integrated heterogeneous multi-scale genomic, transcriptomic, and metabolomic data to redefine the metabolic framework of C_1-utilization in *M. trichosporium* OB3b grown in batch culture under copper, oxygen, and iron sufficiency on methane and nitrate as the sources of carbon and nitrogen, respectively. In this part of our work we present transcriptomic-based analysis of the methanotrophic metabolic network. Metabolomic and ^{13}C-labeling studies are presented in a follow-up paper (Yang et al., 2013).

RESULTS AND DISCUSSION

GENE EXPRESSION STUDIES

Gene expression studies were carried out with *M. trichosporium* OB3b cultures grown on methane at N (10 mM), Cu (9 μM), and Fe (9 μM) sufficiency conditions. The maximum specific growth rate of *M. trichosporium* OB3b in shake flasks during the exponential growth phase was $\mu = 0.038 \pm 0.004\,h^{-1}$. The methane consumption rate during the period of maximum growth rate was 8.95 mmol of $CH_4 h^{-1}$ L culture^{-1} ($OD_{600} = 1$).

All experiments were performed with at least two biological replicates. RNA samples were prepared as described in the Section "Materials and Methods." Illumina sequencing for two biological replicates (BR1 and BR2) returned 28 and 29 million 36-bp reads. The Burrows–Wheeler Aligner (BWA, Li and Durbin, 2009) aligned 98% of the reads to the *M. trichosporium* OB3b genome annotated by MaGE[1] using the default parameters for small genomes. Reads per kilobase of coding sequence per million (reads) mapped (RPKM) (Mortazavi et al., 2008) was calculated to compare gene expression within and across replicates, and no further normalization (other than RPKM) was applied. The samples were in good agreement with each other, with per gene coding sequence RPKM correlations of 0.959 and 0.989 for the Pearson and Spearman correlations, respectively. In total, 4,762 of 4,812 ORFs (CDS, tRNA, and rRNA predicted from the draft genome) were detected. Based on relative expression, genes (omitting rRNAs) could be grouped into six major expression categories (**Table 1**): *very high* (RPKM ≥ 15,000), *high* (RPKM ≥ 1,500), *moderate* (1,500 > RPKM ≥ 500), *modest* (500 > RPKM ≥ 250), *low* (250 > RPKM ≥ 150) *very low* (150 > RPKM ≥ 15), and *not expressed* (RPKM < 15). The majority of genes fell into *low/very low expression* categories (74%). About 14% of genes displayed *moderate/modest* expression and only a small fraction of the genome showed very *high/high expression* (2.7%).

In order to determine whether the draft genome of the strain is missing some functional genes, we performed *de novo* assembly of the transcriptome. Using this approach, a total of 173 genes that are not present in the genome sequence, but have homologs in the non-redundant database were detected. Among those are key subunits of succinate dehydrogenase (*sdhABCD*), 2-oxoglutarate dehydrogenase (E2), and nitric oxide reductase (*norB*) (Table S1 in Supplementary Material). The *de novo* transcriptome assembly provides additional information for highly expressed genes and it was used for verification of some metabolic functions that were predicted by enzymatic studies but were not detected in the draft genome assembly (see below).

[1] https://www.genoscope.cns.fr/agc/microscope/mage

Table 1 | Classification of gene expression level based on replicate averaged RPKMs.

Description of expression level	RPKM range	% of ORFs	Number of ORFs
Very high	>15,000	0.23	11
High	1,500–15,000	2.49	120
Moderate	500–1,500	5.30	255
Modest	250–500	8.61	414
Low	50–250	40.41	1,944
Very low	15–50	23.70	1,140
Not expressed	<15	19.27	927

In addition, the reads obtained from RNA-seq were aligned to the reference genome in order to identify transcription boundaries and transcription start sites for the most highly expressed genes, including the *pmoCAB* operons, *mxaFJGI* operon, *fae1*, *pqqA*, and key genes of the serine cycle (Table S2 in Supplementary Material, see description below). Gene expression data were used to reconstruct central metabolic pathways in *M. trichosporium* OB3b (**Table 2**; **Figure 1**; Table S2 in Supplementary Material). Core functions are described below.

C_1-OXIDATION: METHANE-TO-METHANOL

It has been previously demonstrated that *M. trichosporium* OB3b possesses two types of methane oxidation enzymes: pMMO and sMMO. The expression of the enzymes is determined by copper availability; sMMO is dominant in copper-limited environments while pMMO dominates under copper sufficiency (Hakemian and Rosenzweig, 2007; Semrau et al., 2010). Structures of both enzymes are available (Elango et al., 1997; Hakemian et al., 2008). In this study, *M. trichosporium* OB3b was grown at a copper concentration that has been shown to be sufficient to suppress the expression of sMMO (Park et al., 1991; Phelps et al., 1992; Nielsen et al., 1997; Lloyd et al., 1999; Murrell et al., 2000). Indeed, virtually no expression of the sMMO gene cluster (*mmoXYBZC*) was observed. In contrast, the *pmoCAB* genes were the most highly expressed in the transcriptome, representing about 14% of all reads mapped to the coding regions (**Table 2**). It has previously been shown that PMMO in *M. trichosporium* OB3b is encoded by two copies of the *pmoCAB* operon that appear to be identical (Gilbert et al., 2000). The current genome assembly failed to resolve these closely related duplicated regions. The *pmoCAB* genes were found within one relatively short contig, which includes 320 bp upstream from *pmoC*, and about 66 bp downstream from *pmoB*. It is possible that in the genome assembly, the *pmo* contig represents only those parts of the duplicated regions that are highly similar. Thus, it was not possible to determine relative expression of the two operons with the transcriptomic data.

Previous attempts to identify transcriptional starts of the *pmoCAB* operons in *M. trichosporium* OB3b using a conventional primer extension approach were not successful (Gilbert et al., 2000). The RNA-seq data were used for identification of transcriptional starts for the *pmoCAB* operons. Because the published METTOv1 genome did not contain a complete *pmoCAB* cluster, a separate alignment run was performed using a previously