**RESEARCH**

**Clostridium thermocellum** LL1210 pH homeostasis mechanisms informed by transcriptomics and metabolomics

Jason M. Whitham¹,², Ji-Won Moon¹, Miguel Rodriguez Jr.¹,², Nancy L. Engle¹,², Dawn M. Klingeman¹,², Thomas Rydzak¹,²,³, Malaney M. Abel¹,², Timothy J. Tschaplinski¹,², Adam M. Guss¹,² and Steven D. Brown¹,²,⁴*

**Abstract**

**Background:** *Clostridium (Ruminiclostridium) thermocellum* is a model fermentative anaerobic thermophile being studied and engineered for consolidated bioprocessing of lignocellulosic feedstocks into fuels and chemicals. Engineering efforts have resulted in significant improvements in ethanol yields and titers although further advances are required to make the bacterium industry-ready. For instance, fermentations at lower pH could enable co-culturing with microbes that have lower pH optima, augment productivity, and reduce buffering cost. *C. thermocellum* is typically grown at neutral pH, and little is known about its pH limits or pH homeostasis mechanisms. To better understand *C. thermocellum* pH homeostasis we grew strain LL1210 (*C. thermocellum* DSM1313 Δhpt ΔhydG Δldh Δpfl Δpta ack), currently the highest ethanol producing strain of *C. thermocellum*, at different pH values in chemostat culture and applied systems biology tools.

**Results:** *Clostridium thermocellum* LL1210 was found to be growth-limited below pH 6.24 at a dilution rate of 0.1 h⁻¹. F₁F₀-ATPase gene expression was upregulated while many ATP-utilizing enzymes and pathways were downregulated at pH 6.24. These included most flagella biosynthesis genes, genes for chemotaxis, and other motility-related genes (>50) as well as sulfate transport and reduction, nitrate transport and nitrogen fixation, and fatty acid biosynthesis genes. Clustering and enrichment of differentially expressed genes at pH values 6.48, pH 6.24 and pH 6.12 (washout conditions) compared to pH 6.98 showed inverse differential expression patterns between the F₁F₀-ATPase and genes for other ATP-utilizing enzymes. At and below pH 6.24, amino acids including glutamate and valine; long-chain fatty acids, their iso-counterparts and glycerol conjugates; glycolysis intermediates 3-phosphoglycerate, glucose 6-phosphate, and glucose accumulated intracellularly. Glutamate was 267 times more abundant in cells at pH 6.24 compared to pH 6.98, and intercellular concentration reached 1.8 μmol/g pellet at pH 5.80 (stopped flow).

**Conclusions:** *Clostridium thermocellum* LL1210 can grow under slightly acidic conditions, similar to limits reported for other strains. This foundational study provides a detailed characterization of a relatively acid-intolerant bacterium and provides genetic targets for strain improvement. Future studies should examine adding gene functions used by more acid-tolerant bacteria for improved pH homeostasis at acidic pH values.

**Keywords:** Proton pumping, F₁F₀-ATPase, pH homeostasis, Glutamine synthetase, Glutamate decarboxylase, Urease, GOGAT, Glutamate dehydrogenase, Clostridium thermocellum

*Correspondence: Steve.Brown@lanzatech.com
4 Present Address: LanzaTech, Inc., Skokie, IL, USA
Full list of author information is available at the end of the article
Background

Clostridium (Ruminiclostridium) thermocellum is a model fermentative anaerobic thermophile being studied and engineered for consolidated bioprocessing of lignocellulosic feedstocks into fuels and chemicals [1–3]. C. thermocellum forms specialized biofilms for growth on cellulose [4] and produces mobile [5] and cell-bound enzyme complexes termed cellulosmes [6] that mediate deconstruction of lignocellulosic substrates to short, β1,4-linked glucose oligosaccharides for fermentation to ethanol and organic acids. Wild type C. thermocellum fermentation products include acetic acid, lactic acid, formic acid, H₂, ethanol, and amino acids such as valine [7], with additional products being made at high substrate loadings as a result of overflow metabolism [8].

Organic acid production decreases biofuel yields and acidifies culture medium, each of which are undesirable for an industrial process. Therefore, metabolic engineering efforts have focused on modifying carbon metabolism to decrease production of acidic co-products and increase ethanol production [9–17]. In the highest yielding C. thermocellum strain published to date, strain LL1210, pathways for the synthesis of lactic acid, acetic acid, formic acid, and most H₂ production were eliminated [12], followed by adaptive laboratory evolution to improve growth rate and ethanol titer [18]. This allowed the strain to produce 22 g/L ethanol from 60 g/L cellulose, with a maximum theoretical yield for ethanol of 75% although further advances are required for the bacterium to be industry-ready. Ethanol yields and titers of up to 80% and 38 g/L, respectively, have also been achieved with an engineered C. thermocellum strain in co-culture with an engineered Thermoanaerobacterium saccharolyticum strain [10]. While these strains synthesize essentially no organic acids as end products, fermentation of sugars to reduced compounds such as ethanol or butanol results in the production of a more oxidized compound such as bicarbonate, which will also result in acidification of the medium.

Organisms must be tolerant to the fermentation conditions and the products formed during growth, especially at the high substrate levels required for industrial fermentations. The response to ethanol stress and ethanol-tolerant strains of C. thermocellum have been characterized [19, 20]. C. thermocellum inhibition by pentose sugars [21] and compounds generated within switchgrass fermentations [22] has also been characterized. However, deep understanding of the C. thermocellum response to low pH remains underexplored.

While the anaerobic cellulolytic bacterium Fibrobacter succinogenses has been evolved for steady-state growth at pH 5.75 [23], known anaerobic cellulolytic bacteria typically do not grow at pH values lower than pH 6.0 [2]. High-yielding- and highly productive ethanologens such as Saccharomyces cerevisiae and Zymomonas mobilis, which are not natively cellulolytic, have broad pH ranges (pH 2–6.5 [17] and pH 3.5–7.5 [18], respectively). C. thermocellum is typically cultured in medium maintained at neutral pH and it maintains its intracellular pH between ~7.3–8.5 [24, 25]. For different C. thermocellum strains, growth has been reported between pH 5.9 and 8.1 [26], as low as pH 6.0 [27], and between pH 6.2 and 7.7 [28]. The lower pH limit of the only C. thermocellum strain that is currently genetically tractable, strain DSM1313, has not been reported.

A major mechanism of low pH toxicity is dissipation of the proton gradient across the cytoplasmic membrane. Weak acids such as acetic acid become protonated at the low pH in the supernatant, resulting in an uncharged molecule that can cross the cytoplasmic membrane. The cytoplasm is more alkaline, so the weak acid becomes deprotonated, effectively transporting a proton from outside the cell to the inside, collapsing the ΔpH and acidifying the cytoplasm. Cytoplasmic pH can be maintained by a variety of mechanisms that include proton pumps, antiporters, production of ammonia via urease and arginine deiminase pathways, and organic acid decarboxylation reactions that cleave a carboxylic acid side chain (pKa ~4.7) to bicarbonate (pKa=6.4 and more volatile); tolerance to lower pH can be further enhanced by macromolecule repair or protection, lipid changes and biofilm formation [29–31]. Understanding and enhancing the mechanisms by which C. thermocellum responds to decreased pH will be critical for the industrial application of C. thermocellum for lignocellulosic biofuel production, but those mechanisms are not currently known. Therefore, to better understand the physiological response of C. thermocellum to acidic conditions, we cultured strain LL1210 at different pH values in chemostats and applied systems biology tools.

Results

C. thermocellum LL1210 chemostat growth and lower pH limits in MTC medium

Clostridium thermocellum LL1210 was cultured in duplicate chemostats fed with defined medium (MTC) containing 3 g/L cellulose for 1058 h (Fig. 1, Additional file 1). After initial batch fermentation, steady-state chemostat growth was achieved for pH 7.0 at a dilution rate of 0.1 h⁻¹. The culture pH was then reduced to pH 6.5 at hour 189.5 and steady-state growth was maintained. Culture pH was checked using a regularly calibrated external probe and reported in Additional file 1. A pH adjustment to pH 6.10 resulted in substantial decreases in cell densities (~308 h, Fig. 1), along with lower ethanol concentrations and concomitant increases in cellulbiose.

While the lower pH limit of the only C. thermocellum strain...
concentrations (Additional file 1), which suggests the specific growth rate of strain LL1210 was lower than the dilution rate of 0.1 h\(^{-1}\). Medium flow was temporarily halted for 37.5 h and the culture pH set to 6.25, which restored cellobiose consumption, ethanol production and cell density and permitted medium flow to be resumed at dilution rate of 0.1 h\(^{-1}\). After the equivalent of three fermentor vessel volumes, the chemostats were adjusted from pH 6.25 to 6.15, which coincided with lower culture densities. The mean culture turbidity measurements nearest to the omics sampling points, as measured by optical density (OD\(_{600\text{ nm}}\)) and their standard deviations were 0.655 (0.02), 0.651 (0.04), 0.533 (0.06), and 0.143 (0.05) for pH values pH 7, 6.5, 6.25 and 6.1, respectively. A change in dilution rate to 0.05 h\(^{-1}\) permitted growth at pH 6.10 and 6.00. At pH 5.90, culture turbidity steadily declined; the dilution rate was reduced to 0.01 h\(^{-1}\) and the OD continued to decline for 144 h in one chemostat while the other oscillated, indicative of stress, and eventually recovered. A single chemostat was able to maintain growth at pH 5.90 at a dilution rate of 0.01 h\(^{-1}\), although culture turbidity showed a steady decline at pH 5.80 over a 133 h period and was unable to recover when the flow was stopped for 84 h. Thus, growth-limiting pH for this system is approximately pH 6.25 using strain LL1210 in a chemostat with a dilution rate of 0.1 h\(^{-1}\) and slower dilution rates are required for steady-state growth below this pH.

Samples at dilution rate of 0.1 h\(^{-1}\) were withdrawn from each bioreactor at pH 7 (actual pH = 6.98), pH 6.5 (actual pH = 6.48), pH 6.25 (actual pH = 6.24), and at apparent washout pH 6.15 (actual pH = 6.12) for analysis by transcriptomics and metabolomics (Fig. 1). Additional metabolomics samples were taken at pH 5.80 at a lower dilution rate (0.05 h\(^{-1}\)) to help elucidate the physiological changes.

**pH homeostasis mechanisms inferred by differential gene expression**

A greater number of differentially expressed genes were observed as the pH diverged further from pH 6.98. Using pH 6.98 as a reference, there were 80, 469, and 536 differentially expressed genes, respectively, for cells growing at pH 6.48, 6.24, and pH 6.12 (Additional file 2). To more easily categorize gene expression responses to pH changes, differentially expressed genes were clustered (Fig. 2a), and then gene clusters were checked for enrichment of gene ontology (GO) terms in comparison to the frequency of GO terms for all genes in the *C. thermocellum* genome by Fisher’s exact test (Fig. 2b). Cluster 1 contained only one gene (a 2-isopropylmalate synthase; Clo1313_0857) and cluster 3 only contained tRNA genes and were omitted from enriched function tests. Cluster 2 contained genes for rRNA, tRNA, tRNA modification, purine metabolism, and polyamine transport but no enrichment of GO terms was found. Flagella biosynthesis and chemotaxis, sulfate transport and reduction, nitrate transport and nitrogen fixation, glutamate dehydrogenase (GLDH, Clo1313_1847), and fatty acid biosynthesis (clusters 4, 5, 6) were downregulated, while F0F1-ATP synthase genes (cluster 7) were upregulated at pH 6.24. Diguanylate cyclase genes Clo1313_1339, Clo1313_1404...
transporter and biosynthetic gene were upregulated, and Clo1313_0415) [14]. All but one of the glutamate synthetases (Clo1313_1357 and Clo1313_2038), or the glutamate synthetase (GS, Clo1313_2031), two Type III glutamine synthases (cluster 9) including those for cellulosomal proteins CipA, OlpB, Orf2p, OlpA, CelD, CelH, CelR, CelV, CprA, CtManF, LecA, XynA, and XynD (Additional file 2). Cluster 9 also contained a malate shunt gene, phosphoenolpyruvate carboxykinase (PEPCK, Clo1313_0415) [14]. All but one of the glutamate synthase (GOGAT) genes (Clo1313_2032-3, Clo1313_2035-6) were also in cluster 9.

Previous studies have identified genes commonly expressed during acid stress by Gram-positive neutrophilic bacteria [30], which includes genes for amino acid decarboxylases/antiporter systems, DnaK, GroEL, ClpP protease, urease, F1F0-ATPase, arginine deiminase (ADI), agmatine deiminase (AgDi), and cyclopropane fatty acid (CFA) synthase. Most of these gene systems were not differentially expressed in C. thermocellum and genes for ADI, AgDi, or CFA are not annotated in the genome. Rather than increasing in expression at low pH, several decarboxylases were downregulated at pH 6.24 and 6.12, including aspartate 1-decarboxylase (Clo1313_1318), diaminopimelate decarboxylase (Clo1313_1540), orotidine 5′-phosphate decarboxylase (Clo1313_1266), putative oxaloacetate decarboxylase (Clo1313_1523), and a putative sodium pump decarboxylase γ-subunit (Clo1313_1525). The only decarboxylase that was upregulated at any pH was the S-adenosylmethionine decarboxylase proenzyme (Clo1313_1509), which produces the aminopropyl substrate need for synthesis of the polyamines spermidine and spermine. However, expression of the agmatinase and spermidine synthase (Clo1313_1529-1530) that synthesizes the spermidine and spermine were either not differentially expressed or showed decreased expression at lower pH. Then again, genes for a putative spermidine/putrescine ABC transporter (Clo1313_1471-5) were upregulated at pH 6.24 and 6.12 (Additional file 2). Because a polyamine transporter and biosynthetic gene were upregulated, and polyamines can decrease membrane permeability and may protect cells in acid environments [33], their potential to augment C. thermocellum LL1210 growth at lower pH values was evaluated. Medium was supplemented with 100 μM of the polyamines putrescine, spermidine, spermine, or metabolic precursor arginine and the terminal pH was determined after 144 h. C. thermocellum LL1210 fermentations with and without amendments initiated at pH 7 and 6.75 had similar growth profiles (Additional file 3).

Downregulation of the GLDH (Clo1313_1847) and GOGAT (Clo1313_2032-3, Clo1313_2035-6) genes were observed at pH 6.12, but no significant expression changes were observed for the Type I glutamine synthetase (GS, Clo1313_2031), two Type III glutamine synthetases (Clo1313_1357 and Clo1313_2038), or the glutamate synthase (Clo1313_1849) (Additional file 2). The type III glutamine synthase (Clo1313_2303) was significantly upregulated at pH 6.24 (Additional file 2). To examine the role of nitrogen metabolism in relation to pH and physiology, deletion strains for GLDH, GOGAT, GS, GS–GOGAT, and an annotated NifH (Clo1313_2339) were characterized using unbuffered (MOPS free) carbon-replete medium (10 g/L) (Additional file 4). Consistent with an earlier study [18], the parental strain had a higher specific growth rate compared to strain LL1210 (Additional file 4a). The final pH for the GS-mutant- and LL1210 cultures were about pH 6.3 and pH 5.9, respectively, whereas the range for the other strains was ~pH 5.6–5.8 (Additional file 4b). In the unbuffered system, the GS mutant achieved the highest culture turbidity and produced the second highest titer of ethanol after strain LL1210 (2.0 and 2.7 g/L, respectively) (Additional file 4c). The higher final ethanol titers and pH values for strains GS and LL1210 coincided lower levels of residual substrate.

Genes for two orphan histidine kinases previously predicted to be part of the sporulation cascade in C. thermocellum [34], a pro-σE processing protease and BofA, inhibitor of the stage IV pro-σK processing protease SpoIVFB, were upregulated in C. thermocellum LL1210 at pH 6.24 compared to 6.98 (Additional file 5). To better understand the role of sporulation in the pH response of C. thermocellum, growth and sporulation of strains LL1210, asporogenous strains M1726 (DSM1313 Δhpt Δspo0A1) and M1725 (DSM1313 Δhpt Δspo0A1 Δldh Δpta), and the DSM1313 Δhpt parental strain were

*(See figure on next page.)*

**Fig. 2** K-means clustering of differentially expressed genes (a). GO term enrichment of clustered genes by Fisher’s exact test (b). 866 differentially expressed genes from samples taken at pH 6.48, 6.24, and pH 6.12, compared to samples taken at pH 6.98 were clustered. The % of genes represents the percentage of genes within a cluster that contain the given GO term. Enrichment of GO terms indicates that a significantly higher percentage of genes with GO terms are in clusters (Test Set) than in the C. thermocellum reference genome (α=0.05).
examined. C. thermocellum sporulation was inefficient, consistent with an earlier study [35] and sporulation-deficient mutant strains had neither a growth advantage or disadvantage at lower pH values for the conditions used in this study (Additional file 5).

The molecular chaperone GrpE (Clo1313_0932) and heat shock protein Hsp 20 (Clo1313_0678) were upregulated (2.0- and 2.2-fold) at pH 6.12, which were previously found to be upregulated after furfural shock and heat shock in C. thermocellum ATCC 27405 [36]. Hsp33 (Clo1313_2544) and MutS (Clo1313_1982) were upregulated at pH 6.24 (2.2- and 2.0-fold, respectively). Hsp33 was previously found to be more highly expressed in cellulose-adhered C. thermocellum ATCC 27405 compared to planktonic cells [4]. At pH 6.12, downregulated genes included one of the annotated Clp protease genes (Clo1313_1116) (0.48-fold), an annotated DNA repair proteins RadC (Clo1313_1465) (0.36-fold), and UvrD (Clo1313_2004) (0.31-fold) and a small subunit of an exodeoxyribonuclease VII (Clo1313_1389) (0.50-fold).

Metabolite profiles for low pH conditions
To further investigate the physiological changes that occur at lower pH values, metabolomics was performed to compare intercellular concentrations of metabolites when cells were exposed to pH 6.98, 6.48, 6.24, 6.12, and an additional time point at pH 5.80 (at a lower dilution rate 0.05 h⁻¹). Peak areas for 91 metabolites identified by GC–MS were calculated for comparison (Additional file 6) and 39 of these were significantly higher at one or more of the pH values compared to pH 6.98 (Table 1). Most significantly, higher metabolites in cells at pH 6.24 were nitrogen (N)-containing metabolites that included glutamate, lysine, glycine, and five unknown N-metabolites (Table 1). There were also two unknown phosphate (P)-containing metabolites. Glutamate showed the largest differences in average fold change across all samples tested at lower pH values (Table 1). At pH 6.24, cells contained 267 times more glutamate than cells at pH 6.98 (significant α = 0.05). Intracellular concentrations of glutamate ranged from 0.004 μmol/g pellet at pH 6.98 to 1.798 μmol/g pellet at pH 5.80 (Additional file 6). Concentrations of intercellular valine, alanine, and threonine were also quantified on a μmol/g of pellet basis for comparison with glutamate concentrations. While glutamate and alanine peaked at pH 5.80 (2.660 μmol/g of pellet for alanine), valine threonine peaked at 6.24 with concentrations of 5.312 and 0.267 μmol/g of pellet, respectively. Long-chain fatty acids, their iso-counterparts and glycerol conjugates accumulated at growth-limiting pH values, pH 6.12 and 5.80 (Table 1). Citramalate accumulated at pH 6.24, 6.12, and 5.80 (Table 1) and an accumulation of glycolysis intermediates 3-phosphoglycerate, glucose 6-phosphate, and glucose at pH values 6.12 and 5.80 was consistent with downregulation of genes for major downstream pathways, malate shunt (Clo1313_0415) and the TCA cycle (Clo1313_0708-9) (Additional file 2).

Extracellular amino acid accumulation at growth-limiting pH
At pH 6.24, only valine, alanine, and threonine extracellular concentrations were significantly higher than when pH was 6.98 (α = 0.05) (Additional file 7). Asparagine was significantly lower at pH 6.24 compared to 6.98 (α = 0.05). Extracellular glutamate was not significantly different at pH 6.24 compared to 6.98, although it was significantly lower at pH 6.48 compared to 6.24 (α = 0.05). When cells were shocked with pH 6.12 compared to when they were growing at pH 6.98, there were significantly lower concentrations of all detectable extracellular amino acids except methionine (α = 0.05). The amino acids that had the highest extracellular concentrations in chemostats were valine (1.5 mM at pH 6.24), asparagine (147 μM at pH 6.98), and alanine (104 μM at pH 6.24). The highest average concentration of glutamate was 32 μM (near the limit of detection) at pH 6.24 (Additional file 7). An ABC-type branch-chain amino acid transporter gene (Clo1313_0822) was upregulated at pH 6.24 and 6.12 and ABC-type polar amino acid transport genes (Clo1313_0794-5) were upregulated at pH 6.24 (Additional file 2).

Metabolic inhibition at a cellular level
To gain a global perspective on metabolic inhibition at growth-limiting pH values, at and below pH 6.24, key mechanisms and responses for pH homeostasis were summarized (Fig. 3). A key feature of the physiological response at lower pH values was changes in expression for use and conservation of ATP. At pH 6.24, LL1210 increased expression of its F₁F₀-ATPase (Fig. 2). The F₁F₀-ATPase likely pumps protons to maintain pH homeostasis by ATP hydrolysis as is the case of many other neutrophiles [30]. Competing uses for ATP included valine transport, chaperones protecting cytoplasmic macromolecules from acid damage, and assimilation of ammonium by GS, GOGAT, and GLDH (Table 1, Additional file 2). ATP is also consumed by fatty acid biosynthesis (Table 1, Fig. 2), but there are many metabolites and downregulated genes for this aspect of metabolism. Therefore, it was omitted from Fig. 3, with the exception of biosynthesis pathways for valine and leucine, which are precursors for biosynthesis of iso-fatty acids. The accumulation of fatty acids and iso-derivatives may improve the capacity cytoplasmic membrane to repel protons, but adjustments to membrane lipid composition in response to an increase in proton concentrations are diverse among bacteria and not well understood [37].
Though fatty acid adjustments may have reduced membrane permeability to protons, downregulation of fatty acid biosynthesis genes suggests that ATP conservation was more important. Other gene expression changes that decrease ATP consumption include downregulation of sulfate and nitrate ABC transporter and reduction genes and 80 putative flagella biosynthesis and motility genes (listed in Additional file 2). Reduction of proton influx may also have been achieved through downregulation of gene expression for proton-channeling motility proteins MotAB (Clo1313_0056-7) at pH 6.24 (Fig. 2) [32]. The putative proton-pumping PPi-ase (Clo1313_0823)...
was constitutively and highly expressed at all pH values. Accumulation of intercellular phosphate at pH 5.80 (Table 1) may have been due to PPi-ase activity, ATP hydrolysis or other cellular activities. Another possibility is that collapse of the membrane gradient would have made PPi hydrolysis to free phosphate much more thermodynamically favorable via the membrane-bound PPi-ase. Both reduction of proton influx through MotAB and proton-pumping through the putative PPi-ase would allow more ATP conservation.

Discussion

Clostridium thermocellum strain LL1210 has deletions in key genes for metabolic pathways to produce acetate, lactate, formate, and H2, and it has among the highest reported ethanol yields and titers for this bacterium. LL1210 grew at pH 6.24 at a dilution rate of 0.1 h⁻¹. At pH values at or below 6.24, we observed gene expression patterns and metabolite levels consistent with five pH homeostasis mechanisms used by C. thermocellum LL1210, namely (i) proton pumping by F1F0-ATPase, (ii) macromolecule protection by chaperones, and (iii) prevention of proton influx through MotAB were mechanisms supported by changes in gene expression (Fig. 2, Additional file 2); (iv) proton pumping by PPi-ases was supported by metabolomics (Table 1); and (v) changes in membrane fatty acid composition was supported by both techniques (Fig. 2, Additional file 2, Table 1).

Differential expression analysis at pH 6.24 revealed that the F1F0-ATPase genes were upregulated (Fig. 2), consistent with a model that the bacterium uses F1F0-ATPases to pump-protons at the expense of ATP (Fig. 3). Increased expression and activity of F1F0-ATPases are associated with pH tolerance in several other bacteria challenged with sublethal pH conditions [29, 30, 38]. In addition to the F1F0-ATPase system, proton-pumping may be possible via a constitutively and highly expressed putative proton-pumping PPi-ase (Clo1313_0823) (top 9% of genes) [39, 40]. Rhodospirillum rubrum and plant proton-pumping PPi-ases are reversible [41] and the C. thermocellum proton-pumping PPi-ase has been proposed to be important in energy metabolism in this organism [42]; further studies are required to experimentally test the function in C. thermocellum. Many ATP-utilizing enzymes and pathways were downregulated at lower pH, which included most of the flagella biosynthesis genes, genes for chemotaxis, and other motility-related genes (> 50) as well as sulfate transport and reduction, nitrate transport.
and nitrogen fixation, and fatty acid biosynthesis genes (Fig. 2, Additional file 2). Genes for sulfate transport/reduction and for nitrogen metabolism have shown differential expression in response to stresses [36, 52] and their roles in maintaining homeostasis and for channeling electrons to maintain biofuels productively could better understood through modeling, biochemistry and genetic studies. In this study, clustered and enriched gene ontologies (GO) analyses indicate genes for enzymes using ATP had inverse differential expression patterns compared to F$_{0}$F$_{1}$-ATPase genes, which suggests metabolism is redirected to conserve energy (Fig. 2). Indeed, maintaining ΔpH and protecting macromolecules takes precedence over many reactions since critical enzymes for growth cannot function if they are damaged. In addition to a direct burden on ATP, flagella motility may have an indirect burden since protons channelled into the cell through MotAB would need to be pumped back out by to maintain pH homeostasis. Therefore, motAB downregulation at pH 6.24 would support pH homeostasis (Additional file 2, Fig. 3).

Genes involved in polyamine transport and biosynthesis were differentially expressed at lower pH, suggesting a potential functional role for these compounds in pH tolerance. Specifically, the transporter (Clo1313_1471-5) and SAM decarboxylase gene (Clo1313_1509) were upregulated (Additional file 2). Medium supplementation with polyamines or arginine (a polyamine precursor) did not significantly improve growth or enable C. thermocellum to ferment to a lower terminal pH value compared to unsupplemented fermentations (Additional file 3). Instead, a functional role for stabilization of tRNAs during translation [43] matches well with polyamine biosynthesis and transport gene expression profiles clustering with expression profiles of genes for tRNAs, tRNA modification proteins, rRNA, ribosomal proteins, ribosomal maturation factors, and translational initiation factors (Fig. 2, Additional file 2). Upregulation of translational machinery, tRNAs, and the polyamines that stabilize them, is likely a regulatory response to the abundance of available carbon at pH-limited growth.

Genes involved in sporulation were upregulated at pH 6.24 in chemostat culture. Interestingly, strain LL1210 has a mutation in one of the two copies of the spo0A gene that results in an amino acid change [18]; Spo0A is generally the master regulator of sporulation, and the role of this mutation has not been established. However, examination of sporulation in different genetic backgrounds at acidic pH fermentation demonstrated that C. thermocellum is inefficient at sporulation, and that LL1210 is similar to the wild type strain. Further, a growth advantage for sporulation mutants in acidic, MOPS-free, carbon-replete medium was not observed (Additional file 5), suggesting that a programmed metabolic shutdown via sporulation is not a major driver of sensitivity to low pH.

Nitrogen metabolism was strongly effected at lower pH values, resulting in increased secretion of amino acids including valine and alanine, which has been seen in other studies [8, 18], and also intracellular accumulation of glutamate with minimal secretion (Table 1, Additional files 5, 7). Some neutrophilic bacteria express glutamate decarboxylase in response to pH stress to convert glutamate into γ-aminobutyric acid; the decarboxylation reaction consumes protons from the cytoplasm which reduces intercellular pH [30]. Heterologous expression of a glutamate decarboxylase in C. thermocellum LL1210 could make use of the accumulating intracellular glutamate, possibly enabling steady-state growth at pH values less than 6.25 with a dilution rate of 0.1 h$^{-1}$. Glutamate decarboxylases with pH and temperature ranges suitable for C. thermocellum have already been characterized [44, 45], and heterologous expression of glutamate decarboxylase genes with enhanced production of γ-aminobutyric acid in distantly related organisms has been demonstrated [46]. Urease is another target for improvement of pH tolerance in C. thermocellum. Urease activity in neutrophilic bacteria produces ammonia which is protonated intercellularly or exported and protonated to form ammonium thereby raising intracellular or extracellular pH [30]. The urease genes had low expression in LL1210 at pH values sampled (Additional file 2) in urea-containing growth medium. Previous work has showed that the parental strain (DSM1313 Δhpt) also had low urease expression in urea-containing fermentation medium, but that deletion of the GS gene caused significant upregulation of urease genes [47]. Consistent with higher expression of the urease genes, higher terminal pH was also previously observed for the GS-mutant fermentations compared to the parental strain fermentations [47]. It remains to be determined if higher urease expression increases C. thermocellum’s intracellular pH and enables steady-state growth at pH values below 6.25 with a dilution rate of 0.1 h$^{-1}$. The GS mutant had less residual cellobiose and glucose in unbuffered medium than LL1210 and other nitrogen assimilation mutant strains tested (Additional file 4). Therefore, deletion of the GS in LL1210 might also improve substrate utilization and reduce or eliminate the need for buffering fermentation media.

**Conclusions**

*Clostridium thermocellum* LL1210 can grow under slightly acidic conditions, similar to limits reported for related *C. thermocellum* strains. In nature, *C. thermocellum* may be exposed to acidic pH values, and if it is, metabolism will slow or the bacterium could sporulate, which are not aligned with industrial processes. This
foundational study provides a detailed characterization of a relatively acid-intolerant bacterium and provides genetic targets for strain improvement. Future studies should examine adding gene functions used by more acid-tolerant bacteria for improved pH homeostasis at acidic pH values.

**Methods**

**Bacterial strains, media and growth**

*Clostridium thermocellum* strains DSM1313 Δhpt [10], LL1210 [18], M1726 (LL376) [48], M1725 (LL375) [48], and AG1329 [47] were used in this study. The GLDH (Clo1313_1847), GOGAT (Clo1313_2032-6), GS–GOGAT (Clo1313_2031-6), and a *nifH* (Clo1313_2339) strain were deleted in a *C. thermocellum* DSM1313 Δhpt background to create AG1327, AG1328, AG2068, and AG2078, respectively. Mutants were constructed using the procedures described in detail [49, 50]. Strains are available upon request. Primers used to construct strain AG2078 and other strains used in this study as well as SNPs identified from genome sequencing are provided in Additional file 8. Strains were stored frozen at −80 °C in a modified version of DSM 122 medium, as described earlier [51], except that MOPS (morpholinepropanesulfonic acid) and 10 mM of sodium citrate. Actively growing cultures were transferred twice from freezer stocks prior to batch fermentation experiments and inocula were 2% of the final volume. Cells were routinely cultured at 55 °C in either 50 ml (160 ml serum bottles) or 10 ml (26 ml Balch tubes) of MTC medium, as described earlier [23], except that 1N HCl and 3 N KOH were used for acid and base additions, respectively. Recalibration of the bioreactor probes was performed at 546 h after pH drifted by 0.08 and 0.12. Samples for omics were taken in 50 ml volumes; then harvested and stored as described in [52]. For transcriptomics, two samples were taken immediately after transcriptomic samples, and additional biological duplicate samples were taken at pH 5.8 for metabolomics only. Culture turbidity was measured using optical density readings at 600 nm using a Genesys 20 spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA).

Cells were enumerated using a modified version of the direct cell counting protocol described in [53]. Briefly, 1 ml of culture samples were preserved for up to 1 week in 0.025 mg/ml (final concentration) paraformaldehyde (Sigma-Aldrich,). Cells were diluted 1:1000 with 0.9% (m/v) saline solution and stained with 25 nM 4′,6-diamidino-2-phenylindole (DAPI) (final concentration), immobilized to a 1.5 cm diameter section of a black MilliQ filter, and then imaged with a Zeiss Axio Imager.M1 with a Plan Apochromat 63×/1.4 Oil Immersion objective and X-Cite Series 120 Exfo fluorescence power source (Carl Zeiss, Germany). 8-bit images were imported into ImageJ (http://imagej.nih.gov/ij/) for manual or automated counting. Manual counting was performed with the cell counter plugin, and automated counting was performed with JAVA scripts (Additional file 1) which assumed cells were visible at a % Area (saturation threshold) of 1% (±0.1%) with a particle size greater than 5 pixels after binary images had been processed with Watershed. Spherical morphologies were assumed to have a particle size range of 20–200 pixels and circularity of 0.85–1.00 for conservative estimates, and a particle size range of 50–100 pixels and circularity of 0.95–1.00 for liberal estimates. Automated parameters were validated with manual counts of 24 images. Stepwise calculations of standard deviation of progressively more sample counts were performed to ensure the count variance plateaued. Highly refractive mature spores were also observed by light microscopy after heat fixing to 20 μl of stationary cells to glass slides.

**Substrate and metabolite analyses**

Extracellular fermentation product samples for high-performance liquid chromatography (HPLC) were collected and processed as described in [51]. HPLC data
were generated using a Shimadzu Prominence LC-20A Series system (Shimadzu Scientific Instruments, Columbia, MD) fitted with a refractive index detector (model RID-20A) and an Aminex HPX 87H HPLC column (300 × 7.8 mm) (Bio-Rad, Hercules, Dallas, TX, USA) against known standards for cellulbiose, glucose, ethanol, acetate, lactate, and formate. Intracellular metabolite fold changes and select metabolite concentrations were determined using a 5975C inert XL gas chromatography–mass spectrometer (Agilent, Santa Clara, CA) [22]. Extracellular amino acid concentrations were quantified by an Aracus Amino Acid Analyzer (membraPure, Berlin, Germany) [47]. Two-tailed paired T tests were performed to determine if amino acid concentrations were significantly different. Homoscedasticity was confirmed by performing Breusch–Pagan and Abridged White’s tests prior to performing T tests.

**RNA isolation, cDNA library preparation, sequencing and RNA-seq analysis**

Total RNA was isolated from 50 ml of chemostat sample using the procedures described in [54], and quantity and quality were determined using a Nanodrop instrument (ThermoFisher Scientific Waltham, MA). Depletion, library preparation and sequencing were completed at the Joint Genome Institute (JGI, Walnut Creek, CA). In short, 100 ng of total RNA was depleted of ribosomal material using a Ribo-Zero (TM) rRNA Removal Kit (Epicentre). The rRNA depleted RNA was fragmented and reversed transcribed using random hexamers and SSII (Invitrogen) followed by second strand synthesis. Stranded cDNA libraries were generated using the Illumina TrueSeq Stranded RNA LT kit, following the manufactures protocols, using ten cycles of PCR. qPCR was used to determine the concentration of the libraries. Libraries were pooled and sequenced on an Illumina HiSeq 2500 platform with version 2 chemistry in a 2 × 151 bp configuration (JGI, Walnut Creek, CA; Project ID 503106). Fastq files were verified for integrity by checksum and read quality checked using FastQC [55]. RNaseq reads were trimmed using Trimmomatic MAXINFO method with target length and strictness parameters set to 40 and 0.8, respectively [56]. Reads were mapped to *C. thermocellum* DSM1313 (NC_017304, last modified 2015/07/30) using Bowtie2 with the same parameters as the very-sensitive preset option expect the number of mismatches was set to 1 [57, 58], and reads were counted with HTSeq [59]. Principle component analysis (PCA) was used to check for sample variation using JMP Genomics 8 (SAS Institute, Cary, NC, USA). Differential expression analysis was performed with DESeq 2 [60]. Genes were considered significantly differential expressed when their adjusted p values were less than or equal to 0.05 and log2-fold change values were ≥ 1 or ≤ 1. Raw RNA-seq data have been deposited in NCBI SRA under project number PRJNA395926. PCA plots, raw counts, RPKM normalized gene counts, and differential expression are provided in Additional file 2. Genes found to be differentially expressed at pH 6.48, 6.23, and pH 6.12 compared to pH 6.98 were clustered by their expression patterns using the K-means cluster procedure in JMP Genomics 8. Fisher’s exact test was then performed on 7 of the 9 K-means clusters to identify enriched GO terms. Prior to this analysis, GO terms were identified for the entire *C. thermocellum* DSM1313 genome using the standard workflow in Blast2GO 4.1.5 [61].

**Additional files**

- **Additional file 1.** Data logs, substrate and product concentrations, internal and external pH readings, and media formulation for chemostat reactor cultures of *C. thermocellum* LL1210.
- **Additional file 2.** Raw and processed read counts, alignment statistics, log2-fold changes in the gene expression, and K-means clusters and GO enrichment of differentially expressed genes from samples taken from *C. thermocellum* LL1210 cultured in chemostats at pH values 6.98, 6.48, pH 6.24, and pH 6.12 (washout conditions). Gene expression at pH 6.98 was used as a reference for differential expression at lower pH values.
- **Additional file 3: Figure S1.** Average optical density at wave-length 600nm (A and C) and average terminal pH (B and D) of *C. thermocellum* LL1210 cultured in 48-well plates. OD<sub>600nm</sub> readings were taken automatically every 15 min in a microplate spectrophotometer (Biotek Eon, Winooski, VT) kept in an anaerobic chamber. Only 3-h time points are shown. Nine hundred microliters of inoculated medium was mixed with 100 μl of uninoculated medium supplemented with spermine, spermidine, or putrescine (polymamines), or arginine (polyamine precursor) so that the final concentration was 100 μM. Initial culture pH was 7.00 (A and B) or 6.75 (C and D). Averages were calculated from at least three biological replicates. Error bars indicate standard deviation and are colored the same as the amendments in the legend. Table S1. Average and standard deviation of maximum and terminal optical densities (600nm) and specific growth rate of *C. thermocellum* LL1210 cultured in media with and without amendments and having initial pHs of 7.00 and 6.75.
- **Additional file 4: Figure S2.** Average growth (A), terminal pH (B), and remaining substrates and products at the end of *C. thermocellum*-mutant strain fermentations of cellulbiose in MOPS-free carbon-replete medium (C). Averages were computed with data from four biological replicates. Error bars in each graph indicate standard deviation. Some error bars are too small to see. Deletion mutants are designated as LL1210 (hydrogenase maturation protein, lactate dehydrogenase, pyruvate formate lyase, phosphotransacetylase and acetate kinase), GLDH (glutamate dehydrogenase), GS (glutamine synthetase), GOAT (glutamine synthetase), GS-GOAT (both), and NifH (nitrogenase iron protein). The parental strain designated DSM1313 has a deletion in the hypoxanthine phosphoribosyltransferase.
- **Additional file 5: Figure S3.** Differential expression of genes found in Clostridia sporulation cascades. prd<sup>−</sup> processing protease is a stage III sporulation factor. BofA is an inhibitor of the stage IV prd<sup>−</sup> processing protease SpoIII FB. Table S2. Percentage of spherical morphologies 144 and 216 h after inoculation. Figure S4. Substrates and products (A) and the pH (B) after 144 and 216 h of *C. thermocellum*-mutant fermentations on MOPS-free carbon-replete medium starting with an initial pH of 6.75. Significant differences at α = 0.001 for comparisons with DSM1313 (∆hpt) are indicated with a ** and comparisons with DSM1313 (∆hpt) and LL1210 are indicated with ***. Averages were calculated with six biological replicates. Error bars indicate standard deviation.
Additional file 6: Table S3. Fold changes of intercellular metabolite there were significantly higher or lower in concentration at growth-limiting pHs.

Additional file 7. Extracellular amino acid concentrations in media from C. thermocellum LL1210 chemostats that were sampled when pH values were pH 6.48, pH 6.24, pH 6.12, and below. Demonstrations of data homoscedasticity for 7 tests.

Additional file 8: Table S4. Strains and primers used in this study.

References
1. Blumer-Schueette SE, Brown SD, Sander KB, Bayer EA, Kataeva I, Zurawski JV, et al. Thermophilic lignocellulose deconstruction. Rev: FEMS Microbiol. 2014.
2. Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS. Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev. 2002;66:506–77.
3. Lin PP, Mi L, Moriocka AH, Yoshino KM, Konishi S, Xu SC, et al. Consolidated bioprocessing of cellulose to isobutanol using Clostridium thermocellum. Metab Eng. 2015;31:44–52.
4. Dumitrache A, Klingeman DM, Natzke J, Rodriguez M Jr, Giannone RJ, HetTich RL, et al. Specialized activities and expression differences for Clostridium thermocellum biofilm and planktonic cells. Sci Rep. 2017;7:43583.
5. Xu Q, Resch MG, Podkaminer K, Yang S, Baker JO, Donohoe BS, et al. Dramatic performance of Clostridium thermocellum explained by its wide range of cellulosomal modalities. Sci. Adv. 2016;2:e1501254.
6. Bayer EA, Lamed R, White BA, Flints HJ. From cellulosomes to cellulosomes. Chem Rev. 2008;8:364–77.
7. Ellis LD, Holwerda HK, Hogsett D, Rogers S, Shao X, Tschaplinski TJ, et al. Closing the carbon balance for fermentation by Clostridium thermocellum (ATCC 27405). Bioreour Technol. 2012;103:293–9.
8. Holwerda HK, Thorne PG, Olson DG, Amador-Noguez D, Engle NL, Tschaplinski TJ, et al. The exometabolome of Clostridium thermocellum reveals overflow metabolism at high cellulose loading. Biotechnol Biofuels. 2014;7:155.
9. Tripathi SA, Olson DG, Argyros DA, Miller BB, Barrett TF, Murphy DM, et al. Development of pyrF-based genetic system for targeted gene deletion in Clostridium thermocellum and creation of a pta mutant. Appl Environ Microbiol. 2010;76:6591–9.
10. Argyros DA, Tripathi SA, Barrett TF, Rogers SR, Feinberg LF, Olson DG, et al. High ethanol titers from cellulose by using metabolically engineered thermophilic, anaerobic microbes. Appl Environ Microbiol. 2011;77:8288–94.
11. Rydzak T, Lynd LR, Guss AM. Elimination of formate production in Clostridium thermocellum. J Ind Microbiol Biotechnol. 2015;42:1263–72.
12. Papanek B, Biswas R, Rydzak T, Guss AM. Elimination of metabolic pathways to all traditional fermentation products increases ethanol yields in Clostridium thermocellum. Metab Eng. 2015;32:49–54.
13. Biswas R, Prabhu S, Lynd LR, Guss AM. Increase in ethanol yield via elimination of lactate production in an ethanol-tolerant mutant of Clostridium thermocellum. PLoS ONE. 2014;9:1–7.
14. Olson DG, Hori M, Fuhrer T, Cui J, Zhou J, Maloney MI, et al. Glycolysis without pyruvate kinase in Clostridium thermocellum. Metab Eng. 2017;39:169–80.
15. Biswas R, Wilson CM, Zheng T, Giannone RJ, Dawn M, Olson DG, et al. Elimination of hydrogenase active site assembly blocks H2 production and increases ethanol yield in Clostridium thermocellum. Biotechnol Biofuels. 2015;8:20.
16. Hon S, Olson DG, Holwerda EK, Lanahan AA, Murphy SJL, Maloney MI, et al. The ethanol pathway from Thermoaerobacterium saccharolyticum improves ethanol production in Clostridium thermocellum. Metab Eng. 2017;42:175–84.
17. Lo J, Olson DG, Murphy SJL, Tian L, Hon S, Lanahan A, et al. Engineering electron metabolism to increase ethanol production in Clostridium thermocellum. Metab Eng. 2017;39:71–9.
18. Tian L, Papanek B, Olson DG, Rydzak T, Holwerda EK, Zheng T, et al. Simultaneous achievement of high ethanol yield and titer in Clostridium thermocellum. Biotechnol Biofuels. 2016;9:116.
19. Brown SD, Guss AM, Karpinetis TV, Parks JM, Smolin N, Yang S, et al. Mutant alcohol dehydrogenase leads to improved ethanol tolerance in Clostridium thermocellum. Proc Natl Acad Sci. 2011;108:13752–7.
20. Shao X, Raman B, Zhu M, Mielenz JR, Brown SD, Guss AM, et al. Mutant selection and phenotypic and genetic characterization of
ethanol-tolerant strains of Clostridium thermocellum. Appl Microbiol Biotechnol. 2011;92:641–52.

21. Verbeke TJ, Giannone RJ, Klingeman DM, Engle NL, Rydzak T, Guss AM, et al. Pentose sugars inhibit metabolism and increase expression of an AgrD-type cyclic pentapeptide in Clostridium thermocellum. Sci Rep. 2017;7:43355.

22. Poudel S, Giannone RJ, Rodriguez M, Raman B, Martin MZ, Engle NL, et al. Integrated omics analyses reveal the details of metabolic adaptation of Clostridium thermocellum to lignocellulose-derived growth inhibitors released during the deconstruction of switchgrass. Biotechnol Biofuels. 2017;10:14.

23. Wu C-W, Spike T, Klingeman DM, Rodriguez M, Bremer VR, Brown SD. Generation and characterization of acid tolerant Fibrobacter succinogenes SBS5. Sci Rep. 2017;7:2277.

24. Nochur SV, Demain AL, Roberts MF. Carbohydrate utilization by Clostridium thermocellum: importance of internal pH in regulating growth. Enzyme Microb Technol. 1992;14:338–49.

25. Herrero AA, Gomez RF, Snedecor B, Tolman CJ, Roberts MF. Growth inhibition of Clostridium thermocellum by carboxylic acids: a mechanism based on uncoupling by weak acids. Appl Microbiol Biotechnol. 1985;22:53–62.

26. Mori Y. Characterization of symbiotic coculture of Clostridium thermo‑hydro‑sulfuricum YMY and Clostridium thermocellum YMY. Appl Environ Microbiol. 1990;56:37–42.

27. Koide KE, Zvedov LV, Liebl W, Schwarz WH. Comparative genotyping of Clostridium thermocellum strains isolated from biogas plants: genetic markers and characterization of cellulolytic potential. Microbiol. Syst Appl, 2014.

28. Freier D, Mothershed CP, Wiegler J. Characterization of Clostridium thermocellum JW20. Appl Environ Microbiol. 1988;54:204–11.

29. Cotter PD, Hill C. Surviving the acid test: responses of Gram-positive bacteria to low pH. Microbiol Mol Biol Rev. 2003;67:429–53.

30. Lund P, Tramonti A, De Biase D. Coping with low pH: molecular strategies in neutrophilic bacteria. FEMS Microbiol Rev. 2014;38:1091–125.

31. Kobayashi H, Murakami N, Unemoto T. Regulation of the Cytoplasmic pH in, 1982;257:13246–52.

32. Terahara N, Noguchi Y, Nakamura S, Kami-iike N, Ito M, Namba K, et al. Load- and polysaccharide-dependent activation of the Na+-type MotPS stator in the Bacillus subtilis flagellar motor. Sci Rep. 2017;7:46081.

33. Rhee HI, Kim EJ, Lee JK. Physiological polyaniones: simple primordial stress molecules. J Cell Mol Med. 2007;11:685–703.

34. Mears EB, Lynd LR. The identification of four histidine kinases that influence sporulation of Clostridium thermocellum. Anserobe. 2014;28:109–19.

35. Mears EB, Ziendez JA, Lynd LR. Formation and characterization of non‑growth states in Clostridium thermocellum: spores and L-forms. BMC Microbiol. 2012;12:180.

36. Wilson CM, Yang S, Rodriguez M Jr, Ma Q, Johnson CM, Dice L, et al. Clostridium thermocellum transcriptomic profiles after exposure to furfural or heat stress. Biotechnol Biofuels. 2013;6:131.

37. Søhlenkamp C. Membrane homeostasis in bacteria upon pH challenge. Biogenesis of fatty acids, lipids and membranes. Berlin: Springer; 2017. p. 317–30.

38. Tsuchiya K, Nishimura K, Iwahara M. Purification and characterization of lysyl oxidase from Clostridium thermocellum. J Biol Chem. 1986;261:627–30.

39. Ganev LD, Ganeva VA, Baltsche H, et al. A tightly membrane-bound family. FEBS Lett. 1999;457:527–33.

40. Lynd LR, Guss AM, Himmel ME, Beri D, Herring C, Holwerda EK, et al. Advances in consolidated bioprocessing using Clostridium thermocellum and Thermoanaerobacter saccharolyticum. In: Industrial biotechnology. Weinheim: Wiley-VCH; 2016. p. 365–94.

41. Dumitrache A, Akinoaho H, Rodriguez M, Meng X, Yoo CG, Natzke J, et al. Consolidated bioprocessing of Populus using Clostridium Ruminoclostridium thermocellum: a case study on the impact of lignin composition and structure. Biotechnol Biofuels. 2016;9:31.

42. Sander K, Wilson CM, Rodriguez M, Klingeman DM, Rydzak T, Davison BH, et al. Clostridium thermocellum DSM 1313 transcriptional responses to redox perturbation. Biotechnol Biofuels. 2015;8:211.

43. Dumitrache A, Wolfardt GM, Allen DG, Liss SN, Lynd LR. Tracking the cellulosic activity of Clostridium thermocellum biofilms. Biotechnol Biofuels. 2013;6:175.

44. Wilson CM, Klingeman DM, Schlachter C, Syed MH, Wu C, Guss AM, et al. LacI transcriptional regulatory networks in Clostridium thermocellum. Cytotechnology. 2013;65:13.

45. Andrews SR. FastQC: a quality control tool for high throughput sequence data. 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc. Accessed 5 Feb 2016.

46. Bolger AM, Lohse M, Usadel B. Trimomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.

47. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory‑efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10:R25.

48. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.

49. Anders S. FastQC: a quality control tool for high throughput sequence data. 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc. Accessed 5 Feb 2016.

50. Bologer AM, Lohse M, Usadel B. Trimomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.

51. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10:R25.

52. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.

53. Anders S. FastQC: a quality control tool for high throughput sequence data. 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc. Accessed 5 Feb 2016.

54. Bolger AM, Lohse M, Usadel B. Trimomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.

55. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10:R25.

56. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.