

RECONSTRUÇÃO DA REDE METABÓLICA DA *Gluconacetobacter hansenii* – UM MODELO SIMPLIFICADO

S. S. SOUZA¹, L. M. PORTO¹

¹Universidade Federal de Santa Catarina,
Departamento de Engenharia Química e Engenharia de Alimentos
E-mail para contato: {samara, luismar}@intelab.ufsc.br

RESUMO – Na última década, reconstruções e aplicações de redes metabólicas em escala genômica influenciaram fortemente o campo da biologia de sistemas. A bactéria Gram-negativa *Gluconacetobacter* tem sido extensivamente utilizada para a síntese de celulose, mas só recentemente a sequência da *G. hansenii* ATCC 23769 tornou-se disponível, permitindo estudos deste organismo. A abordagem baseada em restrições tem sido utilizada para realizar reconstruções de redes metabólicas *in silico*. O objetivo foi reconstruir a rede com as principais vias do organismo. Além disso, um conjunto de ferramentas computacionais de biologia sistêmica foi aplicado à rede. Usando o modelo em conjunto com os métodos baseados em restrição, ABF (análise de balanço de fluxo) e AFM (análise de fluxo metabólico), os fluxos metabólicos foram simulados mimetizando três condições ambientais diferentes para se compreender os efeitos da produção de celulose, o crescimento bacteriano e a distribuição dos fluxos de metabólitos.

1. INTRODUÇÃO

Os avanços no sequenciamento de genomas têm progredido rapidamente, acompanhados do alto rendimento das análises e custos cada vez menores. Tais tecnologias estão rotineiramente sendo aplicadas em distintas áreas da biologia, da engenharia e da medicina, permitindo que os pesquisadores abordem questões biológicas relevantes (Metzker, 2010; Zhao e Grant, 2011). A era pós-genômica tem como fundamento relacionar a sequência do genoma anotado com as funções fisiológicas de uma célula, possibilitando reconstituir vias metabólicas completas (Durot *et al.*, 2009; Edwards *et al.*, 2001; Fong e Palsson, 2004; Soon *et al.*, 2013). Assim, surgiu o campo da bioinformática e da biologia computacional que consiste na aplicação de técnicas de computação intensivas para a compreensão dos processos biológicos (Covert *et al.*, 2001). Nos últimos anos, outra área tem se destacado, a biologia de sistemas, ou biologia sistêmica, que tem como um de seus focos as reconstruções metabólicas e a análise de dados baseada em restrições, que fornecem representações matemáticas e computacionais das capacidades metabólicas conhecidas de um dado organismo (Blank e Ebert, 2013; Bordbar e Palsson, 2012).

Através das ferramentas de engenharia metabólica, as principais vias bioquímicas de bactérias podem ser manipuladas e otimizadas para produzir de forma mais eficiente compostos industriais de interesse em diversas áreas (McCloskey *et al.*, 2013). Nessas redes metabólicas a modelagem baseada em restrição é preditiva com relação à identificação de reações, metabólitos e genes. Isto é extremamente útil para a determinação das necessidades metabólicas celulares e pode potencialmente simular o comportamento fenotípico sob diferentes condições genéticas e ambientes fisiológicos variados (Edwards *et al.*, 2001; Fong e Palsson, 2004; Mahadevan *et al.*, 2011).

Gluconacetobacter hansenii tem sido extensivamente utilizada para a síntese de celulose. A *G. hansenii* produz celulose extracelular livre de lignina ou hemicelulose, tornando-a uma excelente fonte de celulose pura. A celulose bacteriana (CB) representa uma alternativa tecnicamente vantajosa em diversas aplicações industriais, assim como na indústria biomédica (Ross *et al.*, 1991; Saxena *et al.*, 1994).

2. METODOLOGIA

O processo de reconstrução de modelo metabólico em escala genômica aqui apresentado envolveu quatro etapas principais: (1) a criação de um modelo *draft*, (2) a cura manual, (3), a conversão para um formato matemático e (4) avaliação da rede. O protocolo recomendado para reconstruções de redes metabólicas, aqui chamado de Protocolo T&P (Thiele e Palsson, 2010), é um protocolo bem estruturado para organismos bem conhecidos e com abundância de informações experimentais podendo, no entanto, ser adaptado e melhorado. A bactéria *G. hansenii*, por outro lado, tem fisiologia relativamente pouco conhecida, o que dificulta uma reconstrução muito abrangente. Para a execução e extensão dos passos do protocolo T&P foram realizadas algumas alterações visando uma reconstrução metabólica mínima para esse organismo.

2.1 Etapa 1 - Reconstrução preliminar – o *draft* do genoma

Os dados anotados da sequência genômica de *G. hansenii* (NCBI ID 714995) foram obtidos do banco de dados GenBank do NCBI (Iyer *et al.*, 2010). A versão mais recente disponível (ACCESSION NZ CM000920, 3636659 bp, de 07/03/2013) foi baixada para garantir que as atualizações e correções desde a publicação original do genoma fossem contabilizadas. Para a montagem do *draft* foram utilizadas duas ferramentas automáticas: o componente PathoLogic do *package* Pathway Tools (SRI International, Menlo Park, CA, USA) versão 16.5 que integra informações de reações, metabólitos e genes, e o recurso Model SEED, que apesar de não conter em seu banco de dados a bactéria *G. hansenii*, permitiu que o genoma fosse automaticamente submetido ao RAST Server, criando um banco do organismo com diversas informações, desde reações, vias, metabólitos e genes.

2.2. Etapa 2 - Cura manual

Essa etapa consistiu em corrigir as eventuais inconsistências, uma vez que a extração das reações metabólicas foi realizada de forma automática e não permite estabelecer certas características

do organismo alvo. Com as informações geradas na Etapa 1, avaliaram-se as reações em cada via e organizou-se as informações em subsistemas para montar o modelo simplificado.

Com o intuito de compreender a produção de celulose bacteriana foram identificadas as principais vias metabólicas envolvidas na sua biossíntese (metabolismo *core*). A escolha das vias baseou-se em bancos de dados online relacionados com o gênero *Gluconacetobacter* como, por exemplo, KEGG (*Kyoto Encyclopedia of Genes and Genomes*), IMG (*Integrated Microbial Genomes*) e os *drafts* gerados no PathoLogic e RAST Server. Isso incluiu a análise de diversas vias como: glicólise, gliconeogênese, ciclo dos ácidos tricarbóxicos, via das pentoses-fosfato, metabolismo do nitrogênio, transporte de elétrons e fosforilação oxidativa, biossíntese da celulose, além das reações de transporte e biomassa.

Nesta etapa algumas restrições foram definidas no modelo, dentre elas restrições estequiométricas (impostas pelos balanços de massa) e restrições termodinâmicas (através da reversibilidade e irreversibilidade dos fluxos). Todas as informações obtidas durante a cura manual foram registradas em planilhas de MS-Excel, que facilitaram o processo de reconstrução.

Um passo essencial na reconstrução do modelo metabólico foi a adição de reações que não são inferidas automaticamente a partir de anotação do genoma. Tais reações incluem: (1) reações espontâneas, (2) reações de transporte, (3) reações de troca e (4) reações de demanda. A adição destas reações diminui o número de metabólitos sem saída (*dead-ends*) e melhora a conectividade da rede. Foi necessária a adição da reação de biomassa para permitir a simulação de crescimento e divisão celular. Para escrever esta reação foram utilizados dados experimentais da bactéria *Gluconacetobacter xylinus*. A partir da demanda de todos os precursores a reação foi escrita e adicionada ao modelo (Zhong *et al.*, 2013).

2.3. Etapa 3 - Conversão da construção metabólica para o formato computacional

O modelo metabólico construído e refinado foi convertido em uma representação matemática (matriz estequiométrica) que permitiu implementar uma grande variedade de ferramentas computacionais para analisar as propriedades da rede. Foram adicionadas as restrições de capacidade que são os limites superiores e inferiores, que definem o fluxo máximo e mínimo permitido das reações. Além disso, foram determinadas as taxas de consumo das fontes de nitrogênio e fosfato de acordo com o meio mínimo definido. Dentre os meios de cultura citados na literatura para esta bactéria, foi escolhido o meio de crescimento Yamanaka (Ruka *et al.*, 2012).

Para a etapa computacional foram utilizados o software MATLAB[®] (Mathworks, Natick, MA, USA) e os *toolboxes* COBRA (UC San Diego) e GEnSys. O GenSys (*Genomic Engineering System*) foi desenvolvido pelo grupo Intelab (UFSC) e foi adicionado ao protocolo para realizar as simulações de análise baseada em restrições: AFM (Análise de Fluxo Metabólico) e ABF (Análise de Balanço de Fluxo). Utilizando programação linear para resolver o modelo estequiométrico, o vetor de fluxos pode ser encontrado através da especificação de uma função objetivo que foi maximizada.

O modelo foi simulado fixando as velocidades de consumo de glicose e manitol de $10 \text{ mmol} \cdot \text{gDW}^{-1} \cdot \text{h}^{-1}$ (gDW: grama de massa seca), e para o glicerol de $20 \text{ mmol} \cdot \text{gDW}^{-1} \cdot \text{h}^{-1}$, ou seja, em equivalência de carbomol, uma vez que glicose e manitol apresentam seis carbonos e o glicerol apenas três. Para a ABF foram definidas as funções objetivo que mimetizam diferentes situações de maximização/minimização que o organismo necessita em determinada condição. Foram avaliados quatro cenários considerados relevantes para o presente estudo. Com o modelo computacional buscou-se investigar as capacidades metabólicas e avaliar as propriedades sistêmicas que elas podem executar sob as restrições conferidas à célula. Sendo assim, a última etapa no processo de reconstrução consistiu na verificação, avaliação e validação da rede.

3. RESULTADOS

Primeiramente o *draft* do modelo foi construído no PathoLogic do Pathway Tools e criado o respectivo PGDB (*Pathway/Genome Database*), aqui denominado de GhaCyc, contendo as vias metabólicas da *G. hanseni*, utilizando como entrada os arquivos GenBank (No formato FASTA), previamente obtidos no banco de dados do NCBI. A Figura 1 apresenta o mapa celular gerado nessa plataforma e fornece informações de subsistemas, genes, enzimas, reações e metabólitos que facilitaram o entendimento da rede metabólica da *G. hanseni*.

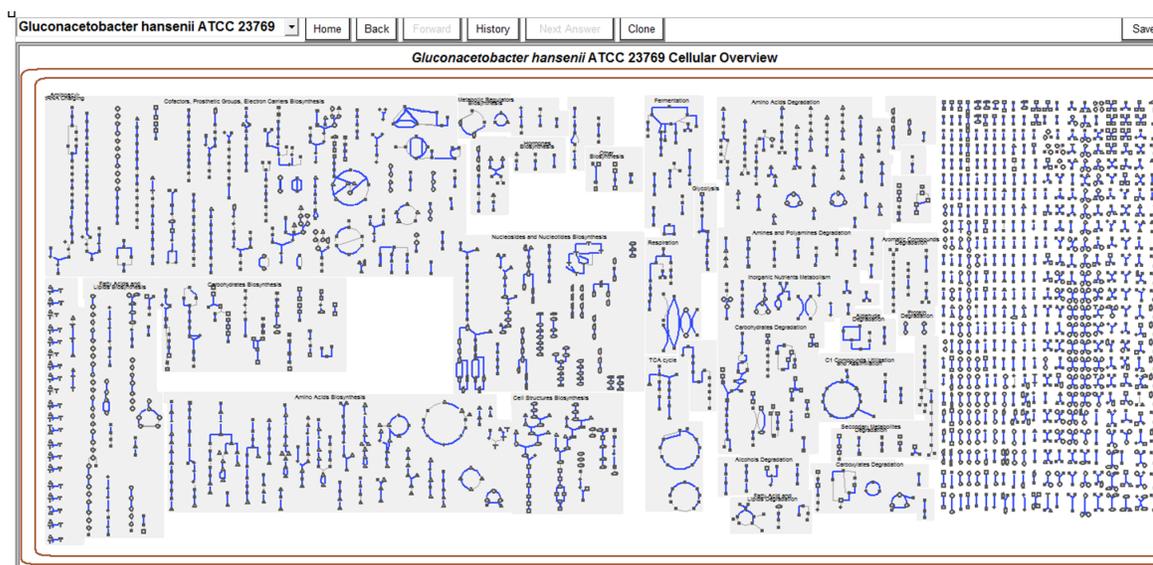


Figura 1 – Mapa celular com as vias da *G. hanseni* no Pathway Tools (GhaCyc).

O *draft* foi exportado em SBML (*System Biology Markup Language*), uma linguagem XML para representações biológicas. Também foram geradas planilhas MS-Excel® (Microsoft Corporation, Redmond, WA, USA). A ferramenta Model SEED complementou as informações obtidas pelo Pathway Tools e gerou um banco de dados com informações úteis sobre o organismo. Os subsistemas da bactéria foram gerados e auxiliaram na etapa de cura manual. A montagem da rede metabólica *core* foi realizada incluindo todas as reações definidas para compor o modelo simplificado da bactéria.

O modelo gerado consiste de uma rede representada matematicamente através da matriz estequiométrica considerando-se as condições de contorno e restrições nas reações envolvidas. O modelo final proposto como *core* consiste em 75 metabólitos e 74 reações (v1-v74). A hipótese de regime estacionário pode ser aplicada para 60 metabólitos, resultando em uma matriz estequiométrica de 60×74.

A AFM forneceu uma primeira distribuição dos fluxos do modelo simplificado para cada fonte de carbono (Figura 2).

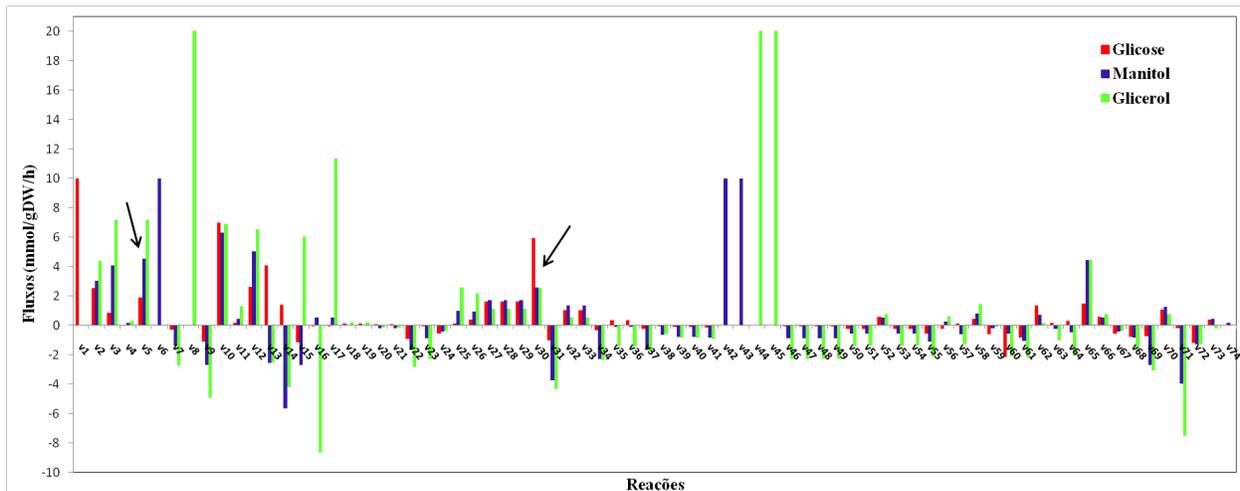


Figura 2 – Análise de fluxo metabólico para as três fontes de carbono. Os fluxos indicados por setas referem-se às reações v5 e v30, relativas à produção de celulose e ácido glucônico, respectivamente.

O resultado das simulações confirma a viabilidade da produção de celulose por *G. hansenii* a partir das três fontes de carbono avaliadas. Analisando o fluxo de produção de celulose (v5) verificou-se que 19%, 45,6% e 35,7% de glicose, manitol e glicerol foram usados para a biossíntese de CB, respectivamente.

Uma vez que o fluxo de produção de celulose com glicose como fonte de carbono foi inferior comparado às outras fontes de carbono, foram analisados outros fluxos, o que levou a um resultado importante em relação à produção de ácido glucônico (v30). Ao utilizar glicose, este fluxo foi igual a $5,94 \text{ mmol} \cdot \text{gDW}^{-1} \cdot \text{h}^{-1}$, quase duas vezes superior ao obtido com a utilização de manitol (2,56) e de glicerol (2,53).

A AFM conseguiu avaliar esse desvio do fluxo metabólico para a produção do ácido glucônico como subproduto nesta via. De acordo com a literatura, isto é resultado das bactérias do gênero *Gluconacetobacter* que oxidam glicose diretamente na forma de ácido glucônico, o qual se acumula no sistema levando à redução do pH.

Durante a ABF foram adicionados os limites superiores e inferiores das reações de troca dos metabólitos que compõem o meio mínimo. Para a simulação do crescimento em condição aeróbica, os seguintes metabólitos externos foram permitidos a entrar e sair da rede livremente: NH_4^+ , PO_4^{3-} , O_2 ,

H^+ , CO_2 e H_2O . Foi obtida a distribuição de fluxos para uma situação de maximização da biomassa, ou seja, a taxa de produção compostos metabólicos precursores foi convertida em componentes da biomassa (Figura 3). A maximização da produção de biomassa indicou que o fluxo de carbonos foi utilizado para o crescimento da bactéria, sem haver produção de celulose. O fluxo referente à produção de celulose (v5) para este cenário foi, portanto, nulo. Este é um caso possível, uma vez que as bactérias não necessitam de celulose para crescer, e é provavelmente uma estratégia que serve para facilitar a manutenção de sua posição perto dos nutrientes, para se proteger contra organismos concorrentes e radiação ultravioleta, e migrar para superfícies de alcance do oxigênio (Iguchi *et al.*, 2000).

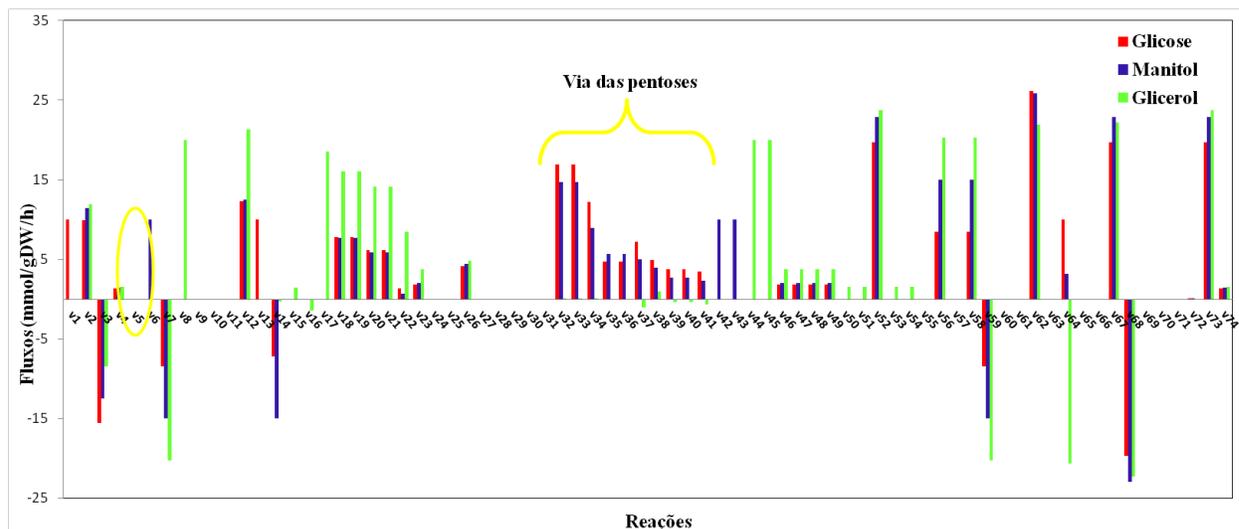


Figura 3 – Análise de balanço de fluxos para as três fontes de carbono. Os fluxos indicados referem-se à produção de celulose (v5) e a via das pentoses, identificadas nas reações v31 a v41.

A ABF mostrou que a utilização de glicerol como fonte de carbono, pela via das pentoses, não foi favorecida teoricamente. As taxas de crescimento específicas calculadas foram: em glicerol, $\mu = 1,54 \text{ h}^{-1}$; em manitol, $\mu = 1,42 \text{ h}^{-1}$; e em glicose, $\mu = 1,33 \text{ h}^{-1}$, mostrando uma dependência com essa fração de carbono direcionada para a via das pentoses (fluxos v31-v41). Se a bactéria tem como objetivo o seu crescimento, a ABF mostrou que a melhor função objetivo é a maximização da biomassa utilizando glicerol como substrato.

Para uma condição de maximização da produção de celulose, os resultados mostraram que a bactéria tem a capacidade de utilizar cada uma das fontes de carbono para as condições de crescimento impostas. Os fluxos foram direcionados para a produção de celulose, e não houve exportação de carbono para produzir biomassa. Isso sugere que o modelo é consistente, apresentando os resultados que seriam teoricamente esperados. Os rendimentos teóricos máximos para produção de celulose foram de: $0,95 \text{ g} \cdot \text{g}^{-1}$ glicose; $0,95 \text{ g} \cdot \text{g}^{-1}$ manitol e $0,93 \text{ g} \cdot \text{g}^{-1}$ glicerol.

Ao maximizar metabólitos externos, a análise do balanço de massa mostrou consistência. Foi avaliada a limitação de alguns nutrientes. Ao definir as taxas de entrada de NH_4^+ , PO_4^{3-} e O_2 para $0 \text{ mmol} \cdot \text{gDW}^{-1} \cdot \text{h}^{-1}$, a bactéria não conseguiu crescer, obtendo-se um distribuição de fluxos nula. Porém,

com limitação de nitrogênio ou de fosfato apenas, isso não foi suficiente para impedir o seu crescimento. É sabido que, se houver limitação de algum nutriente, este organismo continua promovendo a produção de celulose se houver excesso de substrato disponível. Ao incluir uma reação de demanda para cada metabólito da reação da biomassa e otimizar para este fluxo, os resultados mostraram que o modelo produz cada um dos precursores de biomassa.

4. CONCLUSÕES

As modificações no Protocolo T&P recomendado para reconstruções metabólicas (Thiele e Palsson, 2010) permitiram a construção do primeiro modelo *core* da *Gluconacetobacter hansenii*. Este, por sua vez, é um modelo de reconstrução mínimo mas que captura a essência da fisiologia desta bactéria. A descrição do meio mínimo foi fundamental para compor um modelo metabólico consistente com a realidade, e a determinação das funções objetivo que foram consideradas relevantes foi essencial para analisar a rede metabólica. Através das abordagens utilizadas foi realizada a análise *in silico* do comportamento deste organismo. A rede metabólica simplificada obtida é representativa e consistente sob todos os limites e condições de contorno definidos.

5. REFERÊNCIAS

- BLANK, L. M.; EBERT, B. E. From measurement to implementation of metabolic fluxes. *Curr. Opin. Biotech.*, v. 24, n. 1, p. 13-21, 2013.
- BORDBAR, A.; PALSSON, B. O. Using the reconstructed genome-scale human metabolic network to study physiology and pathology. *J. Intern. Med.*, v. 271, n. 2, p. 131-41, 2012.
- COVERT, M. W.; SCHILLING, C. H.; FAMILI, I.; EDWARDS, J. S.; GORYANIN, II; SELKOV, E.; PALSSON, B. O. Metabolic modeling of microbial strains in silico. *Trends Biochem. Sci.*, v. 26, n. 3, p. 179-86, 2001.
- DUROT, M.; BOURGUIGNON, P. Y.; SCHACHTER, V. Genome-scale models of bacterial metabolism: reconstruction and applications. *FEMS Microbiol. Rev.*, v. 33, n. 1, p. 164-90, 2009.
- EDWARDS, J.; IBARRA, R.; PALSSON, B. In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. *Nat. Biotechnol.*, v. 19, n. 2, p. 125-130, 2001.
- FONG, S.; PALSSON, B. Metabolic gene deletion strains of Escherichia coli evolve to computationally predicted growth phenotypes. *Nat. Genet.*, v. 36, n. 10, p. 1056-1058, 2004.
- IGUCHI, M.; YAMANAKA, S.; BUDHIONO, A. Bacterial cellulose—a masterpiece of nature's arts. *J. Mater. Sci.*, v. 35, n. 2, p. 261-270, 2000.
- IYER, P. R.; GEIB, S. M.; CATCHMARK, J.; KAO, T. H.; TIEN, M. Genome sequence of a cellulose-producing bacterium, *Gluconacetobacter hansenii* ATCC 23769. *J. Bacteriol.*, v. 192, n. 16, p. 4256-7, 2010.

- MAHADEVAN, R.; PALSSON, B. O.; LOVLEY, D. R. In situ to in silico and back: elucidating the physiology and ecology of *Geobacter* spp. using genome-scale modelling. *Nat. Rev. Microbiol.*, v. 9, n. 1, p. 39-50, 2011.
- MCCLOSKEY, D.; PALSSON, B. O.; FEIST, A. M. Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*. *Mol. Syst. Biol.*, v. 9, p. 661, 2013.
- METZKER, M. L. Sequencing technologies - the next generation. *Nat. Rev. Genet.*, v. 11, n. 1, p. 31-46, 2010.
- ROSS, P.; MAYER, R.; BENZIMAN, M. Cellulose biosynthesis and function in bacteria. *Microbiol. Rev.*, v. 55, n. 1, p. 35-58, 1991.
- RUKA, D. R.; SIMON, G. P.; DEAN, K. M. Altering the growth conditions of *Gluconacetobacter xylinus* to maximize the yield of bacterial cellulose. *Carbohydr. Polym.*, v. 89, n. 2, p. 613-622, 2012.
- SAXENA, I. M.; KUDLICKA, K.; OKUDA, K.; BROWN, R. M., JR. Characterization of genes in the cellulose-synthesizing operon (acs operon) of *Acetobacter xylinum*: implications for cellulose crystallization. *J. Bacteriol.*, v. 176, n. 18, p. 5735-52, 1994.
- SOON, W. W.; HARIHARAN, M.; SNYDER, M. P. High-throughput sequencing for biology and medicine. *Mol. Syst. Biol.*, v. 9, p. 640, 2013.
- THIELE, I.; PALSSON, B. O. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.*, v. 5, n. 1, p. 93-121, 2010.
- ZHAO, J.; GRANT, S. F. Advances in whole genome sequencing technology. *Curr. Pharm. Biotechnol.*, v. 12, n. 2, p. 293-305, 2011.
- ZHONG, C.; ZHANG, G. C.; LIU, M.; ZHENG, X. T.; HAN, P. P.; JIA, S. R. Metabolic flux analysis of *Gluconacetobacter xylinus* for bacterial cellulose production. *Appl. Microbiol. Biotechnol.*, v. 97, n. 14, p. 6189-99, 2013.