frontiers in PLANT SCIENCE

Plant-Microbe Interaction

Expanded genome sizes across the Pucciniales revealed by flow cytometry

Silvia Tavares, Ana Paula Ramos, Ana Sofia Pires, Helena Gil Azinheira, Patricia Caldeirinha, Tobias Link, Rita Abranches, Maria Do Céu Machado Lavado Silva, Ralf Thomas Voegele, João Loureiro and Pedro Talhinhas

Journal Name:	Frontiers in Plant Science
ISSN:	1664-462X
Article type:	Original Research Article
First received on:	30 May 2014
Frontiers website link:	www.frontiersin.org

Expanded genome sizes across the *Pucciniales* revealed by flow cytometry

Article type: Original Research
Number of words: 4019
Number of figures: 4
Running title: Genome sizes of rust fungi

Sílvia Tavares[1,2], Ana Paula Ramos[3], Ana Sofia Pires[1,2], Helena Gil Azinheira[1], Patrícia Caldeirinha[4], Tobias Link[5], Rita Abranches[2], Maria do Céu Silva[1], Ralf T. Voegele[5], João Loureiro[4], Pedro Talhinhas[1,2,3*]

1Centro de Investigação das Ferrugens do Cafeeiro, BioTrop, Instituto de Investigação Científica Tropical, Oeiras, Portugal
2Plant Cell Biology Laboratory, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal
3CEER-Biosystems Engeneering, Instituto Superior de Agronomia, Universidade de Lisboa, Lisbon, Portugal
4CFE, Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Coimbra, Portugal
5Institut für Phytomedizin, Universität Hohenheim, Stuttgart, Germany

* Correspondence:
Pedro Talhinhas
Centro de Investigação das Ferrugens do Cafeeiro
BioTrop
Instituto de Investigação Científica Tropical
Quinta do Marquês
2784-505 Oeiras
Portugal
ptalhinhas@iict.pt

Abstract

Having in common the capacity to infect plant hosts, rust fungi (phylum Basidiomycota, order Pucciniales) exhibit diverse complexities in their life cycles, host ranges and geographic distributions. Besides revealing specific aspects of the biotrophic life style, the completion of genome sequencing of a few rust fungi revealed the occurrence of large genomes, with a high content of transposable elements. Genome sequencing efforts of other rust fungi have been hampered by uncertainty concerning their genome sizes. In fact, flow cytometry was recently applied to estimate the genome size of a few rust fungi, having confirmed the occurrence of large genomes. In this work we have used an innovative and simple approach to isolate nuclei from the rust and host plant (either including or not a plant DNA reference standard) in order to estimate by flow cytometry the genome sizes of 30 rust species. Monoploid genome sizes varied over 10 fold, from 70 to 893 Mbp, with an average genome size value of 380.2 Mbp, which contrasts with the average value of 37.6 Mbp for all fungi and of 49.9 Mbp for the Basidiomycota. Comparing with over 1500 fungal genome size database entries, at 893.2 Mbp, *Gymnosporangium confusum* possess the largest fungal genome ever reported. Moreover, the vast majority of fungal genomes (95%) is smaller than the smallest rust genome determined in this study. Despite the fact that no correlation could be drawn between the genome sizes and the phylogenomics or the life cycle of the rust fungi analysed, interestingly, rusts with Poaceae hosts presented clearly smaller genomes than those with Fabaceae hosts. Although this study comprises only a very small fraction of the more than 7,000 rust species described so far, it seems already evident that the Pucciniales represent a group where genome size expansion could be a common characteristic, in sharp contrast with sister taxa, making this order of utmost interest for genomic research and for future studies.

Key-words: flow cytometry, genome size, *Gymnosporangium confusum* Dietel, largest fungal genomes, Pucciniales

1. Introduction

Rust fungi (phylum Basidiomycota, sub-phylum Pucciniomycotina, order Pucciniales) are plant pathogens characterized by the orange, brown or red colored spore repositories (Fig. 1) appearing on the leaf surface (Park and Wellings, 2012). Rusts cause major diseases of domesticated plants, being responsible for serious economic damage worldwide on agricultural and forest crops, including cereals and legumes as well as perennial plants. Rust epidemics have impacted the development of human society, such as the early accounts of cereals rust coming from the Bible and from Greek and Roman literatures (Park and Wellings, 2012), or the reports of rust coffee epidemics in Sri Lanka in the 19th century (Silva et al., 2006).

Rusts are obligate biotrophs, without a known saprotrophic phase. They depend entirely on living host cells to complete their biological cycle and only propagate in plant hosts (Duplessis et al., 2011b). Having more than 7000 described species, they constitute the largest group of plant pathogens (Aime, 2006). These fungi are cosmopolitan and almost ubiquitous, parasitising most families of angiosperms, but also ferns, mosses and gymnosperms (Park and Wellings, 2012). Rust fungi are generally highly specialised pathogens frequently having narrow host ranges, and consequently share a common evolutionary history with their host plants (Duplessis et al., 2011b).

The rust fungal life cycle is complex considering that it can involve up to five different spore forms, with varying nuclear composition, and can require two alternate phylogenetically distinct hosts, for completion of sexual and asexual cycles (Fernandez et al., 2013). Karyogamy occurs in the teliospores that germinate to produce the basidia where meiosis takes place, thus responsible for the sexual reproduction, although many rust fungal species have a life cycle conscripted to other spore stages, such as aeciospores or urediniospores (Aime, 2006). Urediniospores are usually accountable for the dissemination of the rust diseases (Aime, 2006).

So far, fungi have shown a remarkable variation in genome size. Indeed variations in chromosome number and size are far from being an exception and ploidy levels ranging from 1x to 50x have already been found (Gregory et al., 2007). For the few Pucciniales species analyzed so far, it has been revealed that genome size can be quite variable, with haploid size-estimates ranging from 64 to 418 Mbp (Eilam et al., 1994). This variation is often considered to be adaptive (Kelkar and Ochman, 2012), since variations in genome size of plant pathogens can have a direct impact in their pathogenicity (D'Hondt et al., 2011).

Most probably due to the smaller genome sizes of fungi in comparison with other organisms, only in the last two decades flow cytometry (FCM) was considered the method of choice for genome size determination studies in fungi, with important impacts in plant pathology (D'Hondt et al., 2011). Genome size is estimated by comparing the fluorescence emitted by an intercalating DNA fluorochrome between the sample and a reference standard with known genome size. Considering that a flow cytometer is readily available, the method provides reliable estimates of genome size in short period of time (in 10 min.) being considered a fast and relatively cheap alternative to other molecular tools (D'Hondt et al., 2011).

Still, rust species may pose technical constraints in the determination of genome size, as it can be especially difficult to extract nuclei in good quantity and quality from spores of obligate parasites.

Therefore, the objective of this study was to report for the first time the genome size of a large number of rust fungi collected from distinct hosts, while describing an innovative approach for obtaining nuclear suspensions of fungal species, in particular of obligate parasites, as rust species. For that, the chopping procedure of Galbraith et al. (1983) developed for plant tissues

was applied for the first time for plant pathogenic fungi, with the analyses being made directly on infected samples and not obligatorily on spores.

2. Material and Methods

2.1 Biological Material

A total of 32 rust samples were obtained from field surveys (during 2013 and 2014) as infected plant material, being subsequently identified by microscopic observation, or retrieved from active collections (as urediniospores), as detailed in Table 1. Infected plant material was preserved as dry herbarium specimens, stored at ISA/UL. Infected plant material was also employed directly for fungal (and plant) nuclear isolation and subsequently for flow cytometric analysis. Urediniospores (ca. 50 mg) were spread in sterile water in Petri dishes and incubated over-night at 25ºC to obtain germ tubes.

Plants used as reference for FCM were grown from seed (*Arabidopsis thaliana* 'Col-0', *Raphanus sativus* 'Saxa' and *Solanum lycopersicum* 'Stupické') and were maintained at the CFE/FCTUC, while urediniospores of the *Melampsora larici-populina* isolate 98AG31 were kindly provided by Dr. Pascal Frey (INRA Centre Nancy, France).

2.2 Fluorescence staining and microscopy

Spore collected from host leaves were stained with 4',6-diamidino- 2-phenylindole (DAPI; Sigma-Aldrich, USA), 1 µg/ml dissolved in water and slides were mounted in vectashield® (Vector Laboratories, Burlingame, CA, USA) an antifading agent (Johnson and Araujo 1981). Samples were observed in an epifluorescence microscope (Leica DMRB, DFC 340FX) equipped with a BP470/40 cube and an excitation wavelength of 340-380 nm. Pictures were captured with MetaMorph® software.

Infected leaf pieces, about 2 to 4 cm^2, were fixed overnight in a 2% solution of glutaraldehyde in 0.1 M sodium phosphate buffer at pH 7.2. Leaf pieces were then sectioned with a freezing microtome (Leica CM1850) and the sections (20-25 µm) were stained using an aqueous solution of 1 µg/ml DAPI, for 2h. The sections were then washed in distilled water, stained with an aqueous solution of 0.3% w/v diethanol for 2-3 seconds, washed again in distilled water and mounted in 50% v/v glycerol.

Leaf material was examined with light microscopes (Leitz Dialux 20 and Leica DM-2500) equipped with mercury bulbs HB 100W, ultra-violet light (excitation filter BP 340-380; barrier filter LP 430).

2.3 Flow cytometry

The nuclear DNA content of rust fungi was estimated by flow cytometry using infected leaf samples (occasionally from spores only), by comparison with the host plant genome size and/or, when the later was either unknown, uncertain or out of range, with healthy leaves of the plant DNA reference standards: *Arabidopsis thaliana* 'Col-0' (2C = 0.32 pg or 313 Mbp; this study after calibration with *Raphanus sativus* 'Saxa') *Raphanus sativus* 'Saxa' (2C= 1.11 pg or 1,086 Mbp; (Doležel et al., 1992) or *Solanum lycopersicum* 'Stupické' (2C = 1.96 pg or 1,917 Mbp; (Doležel et al., 1992) (Table 1). These plant reference standards were preferred to *Melampsora larici-populina*, as they provided unequivocal histograms with better quality of FCM results.

Nuclei were released from infected leaf tissues and/or leaves of the reference standards following the procedure of Galbraith et al. (1983). In brief, approx. 50 mg of both material (internal standardization) was chopped with a razor blade in a Petri dish with 1 mL of Woody

Plant Buffer (WPB; 0.2 M Tris–HCl, 4 mM $MgCl_2$, 1% Triton X-100, 2 mM Na_2EDTA, 86 mM NaCl, 20 mM sodium metabisulfite, 1% PVP-10, pH 7.5; Loureiro et al. 2007). Nuclei from spores were released by grinding approx. 10 mg of spores in a mortar in the presence of 1 mL of WPB. For the latter, nuclei from the plant DNA reference standard were added afterwards (pseudo-internal standardization).

The nuclear suspension was then filtered through a 30 µm nylon filter to remove plant and fungal debris, and 50 µg/mL of propidium iodide (PI; Fluka, Buchs, Switzerland) and 50 µg/mL of RNase (Fluka) were added to stain the nuclei and prevent staining of double-stranded RNA, respectively. After incubation for 5 min at room temperature, the fluorescence intensity of at least 3,000 nuclei per sample was analyzed using a Partec CyFlow Space flow cytometer (Partec GmbH, Görlitz, Germany), equipped with a green solid-state laser emitting at 532 nm for optimal PI excitation. For each rust fungal species, the G_1 peak of the standard species was set to a specific channel (usually between channel positions 500 and 750), with the amplification system kept at a constant voltage and gain throughout the analyses. Each day, prior to analysis, the overall instrument quality was assessed using calibration beads green concentrate (Partec GmbH). For each sample, when possible at least three independent replicate measurements were performed.

2.4 Data analysis

Results were acquired using Partec FloMax software v2.4d (Partec GmbH, Münster, Germany) in the form of four graphics: fluorescence pulse integral in linear scale (FL); forward light scatter (FS) vs. side light scatter (SS), both in logarithmic (log) scale; FL vs. time; and FL vs. SS in log scale. To analyze only intact nuclei, the FL histogram was gated with a polygonal region defined in the FL vs. SS cytogram.

The holoploid genome size in mass units (2C in pg; *sensu* Greilhuber et al., 2005) was assessed using the formula:

$$\frac{\text{Mean G1 fluorescence of sample nuclei}}{\text{Mean G1 fluorescence of reference standard}} \times \text{genome size of the reference standard}$$

The monoploid genome size (1Cx in Mbp; *sensu* Greilhuber et al., 2005) of all species was also calculated by dividing the holoploid genome size by the DNA ploidy level of each taxa. Conversion of mass values into numbers of base pairs was done according to the factor 1 pg = 978 Mbp (Doležel et al., 2005).

The reliability of the genome size measurements was verified by evaluating the quality of the flow cytometry histograms based on the coefficient of variation (CV) of the G_1 peaks and on the background debris, and by the CV of the genome size estimation of each isolate based on the independent measurements.

3. Results

Field surveys conducted in the Lisbon area (Portugal) over one year enabled the identification of several rust-infected plants. Both plants and fungi were identified by experienced botanists and mycologists. Samples collected (Table 1) comprised several botanical and mycological families. Most rusts were retrieved at the urediniosporic infection cycle, but telia and aecia were also readily found in certain rusts (Figs. 1 and 2). The application of the nuclear isolation protocol to rust-infected leaves (Fig 1) enabled the release of intact nuclei from both the host plant and the fungus cells that, according with the clearly defined G_1 peaks of both organisms (Fis. 3), could be efficiently stained with PI. The assignment of each peak to the rust fungi, host plant and plant reference standard was further confirmed by analyzing separately, healthy plant samples and fungal hyphae obtained upon germination of urediniospores in water (Fig. 2). Following this innovative approach, thirty rust fungi species were analyzed by flow cytometry (Table 1).

The genome size determinations based on the fungal G_1 fluorescence peaks had CV values below 10 % (usually between 4 and 7%), which is within the range of accepted values for fungal species (Bourne et al., 2014), and, for each analyzed isolate, CV measures of genome size estimations never exceeded 10 % (Table 1). Cytograms of FS vs. SS (Fig. 3) indicated that the identified nuclei populations were formed by uniform nuclei in size and shape. All parameters considered, high-quality histograms and cytograms were obtained for the analysis of the genome size.

When the genome size of the host plant was known and appropriate (i.e., when it appeared in the same scale set as the fungal species), the host plant itself was used as primary reference standard, otherwise, *Arabidopsis thaliana*, *Raphanus sativus* or *Solanum lycopersicum* were used as reference genomes according with the genome size of the fungal species. The analysis was not affected by the endopolyploid nature of *R. sativus* and especially of *A. thaliana* (Kudo and Kimura, 2001) since the only peak of plant DNA reference standard visible in the scale set was the 2C nuclei and, thus the three plant species were considered adequate for the analysis.

In this study we have analyzed 11 *Puccinia* spp., six *Uromyces* spp., four *Melampsora* spp. and two *Phragmidium* spp. The remaining six genera analysed were represented by a single species (Table 1). The average genome size of the studied rust species belonging to *Melampsora*, *Puccinia* and *Uromyces* was 227.6, 303.6 and 467.5 Mbp respectively. While the five *Melampsora* genomes (four species) were all below the overall average and varied by less than 3x, from 117.8 Mbp (for *M. larici-populina*) to 332.8 Mbp (for *M. ricini*), the 11 *Puccinia* species varied by over than 10x, from the smallest genomes analyzed in this study (76.9 and 77.4 Mbp for *P. triticina* and *P. graminis* f. sp. *tritici*) to the second largest one, *P. chrysanthemi* with 806.5 Mbp. The seven *Uromyces* genomes (six species) varied also by less than 3x, but in most cases with genome sizes higher than the overall average, from 276.8 Mbp in *U. rumicis* to 712.2 Mbp in *U. vignae*.

When clustering rusts according to their hosts family, it is clear that rusts with Poaceae hosts (*Puccinia coronata*, *P. cymbopogonis*, *P. graminis* f. sp. *tritici*, *P. hordei* and *P. triticina*) have clearly smaller genomes (170.6 Mbp on average) than rusts with Fabaceae hosts (556.6 Mbp on average; *Phakopsora pachyrhizi*, *Uromyces appendiculatus*, *U. fabae*, *U. fabae* f. sp. *pisi-sativae*, *U. striatus* and *U. vignae*). The four rust species with Rosaceae hosts (*Gymnosporangium confusum*, *Phragmidium mexicanum*, *Phr. mucronatum* and *Tranzschelia discolor* have varied genome sizes, with estimates below the average (145.5 Mbp in *Phr.*

mucronatum) to the largest estimate discovered so far (893.2 Mbp in *G. confusum*), with an average of 486.4 Mbp.

4. Discussion

The Pucciniales represent the largest group of fungal plant pathogens, and although the high negative impact some species have on food and energy crops, rust fungi are still poorly studied, probably due to the inherent difficulty in studying obligate pathogens with variable and complex life cycles. Recently, genome sequencing of some rust species provided evidence for their large genome sizes (Cantu et al., 2011; Duplessis et al., 2011a), while sequencing efforts of other rusts species have been hampered by uncertainty concerning the genome size of these organisms. Scattered genome size records for 23 rusts (mostly for *Puccinia*, *Melampsora* and *Uromyces* genera) can be found at the Fungal Genome Size database (Kullman et al., 2005). With exception of *U. appendiculatus* and *U. vignae* (400 Mbp) and of *H. vastatrix* (733 Mbp; Carvalho et al., 2014), available genome size values of rust species range from 50 to 170 Mbp. Also, recently, the genome size of *Cronartium quercuum* f. sp. *fusiforme* was estimated as 90 Mbp (Anderson et al., 2010). Under this mixed scenario, the goal of this study was to contribute to the knowledge of the genome sizes amongst the Puccinales, which could be correlated with distinct life cycles, types of spores produced, or host range.

The chopping of fresh tissue material for isolating nuclei suitable in the presence of a nuclei isolation buffer developed for plant tissues was successfully employed for flow cytometric analysis of genome size in rust-infected plant material. This approach proved to be very efficient and could be applied in the future in a variety of plant infected leaves, as it circumvents the need to isolate basidiospores or pycniospores as previously reported for the flow cytometric estimation of genome size in rust fungi (Eilam et al., 1992, 1994; Williams and Mendgen, 1975).

A collection of rust fungi found in nature, together with some of the most economically important rust species, revealed that the variability in the size of the genome was high, ranging over ten-fold, from 76.9 to 893.2 Mbp, with our estimates being even higher than those already made available through the fungal genome size database. Two rust fungi with a genome size higher than the largest rust genome reported recently (*Hemileia vastatrix*, Carvalho et al. 2014, whose genome size was further confirmed by the present study), are presented here, thus constituting the two largest fungal genome sizes reported to date, *Puccinia chrysanthemi* and *Gymnosporangium confusum*, which present 806.5 and 893.2 Mbp, respectively. These estimates surpass the two largest fungal genomes reported so far, *Neottiella vivida* (Ascomycota, Pezizales; Kullman, 2002) and *Scutellospora castanea* (Glomeromycota, Diversisporales; Hijri and Sanders, 2005), with 750 and 795 Mbp, respectively. Remarkably both of these fungi interact closely with plants.

Comparing the 1589 entries in the fungal genome size database (Kullman et al., 2005), the vast majority of fungal genomes (95%) present genome sizes bellow 75 Mbp, with an overall average of 37.6 Mbp. In particular, Basidiomycota estimates averaged 49.9 Mbp and Pucciniales presented genome size values of 126.2 Mbp on average. These estimates markedly contrast with the average genome size of 380.2 Mbp reported from this study (Fig. 4). Even the smallest genomes determined in this study are in the top 5% largest fungal genomes. The few genome sizes available for the Microbotryomycetes, a sister subclass of the Pucciniomycetes, are much smaller, with estimates of 21 Mbp for *Rhodotorula graminis* and *Sporobolomyces roseus*, and of 25 Mbp for *Microbotryum violaceum*.

Rust fungi are able to infect plants from most families, including conifers, ferns and mosses. Also, their life cycle is diverse, both in terms of the number of spore types produced (micro-,

hemi-, demi- or macrocyclic) and concerning the requirement (or not) of two alternate hosts for life cycle completion (autoecious or heteroecious).

In this study, it was revealed that the two *Phragmidium* species analyzed presented completely distinct genomes sizes, despite both being macrocyclic and autoecious. The only distinct characteristic between these *Phragmidium* species is their host range: *Phr. mexicanum* infects only few *Potentilla/Duchesnea* species, while *Phr. mucronatum* colonizes 60 to 65 species of *Rosa*. Also, two *Melampsora hypericorum* samples obtained from *Hypericum calycinum* or *H. androseamum* exhibited distinct genome sizes. This corroborates other reports that have showed the occurrence of intra-specific variation in nuclear content according with the host species, as was the case of *Puccinia hordei* and *P. recondita* (Eilam et al., 1994).

The genome size of *Puccinia graminis* f. sp. *tritici* was estimated as 77.4 Mbp. Eilam et al. (1994) estimated a value of 67 Mbp, while the genome sequence yielded a value of 88.6 Mbp (Duplessis et al., 2011a). These differences could be due to intra-specific variability, although the distinct methodologies adopted may also account for some variation. On the contrary, a relatively small difference of 16.7 Mbp was obtained between the flow cytometry estimate and that obtained from genome sequence (Duplessis et al., 2011a) for *Melampsora larici-populina* isolate 98AG31.

Uromyces appendiculatus and *U. vignae* have been reported to have some of the largest rust genomes (Eilam et al., 1994; Kullman et al., 2005), with 400-418 Mbp. In this work, the genome size of laboratory strains of these two species was estimated as 679.4 and 712.2 Mbp, respectively. Again, such differences could arise from actual intra-specific variability, or to technical advances in the methodologies that were adopted. In fact, the *Uromyces* species analyzed with hosts belonging to the Fabaceae family presented genome sizes all above 300 Mbp. Interestingly, these species constitute a monophyletic group that probably evolved together, and are all autoecious (van der Merwe et al. 2008). Moreover, *Phakopsora pachyrhizi*, another rust that infects a member of the Fabaceae family, also possess a large genome size. This markedly contrasts with the smaller genome sizes of rusts with Poaeceae hosts analyzed in this study (all below 250 Mbp and all belonging to the genus *Puccinia*). Except for *Gymnosporangium confusum*, the species with the largest genomes sizes are either autoecious or hemicyclic with no known alternate host.

A major force conditioning genome size seems to be genetic drift, which was negatively correlated with effective population size (Kelkar and Ochman, 2012). *Puccinia buxi*, the only microcyclic rust fungus analysed in this study, is apparently an unusual rust which is only found in a limited number of locations, due to its specific requirements of shaded and humid microclimatic conditions to develop (Durrieu, 2001; Preece, 2000). This species was found to possess a large genome, which is in accordance with this theory.

The Pucciniales share some common features, such as biotrophy and obligate parasitism. Biotrophy has been highlighted as a lifestyle that leads to increasing genome size as compared to non-biotrophs, oomycetes being the exception to the rule (Spanu, 2012). The very large genome sizes of the 30 rust fungal species revealed by our study strongly reinforce the view that expanded genome sizes occur among biotrophs. From the genome sequencing of rust fungi (e.g., Duplessis and al., 2011a; Nemri et al., 2014) and other biotrophs it is now clear that larger genomes do not imply higher numbers of structural genes, resulting invariably in an increased proliferation of transposable elements (TE) and repetitive DNA. Such a genomic environment can create genetic polymorphisms, especially in the case of sexual abstinence (Spanu, 2012). Although the effect of sex on genome size evolution still remains unclear (Raffaele and Kamoun, 2012), three of the rust fungi with large genome sizes, *Hemileia vastatrix*, *Phakospora pachyrhizi* and to some extent *Puccinia chrysanthemi*, all rely on asexual reproduction. The first two species are demicyclic or at least the aecial host is

unknown and the third also reproduces mainly asexually although is reported to be autoecious in Japan (Alaei et al., 2009). Even for those species which are capable of sexual reproduction, it is expectable that urediniosporic infection cycles may well represent a very important fraction of reproduction, for which TE activity would be potentially an important source of generation of diversity. In this sense, rust species that do not produce urediniospores (demicyclic rusts), such as the autoecious *P. buxi* and *P. smyrnii*, and therefore strictly depend on sexual reproduction for life cycle completion, are of great interest.

Gymnosporagium spp. are unique rust fungi since they comprise the only genus forming teliospores on members of the Cupressaceae. Molecular data (18S and 28S rDNA sequences) question their placement within the Pucciniaceae (Aime, 2006). Now, according to our genome size determinations, the group also has a highlighted position.

A unifying characteristic amongst the species with the larger genome size within the Pucciniales was not found, being more likely that different events have driven the evolution of genome size of particular species or groups of species. Genome variability is considered to be adaptive and host driven resulting in a high capability to overcome the host defenses (Stukenbrock and Croll, 2014). Relationships between genome size and biological parameters is of special interest because they can be linked to the ability of an organism to overcome selection pressure (D'Hondt et al., 2011).

In conclusion, in this work the analysis of the genome size of 30 rust species (representing eight families) revealed the occurrence of very large genome sizes, including the two largest fungal genomes ever reported, *Gymnosporangium confusum* (893.2 Mbp) and *Puccinia chrysanthemi* (806.5 Mbp). Although comprising only a very small fraction of the more than 7,000 rust species described, with many families not represented, this work suggests that the Pucciniales represent a group where genome size expansion could be a common characteristic, in sharp contrast with sister taxa, making this group of organisms a subject of utmost interest for genomic research and for further studies.

Conflict of Interest Statement

All authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

This study was conceived and directed by ST, APR, ASP, HGA, TL, RA, RTV, JL and PT. Collection and identification of field material was performed by APR, HGA and PT. Sample preparation, nuclei isolation and flow cytometry analyses were performed by ST, PC, JL and PT. Microscopy observations and image acquisition were conducted by ST, ASP and MCS. Data analysis and biological interpretation of results were conducted by ST, APR, JL and PT. ST, APR, JL and PT wrote the paper. All authors read and approved the final manuscript.

Acknowledgments

Prof. Arlindo Lima, Eng. Filomena Caetano, Ms. Paula Pais and Eng. Teresa Vasconcelos (ISA/UL, Portugal) are acknowledged for support on the identification of rust and host species. The *Melamspora larici-populina* isolate 98AG31 was kindly provided by Dr. Pascal Frey (INRA Nancy, France). The *Uromyces vignae* isolate CPR-1 was kindly provided by Prof. Kurt Mendgen (Univ. Konstanz, Germany). This work was supported by Fundação para a Ciência e a Tecnologia (FCT, Portugal) through PEst-OE/EQB/LA0004/2011 (at ITQB/UNL) and PTDC/AGR-GPL/114949/2009 (at CIFC/IICT and ITQB/UNL). ST, ASP and PT received postdoctoral grants from FCT (SFRH/BPD/65965/2009, SFRH/BPD/65686/2009 and SFRH/BPD/88994/2012, respectively).

401
402 **References**
403

404 Aime, M.C. (2006). Toward resolving family-level relationships in rust fungi (Uredinales).
405 *Mycoscience* 47, 112-122. doi: 10.1007/s10267-006-0281-0.
406 Alaei, H., De Backer, M., Nuytinck, J., Maes, M., Hofte, M., Heungens, K. (2009).
407 Phylogenetic relationships of *Puccinia horiana* and other rust pathogens of *Chrysanthemum* x
408 *morifolium* based on rDNA ITS sequence analysis. *Mycol. Res.* 113, 668-683.
409 doi:10.1016/j.mycres.2009.02.003.
410 Anderson, C. L., Kubisiak, T. L., Nelson, C. D., Smith, J. A., Davis, J. M. (2010). Genome
411 size variation in the pine fusiform rust pathogen *Cronartium quercuum* f. sp. *fusiforme* as
412 determined by flow cytometry. *Mycologia* 102, 1295-1302. doi: 10.3852/10-040.
413 Bourne, E. C., Mina, D., Gonçalves, S. C., Loureiro, J., Freitas, H., Muller, L. A. H. (2014).
414 Large and variable genome size unrelated to serpentine adaptation but supportive of cryptic
415 sexuality in *Cenococcum geophilum*. *Mycorrhiza* 24, 13-20.
416 Cantu, D., Govindarajulu, M., Kozik, A., Wang, M., Chen, X., Kojima, K. K., Jurka, J.,
417 Michelmore, R. W., Dubcovsky, J. (2011). Next generation sequencing provides rapid access
418 to the genome of *Puccinia striiformis* f. sp. *tritici*, the causal agent of wheat stripe rust. *PLoS*
419 *One* 6, e24230. doi: 10.1371/journal.pone.0024230.
420 Carvalho, G. M. A., Carvalho, C. R., Barreto, R. W., Evans, H. C. (2014). Coffee rust genome
421 measured using flow cytometry: does size matter? *Plant Pathol. in press*
422 doi:10.1111/ppa.12175.
423 D'Hondt, L., Hofte, M., Van Bockstaele, E., Leus, L. (2011). Applications of flow cytometry
424 in plant pathology for genome size determination, detection and physiological status. *Mol*
425 *Plant Pathol.* 12, 815-828. doi: 10.1111/j.1364-3703.2011.00711.x.
426 Doležel, J., Sgorbati, S., Lucretti, S. (1992). Comparison of three DNA fluorochromes for
427 flow cytometric estimation of nuclear DNA content in plants. *Physiol. Plantarum* 85, 625-
428 631.
429 Doležel, J., Greilhuber, J., Lucretti, S., Meister, A., Lysák, M.A., Nardi, L., Obermayer, R.
430 (1998). Plant genome size estimation by flow cytometry: inter-laboratory comparison. *Ann.*
431 *Bot.* 82 (Supplement A), 17-26.
432 Doležel, J., Bartoš, J. A. N. (2005). Plant DNA flow cytometry and estimation of nuclear
433 genome size. *Ann. Bot.* 95, 99-110.
434 Duplessis, S., Cuomo, C. A., Lin, Y. C., Aerts, A., Tisserant, E., Veneault-Fourrey, C., Joly,
435 D. L., Hacquard, S., Amselem, J., Cantarel, B. L., Chiu, R., Coutinho, P. M., Feau, N., Field,
436 M., Frey, P., Gelhaye, E., Goldberg, J., Grabherr, M. G., Kodira, C. D., Kohler, A., Kues, U.,
437 Lindquist, E .A., Lucas, S. M., Mago, R., Mauceli, E., Morin, E., Murat, C., Pangilinan, J. L.,
438 Park, R., Pearson, M., Quesneville, H., Rouhier, N., Sakthikumar, S., Salamov, A. A.,
439 Schmutz, J., Selles, B., Shapiro, H., Tanguay, P., Tuskan, G. A., Henrissat, B., Van De Peer,
440 Y., Rouze, P., Ellis, J. G., Dodds, P. N., Schein, J. E., Zhong, S., Hamelin, R. C., Grigoriev, I.
441 V., Szabo, L. J., Martin, F. (2011a). Obligate biotrophy features unraveled by the genomic
442 analysis of rust fungi. *Proc. Natl. Acad. Sci. USA* 108, 9166-9171. doi:
443 10.1073/pnas.1019315108.
444 Duplessis, S., Joly, D.L., Dodds, P.N. (2011b). Rust Effectors. In: Martin, F., Kamoun, S.
445 (eds) Effectors in Plant-Microbe Interactions. First edn. Wiley-Blackwell, Oxford, UK, pp
446 155-193. doi: 10.1002/9781119949138.ch7.
447 Durrieu, G. (2001). More about box rust. *Mycologist* 15, 144.
448 Eilam, T., Bushnell, W. R., Anikster, Y., McLaughlin, D.J. (1992). Nuclear DNA content of
449 Basidiospores of selected rust fungi as estimated from fluorescence of propidium iodide-
450 stained nuclei. *Phytopathology* 82, 705-712.

Eilam, T., Bushnell, W. R., Anikster, Y. (1994). Relative Nuclear DNA content of rust fungi estimated by flow cytometry of propidium iodide-stained pycniospores. *Phytopathology* 84, 728-735.

Galbraith, D. W., Harkins, K. R., Maddox, J. M., Ayres, N. M., Sharma, D. P., Firoozabady, E. (1983). Rapid flow cytometric analysis of the cell-cycle in intact plant-tissues. *Science* 220, 1049-1051.

Gregory, T. R., Nicol, J. A., Tamm, H., Kullman, B., Kullman, K., Leitch, I. J., Murray, B. G., Kapraun, D. F., Greilhuber, J., Bennett, M. D. (2007). Eukaryotic genome size databases. *Nucleic Acids Res.* 35, D332-338. doi:10.1093/nar/gkl828.

Greilhuber, J., Doležel, J., Lisák, M. A., Bennett, M. D. (2005). The origin, evolution and proposed stabilization of the terms 'Genome Size' and 'C-Value' to describe nuclear DNA contents. *Ann. Bot.* 95, 255–260.

Hijri, M., Sanders, J. R. (2005). Low gene copy number shows that arbuscular mycorrhizal fungi inherit genetically different nuclei. *Nature* 433, 160–163. doi: 10.1038/nature03069.

Johnson, G. D., Araujo, G. M. (1981). A simple method of reducing the fading of immunofluorescence during microscopy. *J. Immunol. Meth.* 43, 349–350.

Kelkar, Y. D., Ochman, H. (2012). Causes and consequences of genome expansion in fungi. *Genome Biol. Evol.* 4, 13-23. doi: 10.1093/gbe/evr124.

Kudo, N., Kimura, Y. (2001). Flow cytometric evidence for endopolyploidy in seedlings of some *Brassica* species. *Theor. Appl. Genet.* 102, 104-110. doi: 10.1007/s001220051624.

Kullman, B. (2002). Nuclear DNA content, life cycle and ploidy in two *Neottiella* species (Pezizales, Ascomycetes). *Persoonia* 18, 103–115.

Kullman, B., Tamm, H., Kullman, K. (2005). Fungal Genome Size Database. http://wwwzbiee/fungal-genomesize.

Laundon, G. F. (1967). Terminology in the rust fungi. *Trans. Brit. Mycol. Soc.* 50, 189-194.

Loureiro, J., Rodriguez, E., Doležel, J., Santos, C. (2007). Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. *Ann. Bot.* 100, 875-888. doi:10.1093/annbot/mcm152.

Nemri, A., Saunders, D. G., Anderson, C., Upadhyaya, N. M., Win, J., Lawrence, G. J., Jones, D. A., Kamoun, S., Ellis, J. G., Dodds, P. N. (2014). The genome sequence and effector complement of the flax rust pathogen *Melampsora lini*. *Frontiers in Plant Sci.* 5, 98. doi:10.3389/fpls.2014.00098.

Park, R. F., Wellings, C. R. (2012). Somatic hybridization in the Uredinales. *Annu. Rev. Phytopathol.* 50, 219-239. doi: 10.1146/annurev-phyto-072910-095405.

Preece, T.F. (2000). The strange story of box rust. *Mycologist* 14, 104-106.

Raffaele, S., Kamoun, S. (2012) Genome evolution in filamentous plant pathogens: why bigger can be better. *Nat. Rev. Microbiol.* 10, 417-430. doi: 10.1038/nrmicro2790.

Silva, M. C., Várzea, V.M., Guerra-Guimarães, L., Azinheira, H., Fernandez, D., Petitot, A.S., Bertrand, B., Lashermes, P., Nicole, M. (2006). Coffee resistance to the main diseases: leaf rust and coffee berry disease. *Braz. J. Plant Physiol.* 18, 119-147. doi: 10.1590/S1677-04202006000100010.

Spanu, P.D. (2012). The genomics of obligate (and nonobligate) biotrophs. *Annu. Rev. Phytopathol.* 50, 91-109. doi: 10.1146/annurev-phyto-081211-173024.

Stukenbrock, E. H., Croll, D. (2014). The evolving fungal genome. *Fungal Biol. Rev.* 28, 1-12. doi:10.1016/j.fbr.2014.02.001.

van der Merwe, M. M., Walker, J., Ericson, L., Burdon, J. J. (2008). Coevolution with higher taxonomic host groups within the *Puccinia/Uromyces* rust lineage obscured by host jumps. *Mycol. Res.* 112, 1387-1408. doi: 10.1016/j.mycres.2008.06.027.

Williams, P. G., Mendgen, K. W. (1975) Cytofluorometry of DNA in uredospores of *Puccinia graminis* f. sp. *tritici*. *Trans. Br. Mycol. Soc.* 64, 23-28.

501

Table 1. List of 32 rust samples analysed for genome size determination, with reference to the rust species and family, source of material [host plant (botanical name, family, location and infection stage), or spores in collection], typical life cycle, plant reference standard used (RS), average (GS, in Mbp), standard deviation (SD, in Mbp) and coefficient of variation (CV, in %) of the monoploid genome size, number of samples (n) and typical life cycle.

Table 1 footnotes

1 Acronyms in brackets refer to host family: Am – Amaryllidaceae; Ap – Apiaceae; As – Asteraceae; Bu – Buxaceae; Eu – Euphorbiaceae; Fa – Fabaceae; Ge – Geraniaceae; Hy – Hypericaceae; Ir – Iridaceae; Ma – Malvaceae; On – Onagraceae; Ox – Oxalidaceae; Pl – Polygonaceae; Po – Poaceae; Ro – Rosaceae; Ru – Rubiaceae; Sa – Salicaceae.

2 Type of spores present in sampled material: 0 - Pycniospores; I – Aeciospores; II – Urediniospores; III - Teliospores; IV – Basidiospores (Laundon, 1967).

3 Plant reference standards used: At – *Arabidopsis thaliana*; H – host plant; Rs – *Raphanus sativus*; Sl – *Solanum lycopersicum*.

4 Typical life cycle: Ma – macrocyclic; Hc – hemicyclic; Dc – demicyclic; Mi – microcyclic; Ae – aotoecious; He – heteroecious.

Figure 1. Examples of rust sporulation. A – *Hemileia vastatrix* uredinia in a *Coffea arabica* leaf; B - *Phragmidium mexicanum* uredinia in a *Duchesnea indica leaf*; C - *Puccinia oxalidis* uredinia in a *Oxalis articulata* leaf; D - *Uromyces transversalis* uredinia in a Gladiolus sp. leaf; E - *Puccinia smyrnii* telia in a *Smyrnium olusatrum* leaf; F - *Puccinia smyrnii* aecia in a *Smyrnium olusatrum* stem; G - *Puccinia buxi* telia in a *Buxus sempervirens* leaf; H - *Gymnosporangium confusum* aecia in a *Crataegus monogyna* leaf.

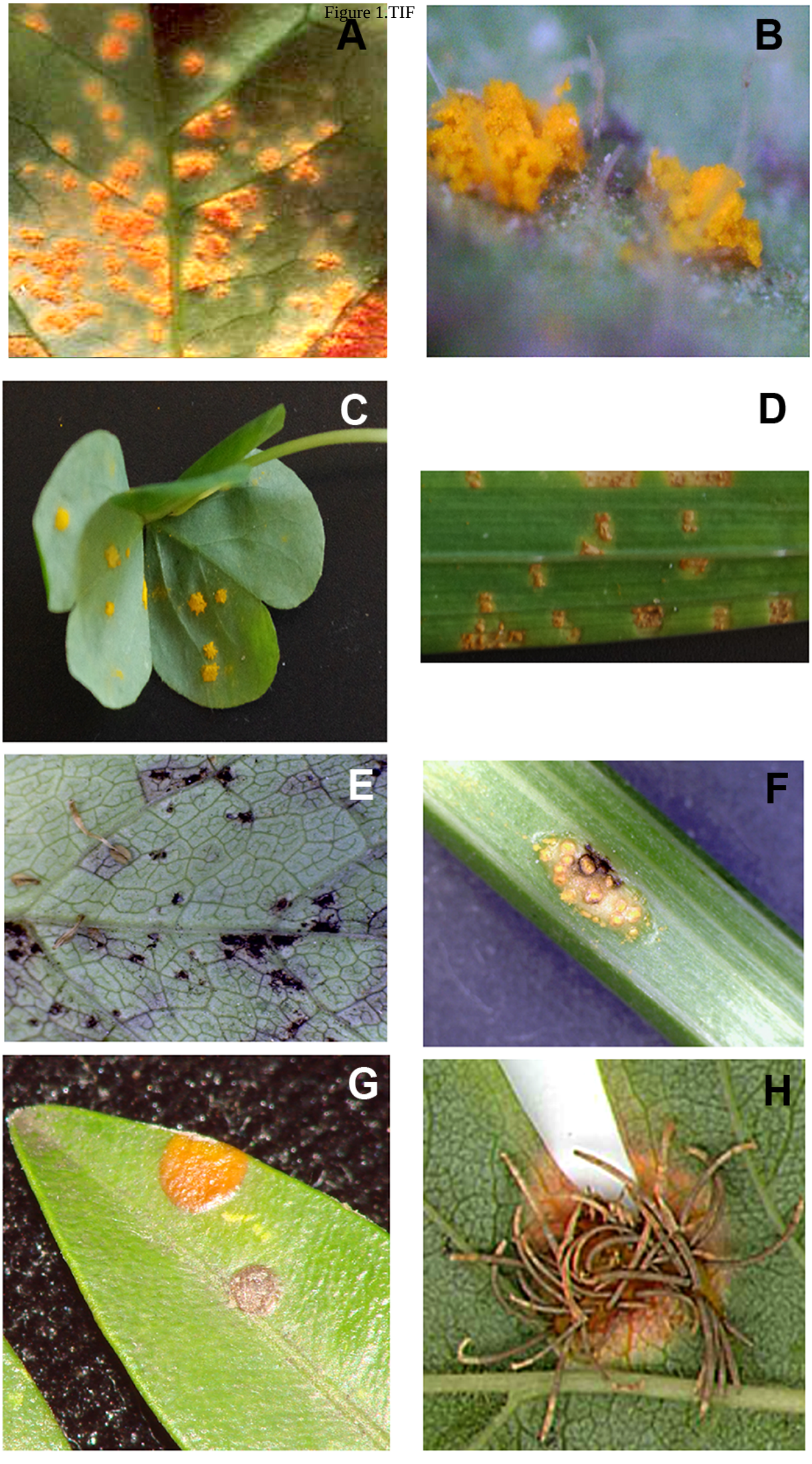
Figure 2. Spores from Pucciniales under phase-contrast microscopy and showing the DAPI-staining nuclei under fluorescence light. A – Teliospore from *Puccinia malvacearum*; B – Teliospore from *Puccinia smyrnii*; C – Aeciospore from *Puccinia smyrnii*; D – Urediniospores from *Phragmidium mexicanum*; E – Urediniospore from *Hemileia vastatrix*; F – Aeciospore from *Gymnosporangium confusum*; G – Teliospores and urediniospores from *Puccinia hordei*. Bars, 10 mm.

Figure 3. Flow cytometric histograms of relative fluorescence intensities of propidium iodide-stained nuclei simultaneously isolated from: A – *Hemileia vastatrix* (*Hv*) and its host plant, *Coffea arabica* (*Ca*); B – *Coffea arabica* (Ca) and the plant DNA reference standard, *Raphanus sativus* (*Rs*, 2C = 1.11 pg DNA); C – *Hemileia vastatrix* (*Hv*) hyphae obtained upon germination of urediniospores in water; D - *Gymnosporangium confusum* (*Gc*), its host plant, *Crataegus monogyna* (*Cm*), and the plant DNA reference standard, *Raphanus sativus* (*Rs*); E – *Puccinia buxi* (*Pb*), the plant DNA reference standard (*Rs*), and its host plant, *Buxus sempervirens* (*Bs*); and F – *Puccinia pelargonii-zonalis* (*Ppz*) and the plant DNA reference standard, *Raphanus sativus* (*Rs*).

Figure 4. Graphical representation of the genome sizes determined in this study for 30 rust species grouped according to genus and family, by comparison with the frequency distribution of genome sizes for other fungi (according to the Fungal Genome Size database (Kullman et al., 2005). Acronyms: Ci - *Coleosporium inulae*; Gc - *Gymnosporangium confusum*; Hv - *Hemileia vastatrix*; Me - *Melampsora euphorbiae*; Mh - *Melampsora hypericorum*; Mlp - *Melampsora larici-populina*; Mr - *Melampsora ricini*; Pa - *Puccinia allii*; Pb - *Puccinia buxi*; Pch - *Puccinia chrysanthemi*; Pco - *Puccinia coronata*; Pcy - *Puccinia*

cymbopogonis; Pe - *Pucciniastrum epilobii*; Pgt - *Puccinia graminis* f. sp. *tritici*; Ph - *Puccinia hordei*; Pma - *Puccinia malvacearum*; Pme - *Phragmidium mexicanum*; Pmu - *Phragmidium mucronatum*; Po - *Puccinia oxalidis*; Pp - *Phakopsora pachyrhizi*; Ppz - *Puccinia pelargonii-zonalis*; Ps - *Puccinia smyrnii*; Pt - *Puccinia triticina*; Td - *Tranzschelia discolor*; Ua - *Uromyces appendiculatus*; Uf - *Uromyces fabae*; Ufp - *Uromyces fabae* f. sp. *pisi-sativae*; Ur - *Uromyces rumicis*; Us - *Uromyces striatus*; Ut - *Uromyces transversalis*; Uv - *Uromyces vignae*.

Figure 1.TIF



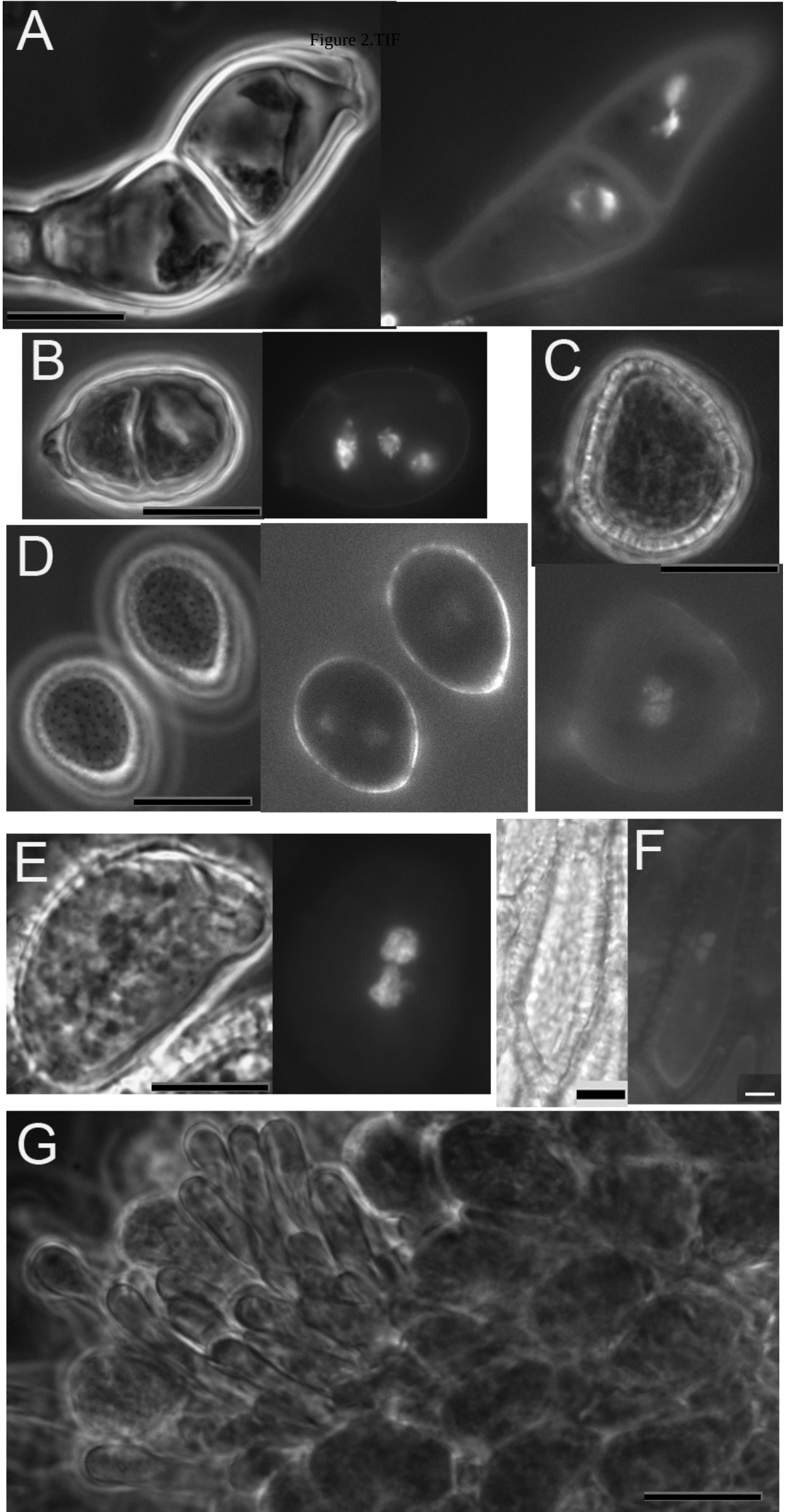


Figure 2.TIF

Figure 3.JPEG

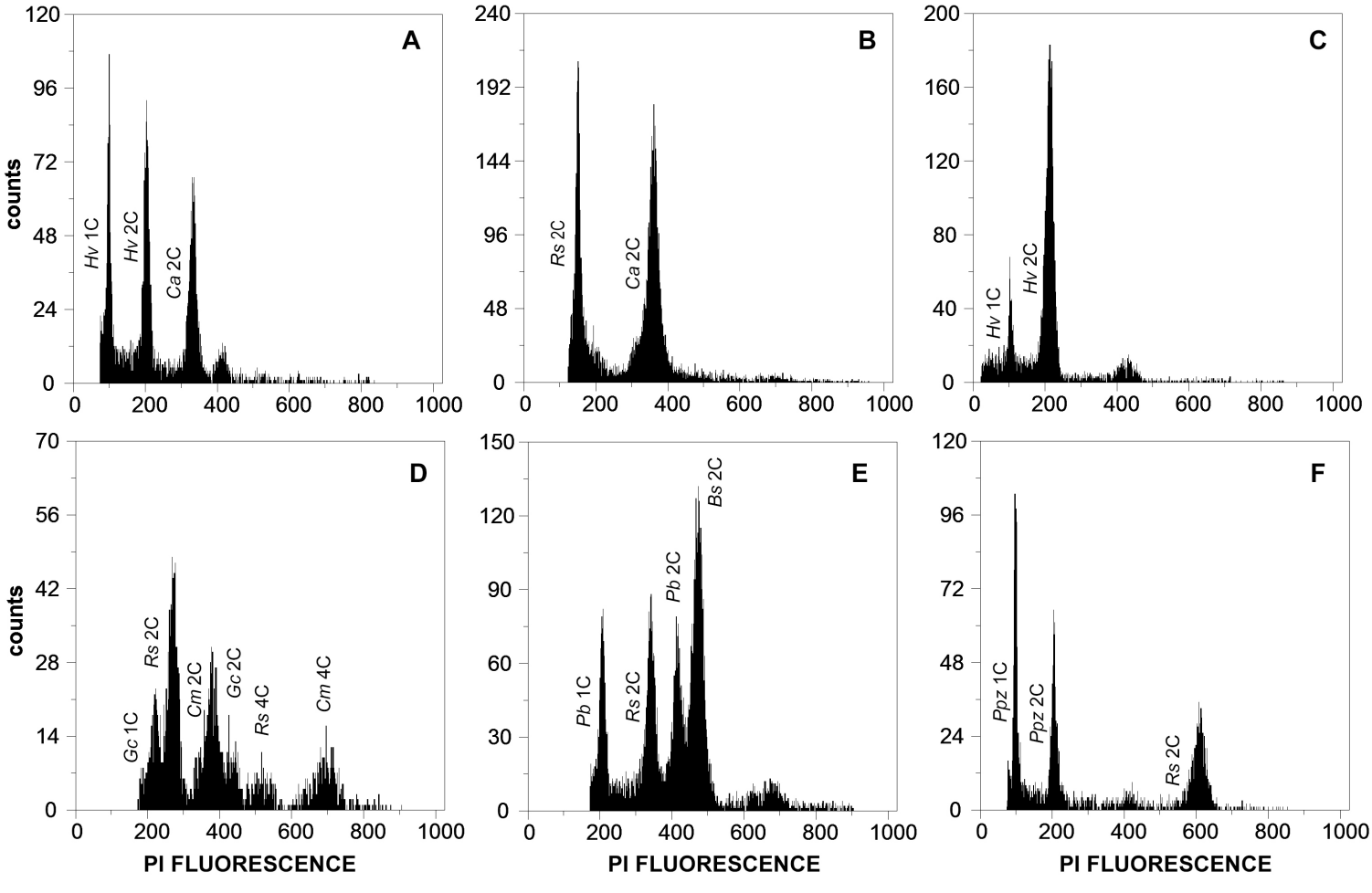


Figure 4.JPEG