DOI:
10.1007/s00294-016-0597-z

Document Version
Peer reviewed version

Development and validation of a custom microarray for global transcriptome profiling of the fungus *Aspergillus nidulans*

Claudine Deloménie [1,*]

Phone (33) 1 46 83 57 85
Email claudine.delomenie@u-psud.fr

Guido Grentzmann [2]

Nathalie Oestreicher [3,4]

Robin Mesnage [5,6]

Christian Vélot [3,4,6]

[1] UMS IPSIT, Plateforme Trans-Prot, Faculté de Pharmacie, Univ. Paris-Sud, Université Paris-Saclay, 5 rue Jean-Baptiste Clément, 92296 Châtenay-Malabry, France

[2] PharmaBioServices, 1 square du Cèdre, 91370 Verrières-Le-Buisson, France

[3] Laboratoire VEAC, Faculté des Sciences, Univ. Paris-Sud, Université Paris-Saclay, Bât. 360, Rue du Doyen André Guinier, 91405 Orsay, France

[4] Pôle Risques MRSH-CNRS, Université de Caen, Esplanade de la Paix, 14032 Caen, France

[5] Gene Expression and Therapy Group, Department of Medical and Molecular Genetics, Faculty of Life Sciences and Medicine, King's College London, 8th Floor, Tower Wing, Guy's Hospital, Great Maze Pond, London, SE1 9RT UK

[6] CRIIGEN, 81 rue Monceau, 75008 Paris, France

Abstract

Transcriptome profiling is a powerful tool for identifying gene networks from whole genome expression analysis in many living species. Here is described the first extensively characterized platform using Agilent microarray technology for transcriptome analysis in the filamentous fungus *Aspergillus (Emericella) nidulans*. We developed and validated a reliable gene expression microarray in 8 × 15 K format, with predictive and experimental data establishing its specificity and sensitivity. Either one or two 60-mer oligonucleotide probes were selected for each of 10,550 nuclear as well as 20 mitochondrial coding sequences. More than 99 % of probes were predicted to hybridize with 100 % identity to their aimed specific *A. nidulans* target only. Probe sensitivity was supported by a highly narrow distribution of melting temperatures together with thermodynamic features, which strongly favored probe–target perfect match hybridization, in comparison with predicted secondary structures. Array quality was evaluated through transcriptome comparison of two *A. nidulans* strains, differing by the presence or not of *Escherichia coli LacZ* transgene. High signal-to-noise ratios were measured, and signal reproducibility was established at intra-probe and inter-probe levels. Reproducibility of microarray performances was assessed by high correlation between two-color dye signals and between technical replicates. Results were confirmed by RT-qPCR analysis on five genes. Though it covers 100 % of the *A. nidulans* targeted coding sequences, this low density array allows limited experimental costs and simplified data analysis process, making it suitable for studying gene expression in this model organism through large numbers of experimental conditions, in basic, biomedical or industrial microbiology research fields.

AQ1

Keywords

Microarray
Aspergillus nidulans
Transcriptome analysis
RT-qPCR
Oligonucleotide probe design

Communicated by M. Kupiec.

Electronic supplementary material

The online version of this article (doi:10.1007/s00294-016-0597-z) contains supplementary material, which is available to authorized users.

Introduction

High throughput “omic” technologies have been revealed as a promising tool for investigating at the molecular level, physiological as well as toxicological mechanisms in many living species. Since its introduction (Schena et al. 1995), the DNA microarray technology markedly evolved and became very popular in life science: a search for the keyword “microarray” in the PubMed database at the time of writing the present report resulted in more than 70,000 citations. Despite their sensitivity and accuracy of detection of large fold-changes in gene expression, distinct microarray platforms used for the same set of biological conditions can result in inconsistent lists of differentially expressed genes (Liu et al. 2012; Zhang et al. 2009). This issue can even be more pronounced for custom microarray designs in species for which commercial microarrays are not available. Here, we aimed to provide to the *Aspergillus* research community, a fully validated and scalable tool for soil filamentous fungus *Aspergillus (Emericella) nidulans* genomic studies.

Aspergillus nidulans is an important model organism used for decades in the field of genetics and cell biology (Morris and Enos 1992; Pontecorvo et al. 1953) for basic and industrial microbiology research (Martinelli and Kinghorn 1994). In particular, it provides an excellent experimental system for the study of gene regulation (Arst and Scazzocchio 1985; Davis and Hynes 1991). For molecular studies, this ascomycete fungus has the advantage of a small genome with only a few repeated sequences (Timberlake 1978), and an efficient transformation system (Tilburn et al. 1983) with an increasing variety of selectable markers. *A. nidulans* genome has been sequenced (Galagan et al. 2005), and its functional annotation remains partial and largely predictive despite a special community effort (Nitsche et al. 2011; Wortman et al. 2008).

AQ2

Behind basic biological research, filamentous fungi are also important in medicine, not only because of the production of antibiotics, but also because some species are pathogenic for humans, such as *Aspergillus fumigatus* responsible for nosocomial infections. They can also be used in industry to produce large amounts of enzymes and low molecular weight compounds. In

agriculture, economically, important groups of plant pathogens are fungi. The effects of intensive agricultural practices on soil health are critical for earth-system functioning (Steffen et al. 2015). Soil microorganisms are key agents in the maintenance of soil quality and resilience. Because of its presence in agricultural soils (Carranza et al. 2014), *Aspergillus* species could be considered as a marker of agricultural soil ecosystems health. The application of genomics technologies, especially transcriptome profiling, has been suggested as an important and reliable tool to assess the biological effects of abiotic factors (van Straalen and Roelofs 2008). Transcriptome profiles bear a signature of the stress to which the organism was exposed. Such changes in gene expression patterns after exposure to specific chemical mixtures could be used as biomarkers to support the risk assessment of soil pollutions.

Microarrays are a convenient and efficient tool for understanding the molecular basis of response to stimuli through exploration of the whole genome expression in a growing number of living species. Despite the development of more recent high throughput sequencing technologies for transcriptomic studies (RNA-SEQ), microarrays are quite cost-effective, and remain widely used for genome expression analysis in various organisms (Fu et al. 2014; Iwahashi 2015; Kleijkers et al. 2015; Liao and Weng 2015; Lin et al. 2014; Xiao et al. 2015; Mesnage et al. 2015). They still offer some advantages compared with the next-generation sequencing (NGS) technologies, such as simple and rapid experimental setup and accessible data analysis procedures. Moreover, the evolution of microarray manufacture has allowed densification of formats into multiarray slides, which considerably improved the ratio of information collected versus cost per sample. Finally, microarray and NGS techniques proved to be complementary and can be combined (Roh et al. 2010).

We provide here the first step-by-step detailed development and validation of the largest set of unique probes collected in a low density 8 × 15 K Agilent platform for *A. nidulans*. By contrast with other recently published *A. nidulans* custom microarrays, we performed an in silico performance analysis of probe specificity and sensitivity, as well as an assessment of reproducibility at intra-probe and inter-probe levels. In addition, probe sensitivity was established by calculating signal-to-noise ratios (SNRs). By restricting the probe number—either one or two probes per coding sequence (CDS) in most cases, we limited the amount of produced data, as well as the risk of observing ambiguous results, thus simplifying and accelerating their analysis. Since

reliable transcriptomic results depend on the selection and quality of designed probes, we finally experimentally evaluated the performance of the microarray, comparing the transcriptome profiles from an *A. nidulans* strain (FGSC A4 lineage) to a transgenic version harboring the *lacZ* bacterial gene. We have thus established the ability of this microarray to produce reliable data, in the aim of sharing it with the scientific community.

Materials and methods

Choice of target sequences and design of oligonucleotide probes

Sets of transcribed sequences for *A. nidulans* (*Emericella nidulans* FGSC A4, taxon identifier 227321) were downloaded and compared from two sources, *i.e.*, GenBank at NCBI (http://www.ncbi.nlm.nih.gov/genbank/) and "*Aspergillus* comparative database" at Broad Institute (http://www.broadinstitute.org/annotation/genome/aspergillus_group /MultiHome.html), which provided 2003 and 2011 releases, respectively. The data provided by the Broad Institute database at the time of this work (now available in AspGD: http://www.aspgd.org/; Arnaud et al. 2012), and allowed to link accession numbers from both sources. From UniprotKB (http://www.uniprot.org/help/uniprotkb), Entrez Gene IDs were used to associate annotation information, *i.e.*, protein name and gene ontology (GO) terms with each CDS. The *A. nidulans* CDS from the "*Aspergillus* comparative database" was uploaded in FASTA format into the free software eArray (https://earray.chem.agilent.com/earray/), working with Java and web interface in Windows. Design settings were chosen for selection of 60-mer sense probes with 3′ bias from each CDS, according to the "base composition" methodology, which considers fusion temperature (Tm), GC %, and cross-hybridization potential of probes. To get the best quality level for *A. nidulans* probes, we selected "best probe" (BP) or alternatively "best distribution" (BD) options. The probe sequences were checked for vector masking and low complexity masking. To obtain a total probe set size suitable for 15 K array format, we designed either one or two probes per CDS. Probes for 19 annotated genes (18 *A. nidulans* genes with "marker" status in and one heterologous control gene from *E. coli*) were designed to be fivefold replicated on the array for probe reproducibility testing.

In silico evaluation of the oligonucleotide probes

A first level of probe quality evaluation was provided by eArray using

proprietary prediction algorithms to assign to each probe a score (base composition—BC-content) reflecting the predicted performance of hybridization (Ferraresso et al. 2008). Further probe quality control was performed by independent calculation of thermodynamic parameters known to influence specificity and sensitivity of hybridization. Potential hybridization of probes to non-specific targets was evaluated using the BLAST algorithm (http://blast.ncbi.nlm.nih.gov/Blast.cgi; Altschul et al. 1997). The number of probes with non-target BLAST hits showing more than 75 % identity on a full size alignment of the probe (Kane et al. 2000) was determined. In addition to Tm calculations provided by eArray, the Tm of each probe and the ΔG of melting of each perfect-match probe–target duplex were calculated according to the nearest-neighbor (NN) model with parameters from Santa Lucia (1998), using Harvard's adnan/tm.pl tool (http://arep.med.harvard.edu/cgi-bin/adnan/tm.pl) with the following settings: temperature of 65 °C, absence of Mg^{2+}, 1 M Na^{+}, and 10 µM DNA. Thermodynamic stability of secondary structures was predicted with AutoDimer (http://www.nist.gov/mml/; Vallone and Butler 2004) using the following settings: temperature of 65 °C, absence of Mg^{2+}, 1 M Na^{+}, 10 µM of total DNA strand, and a score threshold of four for either intramolecular hairpins or intermolecular dimers. Only probes exhibiting secondary structures with ΔG lower than zero were considered for the statistical study. The distance of hybridization of the 5′-end of probe to the 3′-end of target was calculated by subtracting the 5′-end position of the probe on target from total target length.

Microarray manufacture

Oligonucleotide probes were synthesized in situ with the SurePrint™ inkjet technology (Agilent Technologies), according to manufacturer's protocol. Basically, oligonucleotides were synthesized from the 3′-end, base by base on a solid glass surface, following standard phosphoramidite chemistry as previously described (Wolber et al. 2006).

Aspergillus nidulans strains, growth conditions, and RNA isolation

Aspergillus nidulans strains used and/or generated in this study are CV125 (*pabaA1*) and NA1363 (*pabaA1*) harboring a *gpdAp-lacZ-trpCt.riboB* construct integrated in chromosome VIII. The details for the construction of the NA163 transformant strain, including the map of the integrative vector, used are described in Online Resource 1. Briefly~~,~~ this vector ~~is~~, carried out in

a derivative of the pBR322 vector, contained the *E. coli lacZ* sequence (encoding β-galactosidase), fused to the *A. nidulans trpC* terminator, and placed under the control of the strong and constitutive promoter of the *A. nidulans gpdA* gene and the *riboB* gene as selection marker (http://www.fgsc.net/aspergillus/gene_list/loci.html). Media composition, supplements, and basic growth conditions were as described by Cove (1966). Mycelia for RNA preparation were grown at 30 °C for 15 h in 400-mL minimal medium with shaking at 150 rpm, in the presence of fructose (0.1 %) as the carbon source and urea (5 mM) as the nitrogen source. For each strain, four cultures were carried out in each one of two independent experiments. Total RNA was isolated as previously described (Robellet et al. 2010). An additional purification step was performed from 25 % of each RNA extract alcoholic solution (70 % ethanol, 100-mM sodium acetate). After high-speed centrifugation, the pellet was washed three times with 75 % ethanol before solubilization in 250-µL RNase-free water, and remaining impurities were extracted with 900-µL TRIzol® (Life Technologies) according to the manufacturer's instructions. RNA was precipitated in 500-µL isopropanol at room temperature. After high-speed centrifugation, the pellet was washed twice with 75 % ethanol, solubilized in 100-µL RNase-free water, and stored at −80 °C until molecular analyses. RNA purity and concentration were measured by UV absorbance using a BioMate 3S UV–Visible (Thermo Scientific™) spectrophotometer. Genomic DNA content in samples was quantified and found negligible. RNA integrity was checked using RNA 6000 Nano chips and the 2100 Bioanalyzer (Agilent Technologies) with Plant RNA Assay: RNA integrity number (RIN) and 28S/18S ratios were above 6.5 and 1.6, respectively, with low dispersion (coefficient of variation < 10 %).

Microarray hybridization

Target preparation and hybridization were done according to the manufacturer's recommendations. Briefly, 200 ng of total RNA was labelled using the Low Input Quick Amp Labelling Kit, Two-Color (Agilent Technologies), using Cy5-CTP for the reference RNA pool (thereafter called REF) made of a mix of all samples, and Cy3-CTP for individual samples. Internal standards were derived from the Two-Color RNA Spike-in Kit (Agilent Technologies). The labelled targets were purified using the RNeasy® Mini Kit (Qiagen), and their qualities and quantities were confirmed by spectrophotometry and Bioanalyzer 2100 technology (Agilent Technologies). Equal amounts (325 ng) of Cy3- and Cy5-labeled cRNA targets were mixed and incubated on microarray slides at 65 °C for 17 h. After washing, slides

were scanned using the Microarray Scanner (G2565CA) from Agilent at 5-µm resolution and at high and low photomultiplier voltages to optimize the dynamic range of image quantification.

Processing of microarray data

Data were extracted from images using the Agilent Feature Extraction 10.7.3.1. software with default protocols and settings. Data were computed using Genespring GX software version 12.6 (Agilent Technologies). Fluorescence signals were background-subtracted, and a filtering procedure excluded data points considered unreliable as they corresponded to probe sets associated with low-signal intensities or bad quality features. A within-array loess normalization was performed, followed by an inter-array scaling normalization. Normalized log intensity ratios for each spot were obtained. Fold-change and corrected *p*-values were computed following a moderated *t* test with Benjamini–Hochberg multiple testing correction.

Calculation of signal-to-noise ratio

The SNR was calculated by three methods for each probe in Cy3 and Cy5 channels: (1) the signal-to-standard deviation ratio defined as: SSR = [(mean signal intensity − mean background intensity)/standard deviation (SD) of background intensity]; (2) the signal-to-background ratio defined as: SBR = [mean signal intensity/mean background intensity]; and (3) the signal-to-both-standard deviations ratio defined as: SSDR = [(mean signal intensity − mean background intensity)/(SD of signal intensity + SD of background intensity)], as previously described (He and Zhou 2008; Bidard et al. 2010). Three values of SSR, SBR, and SSDR were obtained for each probe, each value corresponding to one strain condition or to the reference.

RT-qPCR

For quantification of mRNA expression, the first strand cDNA was synthesized from 1 µg of total RNA, with random hexamers and oligo-dT priming using the iSCRIPT enzyme (Bio-Rad), according to manufacturer's instructions. PCR primer pairs specific to five target genes and two reference genes (Table 3) were designed using Primer3Plus software (http://primer3plus.com/cgi-bin/dev/primer3plus.cgi or PrimerBLAST (NCBI, http://www.ncbi.nlm.nih.gov/tools/primer-blast/) tools. The cDNA synthesized from 4 ng of total RNA was amplified in a CFX96 real time thermal cycler (Bio-Rad) using the SSoADV Univer SYBR® Green Supermix (Bio-Rad)

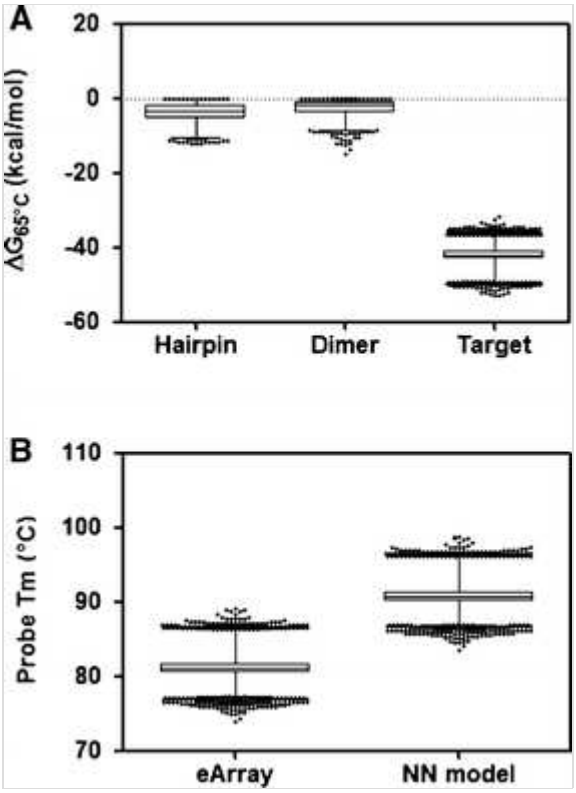
reagent according to manufacturer's instructions, with 500 nM (final concentration) of each primer, in duplicate 10-µl reactions, by 45 two-step cycles (95 °C for 5 s, 61 °C for 10 s). "No RT" controls were amplified on all genes to check for the absence of genomic DNA contamination, and melting curve analysis was performed to assess the purity of the PCR products. All PCR efficiencies were above 90 %. GeNorm in qBase PLUS tool (Vandesompele et al. 2002) was used to select *acnA* and *gpdA* genes as references for gene expression normalization. The normalized relative expression of target genes was determined using the ΔΔCq method with correction for PCR efficiencies (Pfaffl 2001; Hellemans et al. 2007), where NRQ = $E_{\text{Target}}^{-\Delta Cq\ \text{Target}} / E_{\text{Ref}}^{-\Delta Cq\ \text{Ref}}$ and $\Delta Cq = Cq_{\text{sample}} - Cq_{\text{calibrator}}$. Final results were expressed as the mean *n*-fold differences (fold-changes) for target gene expression in transgenic versus control strain, obtained from seven or more individual samples of each group.

Statistical analysis

All probe quality parameters are given as mean ± standard deviation (SD), with median value and minimum–maximum range (in brackets). Thermodynamic features were analyzed by one-way ANOVA with Bonferroni's multiple comparison test in GraphPad Prism®. Data distributions were presented either in box-and-whiskers plots showing interquartile range with 1–99 percentiles (free energy and Tm; Fig. 1), or in Tukey box plots (CV of replicated probes; Fig. 5). RT-qPCR results were compared in GraphPad Prism® by the Student's *t* test, with Welch's correction if unequal variances were found between groups.

Fig. 1

Distribution of thermodynamic features of microarray probes (v1 set). **a** Free energy ($\Delta G_{65\ °C}$, kcal/mol) of perfect probe–target hybrids as well as the most favorable occurring secondary structures (i.e., hairpins and autodimers for 1958 and 2923 probes, respectively) was calculated. All groups differed significantly each other ($p < 0.0001$). **b** Tm of all probes was calculated through two distinct tools, i.e., eArray and classical NN model with Santa Lucia parameters, respectively. *Box- and whiskers-plots* show interquartile range with 1–99 percentile values



Results

Description of the new microarray designed for *A. nidulans* transcriptome analysis

Expressed sequences targeted by oligonucleotide probes

Before starting probe design, we compared *A. nidulans* CDS from two public databases. At the time of this work (*i.e.* 2013), GenBank provided 9541 RefSeq accessions, while the "*Aspergillus* comparative database" (Broad Institute) contained 10,560 unique "ANID" accessions for nuclear genes, besides 20 accessions for mitochondrial genes which did not have a RefSeq accession in Genbank, all these sequences matching with 8746 unique Gene IDs. Nuclear CDS from Broad Institute used here for probe design comprised 5896 (56 %) annotated sequences (5666 relating to an explicit protein function and 230 to a partial function name), and 4664 (44 %) non annotated sequences. Functional information, *i.e.*, GO terms, was associated with 5727 (65 %) of the Uniprot IDs mapping with Gene IDs.

Overview of the microarray

A total of 15,744 oligonucleotide probes were used here to construct a low-density microarray based on the Agilent 8 × 15 K design format. These comprised 15,208 "biological" non-control probes (of which 15,124 unique probes) and 536 control probes included by the manufacturer (Table 1). In a

first design (v1), the biological probes allowed to screen 100 % of the nuclear input CDS from *A. nidulans*, besides the *lacZ* CDS from *E. coli* expressed by the transgenic strain, which was used as a control to validate the microarray. We subsequently upgraded our design (v2) by adding a probe set for mitochondrial CDS. The probe sequences and features are shown in Online Resource 2. Probe set v1 has been deposited in NCBI's Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/gds/) and is accessible through GEO Platform number GPL21331.

Table 1

Probe distribution in custom *A. nidulans* microarray (v1)

Probe group	Type	Number of targets	Number of probes	Purpose
1[a]	All spots	10,607	15,744	*A. nidulans* transcriptome coverage and internal controls
2[b]	*A. nidulans* genes	10,550	15,198	*A. nidulans* transcriptome profiling
3	Genes with two probes (not replicated)	4553	9106	Expression analysis of annotated genes and determination of inter-probe variance
4	Genes with one probe (not replicated)	1106	1106	Expression analysis of annotated genes
5	Genes with one probe (not replicated)	4891	4891	Expression analysis of non-annotated genes
6	Genes with one probe (replicated five times)	17[c]	85	Determination of intra-probe variance
7	Genes with two probes (replicated five times)	1[c]	10	

[a]Group 1 = sum of groups 3–9

[b]Group 2 = sum of groups 3–7

[c]The 19 "marker" genes represented by fivefold replicated probes on the array are: ANID_00651, ANID_00670, ANID_00896, ANID_00973, ANID_01808, ANID_02343, ANID_02734, ANID_04976, ANID_05031, ANID_05634, ANID_05893, ANID_06499, ANID_06542, ANID_06653, ANID_06717, ANID_08041, ANID_08275, and ANID_10973 (*A. nidulans*) and V00296.1 (*E. coli*). Altogether groups 3–8 include the 15,124 distinct biological probes

Probe group	Type	Number of targets	Number of probes	Purpose
8	*E. coli* genes with two probes (replicated five times)	1[c]	10	Exogenous control and determination of intra-probe variance
9	Positive and negative controls	57	536	Manufacturer requirement for intra-array reproducibility control and normalization

[a]Group 1 = sum of groups 3–9

[b]Group 2 = sum of groups 3–7

[c]The 19 “marker” genes represented by fivefold replicated probes on the array are: ANID_00651, ANID_00670, ANID_00896, ANID_00973, ANID_01808, ANID_02343, ANID_02734, ANID_04976, ANID_05031, ANID_05634, ANID_05893, ANID_06499, ANID_06542, ANID_06653, ANID_06717, ANID_08041, ANID_08275, and ANID_10973 (*A. nidulans*) and V00296.1 (*E. coli*). Altogether groups 3–8 include the 15,124 distinct biological probes

Steps of oligonucleotide probe design

The design steps were guided by probe evaluation through the base composition scoring (BC score) associated to the eArray design software. eArray was first launched in “best probe” (BP) mode to design two probes per nuclear annotated CDS (5666), and the resulting probes exhibiting quality score lower than the highest BC1 level were discarded. Then, the 762 input CDS having only one probe after this selection was subjected to a new design session in “best distribution” (BD) mode, providing 762 new probes, which were all included. Unique probes were designed in BP mode for the 4894 not (4664) or poorly (230) annotated CDS. Each one of all the probes described above was synthesized once on the v1 microarray (Online Resource 2). Moreover, for 18 *A. nidulans* annotated CDS chosen as “markers” (because of their expected different expression levels in the studied model), one or two additional BD probes with BC1 score were replicated five times on the array. Two additional control probes were designed in “tilling” (TIL) mode for the V00296 sequence (*lacZ* gene, marker status), encoding β-galactosidase from *E. coli*, and synthesized fivefold each. Finally, we designed (in v2 design version) two probes in BP mode (80 % of them with BC1 quality level) for each of the 20 mitochondrial CDS. Due to the presence of replicated probes, 11 of the 4553 annotated *A. nidulans* mRNA, which were interrogated by more than one probe on v1 array, had three probes (marker genes, seven

spots), one had four probes (ANID_00670, riboB marker gene, 12 spots), and the remaining ones had two probes (six spots for marker genes, two spots for the others). eArray did not produce any probe for ten of the input set of 10,560 transcript sequences, which were strictly identical to ten alternative sequences present in the input set. Consistently, within most of these entity pairs, one sequence was described as an alias of the other in AspGD database at the time of writing (i.e., ANID_05382 and ANID_05381; ANID_05541 and ANID_05540; ANID_08518 and ANID_08515; ANID_09411 and ANID_04288; ANID_09459 and ANID_08193; ANID_09518 and ANID_02539; ANID_09527 and ANID_06576; ANID_10106 and ANID_00643; ANID_08685 and ANID_08686; and ANID_09240 and ANID_05252). Thus, our v1 probe set targeted 100 % of the input nuclear transcript sequences, while v2 probe set targeted 100 % of the nuclear and mitochondrial expressed sequences provided by Broad Institute database at the time of this work.

In silico quality evaluation of microarray oligonucleotide probes

Long oligonucleotide probes, such as the 60-base probes used here, have been shown to provide optimal specificity and sensitivity (Relógio et al. 2002). A first evaluation according to eArray quality scoring (i.e., BC scores) revealed that the distribution of quality levels depended on design mode used, since the BP mode provided overall a majority of probes with the highest quality score (89 % of BC1, 11 % of BC2, and less than 1 % of BC3), while the BD mode generated lower quality levels (35 % of BC1, 64 % of BC2, and 1 % of BC3) on average across all design sessions. However, after selection of probes of the highest quality regarding BC scores, the v1 probe set comprised 99.9 % of probes with the highest quality levels (BC1 or BC2), only 0.1 % being at BC3 level and none at the two lowest levels (BC4 and the worst BCpoor) (Table 2). Above this tool-dependent evaluation, we used three additional criteria for further in silico prediction of probe specificity and sensitivity, *i.e.*, cross hybridization potential, probe location relative to CDS end, and stability of secondary structures, according to previously reported recommendations (Lemoine et al. 2009; Bidard et al. 2010).

Table 2

Probe quality parameters

Criteria	**Measure**	**Quality levels[a] Number of probes (%)**		
eArray quality	BC score	BC1	BC2	BC3

Criteria	Measure	Quality levels[a] Number of probes (%)		
control		14,090 (93.2 %)	1012 (6.7 %)	22 (0.1 %)
Position in CDS	Number of nucleotides from the 3′-end	1–500	501–1000	>1000
		13,356 (88.3 %)	1659 (11.0 %)	109 (0.7 %)
Self-folding: hairpin[b]	$\Delta G_{65\ °C}$ (kcal/mol)	None	≥−1	<−1
		13,166 (87.1 %)	246 (1.6 %)	1712 (11.3 %)
Self-folding: dimer[b]	$\Delta G_{65\ °C}$ (kcal/mol)	None	≥−1	<−1
		12,201 (80.7 %)	782 (5.2 %)	2141 (14.2 %)
Cross-hybridization	% Identity to the best non-target hit	None	≤75 %	>75 %
		15,003 (99.2 %)	13 (0.1 %)	108 (0.7 %)

[a]All 15,124 probes (including two *E. coli* control probes) of v1 set where included in this description. The 40 new probes added in v2 set (not included in this table) had BC1 (for 80 %) or BC2 (for 20 %) quality scores, all of them hybridized within the 500 nucleotides from the 3′-end of targeted transcript, and showed no cross-hybridization potential

[b]The most favorable secondary structure was described for each probe

Predicted probes specificity

The v1 probe set was designed to target nuclear transcript sequences of average length 1466 [99–21,145] nucleotides (median 1242). More than 88 % of these probes hybridized within the 500 bases of the 3′-end of CDS (Table 2), and the median distance between hybridization site of 5′-end of probe and 3′-end of target was 160 [60–1184] bases. This proximity to 3′-end was consistent (1) with the fact that the mRNA 3′-UTR is the less-conserved region in eukaryotic transcripts and, thus, provides the highest specificity toward close paralogues (Tomiuk and Hofmann 2001), and (2) with the priming method used for reverse transcription (poly-dT), which results in 3′ ends over-representation in hybridized samples, thus improving sensitivity as well. Our v1 probe set contained 15,001 probes (99.2 %), which were predicted to hybridize with 100 % identity to their aimed specific *A. nidulans* target only, besides the two *E. coli lacZ*-specific control probes (Table 2).

BLAST of the v1 probe set against the whole set of input CDS showed non intended hits for 121 of the *A. nidulans* probes, of which 108 (targeting 96 unique transcripts, half of which without known function at the time of this work) showed more than 75 % identity with their non-target hit, thus generating a risk of cross hybridization (Kane et al. 2000). Of those probes, ten had 78–98 % identity with their non-intented target~~s~~, while the remaining had 100 % identity, which might be explained by a high degree of sequence identity between the targeted transcript and another one in the input set (paralogues), by the presence of alternative transcripts produced from the same locus, and/or by issues in sequence identification within the CDS set. It has to be noted that within v1 set, the multiple probes designed for *A. nidulans* annotated sequences, including “marker” genes comprised 290 pairs of probes exhibiting identical sequences (each pair targeting one unique CDS), which were generated by distinct eArray design sessions (Online Resource ~~1~~2). Consequently, we removed 40 redundant probes to generate the v2 version of design by including a new set of 20 pairs of probes with distinct sequences targeting specifically each of the 20 described mitochondrial CDS ~~(Table 3)~~.

Table 3

PCR primers used

Accession	Gene description	Forwad primer (5′–3′)	Reverse pri
ANID_08764	Aromatic ring-opening dioxygenase LigB subunit	GGGTGTCAAGAGGGGATTGG	GGACTGCC
ANID_05831	Glutathione transferase	TTGGGCCTTTACAAATTCGGC	CGGCAGGA
ANID_00858	Heat shock protein CLPA	AGACAACGGCGGAGATATGG	AATCGCTC
ANID_02343	Nitroreductase	TTTGAGTCCTCGACCAAGCC	TGATGTCC
ANID_08041	Glyceraldehyde-3-phosphate dehydrogenase	CGACAACGAGTGGGGTTACT	GGCATCAA
ANID_06542	Actin, gamma	GTACGATGAGAGCGGTCCTT	CAGAAAAT
V00296.1	*E. coli* gene lacZ coding for beta-galactosidase (EC 3.2.1.23)	CAGCAGCAGTTTTTCCAGTTC	ACATCCAG

AQ3

Predicted probes sensitivity

The Tm of a probe is an essential criterion for its sensitivity of detection, since an inappropriate Tm can result in a signal saturation (if too high) or, conversely, in a “close-to-background” signal (if too low). Mean Tm of probes (v1 set) calculated according to the nearest-neighbor model was 90.82 ± 1.50 °C [83.60–98.72], with a median value of 90.79 °C, while eArray gave 81.12 ± 1.50 °C [73.87–89.07] with median of 81.09 °C. Both Tm calculations resulted in an interquartile range of 1.23, indicating a highly narrow Tm distribution (Fig. 1b), which is a much more important parameter for probe sensitivity than the absolute Tm value, since formation of all probe-target hybrids occurs simultaneously and at the same temperature on microarray. Consistently, GC percent was 44.8 ± 2.8 % [35.0–60.0], with median value of 45.0 %, which is in accordance with usually accepted range (Lemoine et al. 2009). Sensitivity is ~~being~~ also defined by the strength of

probe–target sequence binding (Rychlik and Rhoads 1989), ~~and~~ therefore we calculated the $\Delta G_{65\ ^{\circ}C}$ of probe–target perfect matches, which was −41.99 ± 2.06 [−52.95 to Please choose an homogeneous way of writing pairs of minimum and maximum values: either [min-max] or [min to max] and apply it everywhere in the paper. −31.75] kcal/mol (median −41.94). Moreover, since the strength of probe's signal is ~~being~~ greatly influenced by the secondary structure formation, free energies of the most probable secondary structures were evaluated (Table 2; Fig. 1a). In the conditions of our calculation, one hairpin structure or more was predicted for 1958 probes (13 % of v1 probe set), while one autodimerization event or more was predicted for 2923 probes (19 % of v1 probe set). Only a minority of probes formed hairpins and/or autodimers with $\Delta G_{65\ ^{\circ}C}$ lower than −1 kcal/mol (11 and 14 % of probe set, respectively), an energy cutoff, which was previously associated with decreased detection signals (Lemoine et al. 2009). However, the formation of perfect-match hybrids was thermodynamically highly favored ($p < 0.0001$) in comparison with any predicted secondary structure described above (Fig. 1a). The inclusion of hairpin events predicted by AutoDimer with more permissive settings for 15,093 out of 15,124 probes (99.8 %) led to similar statistical result (data not shown).

Microarray performances for transcriptome analysis

To evaluate the experimental behavior of our v1 microarray, we conducted a comparative transcriptome analysis of two *A. nidulans* strains: a transgenic one harboring a construct with the *E. coli lacZ* gene, and its isogenic counterpart. Quality of signals generated by the probes for each CDS was assessed according to two experimental criteria: (1) probes sensitivity through their SNRs and (2) signal reproducibility across replicates of the same probe or between alternative probes targeting the same CDS. The overall array quality was evaluated by analysis of technical replicates of fluorescent target labelling and hybridization, as well as through challenging by RT-qPCR as an alternative technology.

Experimental probes sensitivity

The SNR for a nucleic probe reflects its ability to distinguish a true signal from its background according to a threshold value. The SNR of probes was evaluated by three calculation methods, of which SSDR parameter has been shown to generate the lowest level of false positive and false negative spots (He and Zhou 2008). Signals exhibiting SNR values, which were higher than

the following thresholds values, which were chosen on the basis of the high stringency of hybridization used, were considered as specific hybridizations: SSR > 2.0, SBR > 1.6, and SSDR > 0.8. Thus, through all 16 hybridized arrays, specific signals were obtained for 84.1, 77.5, and 83.7 % of probes according to these three parameters, respectively. The average SSR value calculated for all probes on microarray was 372, ranking up to 126,000. Moreover, 71.5 % of all non-control (*i.e.*, "biological", *A. nidulans* or *E. coli*) spots showed SSR > 10, which has been previously considered as indicative of high-quality array signals (Leiske et al. 2006). The probes showing low SNR value may either correspond to genes that were not expressed under the experimental conditions used, or to false positive genes resulting from overannotation (Bidard et al. 2010). The ratio between the signal intensity for each non-control probe and the average signal intensity of negative controls on the array was calculated to determine if a given gene was expressed (Li et al. 2008). Using an arbitrary ratio of 2.0, it was found that 80.2 % of probes corresponded to expressed genes in the studied strains, consistently with SNR evaluation.

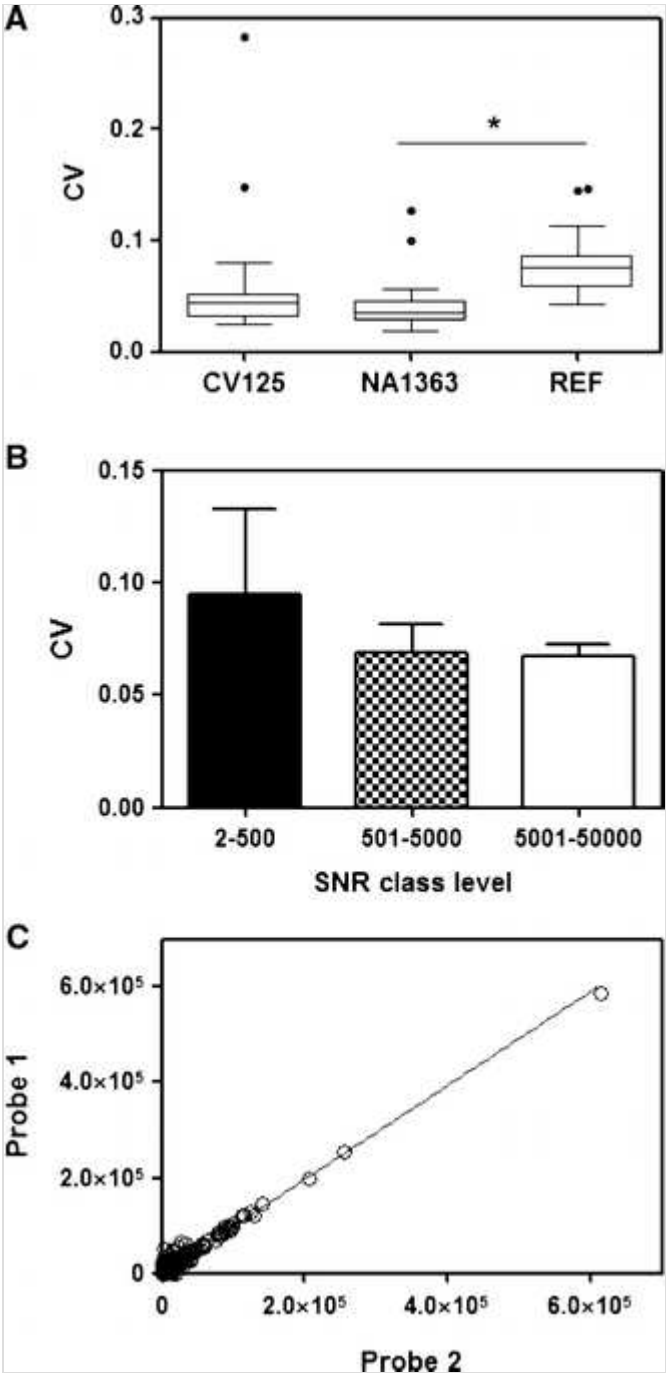
Experimental probes reproducibility

Firstly, intra-probe (and inter-array) reproducibility was determined by measuring the dispersion of signal intensities for identical probes that were replicated on several locations of the array surface. Among the 21 "marker" probes, which were each synthesized five times on the array, the one targeting the ANID_05031 sequence produced fluorescence signal beyond the SNR thresholds defined above. For the 20 probes, which produced detectable signal, the coefficient of variation (CV) of processed signal was calculated within replicates of each dye-condition combination. As shown in Fig. 2a, the replicated probes exhibited similarly low variance levels with Cy3 labelling of CV125 and NA1363 samples (median CV < 4.5 %; $n = 8$ for each), while Cy5 signals provided by the reference RNA pool showed slightly higher variance (median CV of 7.7 %; $n = 16$), which significantly differed from variance of Cy3 signal measured in NA1363 strain. Altogether these values could be considered as an excellent level of probe reproducibility, according to previously reported recommendations (Shi et al. 2006; Bidard et al. 2010). The 20 marker probes corresponding to detectable gene expression provided Cy5 signals distributing across SSR range up to 500 (eight probes), 5000 (nine probes), or 50,000 (three probes) (Fig. 2b). The reproducibility of replicated probes was thus similar and satisfying for low- and high-intensity signals, thus indicating minimal inter-array variation and reproducible target

preparation and hybridization.

Fig. 2

Reproducibility of normalized signals from replicated probes (v1 set). Intra-probe reproducibility. **a** The distribution of median CV of signals provided by 20 biological probes each fivefold replicated is shown for Cy3 ("CV125″ or "NA1363″ samples, $n = 8$ for each) or Cy5 (REF sample, $n = 16$) normalized signals, in Tukey *box*- and *whiskers-plot*, with median values of 4.4, 3.6, and 7.5, respectively ($*p < 0.05$); the replicated probe targeting ANID_05031 was excluded from the analysis, since the signal it produced in our strain and culture conditions was lower than the SNR-based detection threshold. **b** The distribution of median CV values of normalized Cy5 signals (REF sample, $n = 16$) from 20 replicated probes is shown across three classes of SNR levels, as mean ± SD (SSR ranges 2–500, 501–5000, 5001–50,000 included eight, nine, and three probes, respectively). **c** Inter-probe reproducibility: correlation between Cy5 processed signals obtained from the reference pool ($n = 16$ arrays) for 4553 pairs of alternative probes targeting the same transcript in both *A. nidulans* strains ($y = 0.985x + 492$, $R^2 = 0.964$; $p < 0.0001$); similar correlation and statistical significance levels were obtained for Cy3 signals in CV125 ($y = 0.980x - 357$, $R^2 = 0.954$; $p < 0.0001$) or NA1363 ($y = 1.012x + 393$, $R^2 = 0.953$; $p < 0.0001$) strains ($n = 8$ arrays for each), respectively (graphs not shown)



Moreover, inter-probe reproducibility was evaluated through the correlation between signals measured with pairs of probes designed to target the same gene. Indeed, distinct probe sequences may have different affinities for their cRNA targets, thus raising issues of results accuracy. The analysis of alternative probe reproducibility was allowed by the presence on our array of two non-replicated probes for each of 4553 transcripts. As shown in Fig. 2c, alternative probes for this set of targets returned highly correlated signal intensity values either with Cy5 ($R = 0.98$; $p < 0.0001$) or Cy3 labelling.

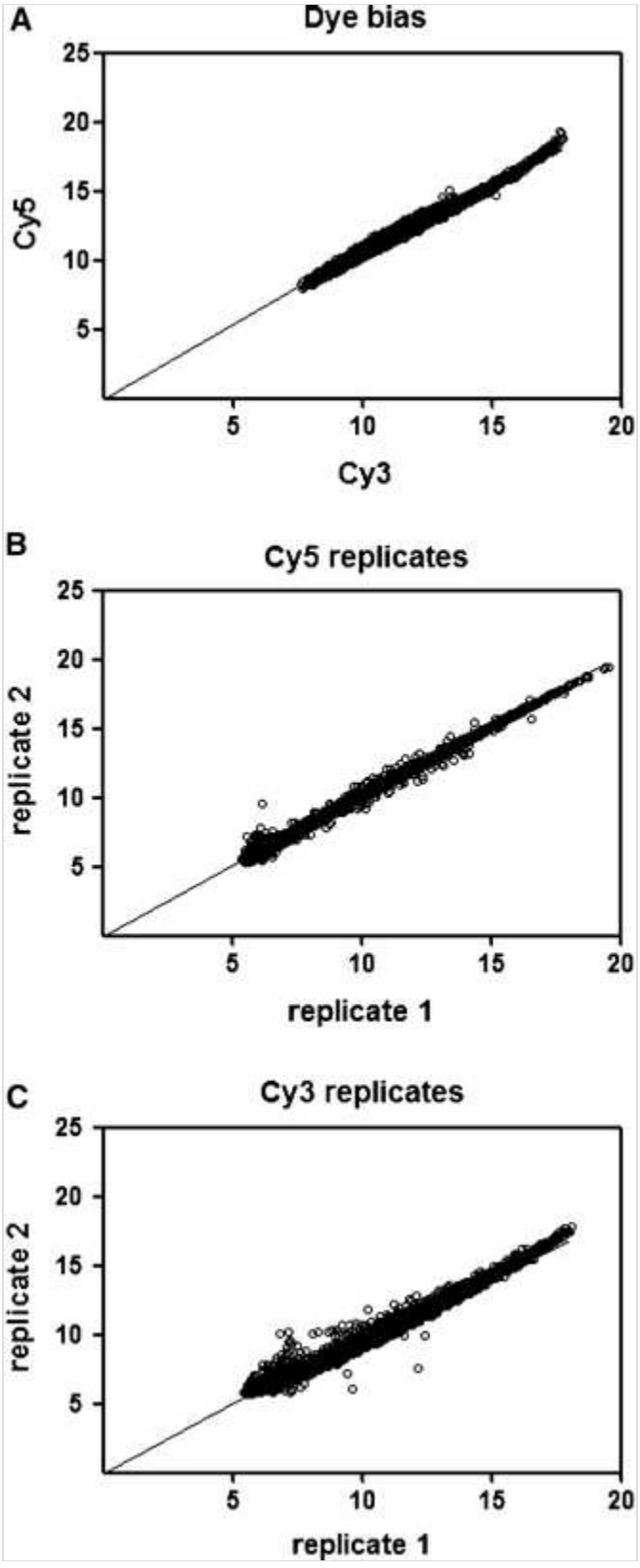
Array quality

To evaluate the array quality, we compared: (1) signals from the same sample

labelled with Cy5 and Cy3 (dye swap), and (2) two technical replicates from the same sample labelled with the same dye, Cy3 or Cy5. The high correlation coefficient observed between both color signals from the same sample assessed the absence of dye bias effect (R = 0.99; Fig. 3a). Moreover, the excellent correlation of signal intensities between two technical replicates labelled with Cy5 or Cy3 hybridized on separate arrays, either within the same slide and from the same labelling reaction (R = 0.99; Fig. 3b) or on distinct slides and from independent labelling reactions (R = 0.99; Fig. 3c), ~~and~~ attested the quality of microarray with regard to its whole hybridization process. Given this high level of technical reproducibility, only biological replicates were performed in the transcriptomic study, and each labelled cRNA sample synthesized from one individual *A. nidulans* culture was hybridized on one array only.

Fig. 3

Correlation of signal intensities between biological replicates or dye swaps. The correlation of $\log_2$ signal intensities is shown between (**a**) two target sets labelled from the same CV125 sample with Cy3 and Cy5, respectively ($y = 0.966x + 0.862$, $R^2 = 0.987$, $p < 0.0001$); **b** two samples of the same Reference pool labelled with Cy5 and hybridized on the same slide ($y = 0.987x + 0.020$, $R^2 = 0.998$, $p < 0.0001$); and **c** two independent Cy3-labeling from the same CV125 RNA sample hybridized on distinct slides ($y = 0.900x + 0.541$, $R^2 = 0.989$, $p < 0.0001$)



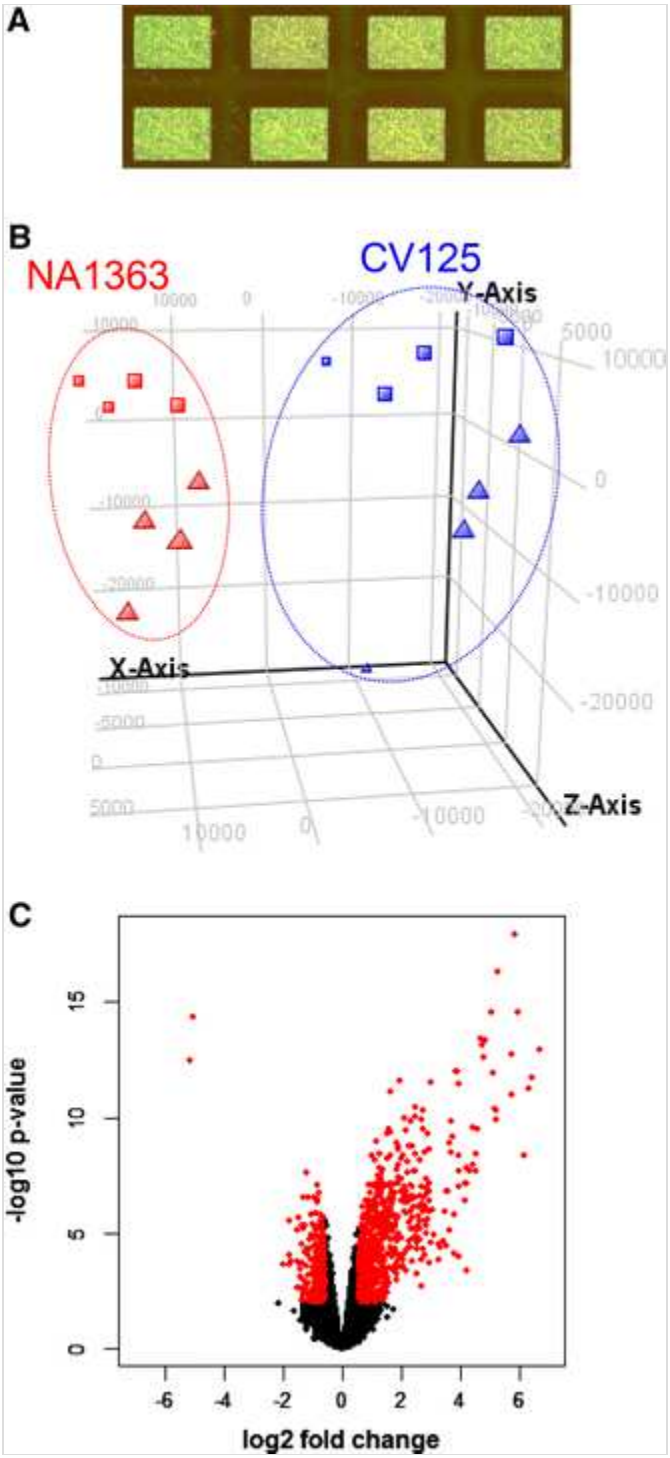
Overall features of comparative transcriptome analysis

The custom *A. nidulans* microarray presented here (Fig. 4a) allowed a good discrimination of overall array signal according to the strain genotype rather than to the culture batch, as shown by principal component analysis (Fig. 4b). The whole transcriptome analysis revealed a set of 1116 probes relating to 865 genes (*i.e.*, 8 % of studied transcriptome) significantly up- or down-regulated

($p < 0.01$, absolute fold-change higher than 1.50) in the NA1363 *lacZ*-harboring transgenic *A. nidulans* strain versus its wild-type counterpart (Fig. 4c). The data discussed in this report have been deposited in NCBI's Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi) and are accessible through GEO Series accession number GSE76918.

Fig. 4

Performance of custom 8 × 15 K *A. nidulans* (v1) microarray for transcriptome analysis. A l*acZ*-harboring transgenic strain (NA1363) of *A. nidulans* was compared with its wild-type counterpart (CV125). **a** Image of the slide after hybridization according to manufacturer's instructions. **b** Principal component analysis of gene expression profiles using Genespring 12.6. Targets from two independent culture series (*square* and *triangle*) were hybridized for each strain. The genotype parameter markedly discriminates the data sets, and the three first principal components explain 23 % (*x*), 16 % (*y*), and 9 % (*z*) of data dispersion, respectively. **c** Volcano plot of statistical results (cutoff: *p* value 0.01; fold-change 1.5) obtained from comparison of both strains



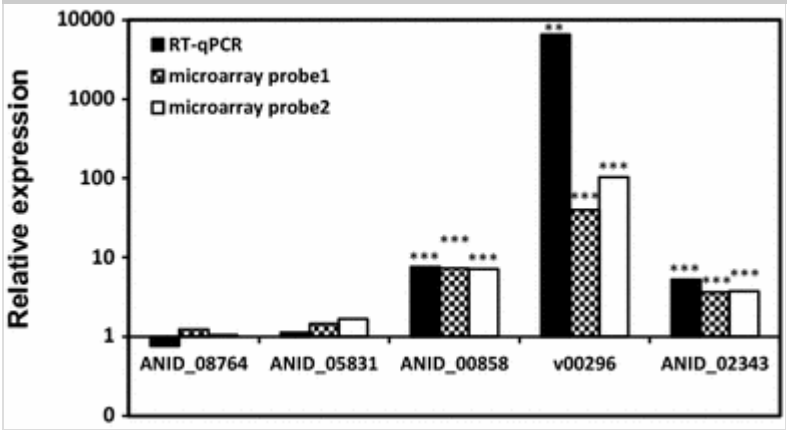
RT-qPCR validation of microarray results

Five genes were chosen in various expression levels and fold-change ranges from the comparative transcriptome analysis performed on the herein described custom microarray, and were subjected to RT-qPCR analysis (Fig. 5). Three of them had shown a differential expression in microarray experiment (mean fold-changes 2.8, 5.9, and 32.5; $p < 0.001$), which was confirmed by RT-qPCR (mean fold-changes 4.1, 7.4, and 2215, respectively; $p < 0.001$). The two tested genes, whose expression was stable according to microarrays, returned no significant expression change by RT-qPCR as well.

Thus, the cross-validation approach successfully confirmed the microarray results, despite a compression effect on high fold-changes in microarray compared with RT-qPCR, as previously described (Ferraresso et al. 2008).

Fig. 5

Validation of microarray data by RT-qPCR analysis. Fives genes were selected, of which three of the 865 genes whose expression was significantly modified in transgenic *A. nidulans* strain harboring *lacZ*, and two non-modulated genes, according to data from v1 microarray experiment. Mean fold-changes measured by either RT-qPCR or by two alternative probe sequences on microarray are shown for NA1363 strain versus CV125 control strain (n = 8 for each) (**$p < 0.01$, ***$p < 0.001$)



Discussion

We provide here the first fully characterized design for Agilent microarray technology in *A. nidulans*, with predictive as well as experimental evaluation. Several custom microarrays have been developed and used for gene expression studies in *A. nidulans* since the early 2000s. Initial designs made of spotted cDNA, which partially covered *A. nidulans* transcriptome (Pocsi et al. 2005), were followed by whole transcriptome oligo arrays, made of 70-mer spotted (Breakspear and Momany 2007) or in situ synthesized oligonucleotides. The latter included designs for high-density array platforms, *i.e.*, Affymetrix including one for three *Aspergillus* species (Andersen et al. 2008) and a Nimblegen 135 K chip (Bromann et al. 2012), as well as medium density arrays, such as 50-mer oligonucleotide (Sarkar et al. 2012) and Agilent arrays. Three Agilent designs were published, including two for use in 4 × 44 K format (Colabardini et al. 2013; Andersen et al. 2013) and recently one in 8 × 15 K format (Piłsyk et al. 2015). However, the reliability and the performance of the above cited microarray designs could be difficult to assess

for the scientific community in the *Aspergillus* research field, because none of them were provided with a detailed characterization.

We designed successively two probe sets, and performed on the first v1 set an extensive validation of the predictive and experimental performance of our protocols for microarray design and hybridization. The high degree of experimental reproducibility of signals that we have established here for replicated probes allowed us to replace 40 probes considered as redundant in v1 design, by 40 new probes targeting specifically the 20 mitochondrial CDS, thus generating an alternative probe set named v2. For a subgroup of 96 CDS (0.9 % of input set), the selected probes exhibited cross-hybridization potential, due to high sequence identity between distinct targets belonging to the set of CDS used. This could be related to sequence redundancy events within the set of CDS sequences, such as the existence of multiple alternative splicing variants reported for 8.3 % of *A. nidulans* genes, and might be improved in future updates of the microarray design, because of progression of genomic knowledge in *A. nidulans* (Sibthorp et al. 2013). The 260 redundant probes remaining in v2 design potentially authorize future improvement of the microarray, since they could be replaced by new probes for newly identified *A. nidulans* transcripts.

Overall, this new *A. nidulans* 8 × 15 K microarray meets technical and scientific requirements, and oligonucleotide probes globally satisfied to the tests for in silico benchmark of oligo sets given the overall good scores displayed for the main parameters (Lemoine et al. 2009; Bidard et al. 2010). This new probe design has been rigorously validated for its specificity, sensitivity, and reproducibility, and showed a good confirmation level by independent RT-qPCR quantitation, so that data are generated with high degree of confidence. Using optimized methods for sample preparation, labelling, and hybridization, this new microarray enables accurate gene expression profiling of *A. nidulans* for biological studies. Its low-density format allows limited cost together with rapid data analysis, making it suitable for studying large numbers of experimental conditions. To the best of our knowledge, our microarray is one of the first publicly available designs for low-density 8 × 15 K Agilent platform. It offers the largest set of unique probes in this format and is also the first Agilent design, whose performances have been extensively characterized.

Our newly developed microarray could be used in biomedical or agricultural as well as in basic research. Microarrays designed for *Candida* species were

used to identify genes, which were associated with nosocomial infectious process in human blood (Fradin et al. 2005). Transcriptome profiles could also be used to understand the resistance mechanisms to antifungal drugs (Sanglard 2003). In agriculture, *Aspergillus* species microarrays have been used to study mycotoxin biosynthesis, which could allow the development of strategies for controlling food contaminations by carcinogenic aflatoxins (Yu et al. 2006). In another study, a microarray analysis has allowed the understanding of infection process and pathogen–host interactions in the rice blast fungus *Magnaporthe oryzae* (Oh et al. 2008). The development of reliable whole genome arrays is also a first step toward the identification of specific gene signatures usable as biomarkers of environmental stress in microbial ecology studies, since they ~~have~~ are a powerful tool for investigating the responses of organisms to normal or adverse environmental conditions (Fuller et al. 2015; Hernández-Oñate and Herrera-Estrella 2015). Indeed, an increasing amount of studies are using microarray analysis to measure the activity of multiple microbial populations in relation to different environmental factors (fertilizers or pesticides metabolism for instance) (Gentry et al. 2006). Because of its ease of use (due to the fact it is not pathogenic), and since it has been extensively studied as a model organism, *A. nidulans* is well characterized from a genomic, proteomic, metabolic, or physiologic point of view. It is, therefore, a highly suitable organism not only to get new insights into the understanding of pathogenic fungi metabolism, but also to better understand the impact of environmental stress on soil filamentous fungi that are relevant indicators of agricultural soil ecosystem health.

In conclusion, although it covers 100 % of the *A. nidulans* targeted coding sequences, this low-density array allows limited experimental costs and simplified data analysis process, making it suitable for studying large numbers of experimental conditions in basic, biomedical as well as agricultural research. We strongly encourage authors publishing studies made with custom microarray designs to complete a detailed analysis of the platform they used, not only by doing a RT-qPCR validation, but also by including an in silico performance analysis of probe specificity and sensitivity, as well as an assessment of reproducibility at intra-probe and inter-probe levels. This is needed to achieve the reproducibility of differential expression discoveries in studies using microarrays.

Acknowledgments

This work was supported by the non-governmental organization “Générations Futures” and the Committee for Independent Research and Information on Genetic Engineering (CRIIGEN), in the framework of a participatory research project. It received funding from the Regional Council Ile-de-France and the University Paris-Sud. We are grateful to Dr Philip de Groot (Netherlands) for assistance in sequence analysis. We thank Nasim Honarmandi (France) and Jérôme Lecardonnel (Platform @BRIDGe, France) for technical help in microarray experiments. We also thank Dr Marco Moroldo (Platform @BRIDGe, France) for helpful discussion. We thank Dr Didier Goidin (Agilent Technologies, France) for helpful contribution to this work.

Electronic supplementary material

Below is the link to the electronic supplementary material.

Supplementary material 1

Construction of the NA1363 *A. nidulans* strain. The map of the ~~pBRgpdAp-LacZ-RiboB~~*pBRgpdAp-lacZ-riboB* plasmid is provided, as well as the sequential steps to generate the studied strain. (DOCX 142 kb)

Supplementary material 2

Probe list. The sequences, names and features of each probe of the 8 x 15 K custom microarray (v1 and v2 probe sets) are presented. For traceability of probe design process, information about design mode (BP, BD or TIL) and eArray quality score (BC1, BC2 or BC3) was included in each probe name. The v2 probe set was generated from the v1 probe set by removing 40 probes (considered as redundant because of sequence identity with another probe or potential cross-hybridization with two TargetIDs) and replacing them by 40 new probes specific of the 20 mitochondrial CDS. The pairs of redundant probes (with identical sequence) are identified in the table. The “marker” status is associated to the probes synthesized 5 times on the array (XLS 7182 kb)

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl Acids Res 25:3389–3402

Andersen MR, Vongsangnak W, Panagiotou G, Salazar MP, Lehmann L,

Nielsen J (2008) A trispecies *Aspergillus* microarray: comparative transcriptomics of three *Aspergillus* species. Proc Natl Acad Sci USA 105:4387–4392

Andersen MR, Nielsen JB, Klitgaard A, Petersen LM, Zachariasen M, Hansen TJ, Blicher LH, Gotfredsen CH, Larsen TO, Nielsen KF, Mortensen UH (2013) Accurate prediction of secondary metabolite gene clusters in filamentous fungi. Proc Natl Acad Sci USA 110:E99–E107

Arnaud MB, Cerqueira GC, Inglis DO, Skrzypek MS, Binkley J, Chibucos MC, Crabtree J, Howarth C, Orvis J, Shah P, Wymore F, Binkley G, Miyasato SR, Simison M, Sherlock G, Wortman JR (2012) The Aspergillus Genome Database (AspGD): recent developments in comprehensive multispecies curation, comparative genomics and community resources. Nucl Acids Res 40((Database issue)):D653-9

Arst HN, Scazzocchio C (1985) Formal genetics and molecular biology of the control of gene expression in *Aspergillus nidulans*. In: Bennett JW, Lasure LL (eds) Gene manipulations in fungi. Academic Press, New York, pp 309–343

Bidard F, Imbeaud S, Reymond N, Lespinet O, Silar P, Clavé C, Delacroix H, Berteaux-Lecellier V, Debuchy RA (2010) General framework for optimization of probes for gene expression microarray and its application to the fungus *Podospora anserina*. BMC Res Notes 18:171

Breakspear A, Momany M (2007) *Aspergillus nidulans* conidiation genes dewA, fluG, and stuA are differentially regulated in early vegetative growth. Eukaryot Cell 9:1697–1700

Bromann K, Toivari M, Viljanen K, Vuoristo A, Ruohonen L, Nakari-Setälä T (2012) Identification and characterization of a novel diterpene gene cluster in *Aspergillus nidulans*. PLoS One 7:e35450

Carranza CS, Bergesio MV, Barberis CL, Chiacchiera SM, Magnoli CE (2014) Survey of *Aspergillus* section *Flavi* presence in agricultural soils and effect of glyphosate on nontoxigenic *A. flavus* growth on soil-based medium. J Appl Microbiol 116:1229–1240

Colabardini AC, Brown NA, Savoldi M, Goldman MH, Goldman GH

(2013) Functional characterization of *Aspergillus nidulans* ypkA, a homologue of the mammalian kinase SGK. PLoS One 8:e57630

Cove DJ (1966) The induction and repression of nitrate reductase in the fungus *Aspergillus nidulans*. Biochim Biophys Acta 113:51–56

Davis MA, Hynes MJ (1991) Regulatory circuits in *Aspergillus nidulans*. In: Bennett JW, Lasure LL (eds) More gene manipula tion in fungi. Academic Press, New York, pp 151–189

Ferraresso S, Vitulo N, Mininni AN, Romualdi C, Cardazzo B, Negrisolo E, Reinhardt R, Canario AV (2008) Development and validation of a gene expression oligo microarray for the gilthead sea bream (*Sparus aurata*). BMC Genom 9:580–594

Fradin C, De Groot P, MacCallum D, Schaller M, Klis F, Odds FC, Hube B (2005) Granulocytes govern the transcriptional response, morphology and proliferation of *Candida albicans* in human blood. Mol Microbiol 56:397–415

Fu SF, Chen PY, Nguyen QT, Huang LY, Zeng GR, Huang TL, Lin CY, Huang HJ (2014) Transcriptome profiling of genes and pathways associated with arsenic toxicity and tolerance in *Arabidopsis*. BMC Plant Biol 14:94

Fuller KK, Loros JJ, Dunlap JC (2015) Fungal photobiology: visible light as a signal for stress, space and time. Curr Genet 61:275–288. doi:10.1007/s00294-014-0451-0

Galagan JE, Calvo SE, Cuomo C et al (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature 438:1105–1115

Gentry TJ, Wickham GS, Schadt CW, He Z, Zhou J (2006) Microarray applications in microbial ecology research. Microb Ecol 52:159–175

Gill RT, Katsoulakis E, Schmitt W, Taroncher-Oldenburg G, Misra J, Stephanopoulos G (2002) Genome-wide dynamic transcrip tional profiling of the light-to-dark transition in *Synechocystis* sp. strain PCC 6803. J Bacteriol 184:3671–3681

He Z, Zhou J (2008) Empirical evaluation of a new method for calculating signal-to-noise ratio for microarray data analysis. Appl Environ Microbiol 74:2957–2966

Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J (2007) qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol 8:R19

Hernández-Oñate MA, Herrera-Estrella A (2015) Damage response involves mechanisms conserved across plants, animals and fungi. Curr Genet 61:359–372. doi:10.1007/s00294-014-0467-5

Iwahashi H (2015) Pressure-dependent gene activation in yeast cells. Subcell Biochem 72:407–422

Kane MD, Jatkoe TA, Stumpf CR, Lu J, Thomas JD, Madore SJ (2000) Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. Nucl Acids Res 28:4552–4557

Kleijkers SH, Eijssen LM, Coonen E, Derhaag JG, Mantikou E, Jonker MJ, Mastenbroek S, Repping S, Evers JL, Dumoulin JC, van Montfoort AP (2015) Differences in gene expression profiles between human preimplantation embryos cultured in two different IVF culture media. Hum Reprod 30:2303–2311

Leiske DL, Karimpour-Fard A, Hume PS, Fairbanks BD, Gill RT (2006) A comparison of alternative 60-mer probe designs in an in situ synthesized oligonucleotide microarray. BMC Genom 7:72–80

Lemoine S, Combes F, Le Crom S (2009) An evaluation of custom microarray applications: the oligonucleotide design challenge. Nucl Acids Res 376:1726–1739

Li X, Chiang HI, Zhu J, Dowd SE, Zhou H (2008) Characterization of a newly developed chicken 44 K agilent microarray. BMC Genom 9:60

Liao BY, Weng MP (2015) Unraveling the association between mRNA expressions and mutant phenotypes in a genome-wide assessment of mice. Natl Acad Sci USA 112:4707–4712

Lin F, Jiang L, Liu Y, Lv Y, Dai H, Zhao H (2014) Genome-wide identification of housekeeping genes in maize. Plant Mol Biol 86:543–554

Liu F, Kuo WP, Jenssen TK, Hovig E (2012) Performance comparison of multiple microarray platforms for gene expression profiling. Methods Mol Biol 802:141–155. doi:10.1007/978-1-61779-400-1_10

Martinelli SD, Kinghorn JR (1994) *Aspergillus*: 50 years on—progress in industrial microbiology, vol 29. Elsevier, London

Masuo S, Terabayashi Y, Shimizu M, Fujii T, Kitazume T, Takaya N (2010) Global gene expression analysis of *Aspergillus nidulans* reveals metabolic shift and transcription suppression under hypoxia. Mol Genet Genom 284:415–424

Mesnage R, Arno M, Costanzo M, Malatesta M, Séralini GE, Antoniou MN (2015) Transcriptome profile analysis reflects rat liver and kidney damage following chronic ultra-low dose Roundup exposure. Environ Health 14:70

Morris NR, Enos AP (1992) Mitotic gold in a mold: *Aspergillus* genetics and the biology of mitosis. Trends Genet 8:32–37

Nitsche BM, Crabtree J, Cerqueira GC, Meyer V, Ram AF, Wortman JR (2011) New resources for functional analysis of omics data for the genus *Aspergillus*. BMC Genom 12:486–496

Oh Y, Donofrio N, Pan H, Coughlan S, Brown DE, Meng S, Mitchell T, Dean RA (2008) Transcriptome analysis reveals new insight into appressorium formation and function in the rice blast fungus *Magnaporthe oryzae*. Genome Biol 9:R85

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT–PCR. Nucl Acids Res 29:e45

Piłsyk S, Natorff R, Sieńko M, Skoneczny M, Paszewski A, Brzywczy J (2015) The *Aspergillus nidulans metZ* gene encodes a transcription factor involved in regulation of sulfur metabolism in this fungus and other Eurotiales. Curr Genet 61:115–125

Pócsi I, Miskei M, Karányi Z, Emri T, Ayoubi P, Pusztahelyi T, Balla G, Prade RA (2005) Comparison of gene expression signatures of diamide, H_2O_2 and menadione exposed *Aspergillus nidulans* cultures–linking genome-wide transcriptional changes to cellular physiology. BMC Genom 6:182

Pontecorvo G, Roper JA, Hemmons LM, Macdonald KD, Bufton AW (1953) The genetics of *Aspergillus nidulans*. Adv Genet 5:141–238

Relógio A, Schwager C, Richter A, Ansorge W, Valcárcel J (2002) Optimization of oligonucleotide-based DNA microarrays. Nucl Acids Res 30:e51

Robellet X, Oestreicher N, Guitton A, Vélot C (2010) Gene silencing of transgenes inserted in the *Aspergillus* nidulans *alcM* and/or *alcS* loci. Curr Genet 56:341–348

Roh SW, Abell GC, Kim KH, Nam YD, Bae JW (2010) Comparing microarrays and next-generation sequencing technologies for microbial ecology research. Trends Biotechnol 28:291–299

Rychlik W, Rhoads RE (1989) A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA. Nucl Acids Res 17:8543–8551

Sanglard D (2003) Resistance and tolerance mechanisms to antifungal drugs in fungal pathogens. Mycologist 17:74–78

Santa Lucia J Jr (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc Natl Acad Sci USA 95:1460–1465

Sarkar A, Funk AN, Scherlach K, Horn F, Schroeckh V, Chankhamjon P, Westermann M, Roth M, Brakhage AA, Hertweck C, Horn U (2012) Differential expression of silent polyketide biosynthesis gene clusters in chemostat cultures of Asp*ergillus nidulans*. J Biotechnol 160:64–71

Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467–470

Shi L, Reid LH, Jones WD, Shippy R, Warrington JA, Baker SC, Collins PJ, de Longueville F, Kawasaki ES, Lee KY et al (2006) The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nat Biotechnol 24:1151–1161

Sibthorp C, Wu H, Cowley G, Wong PW, Palaima P, Morozov IY, Weedall GD, Caddick MX (2013) Transcriptome analysis of the filamentous fungus *Aspergillus nidulans* directed to the global identification of promoters. BMC Genom 14:847

Steffen W, Richardson K, Rockström J, Cornell SE, Fetzer I, Bennett EM, Biggs R, Carpenter SR, de Vries W, de Wit CA, Folke C, Gerten D, Heinke J, Mace GM, Persson LM, Ramanathan V, Reyers B, Sörlin S (2015) Sustainability. Planetary boundaries: guiding human development on a changing planet. Science 347:1259855

Tomiuk S, Hofmann K (2001) Microarray probe selection strategies. Brief Bioinform 2:329–340

Vallone PM, Butler JM (2004) AutoDimer: a screening tool for primer-dimer and hairpin structures. Biotechniques 37:226–231

van Straalen NM, Roelofs D (2008) Genomics technology for assessing soil pollution. J Biol 7:19

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3:1–11

Wolber PK, Collins PJ, Lucas AB, De Witte A, Shannon KW (2006) The agilent in situ-synthesized microarray platform. Methods Enzymol 410:28–57

Wortman JR, Gilsenan JM, The Joardar V et al (2008) Update of the *Aspergillus nidulans* genome annotation: a community effort. Fungal Genet Biol 46:S2–S13

Xiao Y, van Hijum SA, Abee T, Wells-Bennik MH (2015) Genome-wide

transcriptional profiling of *Clostridium perfringens* SM101 during sporulation extends the core of putative sporulation genes and genes determining spore properties and germination characteristics. PLoS One 10:e0127036

Yu J, Cleveland TE, Wilkinson JR, Campbell BC, Kim JH, Kim HS, Bhatnagar D, Payne GA, Nierman WC (2006) *Aspergillus flavus* expressed sequence tags and microarray as tools in understanding aflatoxin biosynthesis. Mycotoxin Res 22:16–21

Zhang M, Zhang L, Zou J, Yao C, Xiao H, Liu Q, Wang J, Wang D, Wang C, Guo Z (2009) Evaluating reproducibility of differential expression discoveries in microarray studies by considering correlated molecular changes. Bioinformatics 25:1662–1668

AQ4