels.

Acknowledgements

We are grateful to Thibaut Malausa, Nicolas Ris and Sylvie Warot (INRA, Institut Sophia Agrobiotech, France) for technical assistance. This work received the financial support of INRA through the Durayam project (No. P10259). The authors declare having no conflict of interests.

Transparency document. Supporting information

Transparency data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2017.05.012>.

References

- [1] P.F. Cannon, U. Damm, P.R. Johnston, B.S. Weir, *Colletotrichum* – current status and future directions, *Stud. Mycol.* 73 (2012) 181–213.
- [2] R. Rodriguez, R. Redman, More than 400 million years of evolution and some plants still can't make it on their own: plant stress tolerance via fungal symbiosis, *J. Exp. Bot.* 59 (2008) 1109–1114.
- [3] B.S. Weir, P.R. Johnston, U. Damm, The *Colletotrichum gloeosporioides* species complex, *Stud. Mycol.* 73 (2012) 115–180.
- [4] L. Cai, D. Udayanga, D.S. Manamgoda, S.S.N. Maharachchikumbura, E.H.C. McKenzie, L.D. Guo, X.Z. Liu, A. Bahkali, K.D. Hyde, The need to carry out re-inventory of plant pathogenic fungi, *Trop. Plant Pathol.* 36 (2011) 205–213.
- [5] C. Garrido, M. Carbú, F.J. Fernández-Acero, G. Budge, I. Vallejo, A. Colyer, J.M. Cantoral, Isolation and pathogenicity of *Colletotrichum* spp. causing anthracnose of strawberry in south west Spain, *Eur. J. Plant Pathol.* 120 (2008) 409–415.
- [6] V.P. Doyle, P.V. Oudemans, S.A. Rehner, A. Litt, Habitat and host indicate lineage identity in *Colletotrichum gloeosporioides* s.l. from wild and agricultural landscapes in North America, *PLoS One* 8 (2013) e62394.
- [7] D.N. Silva, P. Talhinhas, V. Varzea, L. Cai, O.S. Paulo, D. Batista, Application of the *Apn2/MAT* locus to improve the systematics of the *Colletotrichum gloeosporioides* complex: an example from coffee (*Coffea* spp.) hosts, *Mycologia* 104 (2012) 396–409.
- [8] M. Du, C.L. Schardl, E.M. Nuckles, L.J. Vaillancourt, Using mating-type gene sequences for improved phylogenetic resolution of *Colletotrichum* species complexes, *Mycologia* 97 (2005) 641–658.
- [9] P.F. Cannon, U. Damm, P.R. Johnston, B.S. Weir, *Colletotrichum* – current status and future directions, *Stud. Mycol.* 73 (2012) 181–213.
- [10] T.C. Glenn, N.A. Schable, Isolating microsatellite DNA loci method, *Enzymology* 395 (2005) 202–222.
- [11] C. McCarty, Chromaspro 1.34, Free Softw. (1998).
- [12] S. Rozen, H. Skaletsky, Primer3 on the WWW for general users and for biologist programmers, in: S. Misener, S.A. Krawetz (Eds.), *Bioinformatics Methods and Protocols*, Humana Press, Totowa, NJ, 1999, pp. 365–386.
- [13] T.G. Burland, DNASTAR's Lasergene Sequence Analysis Software, in: S. Misener (Ed.), *Bioinformatics Methods and Protocols*, Humana Press, Totowa, NJ, 1999, pp. 71–91.
- [14] L. Penet, E. Barthe, A. Alleyne, J.M. Blazy, Disease risk perception and diversity of management strategies by farmers: the case of anthracnose caused by *Colletotrichum gloeosporioides* on water yams (*Dioscorea alata*) in Guadeloupe, *Crop Prot.* 88 (2016) 7–17.
- [15] L. Penet, D. Cornet, J.M. Blazy, E. Barthe, F. Bussièrre, S. Guyader, C. Pavis, D. Pétro, Varietal Dynamics and Yam Agro-Diversity Demonstrate Complex Trajectories Intersecting Farmers' Strategies, Networks, and Disease Experience, *Front. Plant Sci.* 7 (2016).

- [16] L. Penet, S. Guyader, D. Pétro, M. Salles, F. Bussièrre, Direct Splash Dispersal Prevails over Indirect and Subsequent Spread during Rains in *Colletotrichum gloeosporioides* Infecting Yams, PLoS One 9 (2014) e115757.
- [17] S.N. Rampersad, D. Perez-Brito, C. Torres-Calzada, R. Tapia-Tussell, C.V. Carrington, Genetic structure and demographic history of *Colletotrichum gloeosporioides* sensu lato and *C. truncatum* isolates from Trinidad and Mexico, BMC Evol. Biol. 13 (2013) 1.