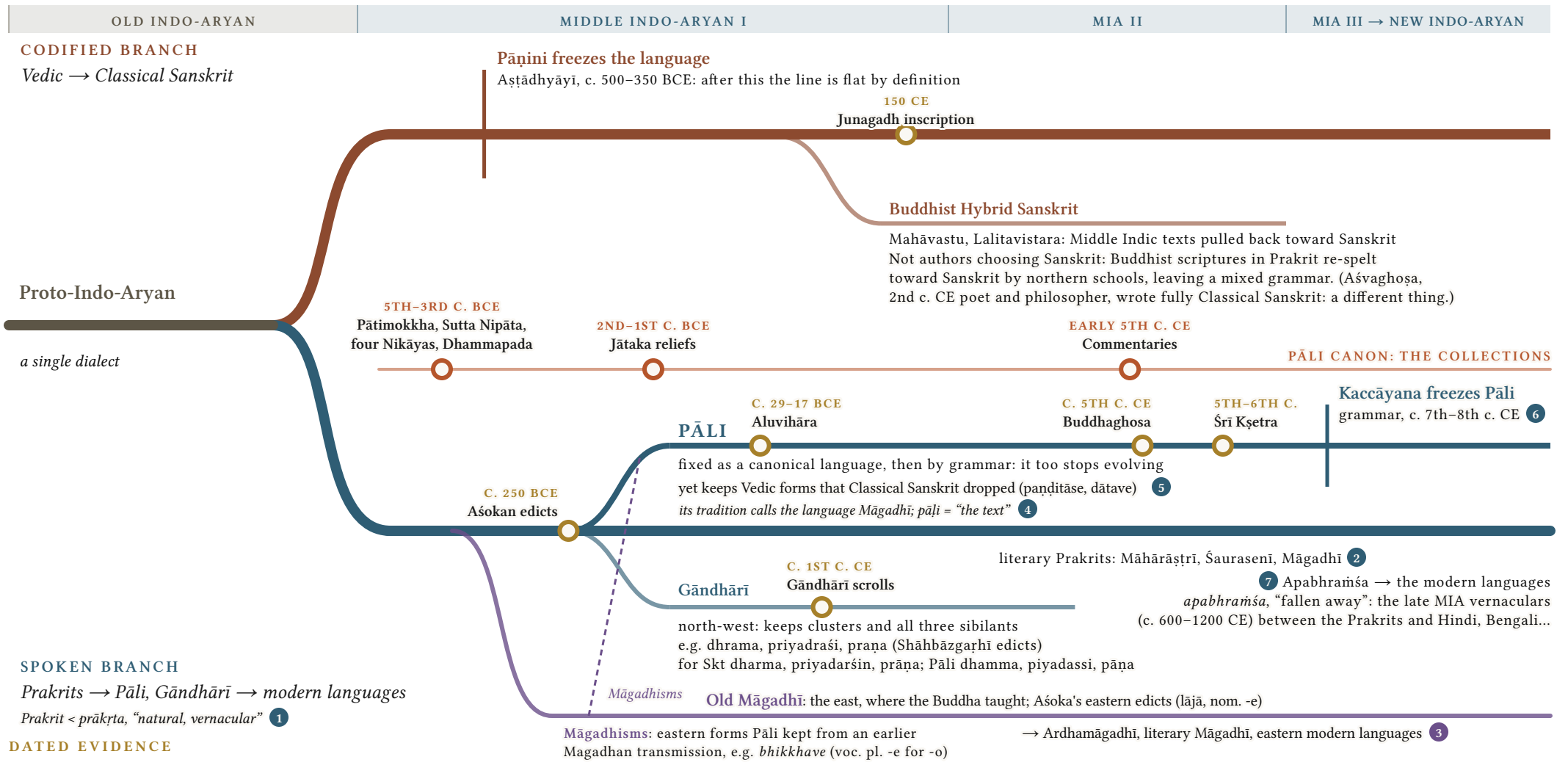
Development of Pāli and Sanskrit from Indo-Aryan origins



C. 250 BCE
Aśokan edicts
The first dated Indic text of any kind, and the anchor everything else hangs on. Nom. sg. -o in the west, -e in the east.
Texts: none quoted, but the Bhabru edict names seven Dhamma texts, usually identified with suttas in the Snp, MN and AN.

C. 29–17 BCE
Aluvihāra
Pāli canon committed to writing (Dīpavaṁsa 20.20–21; Mahāvaṁsa 33.100–101). A tradition, not an artefact.

C. 1ST C. CE
Gāndhārī scrolls
Oldest surviving Buddhist manuscripts: the British Library Kharoṣṭhī birch-bark fragments.
Texts: a Dharmapada, the Rhinoceros Sūtra (= Snp 1.3), Āgama sūtras, avadānas; the Bajaur scrolls add a Prātimokṣa.

150 CE
Junagadh inscription
Rudradāman I, on the Girnār rock at Junagadh, Gujarat (W. India), the same rock as Aśoka's edicts. First long inscription in more or less standard Sanskrit: Sanskrit returns as a public language.
Texts: none Buddhist; a royal eulogy for repairing the Sudarśana lake dam.

C. 5TH C. CE
Buddhaghosa
The commentaries fix Theravāda exegesis, and with it much of how Pāli is read.

MID 5TH–MID 6TH C. CE
Śrī Kṣetra plates
Oldest surviving Pāli text of any kind: twenty gold leaves from the Khin Ba mound at Śrī Kṣetra, the Pyu city near Prome (Pyay), Burma, dug 1926–27. Eight canonical excerpts: praises of the Buddha, dependent origination, factors of awakening.

PĀLI CANON: WHEN THE COLLECTIONS TOOK SHAPE

5TH–4TH C. BCE
Pātimokkha
The monks' rules, by tradition recited by Upāli at the First Council. The same code is in every surviving Vinaya, so its core predates the first schisms.
Finds: a Gāndhārī Prātimokṣa on birch bark in the Bajaur scrolls (NW Pakistan, 2nd c. CE), in a version matching no known school.

5TH–3RD C. BCE
Sutta Nipāta
Discourses in verse. Its Aṭṭhakavagga and Pārāyanavagga are among the oldest Buddhist texts: the canon comments on them in the Niddesa.
Finds: Aśoka's Bhabru edict names a Muni-gāthā, Moneya-sūte and Upatisa-pasine, usually identified with Snp suttas; a Gāndhārī Rhinoceros Sūtra (= Snp 1.3) in the British Library scrolls.

5TH–3RD C. BCE
Four Nikāyas
Dīgha, Majjhima, Saṁyutta, Aṅguttara. Largely fixed before Aśoka; written down in Sri Lanka in the 1st c. BCE. An early core with later additions.
Finds: Bhabru names a Lāghulovāda (cf. MN 61) and Anāgata-bhayāni (cf. AN 5.77–80); Gāndhārī Saṁyukta sūtras in the Senior scrolls (mid-2nd c. CE); Nikāya formulas on the Śrī Kṣetra gold.

5TH–3RD C. BCE
Dhammapada
Verses ascribed to the Buddha, gathered as a collection by about the 3rd c. BCE. Other schools' versions share a core of some 330 verses.
Finds: the Khotan Dharmapada, Gāndhārī birch bark found near Khotan, Xinjiang (1st–2nd c. CE); fragments in the British Library scrolls; Sanskrit Udānavarga leaves from Turfan.

2ND–1ST C. BCE
Jātaka
Only the verses are canonical; some go back to the 5th c. BCE. The prose was fixed in a commentary of about the 5th c. CE.
Finds: the Bhārhut railing (Madhya Pradesh, c. 125–100 BCE) carves jātaka scenes with Brāhmī labels naming them; the Sāñcī gateways show the Vessantara and Chaddanta stories.

EARLY 5TH C. CE
Commentaries
Buddhaghosa renders the old Sinhala commentaries of the Mahāvihāra into Pāli under King Mahānāma (c. 410–434). The Khuddaka commentaries (Dhammapāla and others) are later.
Finds: no early manuscripts; anchored by a document instead: the Chinese translation of the Vinaya commentary (*Shanjian lü piposha*, 488/489 CE).

Dates are conventional and in most cases contested; see the accompanying notes.

Scale is not uniform: 600 BCE – 600 CE is expanded. Dashed rules mark the breaks.

common ancestor · codified / Sanskrit · spoken / Middle Indic · dated evidence · eastern / Māgadhī · eastern forms carried into Pāli · Pāli canon collections

NOTES — *numbered markers in the diagram refer to these*

1 Prakrits
prākṛta, "natural, ordinary", as against *saṁskṛta*, "refined, put together": the vernacular Middle Indo-Aryan languages that grew out of Old Indo-Aryan speech. In the broad sense every MIA language is a Prakrit, Pāli and Gāndhārī included; in the narrow sense the word means the literary Prakrits of note 2.

2 Literary Prakrits
Prakrits standardised as languages of literature rather than of speech, and codified by grammarians (Vararuci; Hemacandra, 12th c.). In Sanskrit drama kings and brahmins speak Sanskrit, women Śaurasenī, low-status characters Māgadhī; songs are in Māhārāṣṭrī. The Jain canon is in Ardhamāgadhī. All are MIA II: intervocalic stops are lost (Skt *loka* → Prakrit *loa*, "world").

3 Māgadhī: two things with one name
(a) The eastern dialect of Magadha, where the Buddha taught. Aśoka's eastern edicts show it: *lājā* for *rājā*, nom. sg. *-e* for *-o*. Its line runs on into Ardhamāgadhī, the literary Māgadhī Prakrit, and the eastern modern languages (Bengali, Odia, Maithili, Magahi).
(b) The Theravāda name for Pāli itself: the commentaries call the language of the canon Māgadhī, the "root language" (*mūlabhāsā*). Linguistically Pāli looks western (*-o*, *r*), but it passed through eastern hands and keeps Māgadhisms such as *bhikkhave* (dashed purple line).

4 "Pāli" means "the text"
In the commentaries *pāḷi* means "line, row", hence "the canonical text" as against its commentary (*aṭṭhakathā*): *pāḷiyaṁ*, "in the text". So in what the tradition calls Māgadhī, *pāḷi* names the text, not the language. "Pāli" as the language's name comes from *pāḷibhāsā*, "the language of the texts", a later usage.

5 Pāli forms older than Sanskrit
Some Pāli forms, mostly in old verse (Sutta Nipāta), are older than Classical Sanskrit: Vedic has them, and Pāṇini's standard dropped them.

Pāli	Vedic	Classical Skt	English
paṇḍitāse	*paṇḍitāsaḥ**	*paṇḍitāḥ*	the wise
nom. pl. masc., a-stem · Snp 4.11			
asmase	*smasi*	*smaḥ*	we are
1st pl. present of √as "be" · Snp 3.9 (kevalino 'smase); RV smasi			
devehi	*devebhiḥ*	*devaiḥ*	by the gods
instr. pl.; verse keeps -ebhi (dantebhi, MN 36)			
dātave	*dātave*	*dātum*	to give
infinitive; Vedic -tave is a dative · Snp 2.7			
dissare	*dadṛśre*	*dṛśyante*	they are seen
3rd pl. middle in -re · AN 5.55; Vedic perfect, RV 1.24.10			

* Vedic ending, as in *devāsaḥ*; the word *paṇḍita* itself is post-Vedic.

devaiḥ has no *bh* for *devehi* (← *-ebhi*) to come from: Pāli descends from a Vedic-shaped ancestor, not from Classical Sanskrit. The spoken branch split off before the freeze.

6 When Pāli froze
Pāli had no single Pāṇini moment; it set in stages. The canon was fixed by recitation, written down at Aluvihāra, read through Buddhaghosa's commentaries (5th c.), then codified by Kaccāyana's grammar, c. 7th–8th c. CE (Pind; tradition ascribes it to the disciple Mahākaccāna). It is modelled chiefly on the Kātantra, not on Pāṇini. Later: Moggallāna (12th c.), Aggavaṁsa's Saddanīti (1154).

7 New Indo-Aryan
MIA and NIA in the header are Middle and New Indo-Aryan. NIA is the modern stage, from about 1000 CE: Hindi-Urdu, Bengali, Marathi, Gujarati, Punjabi, Odia, Nepali and many more. They grew out of the regional Apabhraṁśas, not out of Sanskrit or Pāli (Sinhala took its own route from early Middle Indic). Both lent them learned words borrowed whole (*tatsama*) beside words inherited through Prakrit (*tadbhava*): Hindi *karm* beside *kām* < Prakrit *kamma*.