Box GI.1. Carbohydrate-active enzymes (CAZymes).

CAZymes are enzymes that assemble, break down, and modify glycoconjugates, oligo- and polysaccharides (Cantarel *et al.*, 2009, Drula *et al.*, 2022, Sun *et al.*, 2023). A list of all CAZymes and their classification is available on the CAZy website (http://www.cazy.org/) (Drula *et al.*, 2022). CAZymes are divided into 6 **classes** based on their activities:

- **Glycoside hydrolases** (GHs), including glycosidases and transglycosidases. They are responsible for the hydrolysis and/or transglycosylation of glycosidic bonds (Cantarel *et al.*, 2009).
- **Glycosyltransferases** (GTs). They are responsible for the biosynthesis of glycosidic bonds from phospho-activated sugar donors (Cantarel *et al.*, 2009).
- **Polysaccharide lyases** (PLs). They cleave the glycosidic bonds of uronic acid-containing polysaccharides by a β-elimination mechanism (Cantarel *et al.*, 2009).
- **Carbohydrate esterases** (CEs). They remove ester-based modifications present in mono-, oligo- and polysaccharides and thereby facilitate the action of GHs on complex polysaccharides (Cantarel *et al.*, 2009).
- **Carbohydrate-binding modules** (CBMs). They are autonomously folding and functioning protein fragments. While they do not have any enzymatic activity by themselves, they enhance the activity of the enzymes that belong to the other CAZymes classes by targeting them to their substrate, or by promoting a longer interaction with it (Cantarel *et al.*, 2009).
- **Auxiliary Activities** (AAs). They are a widespread collection of catalytic modules involved in plant cell wall degradation (Levasseur *et al.*, 2013). AAs group ligninolytic enzymes and lytic polysaccharide mono-oxygenases (LPMOs). Lignin breakdown enzymes may not act on carbohydrates; however, as the degradation of lignin may grant access to the carbohydrates to which it is intimately associated in the plant cell wall, these enzymes were integrated to the CAZy database.

Within classes, CAZymes are further divided into **families**. Each family is built around at least one biochemically characterized member, and is then populated with sequences that exhibit similar amino acid sequences (Cantarel *et al.*, 2009, Drula *et al.*, 2022). The substrate specificity is variable within the families, while the catalytic mechanism is extremely well conserved and thus predictable once established for a family member (Drula *et al.*, 2022).