



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

A draft genome of the ginger species *Alpinia nigra* and new insights into the genetic basis of flexistyly

Citation for published version:

Ranavat, S, Becher, H, Newman, MF, Gowda, V & Twyford, AD 2021, 'A draft genome of the ginger species *Alpinia nigra* and new insights into the genetic basis of flexistyly', *Genes*, vol. 12, no. 9, 1297. <https://doi.org/10.3390/genes12091297>

Digital Object Identifier (DOI):

[10.3390/genes12091297](https://doi.org/10.3390/genes12091297)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Genes

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Article

A Draft Genome of the Ginger Species *Alpinia nigra* and New Insights into the Genetic Basis of Flexistyly

Surabhi Ranavat ^{1,2,*} , Hannes Becher ¹, Mark F. Newman ², Vinita Gowda ³ and Alex D. Twyford ^{1,2} 

¹ Institute of Evolutionary Biology, University of Edinburgh, Charlotte Auerbach Road, Edinburgh EH9 3FL, UK; H.Becher@ed.ac.uk (H.B.); Alex.Twyford@ed.ac.uk (A.D.T.)

² Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, UK; MNewman@rbge.org.uk

³ Department of Biological Sciences, Indian Institute of Science Education and Research Bhopal, Bhopal Bypass Road, Bhauri, Bhopal 462 066, Madhya Pradesh, India; gowdav@iiserb.ac.in

* Correspondence: s.ranavat@sms.ed.ac.uk

Abstract: Angiosperms possess various strategies to ensure reproductive success, such as stylar polymorphisms that encourage outcrossing. Here, we investigate the genetic basis of one such dimorphism that combines both temporal and spatial separation of sexual function, termed flexistyly. It is a floral strategy characterised by the presence of two morphs that differ in the timing of stylar movement. We performed a *de novo* assembly of the genome of *Alpinia nigra* using high-depth genomic sequencing. We then used Pool-seq to identify candidate regions for flexistyly based on allele frequency or coverage differences between pools of anaflexistylous and cataflexistylous morphs. The final genome assembly size was 2 Gb, and showed no evidence of recent polyploidy. The Pool-seq did not reveal large regions with high F_{ST} values, suggesting large structural chromosomal polymorphisms are unlikely to underlie differences between morphs. Similarly, no region had a 1:2 mapping depth ratio which would be indicative of hemizygosity. We propose that flexistyly is governed by a small genomic region that might be difficult to detect with Pool-seq, or a complex genomic region that proved difficult to assemble. Our genome will be a valuable resource for future studies of gingers, and provides the first steps towards characterising this complex floral phenotype.

Keywords: flexistyly; *Alpinia*; stylar polymorphism; Zingiberaceae; Pool-seq; genome assembly; diploidisation



Citation: Ranavat, S.; Becher, H.; Newman, M.F.; Gowda, V.; Twyford, A.D. A Draft Genome of the Ginger Species *Alpinia nigra* and New Insights into the Genetic Basis of Flexistyly. *Genes* **2021**, *12*, 1297. <https://doi.org/10.3390/genes12091297>

Academic Editor: Roberta Bergero

Received: 2 July 2021

Accepted: 20 August 2021

Published: 24 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Angiosperms possess a remarkable range of floral traits and forms that enable them to reproduce successfully [1]. The evolution of these reproductive strategies may be driven by a range of factors. These traits are likely to be adaptations to ensure pollination by means of a pollen vector [2,3] or to prevent conflicts that interfere with pollen deposition on the stigma and pollen removal from the anthers [4]. Alternatively, as flowering plants are mainly hermaphroditic in nature, opportunities for self-fertilisation can occur. This can lead to reduction in viability and fertility of offspring, or inbreeding depression [5]. Therefore, strategies to avoid self-pollination are widespread in angiosperms, and many of these reproductive strategies involve traits which promote outcrossing through pollen vectors [3].

Among many such floral traits are stylar polymorphisms that promote insect-mediated cross-pollination and prevent pollen wastage due to interference between male and female function within a bisexual flower [3]. The most widely studied stylar polymorphism is heterostyly, where each population comprises two (distyly) or three (tristyly) floral morphs that differ reciprocally in style and stamen lengths. Recognised as early as the 16th century in the genus *Primula* [6], heterostyly was also one of Darwin's research interests. His work led to the publication of *The different forms of flowers on plants of the same species* [7] which has laid a strong foundation for the research on heterostyly over the years.

Significant progress has been made into understanding the genetic basis of heterostyly. Distyly was thought to be a diallelic trait, with the short-styled morph dominant (S/s) and the long-styled morph recessive (s/s). Three genes at the S locus were thought to control the trait— G (style length, (Griffel–German: style)), P (pollen size and number) and A (anther position) [8] in tight linkage due to suppressed recombination [9]. Recent work on the genetics of heterostyly in *Primula* has revealed that the short-styled morphs are hemizygous and not heterozygous for the S locus. The s haplotype lacks a 278-kb sequence that contain genes such as *CYP734A50* (cytochrome P450) that controls style length and *GLO2* (short-morph-specific *GLOBOSA* gene) that controls anther height [10–12]. The hemizygous nature of distyly has also been found in other systems such as *Fagopyrum* [13], *Linum* [14] and *Turnera* [15]. While we are now unravelling the genetic basis of distyly in a range of plant groups, such knowledge is lacking for other stylar polymorphisms.

Among the range of stylar polymorphisms, only flexistyly involves stylar movement. The term flexistyly was first coined by Li et al. [16] who observed this mechanism in nine species of the ginger genus *Alpinia*. Flexistylous species possess two morphs that exhibit spatial and temporal separation of sexual functions. The morphs are cataflexistylous (protandrous, cata–Greek: downwards), where the anthers disperse pollen in the morning, while the style is curved above the anther, and anaflexistylous (protogynous, ana–Greek: upwards), where the style is curved below the anther while the anther does not disperse pollen (Figure 1). Around midday, the style of the cataflexistylous flower curves below the anther and receives cross pollen while the anther halts pollen release, and the style of the anaflexistylous flower curves above the anther which begins to release pollen [16–18]. Therefore, such species are heterodichogamous in nature, where some individuals in a population are first male in function and then female, and vice versa [19]. Flexistyly is known to reduce pollen–pistil interference and promote outcrossing [16–18,20]. This phenomenon was not observed in detail until 1996 when Cui et al. [21] discovered reciprocal stylar movement in *Amomum tsao-ko* (now accepted as *Lanxangia tsao-ko*). They observed the two floral morphs and found that only one morph type can be borne on a particular plant. The response of flowers was found to be dependent on photoperiod and the flowers were found to be self-compatible when artificially pollinated. The fruit set in the presence of only one morph type was found to be low hence it is pollinator-dependent for reproductive success.

Flexistyly is now known to occur in more than 24 species in the plant family Zingiberaceae, subfamily Alpinioideae, mostly in *Alpinia* s.l. [22]. It is also found in several other ginger genera such as *Amomum* s.l., *Etilingera*, *Paramomum*, *Plagiostachys* and *Siliquamomum* [21–25]. The two stylar morphs are always present in a 1:1 ratio in natural populations and hence it is presumed that two alleles segregating at a single locus control flexistyly [19]. A similar pattern is seen in other heterodichogamous species such as walnuts [26] and pecans [27], where protandry is recessive and protogyny is dominant homozygote, but the heterozygote does not occur due to the lack of selfing.

Whole-genome sequencing of pools of individuals (Pool-seq) is a powerful and cost-effective approach to determining allele frequency differences across the genome between natural populations or experimental lines differing in phenotype [28]. This approach became popular as the costs for library preparation and sequencing are greatly reduced relative to individual genome resequencing. Most notably, Pool-seq can be used to find the genes underlying a phenotype, without the need for prior candidate genes [29] and without carrying out experimental crosses and growing a large number of progeny (like QTL mapping). This approach has now been widely used in many different species and there are many established bioinformatic pipelines available [29]. A limitation of Pool-seq, however, is that pools of individuals must be compared to a reference genome. So far however, there is no complete genome assembly available for *Alpinia* or the wider Zingiberaceae. The nearest available genomes are of *Curcuma longa* [30] and *Musa* spp. from the banana family (Musaceae, Zingiberales) [31–34]. The generation of a complete genome for a representative ginger species would fill an important phylogenetic gap,

and enable comparative evolutionary genomic analyses such as Pool-seq comparisons of natural populations, in this key lineage.



Figure 1. Flexistylous in *Alpinia*. (A) Cataflexistylous morph (style curved above the anther along with anther dehiscence) (*Alpinia* cf. *malaccensis*). Picture taken at 10:22 h. (B) Anaflexistylous morph (style curved below the anther and no anther dehiscence) (*Alpinia roxburghii*). Picture taken at 10:30 h. Accessions photographed at the Royal Botanic Garden Edinburgh.

As little is known about the genetic basis of flexistylous, the aims of this study were: (1) to generate a complete genome for a representative *Alpinia* species and (2) to identify candidate genomic regions underlying the stylar dimorphism observed in *Alpinia* using a Pool-seq approach.

As a complex phenotypic trait combining stylar movement and response to different photoperiods, we expect flexistylous to be controlled by multiple genes. However, tight linkage between genes is necessary for the maintenance of this dimorphism in the presence of recombination. Based on this, our expectation is that there will be few regions of the genome differentiated between morphs. Such regions may be detected by outlier scans for allele frequency differences, measured as F_{ST} between pools of samples of ana- and cataflexistylous morphs. Tight clusters of genetic variants differentially fixed between sample pools may correspond to structural chromosomal differences such as chromosomal inversions, that protect multiple loci from recombination [35–38]. Alternatively, if this trait is hemizygous, we would expect to see a coverage difference between the pools in regions of the genome underlying flexistylous. Finally, flexistylous may be controlled by an individual locus, which may be more difficult to detect in genome-wide surveys of differentiation.

2. Materials and Methods

2.1. Study System

Alpinia Roxb. is the largest and one of the most taxonomically challenging genus in the Zingiberaceae. It comprises more than 230 species that are widespread throughout tropical and subtropical Asia and extend as far as the Western Pacific islands and northern Australia [39]. It is an economically important genus, with species rich in compounds with pharmacological activity [40]. *Alpinia* species are thought to be tetraploids with $2n = 4x = 48$ [41,42]. Previous genome size estimates in *Alpinia* include a 1C value of 2.29 pg for *Alpinia nigra* [43]. Many species in this genus are flexistylous, one being *Alpinia nigra* (Gaertn.) B.L.Burtt, which is distributed from the Indian subcontinent to China (South

Yunnan) (Figure 2). This species was chosen to study the genetics of flexistyly as large populations can be easily found in its native range.

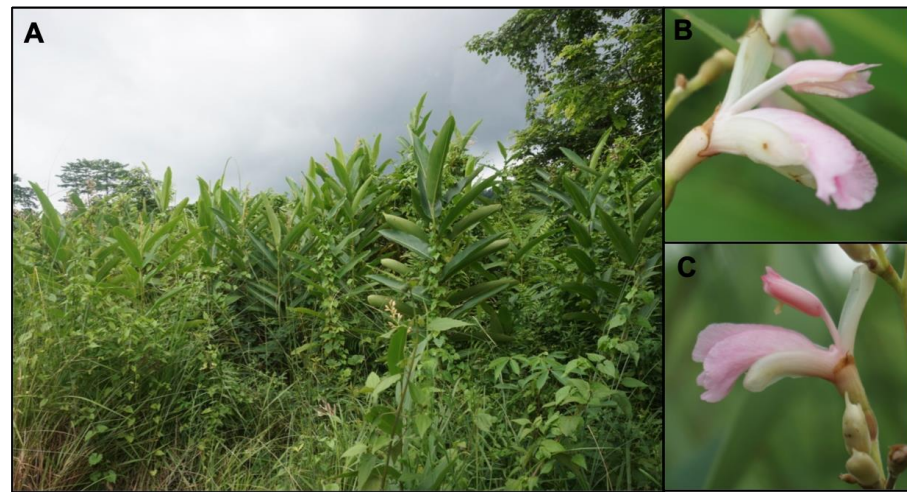


Figure 2. Habitat and floral details of *Alpinia nigra* at Pakke Tiger Reserve, Arunachal Pradesh, India. (A) Population of *Alpinia nigra* at Pakke Tiger Reserve, (B) cataflexistylous (protandrous) morph (picture taken at 08:45 h) and (C) anaflexistylous (protogynous) morph (picture taken at 09:40 h).

2.2. Sample Collection

Samples were collected in July 2018 from Pakke Tiger Reserve near the Seijosa region (26°57' N 92°59' E) located in East Kameng district, Arunachal Pradesh (collection permit number, CWL/G/13(95)/2011-12/Pt.V/342/-26) in India. The populations of *A. nigra* were found at an elevation of 100–200 m in the periphery of the reserve (Figure 2). They were found in marshy regions and were widespread in the reserve.

Leaf tissues of 51 anaflexistylous (ana-morph) and 63 cataflexistylous (cata-morph) individuals of *Alpinia nigra* were collected, along with one individual for the reference genome that was cataflexistylous in nature (Supplementary Materials Table S1). As *Alpinia* and other species in the Zingiberaceae reproduce vegetatively by underground rhizomes, care was taken to minimize sampling clones (i.e., same genet) by collecting individuals at least 1 m apart. Samples were collected across a total distance of 2 km (Figure 3). The tissues were cut into 1 cm² pieces and dried in silica gel [44]. The voucher specimens and tissues were deposited at BHPL, IISER Bhopal.

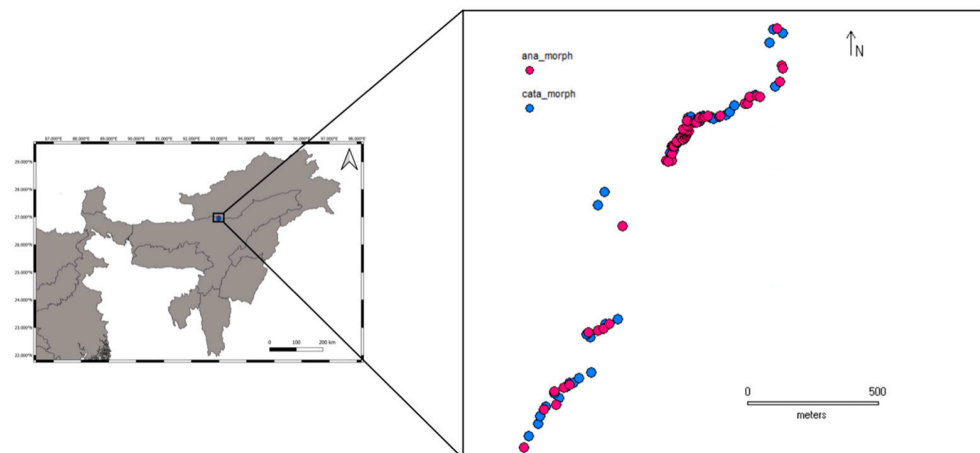


Figure 3. Sample collection map of the ana- and cata-morphs of *Alpinia nigra* from Pakke Tiger Reserve, Arunachal Pradesh, North-East India.

2.3. Reference Genome Assembly

We adopted a sequencing and assembly strategy aimed at giving the highest assembly completeness possible, while accounting for the necessary limitations of working with our field-collected samples. As *Alpinia nigra* is a tropical plant found in remote locations, the tissue was collected in silica gel, which preserves the tissue by drying it. This DNA preservation method prevents the utilisation of long-read sequencing due to difficulties in obtaining high molecular weight DNA. However, a combination of a high-coverage, longer Illumina reads (250 bp paired end, PE), combined with suitable genome assembly software such as DISCOVAR *de novo*, can produce a highly complete genome suitable for downstream comparative genomic analyses.

2.3.1. DNA Extraction (Genome)

DNA was extracted from a silica-dried leaf tissue sample from a cata-morph individual collected from Pakke Tiger Reserve, using a modified CTAB protocol at TrEE Lab, IISER Bhopal [45]. Eight such extractions were performed using leaf tissue from the same individual followed by gel extraction using a QIAEX II kit (Qiagen, Valencia, CA, USA) following the manufacturer's protocol. The DNA from all the samples was pooled and purified using Genomic Tips kit (Qiagen, Valencia, CA, USA). The sample was then concentrated using a Savant SpeedVac (Thermo Fisher Scientific, Waltham, MA, USA). An Illumina Tru-Seq Nano gel-free library with approximately 550 bp insert size was prepared from the DNA sample. The genome was finally sequenced at Edinburgh Genomics on a single lane of an SP flowcell on the NovaSeq 6000 (Illumina, San Diego, CA, USA) with 250 bp paired-end reads, with the aim of generating 100x sequencing coverage.

2.3.2. Genome Assembly

The quality of the raw reads was checked using FastQC v0.11.7 [46]. We used the k-mer analysis toolkit (KAT) [47] to generate both a k-mer spectrum of the raw sequencing data and a completeness plot comparing raw data and assembly (using the spectra cn tool). We obtained an estimate of the sample's genome size from the k-mer spectrum. Smudgeplot [48] was used to estimate the ploidy and copy number of genomic regions from the reads. The genome was subsequently assembled with default parameters using DISCOVAR *de novo*, which can rapidly generate highly complete assemblies for large genomes using paired reads of length 250 bp or greater [49].

The presence of DNA contamination from bacteria and fungi was assessed using BlobTools v1.1.1 [50], with per-contig mapping depths from reads mapped back to the assembly using BWA MEM v0.7.17 [51], and sequence similarities to the NCBI nt database using blastn (v2.9.0+) and a Diamond (v0.9.24) blastx [52] search against the Uniprot database. The completeness of the assembly was evaluated using BUSCO v3.0.2 [53] with the embryophyta lineage dataset (downloaded in 2019). The quality of the assembly was evaluated using QUAST v5.0.2 [54].

2.4. Pool-Seq

2.4.1. DNA Extraction

DNA was extracted separately from 51 ana-morph and 63 cata-morph individuals using the Qiagen DNeasy Plant Mini kit (Qiagen, Valencia, CA, USA) following the manufacturer's protocol. The DNA was quantified using the Qubit dsDNA BR Assay kit on the Qubit 2.0 Fluorometer. Two pools of the respective ana- and cata-morphs were prepared, with equimolar quantities of each individual sample. Low-molecular-weight DNA fragments were removed from the DNA pools by gel extraction using the ZymoClean Large Fragment DNA Recovery kit (Zymo Research, Irvine, CA, USA). The DNA was then concentrated using the Savant SpeedVac (Thermo Fisher Scientific, Waltham, MA, USA) to reach a suitable final concentration. The samples were sequenced by Novogene (UK) with 150 bp paired-end reads on an Illumina NovaSeq 6000 instrument (Illumina, San Diego, CA, USA), aiming at generating 50x sequencing coverage per pool.

2.4.2. Genome Mapping and Analysis

Our comparative genomic analyses aimed to detect allele frequency differences between the ana- and cata-morph pools using F_{ST} calculated in sliding windows across the genome. A high F_{ST} value between morph types would be indicative of differentiated regions associated with the phenotype.

The raw reads were quality checked using FastQC v0.11.8 [46] and were aligned to our reference assembly using BWA MEM v.0.7.17 [51], to generate a SAM file. Samtools v1.9 [55] was then used for conversion to BAM, sorting and indexing. PoPoolation2 [56] was used to compare the allele frequency differences between the two pools as follows: Ambiguously mapped reads were removed using Samtools v1.9 using the option view, -q 20 was set to include reads with mapping quality > 20, and then the file was sorted. Samtools was used to create an mpileup with the BAM files from the two different pools, which was then synchronised. F_{ST} values between the two sample pools were calculated using a sliding window approach with window sizes 1 kbp and 10 kbp with a step size of 1 kbp for both windows with min-coverage 10, max-coverage 500 and min-count 10. We used two different window sizes as larger window sizes smooth out the noise but reduce the potential signal from individual loci. The F_{ST} files were converted into an IGV (Integrated Genomics Viewer) compatible format. Manhattan plots of F_{ST} values were made using R v3.6.3. The IGV files were converted to be compatible with R using the perl script Genome_R_script.pl (https://github.com/Gammerdinger/Manhattan_plots, accessed on 16 October 2020). For plotting, all contigs were arranged in descending order of length. Per-contig average mapping depths were computed using a custom python script (Supplementary Materials File S1) from the output of Samtools's subroutine "depth" (samtools depth, aa <IN.BAM>).

3. Results

3.1. *Alpinia nigra* Genome Assembly

A reference genome of wild-collected *Alpinia nigra* was generated from a cataflexistylous (cata) individual using 528 million paired Illumina reads. Based on a previous estimate of 2.29 pg [43], this would correspond to approximately 118x coverage. The pre-assembly quality check indicated that most sequences were of high quality (Phred score > 30) and did not indicate adapter contamination. The 27-mer spectrum had two main peaks indicating the presence of heterozygosity, and suggesting that the genome was diploid (Figure 4). An additional, smaller third peak at $x = 222$ indicated some degree of sequence duplication.

The k-mer comparison plot (Figure 4) indicates very high completeness of this reference assembly with respect to low-copy sequences. There were very few k-mers absent from the assembly, which were present in the sequencing data at diploid multiplicity (thin black line at 118x, Figure 4). Based on the k-mer spectrum, the genome size was estimated as 2.01 Gbp, which corresponds to 2.056 pg [57]. As expected for an outcrossing species, the k-mer spectrum showed moderate heterozygosity, estimated as 1.06%. Additional sequence filtering was not required as the blobplot showed very low contamination (Figure S1).

Assembly metrics also revealed a high level of assembly completeness, with complete (single and duplicated) BUSCOs over 91% (Table 1). The overall assembly completeness was relatively high, with a scaffold N50 of 48.9 kbp, though many contigs smaller than 1000 bp were present.

Graphical inspection of the multiplicity of heterozygous k-mer pairs, implemented in Smudgeplots (Figure 5), revealed that 84% of the k-mer pairs were in 1:1 ratio, which is a strong signal of diploidy. Haplotype ratios of 2:1 and 3:1 were also observed albeit in low proportions suggesting historical polyploidy and/or the retention of duplicate regions.

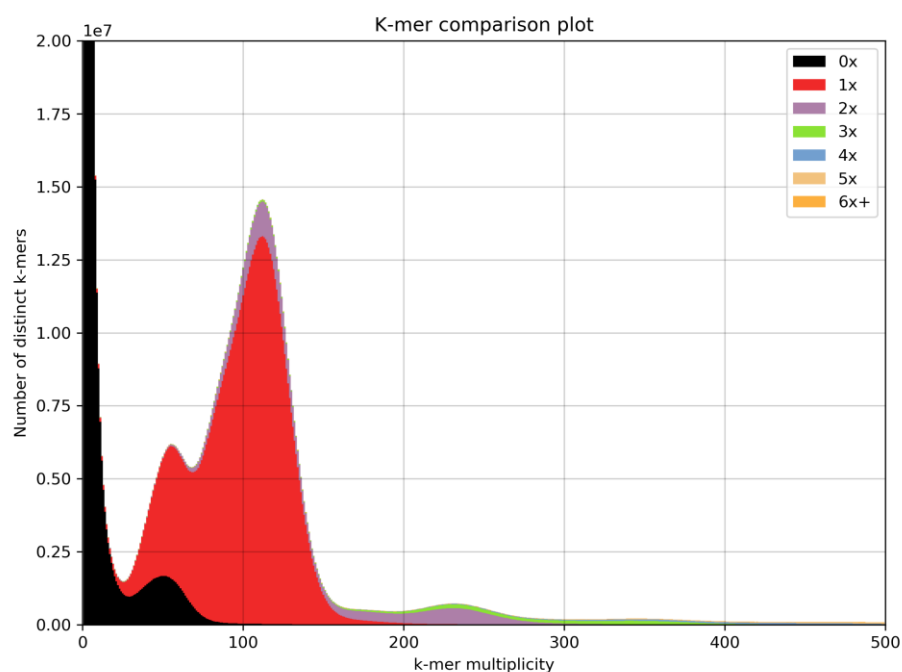


Figure 4. A k-mer comparison plot for the genome of *Alpinia nigra* ($k = 27$). The colour of the plots indicates the number of times k-mers from the reads appear in the assembly. Black region indicates the k-mers missing from the assembly, red indicates k-mers that are present once in the assembly, purple twice, and so on.

Table 1. Assembly metrics for the *Alpinia nigra* genome assembled using DISCOVAR *de novo*. The assembly metrics were estimated using QUASt v5.0.2 [54] and the completeness of the assembly was estimated with BUSCO v3.0.2 using the embryophyta_odb10 BUSCO set.

Total Length (≥ 0 bp)	2389 Mb
Total length (≥ 1000 bp)	2119 Mb
No. of contigs (≥ 0 bp)	1,072,070
No. of contigs (≥ 1000 bp)	158,389
N50 Length	48.9 Kb
Longest contig	591.2 Kb
BUSCO score ¹	C: 91.1%, (S: 84.1%, D: 7.0%), F: 4.9%, M: 4.0%, n : 1375

¹ C= complete (S = single copy, D = duplicated), F = fragmented, M = missing, n = number of orthologues in comparison.

3.2. Pool-Seq

Two pools of 51 ana-morph and 63 cata-morph individuals, respectively, from a single wild population (the same as the reference genome) were sequenced using Illumina short reads. A total of 392 million paired reads for ana-morph pool and 387 million paired reads for cata-morph pool were generated. About 92% of reads from the ana-morph pool and 90% of reads from the cata-morph pool mapped to the reference genome. Ambiguously mapped reads were removed, leaving only reads mapping to a single location for downstream analysis. The peak mapping depths were 55x for the cata and 57x for the ana pool. There were no obvious coverage differences in the first 200,000 contigs between the ana and cata pools that were ordered by size (summed length, 2.15 Gbp; largest contig- 591 kbp; smallest contig, 700 bp) (Supplementary Materials Figure S1) nor between windows of 10 kb length (Figure 6).

F_{ST} values from the Popoolation2 analysis revealed low genome-wide F_{ST} between ana and cataflexistylous pools, with a mean value, based on 10 kbp windows, of 0.0432. The distribution of F_{ST} values showed that over 99.6% of the windows had F_{ST} values of 0.2 or lower. The number of windows greater than $F_{ST} > 0.5$ and > 0.8 was 152 and 7, respectively,

from a total of 1.66 million windows. All outlier windows with F_{ST} values > 0.2 had lower mapping depth than expected, indicating potential pool-specific sequence contamination (discussed below).

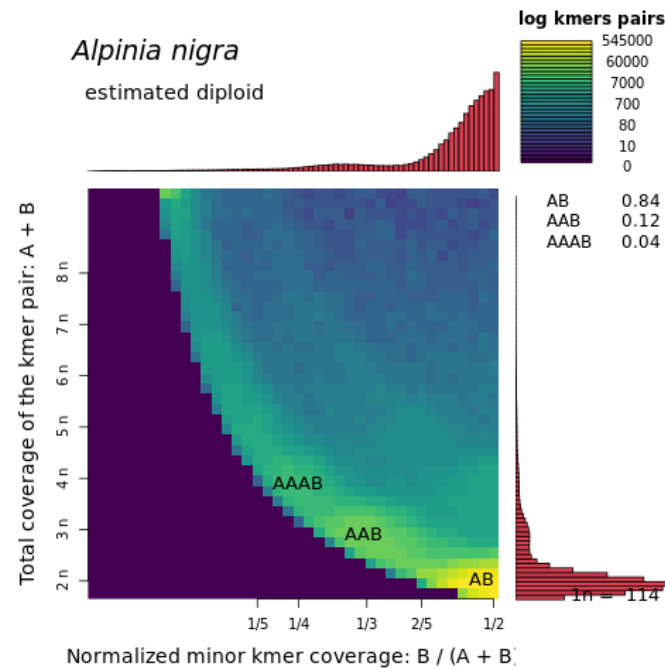


Figure 5. Smudgeplot for *Alpinia nigra* for $k = 21$. The colour intensity of the smudge indicates how frequently the haplotype structure is represented in the genome. The bar plots represent the sequencing coverage.

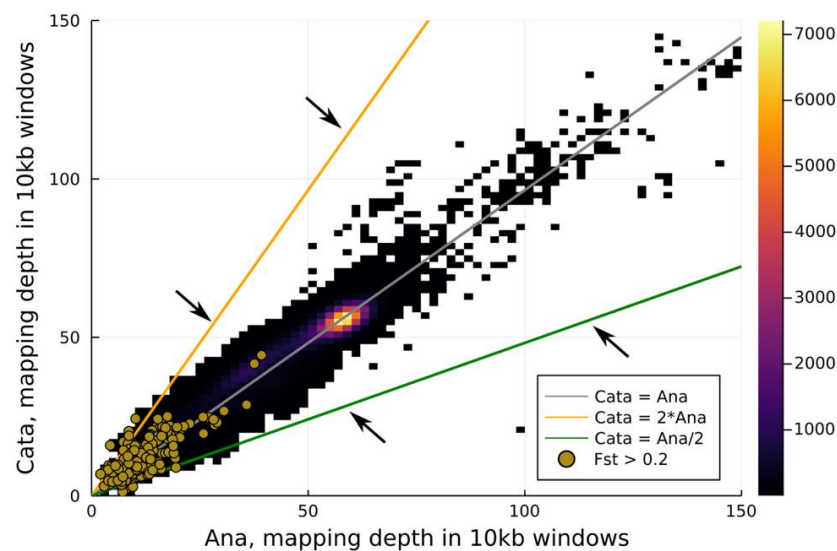


Figure 6. Histogram of average mapping depths in windows of 10 kb between the ana and cata pools. The colour bar on the right corresponds to tile colours in the main plot and indicates the number of windows in a certain mapping depth class. Windows with F_{ST} values > 0.2 are indicated by gold dots. The orange and green lines indicate mapping depth ratios between the ana and cata pools of 2 and $1/2$, respectively. The position of the arrows indicate the points where hemizygous sequences are expected in either of the morphs.

To visualise F_{ST} values across scaffolds, a sliding-window analysis was done for window size 10 kbp (Figure 7A). Overall, no large region stood out in terms of densely packed high F_{ST} values. Therefore, we inspected more closely 500 Mp lengths to check whether any moderate regions of genomic differences could be detected (Figure 7B–F).

Most of the genome showed low differentiation and there was no evidence for clustering of any outliers, which would be indicative of a large inversion or region of divergence. No outliers were seen for the analysis based on a window size of 1 kbp, either (Figure 8).

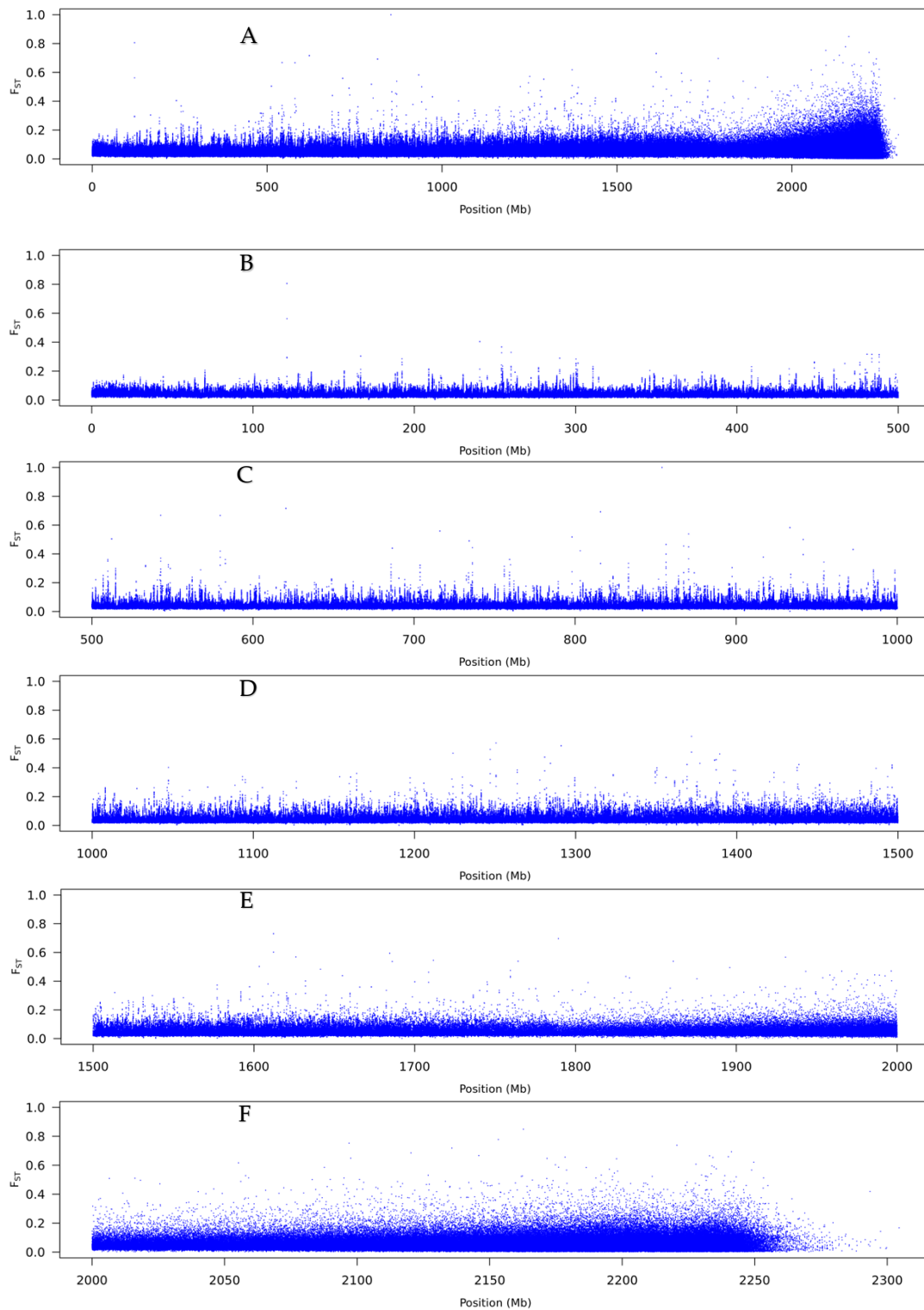


Figure 7. F_{ST} summary from the comparison of the pools of anaflexistylous and cataflexistylous individuals of *Alpinia nigra*, for window sizes of 10 kbp and step sizes of 1 kbp. (A) Genome-wide F_{ST} scan, (B) scan across the first 500 Mbp, (C) scan from 500 to 1000 Mbp, (D) scan from 1000 to 1500 Mbp, (E) scan from 1500 to 2000 Mbp and (F) scan from 2000 Mbp to end. The scaffolds are ordered by length starting from the largest.

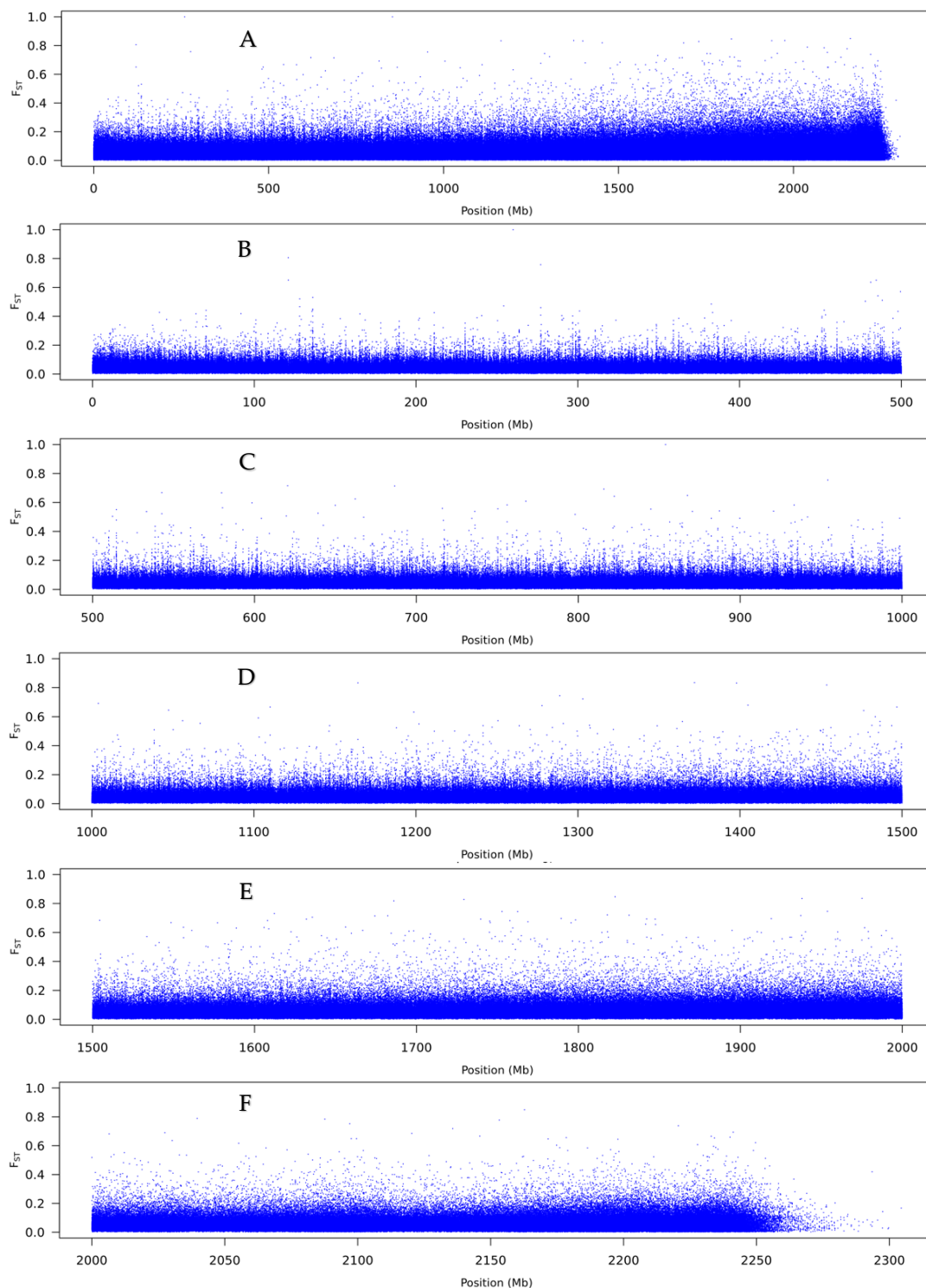


Figure 8. F_{ST} summary from the comparison of the pools of anaflexistylous and cataflexistylous individuals of *Alpinia nigra*, for window size of 1 kbp and step size of 1 kbp. (A) Genome-wide F_{ST} scan, (B) scan across the first 500 Mbp, (C) scan from 500 to 1000 Mbp, (D) scan from 1000 to 1500 Mbp, (E) scan from 1500 to 2000 Mbp and (F) scan from 2000 Mbp to end. The scaffolds are ordered by length starting from the largest.

4. Discussion

In this study, we aimed to find the genomic regions that govern flexistylous, a complex floral trait found across several genera in the Zingiberaceae. We produced a highly complete draft genome assembly for *Alpinia nigra*, which served as a reference genome for

comparative genomic analysis. Using a Pool-seq approach, we carried out a genome-wide sliding window F_{ST} scan of the pools of ana-morph and cata-morph. We found no clear genetic difference between the two morphs, and relatively few windows had high F_{ST} values. On comparing the mapping depths, no regions stood out with a 1:2 mapping depth ratio as would be expected in the case of hemizygoty in one morph. Here, we discuss the utility of the draft genome produced and the possible reasons why it might be challenging to detect the genomic region that governs this trait.

4.1. Draft Genome Assembly of *Alpinia nigra*

The metrics of our *Alpinia nigra* draft assembly indicate that this genome is of high completeness. It has a high contiguity (N50 = 48 kbp) for a genome assembled using short-reads, that is comparable or better than other assemblies produced using DISCOVAR *de novo*, such as those of malaria mosquito (N50 = 22.3 kb) [58], Guinea yam (N50 = 2.7–3.3 kb) [59], grass pea (N50 = 5.7 kbp) [60] and wheat (N50 = 16.7 kbp) [61]. Given a 91% BUSCO completeness score and high completeness indicated by the k-mer comparison plot (Figure 4), it is likely to have a good representation of genic regions. K-mer and Smudgeplot analyses further indicated that this species is diploid in nature, and not a tetraploid, as has always been suggested for *Alpinia*.

Within the Zingiberales, the only other genomes sequenced so far are of *Curcuma longa* (turmeric) [30] and banana species such as *Musa acuminata*, *M. balbisiana*, *M. itinerans* and *M. schizocarpa* [31–34]. As banana is one of the most economically important fruits in the world, these genomes have served as an invaluable resource for breeding strategies for new cultivars with improved yield [33]. The availability of these high-quality genomes has also provided an insight into the evolution of monocot lineages [31]. Similarly, many species in the Zingiberaceae are of agricultural importance such as ginger (*Zingiber officinale*), cardamom (*Elettaria cardamomum*) and turmeric (*Curcuma longa*). Therefore, the availability of a draft genome from the ginger family can help in breeding efforts to improve yields and disease resistance, and in investigating the potential of wild relatives [62] and the biosynthetic pathways of essential compounds [63]. The draft genome presented here will also be a useful resource for ecological, systematic and evolutionary studies in the Zingiberaceae. While this genome will be useful for a range of sequence-based analyses, future work on annotation, informed by transcriptomic sequencing, will greatly improve the utility of this genome. While our assembly has high completeness, the contiguity of structural variants will be limited and scaffolds will be broken by genomic repeats longer than the read length [64]. Here, long-read sequencing such as PacBio HiFi, combined with long-range scaffolding with technologies such as HiC or BioNano genome mapping, could be used in the future to improve genome contiguity and assembly.

4.2. The Genetic Basis of Flexistyly

We used a Pool-seq approach to investigate the allele frequency differences between ana- and cataflexistyly individuals of *Alpinia nigra*. Pool-seq has been used to identify the genes underlying many traits such as the dwarfism gene in watermelon [65], to identify genomic regions that are involved in the maintenance of ecotypic variation in *Mimulus guttatus* [66] and to compare genomic diversity in populations of non-model species such as brown trout [67].

Flexistyly is a complex phenotype, and we thus hypothesised it may be controlled by multiple genes (as is, for instance, distyly). These genes are presumably tightly linked as the morph types are always found in a 1:1 ratio within a population [16–18,21,23,68]. These loci would be expected to reside in a region of reduced recombination and inherited together, such as a large polymorphic chromosomal inversion or a supergene [9,69,70]. However, the sliding window F_{ST} scans did not reveal a signal consistent with this expectation, with a large accumulation of F_{ST} values present only at the tail end of the genome where the reads do not map so well to the short sequence scaffolds. As such, there is no clear evidence for a large chromosomal inversion sheltering multiple different loci against recombination.

However, it is also plausible that a complex genomic region such as an inversion could be overlooked if it is difficult to assemble due to active transposable elements often found at inversion breakpoints [71].

In genomic analysis of heterostylous *Primula* spp. there was a notable coverage difference in the thrum-specific region as compared to the pin, indicating hemizyosity [11,72]. However, such a pattern was not seen upon comparing the read depths of the ana-morph and cata-morph pools of *Alpinia*. This implies that the genes that underlie flexistyly may not be hemizygous, though this prediction should be re-evaluated with improved genome assemblies. Notably, during our initial coverage analyses, we did find a number of regions with increased coverage in one sequencing pool. However all such contigs showed strong sequence similarity to common contaminants. This highlights the need for careful interpretation of sequencing coverage analyses when testing for hemizyosity.

While many genomic analyses have detected major structural variants underlying important phenotypes [73], if our findings hold with an improved reference genome, they would point towards flexistyly being governed by a single small genomic region or a single gene. Such a signal would be difficult to detect in our analyses. While we did find a handful of nearly fixed SNP differences between pools, these SNPs appeared in isolation, rather than being clustered as would be expected from highly divergent alleles. However, single SNPs or smaller structural variants can have major impact on phenotype [74,75] and should be tested in future studies. An alternative route to pooled sequencing for finding the gene(s) for flexistyly would be to perform comparative transcriptomics, particularly tissue-specific sequencing of stilar tissue.

Understanding the phenotype in more detail can also give clues to the genes that underlie flexistyly. All the flowers of the same morph type that open on the same day are synchronous [16–18]. Studies have revealed that this stilar movement is dependent on factors such as light and auxin transport [76–78]. When exposed to light before anthesis, the style of the anaflexistylous morph curves above the anther instead of downwards. The second curvature was not affected by light in either morph type hence both the curvatures of cata-morph and the second curvature of the ana-morph are controlled endogenously [77,78]. It is presumed that light activates auxin transport which in turn promotes stilar movement in *Alpinia*. The effects of these factors are different for different morph types [76] and the same might be the case for anther dehiscence. Therefore, an annotated genome would be useful to see if any SNP differences between morphs reside in loci involved in light regulation.

4.3. The Evolution of Flexistyly

There are several evolutionary explanations why flexistyly may be present in *Alpinia* and other closely related genera. *Alpinia* flowers are hermaphroditic and are open only for a day. This differs from other taxa with similar zygomorphic structure, such as orchids, where the flowers are open for days or even weeks [79]. The risk of selfing is likely to be higher in gingers in which pollination must be achieved in a much shorter period than in most orchids. Therefore, this stilar movement may have evolved to prevent self-pollination and geitonogamy as well as sexual interference within the flower [17,20,80]. As flexistyly has been reported from many genera within Alpinieae, this dimorphism may have evolved in the common ancestor of the tribe or may have evolved independently several times [22]. As more evolutionary research on the genomic basis of flexistyly is performed, the origins of this trait will become clear.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/genes12091297/s1>, Table S1: Collection details of the anaflexistylous and cataflexistylous tissue samples of *Alpinia nigra* from Pakke Tiger Reserve, Arunachal Pradesh, India. File S1: Python script for computation of windowed mapping depths. Figure S1: Blobplot for the *Alpinia nigra* reads using Blobtools v1.1.1. Figure S2: A comparison of the paired mapping depth per scaffold of the anaflexistylous and cataflexistylous pool of the 200,000 longest contigs ordered by length.

Author Contributions: Conceptualization, A.D.T.; formal analysis, S.R. and H.B.; writing—original draft preparation, S.R.; writing—review and editing, all authors. All authors have read and agreed to the published version of the manuscript.

Funding: S.R. is supported by The Darwin Trust of Edinburgh. The RBGE is supported by the Scottish Government’s Rural and Environmental Science and Analytical Services Division. V.G. is supported by the Science and Engineering Research Board (SERB), the Ministry of Human Resource Development (MHRD) and National Geographic. A.D.T. is supported by a Natural Environment Research Council Fellowship (NE/L011336/1). Fieldwork was funded by the Davis Expedition Fund and Heredity Fieldwork Grant.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The whole genome sequencing data and Pool-seq data of the anaflexistylous and cataflexistylous morphs have been deposited under NCBI BioProject ID PRJNA756475.

Acknowledgments: S.R. would like to thank Susnata Salony, Pakke Tiger Reserve forest personnel and members of Nature Conservation Foundation for help with fieldwork, and Sujai Kumar for access to Club Ashworth Cluster. All the authors would like to thank Roberta Bergero for the invitation and two anonymous reviewers for useful comments.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- Barrett, S.C.H. Understanding plant reproductive diversity. *Philos. Trans. R. Soc. B Biol. Sci.* **2010**, *365*, 99–109. [[CrossRef](#)] [[PubMed](#)]
- Waser, N.M.; Ollerton, J. Geographical variation in diversity and specificity of pollination systems. In *Plant-Pollinator Interactions: From Specialization to Generalization*; Ollerton, J., Johnson, S.D., Hingston, A.B., Eds.; University of Chicago Press: Chicago, IL, USA, 2006; pp. 283–308.
- Barrett, S.C.H. Darwin’s legacy: The forms, function and sexual diversity of flowers. *Philos. Trans. R. Soc. B Biol. Sci.* **2010**, *365*, 351–368. [[CrossRef](#)]
- Barrett, S.C.H. Sexual interference of the floral kind. *Heredity* **2002**, *88*, 154–159. [[CrossRef](#)] [[PubMed](#)]
- Charlesworth, D.; Charlesworth, B. Inbreeding depression and its evolutionary consequences. *Annu. Rev. Ecol. Syst.* **1987**, *18*, 237–268. [[CrossRef](#)]
- Ornduff, R. Historical Perspectives on Heterostyly. In *Evolution and Function of Heterostyly. Monographs on Theoretical and Applied Genetics*; Barrett, S.C.H., Ed.; Springer: Berlin/Heidelberg, Germany, 1992; Volume 15, pp. 31–39.
- Darwin, C. *The Different Forms of Flowers on Plants of the Same Species*; John Murray: London, UK, 1877.
- Lewis, D.; Jones, D.A. The genetics of heterostyly. In *Evolution and Function of Heterostyly*; Barrett, S.C.H., Ed.; Springer: New York, NY, USA, 1992; pp. 129–150.
- Charlesworth, D. The status of supergenes in the 21st century: Recombination suppression in Batesian mimicry and sex chromosomes and other complex adaptations. *Evol. Appl.* **2016**, *9*, 74–90. [[CrossRef](#)]
- Huu, C.N.; Kappel, C.; Keller, B.; Sicard, A.; Takebayashi, Y.; Breuninger, H.; Nowak, M.D.; Baurle, I.; Himmelbach, A.; Burkart, M. Presence versus absence of CYP734A50 underlies the style-length dimorphism in primroses. *E Life* **2016**, *5*, e17956. [[CrossRef](#)] [[PubMed](#)]
- Li, J.; Cocker, J.M.; Wright, J.; Webster, M.A.; McMullan, M.; Dyer, S.; Swarbreck, D.; Caccamo, M.; van Oosterhout, C.; Gilmartin, P.M. Genetic architecture and evolution of the S locus supergene in *Primula vulgaris*. *Nat. Plants* **2016**, *2*, 16188. [[CrossRef](#)]
- Huu, C.N.; Keller, B.; Conti, E.; Kappel, C.; Lenhard, M. Supergene evolution via stepwise duplications and neofunctionalization of a floral-organ identity gene. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 23148–23157. [[CrossRef](#)] [[PubMed](#)]
- Yasui, Y.; Mori, M.; Aii, J.; Abe, T.; Matsumoto, D.; Sato, S.; Hayashi, Y.; Ohnishi, O.; Ota, T. S-LOCUS EARLY FLOWERING 3 is exclusively present in the genomes of short-styled buckwheat plants that exhibit heteromorphic self-incompatibility. *PLoS ONE* **2012**, *7*, e31264. [[CrossRef](#)]
- Ushijima, K.; Nakano, R.; Bando, M.; Shigezane, Y.; Ikeda, K.; Namba, Y.; Kume, S.; Kitabata, T.; Mori, H.; Kubo, Y. Isolation of the floral morph-related genes in heterostylous flax (*Linum grandiflorum*): The genetic polymorphism and the transcriptional and post-transcriptional regulations of the S locus. *Plant J.* **2012**, *69*, 317–331. [[CrossRef](#)]
- Shore, J.S.; Hamam, H.J.; Chafe, P.D.J.; Labonne, J.D.J.; Henning, P.M.; McCubbin, A.G. The long and short of the S-locus in *Turnera* (Passifloraceae). *New Phytol.* **2019**, *224*, 1316–1329. [[CrossRef](#)]
- Li, Q.J.; Xu, Z.F.; Kress, W.J.; Xia, Y.M.; Zhang, L.; Deng, X.B.; Gao, J.Y.; Bai, Z.L. Flexible style that encourages outcrossing. *Nature* **2001**, *410*, 432. [[CrossRef](#)] [[PubMed](#)]

17. Zhang, L.; Li, Q.J.; Deng, X.B.; Ren, P.Y.; Gao, J.Y. Reproductive biology of *Alpinia blepharocalyx* (Zingiberaceae): Another example of flexistyly. *Plant Syst. Evol.* **2003**, *241*, 67–76. [CrossRef]
18. Li, Q.J.; Kress, W.J.; Xu, Z.-F.; Xia, Y.-M.; Zhang, L.; Deng, X.-B.; Gao, J.Y. Mating system and stigmatic behaviour during flowering of *Alpinia kwangsiensis* (Zingiberaceae). *Plant Syst. Evol.* **2002**, *232*, 123–132. [CrossRef]
19. Renner, S.S. How common is heterodichogamy? *Trends Ecol. Evol.* **2001**, *16*, 595–597. [CrossRef]
20. Sun, S.; Gao, J.Y.; Liao, W.J.; Li, Q.J.; Zhang, D.Y. Adaptive significance of flexistyly in *Alpinia blepharocalyx* (Zingiberaceae): A hand-pollination experiment. *Ann. Bot.* **2007**, *99*, 661–666. [CrossRef] [PubMed]
21. Cui, X.; Wei, R.; Huang, R. A study on the breeding system of *Amomum tsao-ko*. In *Proceedings of the Second Symposium on the Family Zingiberaceae*; Wu, T.L., Wu, Q.G., Chen, Z.Y., Eds.; South China Institute of Botany: Guangzhou, China, 1996; pp. 288–296.
22. Kress, W.J.; Liu, A.-Z.; Newman, M.; Li, Q.J. The molecular phylogeny of *Alpinia* (Zingiberaceae): A complex and polyphyletic genus of ginger. *Am. J. Bot.* **2005**, *92*, 167–178. [CrossRef]
23. Jia, X.-C.; Li, J.; Lu, G.-H.; Wang, Y.-Q. The concrete evidence of flexistyly in *Plagiostachys*: Pollination biology of a wild ginger on Hainan Island, China. *Ecol. Evol.* **2015**, *5*, 5364–5371. [CrossRef]
24. Leong-Škorničková, J.; Tran, H.Đ.; Nguyen, Q.B.; Šída, O. *Siliquamomum alcornone* (Zingiberaceae: Alpinioideae), a new species from central Vietnam. *Gard. Bull.* **2014**, *66*, 39–46.
25. Poulsen, A.D.; Phonsena, P. Morphological variation and distribution of the useful ginger *Etilingera pavieana* (Zingiberaceae). *Nord. J. Bot.* **2017**, *35*, 467–475. [CrossRef]
26. Gleeson, S.K. Heterodichogamy in walnuts: Inheritance and stable ratios. *Evolution* **1982**, *36*, 892. [CrossRef]
27. Thompson, T.E.; Romberg, L.D. Inheritance of heterodichogamy in pecan. *J. Hered.* **1985**, *76*, 456–458. [CrossRef]
28. Futschik, A.; Schlötterer, C. The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. *Genetics* **2010**, *186*, 207–218. [CrossRef] [PubMed]
29. Schlötterer, C.; Tobler, R.; Kofler, R.; Nolte, V. Sequencing pools of individuals—Mining genome-wide polymorphism data without big funding. *Nat. Rev. Genet.* **2014**, *15*, 749–763. [CrossRef] [PubMed]
30. Chakraborty, A.; Mahajan, S.; Jaiswal, S.K.; Sharma, V.K. Genome sequencing of turmeric provides evolutionary insights into its medicinal properties. *bioRxiv* **2020**. [CrossRef]
31. D’Hont, A.; Denoeud, F.; Aury, J.-M.; Baurens, F.-C.; Carreel, F.; Garsmeur, O.; Noel, B.; Bocs, S.; Droc, G.; Rouard, M.; et al. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* **2012**, *488*, 213–217. [CrossRef]
32. Davey, M.W.; Gudimella, R.; Harikrishna, J.A.; Sin, L.W.; Khalid, N.; Keulemans, J. A draft *Musa balbisiana* genome sequence for molecular genetics in polyploid, inter- and intra-specific *Musa* hybrids. *BMC Genom.* **2013**, *14*, 683. [CrossRef]
33. Wu, W.; Yang, Y.-L.; He, W.-M.; Rouard, M.; Li, W.-M.; Xu, M.; Roux, N.; Ge, X.-J. Whole genome sequencing of a banana wild relative *Musa itinerans* provides insights into lineage-specific diversification of the *Musa* genus. *Sci. Rep.* **2016**, *6*, 31586. [CrossRef]
34. Belser, C.; Istace, B.; Denis, E.; Dubarry, M.; Baurens, F.-C.; Falentin, C.; Genete, M.; Berrabah, W.; Chèvre, A.-M.; Delourme, R.; et al. Chromosome-scale assemblies of plant genomes using nanopore long reads and optical maps. *Nat. Plants* **2018**, *4*, 879–887. [CrossRef] [PubMed]
35. Lowry, D.B.; Willis, J.H. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.* **2010**, *8*, e1000500. [CrossRef] [PubMed]
36. Fang, Z.; Pyhäjärvi, T.; Weber, A.L.; Dawe, R.K.; Glaubitz, J.C.; González, J.d.J.S.; Ross-Ibarra, C.; Doebley, J.; Morrell, P.L.; Ross-Ibarra, J. Megabase-scale inversion polymorphism in the wild ancestor of maize. *Genetics* **2012**, *191*, 883–894. [CrossRef]
37. Hooper, D.M.; Price, T.D. Chromosomal inversion differences correlate with range overlap in passerine birds. *Nat. Ecol. Evol.* **2017**, *1*, 1526–1534. [CrossRef] [PubMed]
38. Joron, M.; Frezal, L.; Jones, R.T.; Chamberlain, N.L.; Lee, S.F.; Haag, C.R.; Whibley, A.; Becuwe, M.; Baxter, S.W.; Ferguson, L.; et al. Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* **2011**, *477*, 203–206. [CrossRef] [PubMed]
39. Smith, R.M. *Alpinia* (Zingiberaceae): A proposed new infrageneric classification. *Edinb. J. Bot.* **1990**, *47*, 1–75. [CrossRef]
40. Ma, X.-N.; Xie, C.-L.; Miao, Z.; Yang, Q.; Yang, X.-W. An overview of chemical constituents from *Alpinia* species in the last six decades. *RSC Adv.* **2017**, *7*, 14114. [CrossRef]
41. Raghavan, T.S.; Venkatasubban, K.R. Cytological studies in the family Zingiberaceae with special reference to chromosome number and Cyto-Taxonomy. *Proc. Indian Acad. Sci.-Sect. B* **1943**, *17*, 118–132. [CrossRef]
42. Chen, Z. Cytotaxonomy of the tribe Alpineae. In *Proceedings of the Second Symposium on the Family Zingiberaceae*; Zhongshan University Press: Guangzhou, China, 1996; pp. 112–121.
43. Basak, S.; Rangan, L. New record of nuclear DNA amounts of some Zingiberaceae species from North east India. *Data Br.* **2018**, *17*, 66–70. [CrossRef] [PubMed]
44. Wilkie, P.; Poulsen, A.D.; Harris, D.; Forrest, L.L. The collection and storage of plant material for DNA extraction: The Teabag Method. *Gard. Bull. Singap.* **2013**, *65*, 231–234.
45. Doyle, J.J.; Doyle, J.L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **1987**, *19*, 11–15.
46. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010. Available online: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> (accessed on 2 June 2021).
47. Mapleson, D.; Garcia Accinelli, G.; Kettleborough, G.; Wright, J.; Clavijo, B.J. KAT: A K-mer analysis toolkit to quality control NGS datasets and genome assemblies. *Bioinformatics* **2016**, *33*, 574–576. [CrossRef]

48. Ranallo-Benavidez, T.R.; Jaron, K.S.; Schatz, M.C. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat. Commun.* **2020**, *11*, 1432. [[CrossRef](#)]
49. Weisenfeld, N.I.; Yin, S.; Sharpe, T.; Lau, B.; Hegarty, R.; Holmes, L.; Sogoloff, B.; Tabbaa, D.; Williams, L.; Russ, C.; et al. Comprehensive variation discovery in single human genomes. *Nat. Genet.* **2014**, *46*, 1350–1355. [[CrossRef](#)] [[PubMed](#)]
50. Laetsch, D.R.; Blaxter, M.L. BlobTools: Interrogation of genome assemblies. *F1000Research* **2017**, *6*, 1287. [[CrossRef](#)]
51. Li, H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
52. Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **2015**, *12*, 59–60. [[CrossRef](#)] [[PubMed](#)]
53. Simão, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **2015**, *31*, 3210–3212. [[CrossRef](#)]
54. Gurevich, A.; Saveliev, V.; Vyahhi, N.; Tesler, G. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **2013**, *29*, 1072–1075. [[CrossRef](#)]
55. Danecek, P.; Bonfield, J.K.; Liddle, J.; Marshall, J.; Ohan, V.; Pollard, M.O.; Whitwham, A.; Keane, T.; McCarthy, S.A.; Davies, R.M.; et al. Twelve years of SAMtools and BCFtools. *Gigascience* **2021**, *10*, giab008. [[CrossRef](#)]
56. Kofler, R.; Pandey, R.V.; Schlötterer, C. PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* **2011**, *27*, 3435–3436. [[CrossRef](#)]
57. Doležel, J.; Bartoš, J.; Voglmayr, H.; Greilhuber, J. Letter to the editor. *Cytom. Part A* **2003**, *51A*, 127–128. [[CrossRef](#)]
58. Love, R.R.; Weisenfeld, N.I.; Jaffe, D.B.; Besansky, N.J.; Neafsey, D.E. Evaluation of DISCOVAR *de novo* using a mosquito sample for cost-effective short-read genome assembly. *BMC Genom.* **2016**, *17*, 187. [[CrossRef](#)]
59. Tamiru, M.; Natsume, S.; Takagi, H.; White, B.; Yaegashi, H.; Shimizu, M.; Yoshida, K.; Uemura, A.; Oikawa, K.; Abe, A.; et al. Genome sequencing of the staple food crop white Guinea yam enables the development of a molecular marker for sex determination. *BMC Biol.* **2017**, *15*, 86. [[CrossRef](#)]
60. Emmrich, P.M.F.; Sarkar, A.; Njaci, I.; Kaithakottil, G.G.; Ellis, N.; Moore, C.; Edwards, A.; Heavens, D.; Waite, D.; Cheema, J.; et al. A draft genome of grass pea (*Lathyrus sativus*), a resilient diploid legume. *bioRxiv* **2020**. [[CrossRef](#)]
61. Clavijo, B.J.; Venturini, L.; Schudoma, C.; Accinelli, G.G.; Kaithakottil, G.; Wright, J.; Borrill, P.; Kettleborough, G.; Heavens, D.; Chapman, H.; et al. An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. *Genome Res.* **2017**, *27*, 885–896. [[CrossRef](#)] [[PubMed](#)]
62. Edwards, D.; Batley, J. Plant genome sequencing: Applications for crop improvement. *Plant Biotechnol. J.* **2010**, *8*, 2–9. [[CrossRef](#)]
63. Lichman, B.R.; Godden, G.T.; Hamilton, J.P.; Palmer, L.; Kamileen, M.O.; Zhao, D.; Vaillancourt, B.; Wood, J.C.; Sun, M.; Kinser, T.J.; et al. The evolutionary origins of the cat attractant nepetalactone in catnip. *Sci. Adv.* **2020**, *6*, eaba0721. [[CrossRef](#)] [[PubMed](#)]
64. Thomma, B.P.H.J.; Seidl, M.F.; Shi-Kunne, X.; Cook, D.E.; Bolton, M.D.; van Kan, J.A.L.; Faino, L. Mind the gap; seven reasons to close fragmented genome assemblies. *Fungal Genet. Biol.* **2016**, *90*, 24–30. [[CrossRef](#)] [[PubMed](#)]
65. Dong, W.; Wu, D.; Li, G.; Wu, D.; Wang, Z. Next-generation sequencing from bulked segregant analysis identifies a dwarfism gene in watermelon. *Sci. Rep.* **2018**, *8*, 2908. [[CrossRef](#)] [[PubMed](#)]
66. Gould, B.A.; Chen, Y.; Lowry, D.B. Pooled ecotype sequencing reveals candidate genetic mechanisms for adaptive differentiation and reproductive isolation. *Mol. Ecol.* **2017**, *26*, 163–177. [[CrossRef](#)] [[PubMed](#)]
67. Kurland, S.; Wheat, C.W.; Paz Celorio Mancera, M.; Kutschera, V.E.; Hill, J.; Andersson, A.; Rubin, C.; Andersson, L.; Ryman, N.; Laikre, L. Exploring a Pool-seq-only approach for gaining population genomic insights in nonmodel species. *Ecol. Evol.* **2019**, *9*, 11448–11463. [[CrossRef](#)]
68. Ren, P.Y.; Liu, M.; Li, Q.J. An example of flexistylly in a wild cardamom species (*Amomum maximum*, (Zingiberaceae)). *Plant Syst. Evol.* **2007**, *267*, 147–154. [[CrossRef](#)]
69. Jay, P.; Whibley, A.; Frézal, L.; Rodríguez de Cara, M.Á.; Nowell, R.W.; Mallet, J.; Dasmahapatra, K.K.; Joron, M. Supergene evolution triggered by the introgression of a chromosomal inversion. *Curr. Biol.* **2018**, *28*, 1839–1845.e3. [[CrossRef](#)]
70. Twyford, A.D.; Friedman, J. Adaptive divergence in the monkey flower *Mimulus guttatus* is maintained by a chromosomal inversion. *Evolution* **2015**, *69*, 1476–1486. [[CrossRef](#)]
71. Delprat, A.; Negre, B.; Puig, M.; Ruiz, A. The transposon *Galileo* generates natural chromosomal inversions in *Drosophila* by ectopic recombination. *PLoS ONE* **2009**, *4*, e7883. [[CrossRef](#)]
72. Cocker, J.M.; Wright, J.; Li, J.; Swarbreck, D.; Dyer, S.; Caccamo, M.; Gilmartin, P.M. *Primula vulgaris* (primrose) genome assembly, annotation and gene expression, with comparative genomics on the heterostylly supergene. *Sci. Rep.* **2018**, *8*, 17942. [[CrossRef](#)]
73. Wellenreuther, M.; Bernatchez, L. Eco-evolutionary genomics of chromosomal inversions. *Trends Ecol. Evol.* **2018**, *33*, 427–440. [[CrossRef](#)]
74. Thompson, N.F.; Anderson, E.C.; Clemento, A.J.; Campbell, M.A.; Pearse, D.E.; Hearsey, J.W.; Kinziger, A.P.; Garza, J.C. A complex phenotype in salmon controlled by a simple change in migratory timing. *Science* **2020**, *370*, 609–613. [[CrossRef](#)]
75. Wang, J.; Ding, J.; Tan, B.; Robinson, K.M.; Michelson, I.H.; Johansson, A.; Nystedt, B.; Scofield, D.G.; Nilsson, O.; Jansson, S.; et al. A major locus controls local adaptation and adaptive life history variation in a perennial plant. *Genome Biol.* **2018**, *19*, 72. [[CrossRef](#)]
76. Luo, Y.-L.; Bi, T.-J.; Li, D.; Su, Z.-L.; Tao, C.; Luo, Y.-J.; Li, Q.-J. Effects of indole-3-acetic acid and auxin transport inhibitors on the style curvature of three *Alpinia* species (Zingiberaceae). *Acta Physiol. Plant.* **2012**, *34*, 2019–2025. [[CrossRef](#)]

-
77. Luo, Y.-L.; Li, Q.-J. Effects of light and low temperature on the reciprocal style curvature of flexistylous *Alpinia* species (Zingiberaceae). *Acta Physiol. Plant.* **2010**, *32*, 1229–1234. [[CrossRef](#)]
 78. Su, Z.-L.; Cui, X.-L.; Li, L.; Li, Q.-J.; Luo, Y.-L. Similar behavior, different mechanisms: The research of the style bending of *Alpinia* species. *Acta Physiol. Plant.* **2017**, *39*, 143. [[CrossRef](#)]
 79. Primack, R.B. Longevity of Individual Flowers. *Annu. Rev. Ecol. Syst.* **1985**, *16*, 15–37. [[CrossRef](#)]
 80. Sun, S.; Zhang, D.Y.; Ives, A.R.; Li, Q.J. Why do stigmas move in a flexistylous plant? *J. Ecol. Biol.* **2011**, *24*, 497–504. [[CrossRef](#)]